Phononic bandgap nano-acoustic cavity with ultralong phonon lifetime

Gregory S. MacCabe,[1, 2, *] Hengjiang Ren,[1, 2, *] Jie Luo,[1, 2] Justin D. Cohen,[1, 2]
Hengyun Zhou,[1, 2, †] Alp Sipahigil,[1, 2] Mohammad Mirhosseini,[1, 2] and Oskar Painter[1, 2]

[1]*Kavli Nanoscience Institute, California Institute of Technology, Pasadena, California 91125, USA*
[2]*Institute for Quantum Information and Matter and Thomas J. Watson, Sr.,*
Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA
(Dated: January 15, 2019)

We present measurements at millikelvin temperatures of the microwave-frequency acoustic properties of a crystalline silicon nanobeam cavity incorporating a phononic bandgap clamping structure for acoustic confinement. Utilizing pulsed laser light to excite a co-localized optical mode of the nanobeam cavity, we measure the dynamics of cavity acoustic modes with single-phonon sensitivity. Energy ringdown measurements for the fundamental 5 GHz acoustic mode of the cavity shows an exponential increase in phonon lifetime versus number of periods in the phononic bandgap shield, increasing up to $\tau_{\mathrm{ph,0}} \approx 1.5$ seconds. This ultralong lifetime, corresponding to an effective phonon propagation length of several kilometers, is found to be consistent with damping from non-resonant two-level system defects on the surface of the silicon device. Potential applications of these ultra-coherent nanoscale mechanical resonators range from tests of various collapse models of quantum mechanics to miniature quantum memory elements in hybrid superconducting quantum circuits.

In optics, geometric structuring at the nanoscale has become a powerful method for modifying the electromagnetic properties of a bulk material, leading to metamaterials [1] capable of manipulating light in unprecedented ways [2]. In the most extreme case, photonic bandgaps can emerge in which light is forbidden from propagating, dramatically altering the emission of light from within such materials [3–5]. More recently, a similar phononics revolution [6] in the engineering of acoustic waves has led to a variety of new devices, including acoustic cloaks that can shield objects from observation [7], thermal crystals for controlling the flow of heat [8, 9], optomechanical crystals that couple photons and phonons via radiation pressure [10], and phononic topological insulators whose protected edge states can transport acoustic waves with minimal scattering [11, 12].

Phononic bandgap structures, similar to their electromagnetic counterparts, can be used to modify the emission or scattering of phonons. These ideas have recently been explored in quantum optomechanics [13–17] and electromechanics [18] experiments to greatly reduce the mechanical coupling to the thermal environment through acoustic radiation. At ultrasonic frequencies and below, one can combine phononic bandgap clamping with a form of 'dissipation dilution' in high stress films [19] to realize quality (Q) factors in excess of 10^8 in two-dimensional nanomembranes [16] and approaching 10^9 in one-dimensional strain-engineered nanobeams [17]. At higher, microwave frequencies the benefit of stress-loading of the film fades as local strain energy dominates [17] and one is left once again to deal with intrinsic material absorption.

To date, far less attention has been paid to the impact of geometry and phononic bandgaps on acoustic material absorption [20, 21]. Fundamental limits to sound absorption in solids are known to result from the anharmonicity of the host crystal lattice [22–24]. At low temperatures T, in the Landau-Rumer regime ($\omega \tau_{\mathrm{th}} \gg 1$) where the thermal phonon relaxation rate (τ_{th}^{-1}) is much smaller than the acoustic frequency (ω), a quantum model of three-phonon scattering can be used to describe phonon-phonon mixing that results in damping and thermalization of acoustic modes [22, 23]. Landau-Rumer damping scales approximately as T^{α}, where $\alpha \approx 4$ depends upon the phonon dispersion and density of states (DOS) [23]. At the very lowest lattice temperatures ($\lesssim 10$ K), where Landau-Rumer damping has dropped off, a residual damping emerges due to material defects. These two-level system (TLS) defects [25], typically found in amorphous materials, correspond to a pair of nearly degenerate local arrangements of atoms in the solid which can have both an electric and an acoustic transition dipole, and couple to both electric and strain fields. Recent theoretical analysis shows that TLS interactions with acoustic waves can be dramatically altered in a structured material [20].

Here we explore the limits of acoustic damping and coherence of a microwave acoustic nanocavity with a phononic crystal shield that possesses a wide bandgap for all polarizations of acoustic waves. Our nanocavity, formed from an optomechanical crystal (OMC) nanobeam resonator [10, 14], supports an acoustic breathing mode at $\omega_{\mathrm{m}}/2\pi \approx 5$ GHz and a co-localized optical resonant mode at $\omega_{\mathrm{c}}/2\pi \approx 195$ THz ($\lambda_{\mathrm{c}} \approx 1550$ nm) which allows us to excite and readout mechanical motion using radiation pressure from a pulsed laser source. This minimally invasive pulsed measurement technique avoids a slew of parasitic damping effects − typically associated with electrode materials and mechanical contact [26], or probe fields for continuous readout − and

---

* These authors contributed equally to this work.

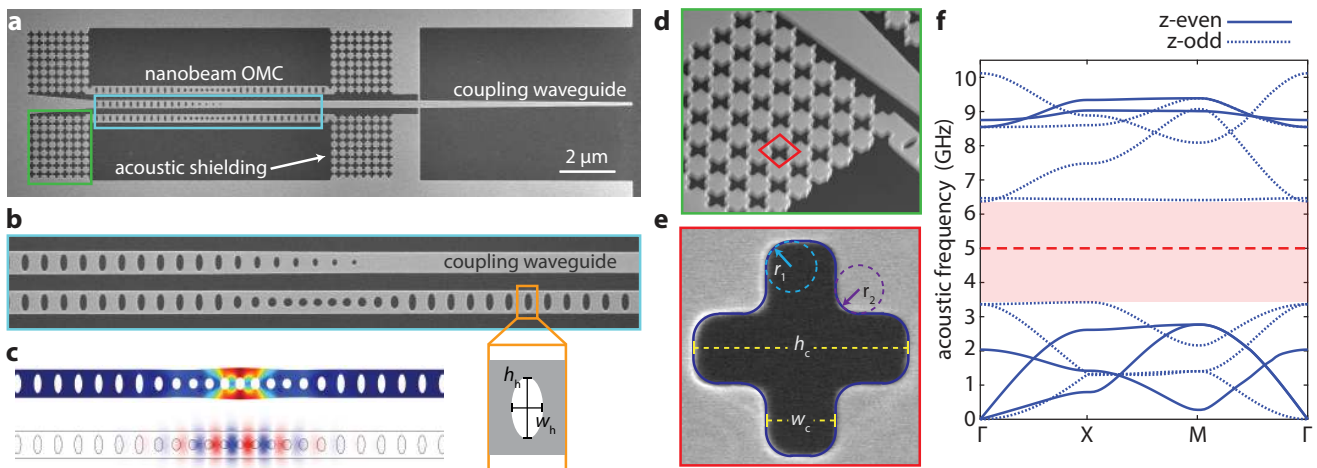† Current Address: Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

FIG. 1. **Nanobeam optomechanical crystal and phononic shield design. a**, Scanning electron microscope (SEM) image of a full nanobeam optomechanical crystal (OMC) device fabricated on SOI with $N = 7$ periods of acoustic shielding. A central coupling waveguide allows for fibre-to-chip optical coupling as well as side-coupling to individual nanobeam OMC cavities. **b**, SEM image of an individual nanobeam OMC and the coupling waveguide, with enlarged illustration of an individual unit cell in the end-mirror portion of the nanobeam. **c**, FEM simulations of the mechanical (top; total displacement) and optical (bottom; transverse electric field) modes of interest in the nanobeam. Distortion of the mechanical displacement profile is exaggerated for clarity. **d**, SEM image showing the nanobeam clamping geometry. **e**, SEM image of an individual unit cell of the cross-crystal acoustic shield. The dashed lines show fitted geometric parameters used in simulation, including cross height ($h_c = 474$ nm), cross width ($w_c = 164$ nm), inner fillet radius (r_1), and outer fillet radius (r_2). **f**, Simulated acoustic band structure of the realized cross-crystal shield unit cell, with the full acoustic bandgap highlighted in pink. Solid (dotted) lines correspond to modes of even (odd) symmetry in the direction normal to the plane of the unit cell. The dashed red line indicates the mechanical breathing-mode frequency at $\omega_m/2\pi = 5.0$ GHz.

allows for the sensitive measurement of motion at the single phonon level [27]. The results of acoustic ring-down measurements at millikelvin temperatures show that damping due to radiation is effectively suppressed by the phononic shield, with breathing mode quality factors reaching $Q = 4.9 \times 10^{10}$, corresponding to an unprecedented frequency-Q product of f-$Q = 2.6 \times 10^{20}$. The temperature and amplitude dependence of the residual acoustic damping is consistent with relaxation damping of non-resonant TLS, modeling of which indicates that not only does the phononic bandgap directly eliminate the acoustic radiation of the breathing mode but it also reduces the phonon damping of TLS in the host material.

The devices studied in this work are fabricated from the 220 nm device layer of a silicon-on-insulator (SOI) microchip. Details of the fabrication process are provided in App. A. In Figs. 1(a-b) we show scanning electron microscope images of a single fabricated device, which consists of a coupling optical waveguide, the nanobeam OMC cavities that support both the microwave acoustic and optical resonant modes, and the acoustic shield that connects the cavity to the surrounding chip substrate. Fig. 1(c) shows finite-element method (FEM) simulations of the microwave acoustic breathing mode and fundamental optical mode of the nanobeam cavity. We use the on-chip coupling waveguide to direct laser light to the nanobeam OMC cavities. A pair of cavities with slightly different optical mode frequencies are

evanescently coupled to each waveguide. An integrated photonic crystal back mirror in the waveguide allows for optical measurement in a reflection geometry. The design of the OMC cavities, detailed in Ref. [14], uses a tapering of the etched hole size and shape in the nanobeam to provide strong localization and overlap of the breathing mode and the fundamental optical mode, resulting in a vacuum optomechanical coupling rate [28] between photons and phonons of $g_0/2\pi \approx 1$ MHz.

In order to minimize mechanical clamping losses, the nanobeam is anchored to the Si bulk with a periodic cross structure which is designed to have a complete phononic bandgap at the breathing mode frequency [14]. Through tuning of the cross height h_c and width w_c (c.f., Figs. 1(d-e)), bandgaps as wide as ~ 3 GHz can be achieved as shown in Fig. 1. We analyze SEM images of realized structures to provide accurate structure dimensions for our FEM models, and in particular we include in our modeling a filleting of the inner and outer corners (r_1 and r_2 in Fig. 1e) of the crosses arising from technical limitations of the patterning of the structure. To investigate the efficacy of the acoustic shielding we fabricate and characterize arrays of devices with a scaling of the cross period number from $N_{shield} = 0$ to 10, with all other design parameters held constant. FEM modeling indicates (see App. B) that the addition of the cross shield provides significant protection against nanometer-scale disorder which is inherently introduced during device fabri-
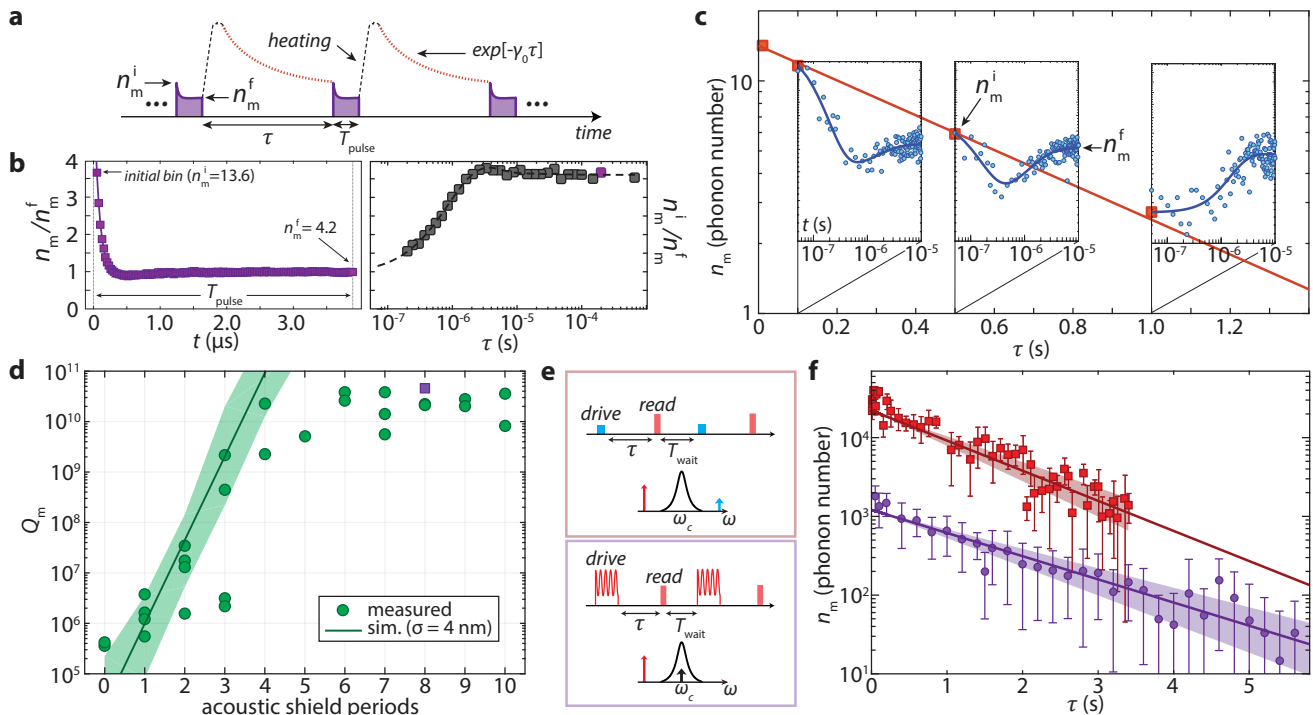
FIG. 2. **Ringdown measurements of the acoustic breathing mode.** **a**, Illustration of the ringdown measurement performed using a red-detuned ($\Delta = +\omega_m$) pulsed laser for excitation and readout. **b**, Normalized phonon occupancy measured during (left) and after (right) the laser readout pulse ($n_c = 569$; optomechanical back-action rate $\gamma_{OM}/2\pi = 1.07$ MHz) for a 6-shield device (device B). Squares are measured data points. Solid and dashed lines are a best fit to the dynamical model of the hot bath (see App. D). The displayed pulse-on-state plot (left) corresponds to a delay of $\tau = 200$ μs, with $n_m^i = 4.2$ and $n_m^f = 13.6$ phonons. **c**, Ringdown measurements of a 7-shield device (device C) for readout pulse amplitude of $n_c = 320$. The series of inset panels show the measured (and fit; solid blue curve) anti-Stokes signal during the optical pulse at a series of pulse delays. **d**, Plot of the measured breathing mode Q-factor versus number of acoustic shield periods N_{shield}. The solid green line is a fit to the corresponding simulated radiation-limited Q-factor (see App. B) for devices with standard deviation (SD) $\sigma = 4$ nm disorder in hole position and size, similar to the value measured from device SEM image analysis. The shaded green region corresponding to the range of simulated Q values (ensemble size 10) within one SD of the mean. The square purple data points represents the measured Q in (**f**). **e**, Acoustic excitation is performed coherently by using either a blue-detuned pump (upper diagram) to drive the breathing mode into self-oscillation, or using an RF-modulated red-detuned pump [29] (lower diagram). See App. E for details of the coherent excitation and readout parameters. **f**, Ringdown measurements performed on an eight-shield device (device D) at large phonon amplitude. For blue-detuned driving (red squares) the fit decay rate is $\gamma_0/2\pi = (0.122 \pm 0.020)$ Hz. For modulated-pump driving (purple circles) the fit decay rate is $\gamma_0/2\pi = (0.108 \pm 0.006)$ Hz. The error bars are 90% confidence intervals of the measured values of n_m^i. The shaded regions are the 90% confidence intervals for the exponential fit curves.

cation.

Optical measurements of the acoustic properties of the OMC cavity are performed at millikelvin temperatures in a dilution refrigerator. The sample containing an array of different OMC devices is mounted directly on a copper mount attached to the mixing chamber stage of the fridge, and a single lensed optical-fiber is positioned with a 3-axis stage to couple light into and out of each device [27]. In a first set of measurements of acoustic energy damping, we employ a single pulsed laser scheme to perform both excitation and readout of the breathing mode. In this scenario, depicted in Fig. 2(a), the laser frequency (ω_l) is tuned to the red motional sideband of the OMC cavity optical resonance, $\Delta \equiv \omega_c - \omega_l \approx +\omega_m$, and is pulsed on for a duration T_{pulse} and then off for a

variable time τ. This produces a periodic train of photon pulses due to anti-Stokes scattering of the probe laser which are on-resonance with the optical cavity. The anti-Stokes scattered photons are filtered from the probe laser and sent to a single photon detector producing a photon count rate proportional to the number of phonons in the acoustic resonator (see Apps. C for details of the measurement set-up and phonon number calibration methods).

We display in Fig. 2(b) a typical readout signal, showing the normalized phonon occupancy during and immediately after the application of a 4 μs pulse. The initial optomechanical back-action cooling of the acoustic breathing mode is followed by a slower turn-on of heating of the mode during the pulse. After the pulse, with

the back-action cooling turned off, a transient heating of the acoustic mode occurs over several microseconds. The parasitic heating is attributable to very weak optical absorption of the probe pulse in the Si cavity which produces a hot bath coupled to the breathing mode [27]. Here we use the transient heating of the acoustic mode to perform ringdown measurements of the stored phonon number. A phenomenological model of the dynamics of the induced damping (γ_p) and effective occupancy (n_p) of the hot bath (see App. D) allows us to fit the anti-Stokes decay signal. Plotting the initial mode occupancy at the beginning of the fit readout pulse (n_m^i) versus delay time τ between pulses (c.f., Fig. 2(a)), we plot the ringdown of the stored phonon number in the the breathing mode as displayed in Fig. 2(c) for a device with $N_\mathrm{shield} = 7$.

Performing a series of ringdown measurements over a range of devices with varying N_shield, and fitting an exponential decay curve to each ringdown we produce the Q-factor plot in Fig. 2(d). We observe an initial trend in Q-factor versus shield number which rises on average exponentially with each additional shield period, and then saturates for $N_\mathrm{shield} \geq 5$ to $Q_\mathrm{m} \gtrsim 10^{10}$. As indicated in Fig. 2(c) these Q values correspond to ringdown of small, near-single-phonon level amplitudes. We also perform ringdown measurements at high phonon amplitude using two additional methods displayed schematically in Fig. 2(e) and described in detail in App. E. These methods use two laser tones to selectively excite the acoustic breathing mode using a $\times 1000$ weaker excitation and readout optical pulse amplitude ($n_\mathrm{c} \lesssim 0.3$). The measured ringdown curves, displayed in Fig. 2(f), show the decay from initial phonon occupancies of 10^3-10^4 of an 8-shield device (device D; square purple data point in Fig. 2(d)). The two methods yield similar breathing mode energy decay rates of $\gamma_0/2\pi = 0.108$ Hz and 0.122 Hz, the smaller of which corresponds to a Q-value of $Q_\mathrm{m} = 4.92^{+0.39}_{-0.26} \times 10^{10}$ and a phonon lifetime of $\tau_\mathrm{ph,0} = 1.47^{+0.09}_{-0.08}$ s. Comparing all three excitation methods with widely varying optical-absorption-heating and phonon amplitude, we consistently measure $Q_\mathrm{m} \gtrsim 10^{10}$ for devices with $N_\mathrm{shield} \geq 5$.

In order to understand the origin of the residual damping for large N_shield we also measured the temperature dependence of the energy damping rate, breathing mode frequency, and full width at half maximum (FWHM) linewidth of the breathing mode for the highest Q 8-shield device (device D). In Fig. 3(a) we plot the energy damping rate which shows an approximately linear rise in temperature up to $T_\mathrm{f} \approx 100$ mK, and then a much faster $\sim (T_\mathrm{f})^{2.4}$ rise in the damping. Using the two-tone coherent excitation method [29], we plot in Fig. 3(b) the measured breathing mode acoustic spectrum at $T_\mathrm{f} = 7$ mK and pump power $n_\mathrm{c} = 0.1$. The top plot shows rapid spectral scans (40 ms per scan) in which the probe frequency is swept across the acoustic resonance. These rapid scans show a jittering acoustic line with a roughly $\Delta\omega_\mathrm{m}/2\pi \approx 1$ kHz linewidth, consistent with the predicted magnitude of optical back-action
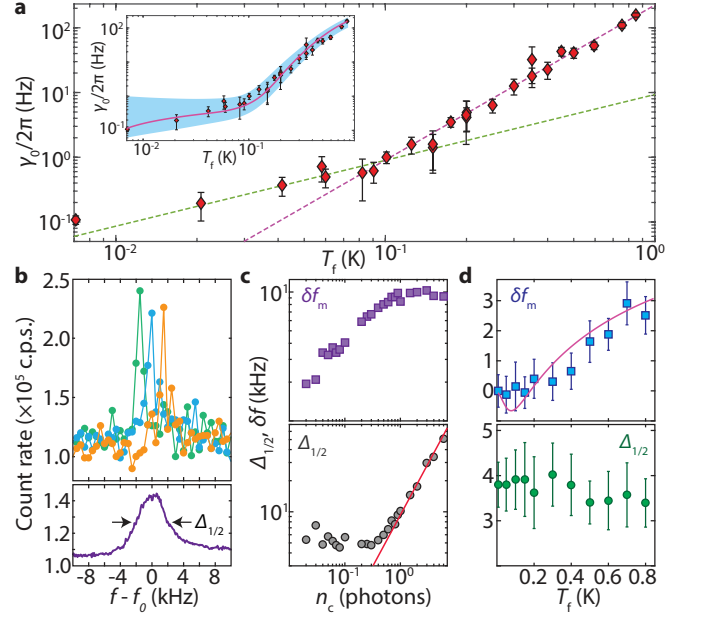

FIG. 3. **Temperature dependence of acoustic damping, frequency, and frequency jitter.** a, Plot of the measured breathing mode energy damping rate, $\gamma_0/2\pi$, as a function of fridge temperature (T_f). Dashed green (magenta) curve is a fit with temperature dependence $\gamma_0 \sim T_\mathrm{f}^{1.01}$ ($\gamma_0 \sim T_\mathrm{f}^{2.39}$). Error bars are 90% confidence intervals of the exponential fit to measured ring down curves. Inset: Plot of measured damping data with estimated energy damping from a TLS model (see App. G). The shaded blue region corresponds to the standard deviation of $\log(\gamma_0/2\pi)$ for 100 different random TLS distributions. **b**, Two-tone coherent spectroscopy signal. Upper plot: three individual spectrum of rapid frequency sweeps with a frequency step size of 500 Hz and dwell time of 1 ms (RBW ≈ 0.5 kHz). Lower plot: average spectrum of rapid scan spectra taken over minutes, showing broadened acoustic response with FWHM linewidth of $\Delta_{1/2}/2\pi = 4.05$ kHz. The large on-resonance response corresponds to an estimated optomechanical cooperativity of $C \equiv \gamma_\mathrm{OM}/(\gamma_0 + \gamma_\mathrm{p}) \gtrsim 1.1$, consistent with the predicted magnitude of back-action damping $\gamma_\mathrm{OM}/2\pi \approx 817$ Hz and bath-induced damping $\gamma_\mathrm{p}/2\pi \approx 120$ Hz at the measurement pump power level $n_\mathrm{c} = 0.1$. **c**, Breathing mode resonance frequency shift δf and ensemble average FWHM-linewidth $\Delta_{1/2}$ as a function of pump photon number n_c. Solid red curve is fit to back-action limited linewidth, yielding $g_0/2\pi = 1.15$ MHz. **d**, Measured δf (upper plot) and $\Delta_{1/2}$ (lower plot) versus T_f. Error bars are 90% confidence intervals from Voigt fit to measured spectra. Data presented in (**a-d**) are for device D. The solid magenta curves in the inset of (**a**) and top panel of (**d**) corresponds to simulations of a single random TLS distribution.

($\gamma_\mathrm{OM}/2\pi \approx 820$ Hz) and hot bath damping ($\gamma_\mathrm{p}/2\pi \approx 120$ Hz) at the $n_\mathrm{c} = 0.1$ measurement power. An ensemble average of these scans, taken over several minutes, yields a broadened and reduced contrast acoustic line of FWHM $\Delta_{1/2}/2\pi = 4.05$ kHz.

Note that in Fig. 3(b) we are measuring the acous-

tic line with the laser light on, as opposed to the ring-down measurements of Fig. 2 in which the laser light is off. Lowering the optical pump power to reduce back-action and absorption-induced damping limits further the already low signal-to-noise ratio, and scanning more slowly begins to introduce frequency jitter into the measured line. As such, we can only bound the intrinsic low temperature coherence time of the breathing mode to $\tau_{\rm coh,0} \gtrsim 2/\Delta\omega_{\rm m} \approx 0.3$ ms. Further information can, however, be gleaned by measuring the linewidth and center frequency of the ensemble averaged spectrum as a function of $n_{\rm c}$ (Fig. 3(c)) and $T_{\rm f}$ (Fig. 3(d)). The width of the frequency jitter spectrum, averaged over minutes, is roughly independent of optical pump power and temperature down to the lowest measurable pump powers ($n_{\rm c} = 0.02$) and up to $T_{\rm f} = 800$ mK. The center frequency, on the other hand, shifts up in frequency with both temperature and optical power. The frequency shift versus $T_{\rm f}$ is consistent with the frequency shift versus $n_{\rm c}$ if the hot bath temperature (see App. D) is used as a proxy for the fridge temperature.

Estimates of the magnitude of Landau-Rumer damping of the breathing mode (see App. F) indicate that 3-phonon scattering in Si is far too weak at $T_{\rm f} \lesssim 1$ K to explain the measured damping. Analysis of the interactions of TLS with the localized acoustic modes of the confined geometry of the OMC cavity structure, however, show that TLS interactions can explain all of the observed breathing mode behavior. In this analysis, detailed in App. G, FEM simulation is used to find the frequencies and radiation-limited damping rates of the acoustic quasi-normal modes of the OMC cavity structure. An estimate of the spectral density of TLS within the breathing mode volume ($V_m \approx 0.11$ $(\mu m)^3$) is ascertained from estimated surface oxide (~ 0.25 nm [30]) and etch-damage (~ 15 nm [31]) layer thicknesses in the Si device, and bulk TLS density found in amorphous materials [25, 32]. Using the resulting effective spectral density of interacting TLS, $n_{0,m} \approx 20$ GHz^{-1}, and average TLS transverse and longitudinal deformation potentials of $\bar{M} \approx 0.04$ eV and $\bar{D} \approx 3.2$ eV, respectively, yields breathing mode damping and frequency shifts which are in excellent agreement with the measured data (see Fig. 3). The estimated level of frequency jitter is also found in agreement with the measured value, assuming all TLS are being pumped via the same optical absorption that drives the hot bath.

Several key observations can be drawn from the TLS damping modeling. The first is that the typical T^3 dependence of TLS relaxation damping of acoustic waves is dependent on the phonon bath DOS into which the TLS decay [20, 21]. In the OMC cavity the phonon DOS is strongly modified from a three-dimensional bulk material. This directly results in the observed weak temperature dependence of the acoustic damping for $T_{\rm f} \lesssim 100$ mK, where the thermally acitvated TLS interact resonantly with an approximately one-dimensional phonon

DOS. A second point to note is that the TLS resonant damping is strongly suppressed due to the phononic bandgap surrounding the OMC cavity. Estimates of the phonon-induced spontaneous decay rate of TLS in the bandgap is on the order of Hz; combined with the discrete number of TLS in the small mode volume of the breathing mode, acoustic energy from the breathing mode cannot escape via resonant coupling to TLS. The observed lack of saturation of the breathing mode energy damping with either temperature or phonon amplitude is further evidence that non-resonant relaxation damping − due to dispersive coupling to TLS − is dominant [25]. Finally, the small average number of estimated TLS in V_m which are thermally activated at the lowest temperatures (~ 2), leads to significant variation in the simulated TLS relaxation damping at $T_{\rm f} \sim 10$ mK (see shaded blue region of the inset to Fig. 3(a)). This is consistent with the observed fluctuations from device-to-device in the low-temperature $Q_{\rm m}$ for devices with $N_{\rm shield} > 5$ (see Fig. 2(d)).

Utilizing the advanced methods of nanofabrication and cavity optomechanics has provided a new toolkit to explore quantum acoustodynamics in solid-state materials. Continued studies of the behavior of TLS in similar engineered nanostructures to the OMC cavity of this work may lead to, among other things, new approaches to modifying the behavior of quasi-particles in superconductors [33], mitigating decoherence in superconducting [34, 35] and color center [36, 37] qubits, and even new coherent TLS-based qubit states in strong coupling with an acoustic cavity [38]. The extremely small motional mass ($m_{\rm eff} = 136$ fg [14]) and narrow linewidth of the OMC cavity also make it ideal for precision mass sensing [39] and in exploring limits to alternative quantum collapse models [40]. Perhaps most intriguing is the possibility of creating a hybrid quantum architecture consisting of acoustic and superconducting quantum circuits [41–48], where the small scale, reduced cross-talk, and ultralong coherence time of quantum acoustic devices may provide significant improvements in connectivity and performance of current quantum hardware.

ACKNOWLEDGMENTS

[1] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, Phys. Rev. Lett. **84**, 4184 (2000).

[2] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, 2008).

[3] E. Yablonovitch, T. Gmitter, J. P. Harbison, and R. Bhat, Appl. Phys. Lett. **51**, 2222 (1987).

[4] S. John and J. Wang, Physical Review B **43**, 12772 (1991).

[5] M. Fujita, S. Takahashi, Y. Tanaka, T. Asano, and S. Noda, Science **308**, 1296 (2005).

[6] M. Maldovan, Nature **503**, 209 (2013).

[7] S. Zhang, C. Xia, and N. Fang, Phys. Rev. Lett. **106**, 024301 (2011).

[8] S. Narayana and Y. Sato, Phys. Rev. Lett. **108**, 214303 (2012).

[9] M. Maldovan, Phys. Rev. Lett. **110**, 025902 (2013).

[10] M. Eichenfield, R. Camacho, J. Chan, K. J. Vahala, and O. Painter, Nature **459**, 550 (2009).

[11] C. Brendel, V. Peano, O. Painter, and F. Marquardt, Phys. Rev. B **97**, 020102(R) (2018).

[12] J. Cha, K. W. Kim, and C. Daraio, Nature **564**, 229 (2018).

[13] T. P. M. Alegre, A. Safavi-Naeini, M. Winger, and O. Painter, Opt. Express **19**, 5658 (2011).

[14] J. Chan, A. H. Safavi-Naeini, J. T. Hill, S. Meenehan, and O. Painter, Appl. Phys. Lett. **101**, 081115 (2012).

[15] P.-L. Yu, K. Cicak, N. S. Kampel, Y. Tsaturyan, T. P. Purdy, R. W. Simmonds, and C. A. Regal, App. Phys. Lett. **104**, 023510 (2014).

[16] A. B. Y. Tsaturyan, E. S. Polzik, and A. Schliesser, Nature Nanotech. **12**, 776 (2017).

[17] A. H. Ghadimi, S. A. Fedorov, N. J. Engelsen, M. J. Bereyhi, R. Schilling, D. J. Wilson, and T. J. Kippenberg, Science **360**, 764 (2018).

[18] M. Kalaee, M. Mirhosseini, P. B. Dieterle, M. Peruzzo, J. M. Fink, and O. Painter, arXiv:1808.04874 (2018).

[19] Q. P. Unterreithmeier, T. Faust, and J. P. Kotthaus, Phys. Rev. Lett. **105**, 027205 (2010).

[20] R. O. Behunin, F. Intravaia, and P. T. Rakich, Phys. Rev. B **93**, 224110 (2016).

[21] B. D. Hauer, P. H. Kim, C. D. F. Souris, and J. P. Davis, Phys. Rev. B **98**, 214303 (2018).

[22] L. Landau and G. Rumer, Phys. Z. Sowjet. **11** (1937).

[23] G. P. Srivastava, *The Physics of Phonons* (Taylor & Francis Group, 1990).

[24] T. O. Woodruff and H. Ehrenreich, Phys. Rev. **123**, 1553 (1961).

[25] W. A. Phillips, Rep. Prog. Phys. **50**, 1657 (1987).

[26] S. Galliou, M. Goryachev, R. Bourquin, P. Abbe, J. P. Aubry, and M. E. Tobar, Scientific Reports **3**, 2132 (2013).

[27] S. M. Meenehan, J. D. Cohen, G. S. MacCabe, F. Marsili, M. D. Shaw, and O. Painter, Phys. Rev. X **5**, 041002 (2015).

[28] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Cavity Optomechanics: Nano- and Micromechanical Resonators Interacting with Light* (Springer-Verlag, 2014).

[29] A. H. Safavi-Naeini, T. P. M. Alegre, J. Chan, M. Eichenfield, M. Winger, Q. Lin, J. T. Hill, D. Chang, and O. Painter, Nature **472**, 69 (2011).

[30] N. Yabumoto, M. Oshima, O. Michikami, and S. Yoshii, Jpn. J. Appl. Phys. **20**, 893 (1981).

[31] G. S. Oehrlein, Materials Science and Engineering: B **4**, 441 (1989).

[32] R. N. Kleiman, G. Agnolet, and D. J. Bishop, Phys. Rev. Lett. **59**, 2079 (1987).

[33] K. Rostem, P. J. de Visser, and E. J. Wollack, Phys. Rev. B **98**, 014522 (2018).

[34] J. Gao, *The Physics of Superconducting Microwave Resonators*, Ph.D. thesis, California Institute of Technology (2008).

[35] J. M. Martinis and A. Megrant, arXiv:1410.5793 (2014).

[36] Y.-I. Sohn, S. Meesala, B. Pingault, H. A. Atikian, J. Holzgrafe, M. Gündoğan, C. Stavrakas, M. J. Stanley, A. Sipahigil, J. Choi, M. Zhang, J. L. Pacheco, J. Abraham, E. Bielejec, M. D. Lukin, M. Atatüre, and M. Lončar, Nat. Commun. **9**, 2012 (2018).

[37] T. Astner, J. Gugler, A. Angerer, S. Wald, S. Putz, N. J. Mauser, M. Trupke, H. Sumiya, S. Onoda, J. Isoya, J. Schmiedmayer, P. Mohn, and J. Majer, Nat. Mat. **17**, 313 (2018).

[38] T. Ramos, V. Sudhir, K. Stannigel, P. Zoller, and T. J. Kippenberg, Phys. Rev. Lett. **110**, 193602 (2013).

[39] M. S. Hanay, S. I. Kelber, C. D. OConnell, P. Mulvaney, J. E. Sader, and M. L. Roukes, Nature Nanotech. **10**, 339 (2015).

[40] S. Nimmrichter, K. Hornberger, and K. Hammerer, Phys. Rev. Lett. **113**, 020405 (2014).

[41] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169 (2013).

[42] J.-M. Pirkkalainen, S. U. Cho, J. Li, G. S. Paraoanu, P. J. Hakonen, and M. A. Sillanpää, Nature **494**, 211 (2013).

[43] M. V. Gustafsson, T. Aref, A. F. Kockum, M. K. Ekstrm, G. Johansson, and P. Delsing, Science **346**, 207 (2014).

[44] Y. Chu, P. Kharel, W. H. Renninger, L. D. Burkhart, L. Frunzio, P. T. Rakich, and R. J. Schoelkopf, Science **358**, 199 (2017).

[45] R. Manenti, A. F. Kockum, A. Patterson, T. Behrle, J. Rahamim, G. Tancredi, F. Nori, and P. J. Leek, Nat. Commun. **8**, 975 (2017).

[46] K. J. Satzinger, Y. P. Zhong, H.-S. Chang, G. A. Peairs, A. Bienfait, M.-H. Chou, A. Y. Cleland, C. R. Conner, E. Dumur, J. Grebe, I. Gutierrez, B. H. November, R. G. Povey, S. J. Whiteley, D. D. Awschalom, D. I. Schuster, and A. N. Cleland, arXiv:1804.07308 (2018).

[47] B. A. Moores, L. R. Sletten, J. J. Viennot, and K. Lehnert, Phys. Rev. Lett. **120**, 227701 (2018).

[48] P. Arrangoiz-Arriola, E. A. Wollack, M. Pechal, J. D. Witmer, J. T. Hill, and A. H. Safavi-Naeini, Phys. Rev. X **8**, 031007 (2018).

[49] J. D. Cohen, S. M. Meenehan, G. S. MacCabe, S. Gröblacher, A. H. Safavi-Naeini, F. Marsili, M. D. Shaw, and O. Painter, Nature **520**, 522 (2015).

[50] S. M. Meenehan, J. D. Cohen, S. Gröblacher, J. T. Hill, A. H. Safavi-Naeini, M. Aspelmeyer, and O. Painter, Phys. Rev. A **90**, 011803 (2014).

[51] P. S. Zyryanov and G. G. Taluts, Soviet Physics JETP-USSR **22**, 1326 (1966).

[52] M. G. Holland, Phys. Rev. **132**, 2461 (1963).

[53] J. Callaway, Phys. Rev. **113**, 1046 (1959).

[54] D. ter Haar, *Collected Papers of L. D. Landau*, 2nd ed. (Elsevier, 2013).

[55] A. Akhiezer, J. Phys. (Moscow) **1**, 277 (1939).

[56] R. Lifshitz and M. L. Roukes, Phys. Rev. B **61**, 5600 (2000).

[57] N. Moiseyev, *Non-Hermitian Qauntum Mechanics* (Cambridge University Press, 2011).

[58] M. Field, C. G. Smith, M. Pepper, D. A. Ritchie, J. E. F. Frost, G. A. C. Jones, and D. G. Hasko, Phys. Rev. Lett. **70**, 1311 (1993).

[59] H. Lefebvre-Brion and R. W. Field, *The Spectra and Dynamics of Diatomic Molecules: Revised and Enlarged Edition* (Elsevier Academic Press, 2004).

[60] W. B. Gauster, Phys. Rev. B **4**, 1288 (1971).

[61] J. J. Hall, Phys. Rev. **161**, 756 (1967).

[62] H. J. McSkimin, J. Appl. Phys. **24**, 988 (1953).

[63] G. S. Oehrlein and J. F. Rembetski, IBM J. Res. Develop. **36**, 140 (1992).

[64] Y. H. Lee, G. S. Oehrlein, and C. Ransom, Radiation Effects and Defects in Solids: Incorporating Plasma Science and Plasma Technology **111-112**, 221 (1989).

[65] M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and K. Suma, Appl. Phys. Lett. **55**, 562 (1989).

[66] M. Morita and T. Ohmi, Jpn. J. Appl. Phys. **33**, 370 (1994).

[67] W. A. Phillips, Phys. Rev. Lett. **61**, 2632 (1988).

[68] J. L. Black and B. I. Halperin, Phys. Rev. B **16**, 2879 (1977).

Appendix A: Device Fabrication

The devices were fabricated using a silicon-on-insulator wafer with a silicon (Si) device layer thickness of 220 nm and buried-oxide layer thickness of 3 μm. The device geometry was defined by electron-beam lithography followed by inductively coupled plasma reactive ion etching (ICP-RIE) to transfer the pattern through the 220 nm Si device layer. Photoresist was then used to define a 'trench' region of the chip to be etched and cleared for fiber access to device waveguides. In the unprotected trench region of the chip, the buried-oxide layer is etched using a highly anisotropic plasma etch, and the handle Si layer is cleared to a depth of 100 μm using an isotropic plasma etch. The devices were then undercut using a vapor-HF etch and cleaned in a piranha solution before a final vapor-HF etch to remove the chemically-grown oxide. In fabrication, devices were spatially grouped into arrays in which the number of acoustic radiation shield periods is scaled from zero to ten while all other geometric parameters are held nominally identical.

Appendix B: Modeling of Disorder in the OMC Cavity

The use of a phononic bandgap shield is necessitated by the lack of a full gap for the nanobeam cavity. In Fig. 4 we present a numerical study of the effects of random fabrication imperfections on the radiative damping of the shielded OMC cavity mode. We compare in Fig. 4(b) the simulated acoustic Q-factor of the ideal, unperturbed cavity structure to that of cavity structures with a fixed level of disorder in the hole sizes (standard deviation, $\sigma_{\text{size}} = 4$ nm) and varying levels of disorder in the hole centers ($\sigma_{\text{pos.}} = 2, 4, 8$ nm). An absence of perturbations to the nanobeam
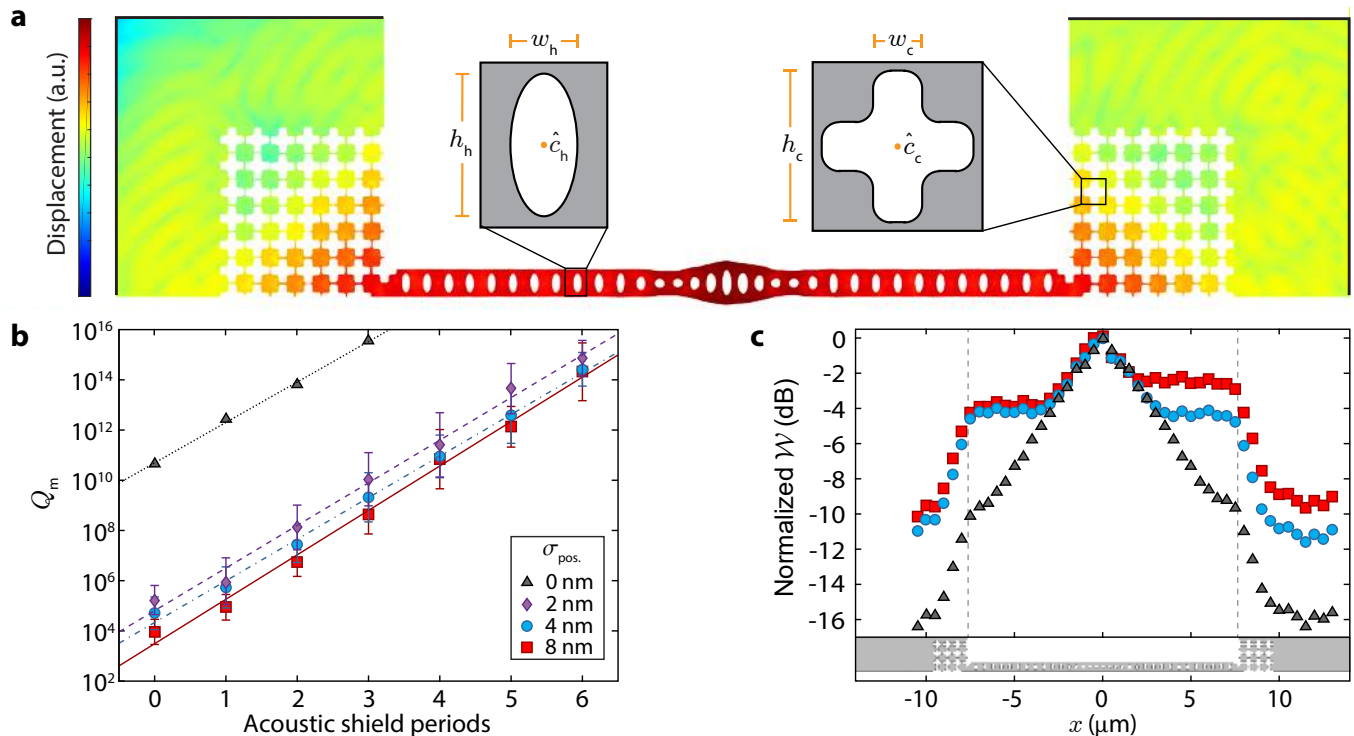

FIG. 4. **Impact of fabrication imperfections. a**, FEM simulation of the breathing-mode mechanical displacement field for a nanobeam OMC with $N = 6$ periods of acoustic shielding, illustrating localization of the vibrational energy. The geometry in the simulation consists of the nanobeam OMC, acoustic shielding, and the surrounding Si substrate. The borders of the simulation geometry are modeled as an absorbing perfectly-matched layer (PML, outlined with solid black lines). The insets show critical parameters of the device geometry. To introduce disorder into the simulations, each of these geometric parameters is drawn from independent Gaussian distributions centered on the nominal design parameter value with standard deviation $\sigma_{\text{pos.}}$ for the center positions and σ_{size} for the diameter or length of holes. **b**, Plot of the simulated mechanical Q-factor due to acoustic radiation from the cavity through the acoustic shielding. The straight lines are exponential fits to the mean data points of the simulation (the error bars indicate the standard deviation of ensemble of simulations for each shield number and disorder level). **c**, Plots of the normalized acoustic energy density $\mathcal{W}$ along a line cut through the center of the beam for $\sigma_{\text{pos.}} = 0$ nm (black triangles), 4 nm (blue circles), and 8 nm (red squares).

cavity, even without any shielding, yields large radiation-limited Q-factors in excess of 10^{10}. This is a result of the quasi-bandgap that exists in the nanobeam mirror section for modes of a specific symmetry about the center-line of the beam; however, any perturbation that breaks this symmetry results in a compromised quasi-bandgap in the nanobeam (Q drops from $\gtrsim 10^{10}$ to $\lesssim 10^5$ for nanometer-scale perturbations). Conversely, the exponential trend of the radiation-limited Q-factor with the number of shield periods is consistently a factor of $\times 5.5$ per additional period, independent of the disorder level. This interpretation is further bolstered by the plots in Fig. 4(c) comparing the linear acoustic energy density along the axis of the nanobeam, $\mathcal{W}$, for the mode of the unperturbed cavity and the modes of two different realizations of disordered cavities.

Appendix C: Measurement Setup, Optical Characterization, and Optomechanical Calibration

1. Measurement Setup

The full measurement setup used for device characterization is shown in Fig. 5. The light source is a fiber-coupled tunable external-cavity diode laser, of which a small portion is sent to a wavemeter (λ-meter) for frequency stabilization. The light is then sent to high-finesse tunable fiber Fabry-Perot filter (Micron Optics FFP-TF2, bandwidth 50 MHz, FSR 20 GHz) to reject laser phase noise at the mechanical frequency, which can contribute to noise-photon counts on the SPDs. After this prefiltering, the light is routed to an electro-optic phase modulator (ϕ-mod) which is driven by an RF signal generator at the mechanical frequency to generate optical sidebands used for locking the detection-path filters. The light is then directed via 2×2 mechanical optical switches into a "high-extinction" path consisting of a series of modulator components which are driven by a digital pulse generator to generate high-extinction-ratio optical pulses. The digital pulse generator is used to synchronize the switching of the modulation components as well as to trigger the time-correlated single-photon-counting (TCSPC) module. Of these modulation components, two are electro-optic intensity modulators which together provide ~ 60 dB of fast extinction (~ 20 ns rise and fall times), and two are Agiltron NS 1×1 switches (rise time 100 ns, fall time ~ 30 μs) which provide a total of 36 dB of additional extinction. The total optical extinction used to generate our optical pulses is approximately 96 dB, which is greater than the cross-talk specification of our mechanical optical switches. For this reason we use two 2×2 switches in parallel to isolate the high-extinction path to ensure that our off-state optical power is limited by our high-extinction modulation components rather than by cross-talk through the mechanical switches. The light is then passed through a variable optical attenuator (VOA) to control the input pulse on-state power level to the cavity, and sent to a circulator which directs the light to a lensed-fiber tip for end-fire coupling to devices inside a dilution refrigerator. The reflected signal is then routed back to either one of two detection setups. The first includes an erbium-doped fiber amplifier (EDFA) and a high-speed photodetector (PD) connected to a spectrum analyzer (SA) and a vector network analyzer (VNA). The second detection path is used for the phonon counting measurements. Here the light passes through three cascaded high-finesse tunable fiber Fabry-Perot filters (Micron Optics FFP-TF2) inside an insulating housing and then to the SPD inside the dilution refrigerator.

The cascaded fiber Fabry-Perot (FP) filters are aligned to the optical cavity resonance frequency ω_{c} during measurement such that the signal reaching the SPDs consists of sideband-scattered photons and a small contribution of laser-frequency pump-bleed-through. In total the filters suppress the pump by >100 dB. This bleed-through is calibrated by positioning the laser far off-resonance of the optical cavity, such that the device acts simply as a mirror, while fixing the relative detuning of the filters and the pump laser at the mechanical frequency $\omega_{\mathrm{m}}/2\pi$ and measuring the photon count rate on the SPDs as a function of laser power.

Additionally, both the FP-filters and the EOMs will drift during measurement and must be periodically re-locked. We therefore regularly stop the measurement and perform a re-locking routine. First, we re-lock the EOMs by applying a sinusoidal dithering signal of ~ 1 V to them while monitoring the optical transmission, then decrease the dithering amplitude gradually to lock to the minimum of transmission. Next we switch out of the high-extinction pulse path (SW-2A,2B) and out of the SPD path (SW-5), drive the phase modulator with a large RF power at $\omega_{\mathrm{m}}/2\pi$ to generate large optical sidebands at the cavity resonance frequency, and send this light into the FP-filter stack. The transmission through each filter is monitored while a dithering sinusoidal voltage is applied to each filter successively, and the amplitude and DC offset of the dithering signal are adjusted until the optical transmission signal at the desired sideband is maximized. The offset voltage is then held fixed during the subsequent measurement run. The filters will drift due to both thermal fluctuations and acoustic disturbances in their environment, so in order to further improve the filters' stability we have placed them inside a custom-built insulated housing.

The SPDs used in this work are amorphous WSi-based superconducting nanowire single-photon detectors developed in collaboration between the Jet Propulsion Laboratory and NIST. The SPDs are mounted on the still stage of the dilution refrigerator at ~ 700 mK. Single-mode optical fibers are passed into the refrigerator through vacuum feedthroughs and coupled to the SPDs via a fiber sleeve attached to each SPD mount. The radio-frequency output of
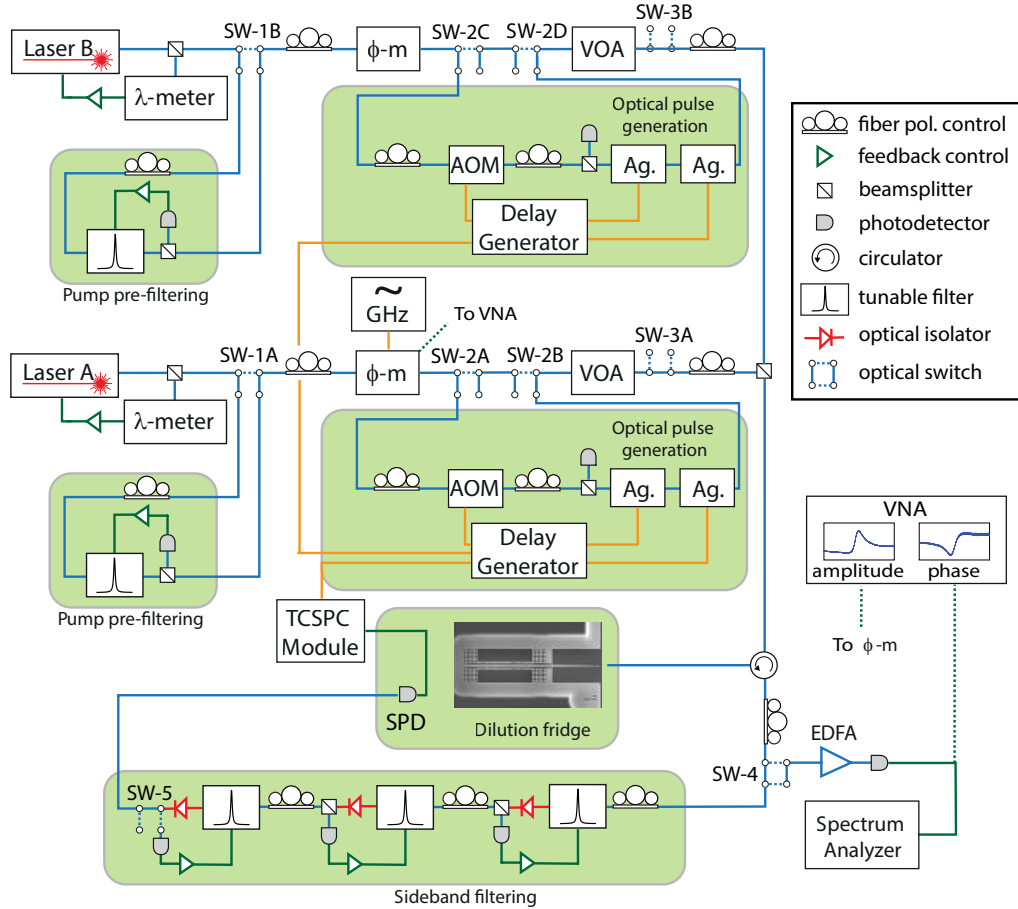
FIG. 5. **Pulsed-excitation phonon counting measurement setup.** Simplified diagram of the experimental setup used for low-temperature optomechanical device characterization and phonon-counting measurements. Lasers A and B are passed through 50 MHz-bandwidth filters to suppress broadband spontaneous emission noise. Both lasers are equipped with modulation components (AOM, Ag.) for generating high-extinction optical pulses. The modulation components are triggered by a digital delay generator (Laser B components are triggered by the 'master' Laser A generator). Upon reflection from the device under test, a circulator routes the outgoing light to either (1) an EDFA and spectrum analyzer, or (2) a sideband-filtering bank consisting of three cascaded fiber Fabry-Perot filters (Micron Optics FFP-TF2) and the SPD operated at ~ 760 mK. λ-meter: wavemeter, ϕ-m: electro-optic phase modulator, EOM: electro-optic intensity modulator, AOM: acousto-optic modulator, Ag.: Agiltron 1x1 MEMS switch, SW: optical 2×2 switch, VOA: variable optical attenuator, EDFA: erbium-doped fiber amplifier, VNA: vector network analyzer, SPD: single photon detector, TCSPC: time-correlated single photon counting module (PicoQuant PicoHarp 300).

each SPD is amplified by a cold-amplifier mounted on the 50 K stage of the refrigerator as well as a room-temperature amplifier, then read out by a triggered PicoQuant PicoHarp 300 time-correlated single photon counting module. We have observed SPD dark count rates as low as ~ 0.6 c.p.s. and SPD quantum efficiency $\eta_{\mathrm{SPD}} \simeq 60\%$.

2. Optical Characterization

Each device we have measured in this work was characterized optically in order to determine its optical resonance frequency ω_{c}, total optical linewidth κ, waveguide-cavity coupling rate κ_{e}, waveguide-cavity coupling efficiency $\eta_{\kappa} = \kappa_{\mathrm{e}}/\kappa$, and fiber-to waveguide coupling efficiency η_{cpl}. In particular, the waveguide-cavity coupling efficiency η_{κ} is measured by placing the laser far off-resonance and using the VNA to drive an intensity modulator to sweep an optical sideband through the cavity frequency and measure the optical response on a high-speed PD (after amplification by the EDFA) connected to the VNA signal port. From this we obtain the amplitude and phase response of the cavity, which are fitted to determined η_{κ}.
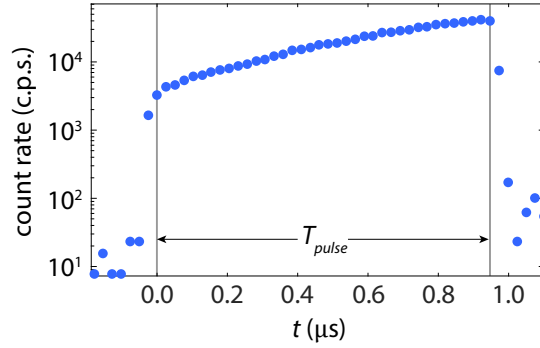
FIG. 6. **Blue-detuned calibration of sideband photon scattering rate. a**, Plot of photon count rate during optical pulse for $\Delta = -\omega_m$, $n_c = 101$ and $T_{per} = 0.2$ ms. Measurements performed on device A with parameters $(\kappa, \kappa_e, g_0, \omega_m, \gamma_0) = 2\pi(1.507 \text{ GHz}, 778 \text{ MHz}, 713 \text{ kHz}, 5.053 \text{ GHz}, 14.1 \text{ kHz})$, and zero acoustic shielding periods.

3. Vacuum Optomechanical Coupling Rate and Mode Occupancy Calibration

The measurements presented in this work rely on calibration of the vacuum optomechanical coupling rate g_0. After fitting the total optical linewidth κ from an optical reflection spectrum and calibrating η_κ as described above, the intracavity photon number n_c for a specified detuning is known and is proportional to the optical power input to the cavity:

$$n_c = \langle \hat{a}^\dagger \hat{a} \rangle = \frac{P_{in}}{\hbar \omega_l} \frac{\kappa_e}{\Delta^2 + (\kappa/2)^2}, \tag{C1}$$

where ω_l is the applied laser frequency. In these phonon-counting measurements, a critical calibration parameter is the photon scattering rate per phonon in the mechanical mode. In particular, from Ref. [27, 49, 50] the photon count rate at the SPD for a red- or blue-detuned pump is:

$$\Gamma(\Delta = \pm\omega_m) = \Gamma_{dark} + \Gamma_{pump} + \Gamma_{SB,0}(n_m + \frac{1}{2}(1 \mp 1)), \tag{C2}$$

where n_m is the average phonon occupancy of the breathing mode, $\Gamma_{SB,0} = \eta_{det}\eta_{cpl}\eta_\kappa\gamma_{OM}$ is the *detected* photon scattering rate per phonon (including experimental set-up efficiencies) and $\gamma_{OM} = 4g_0^2 n_c/\kappa$ is the optomechanical damping rate. In the absence of mechanical occupancy, pump photons may be spontaneously scattered by the mechanics owing to the nonzero mechanical susceptibility at the pump frequency. These real photons scattered by the mechanical vacuum noise in the presence of a pump laser are detected at a rate $\Gamma_{SB,0}$, providing a calibration of the per-phonon count rate directly to the vacuum noise. Here η_{det} is the measured overall detection efficiency of the setup, including losses in the fiber runs and circulator, insertion losses in the filters, the fiber run inside the dilution refrigerator, and the detection efficiency of the SPD. To calibrate $\Gamma_{SB,0}$ (for a fixed intracavity photon number n_c), a blue-detuned pump ($\Delta = -\omega_m$) pulsed with a repetition time of T_{per} drives the mechanics. In the initial time bin during the pulse, $n_m \ll 1$ (if $1/T_{per} \ll \gamma_0$) and we can approximate the sideband photon count rate $\Gamma \approx \Gamma_{SB,0}$. Including detection non-idealities such as bleed-through of the pump laser to the SPD and dark counts on the SPD, the detected count rate is $\Gamma(T_{pulse} = 0) = \Gamma_{DCR} + \Gamma_{pump} + \Gamma_{SB,0}$. This measurement provides an absolute calibration of the detection photon count rate to the mechanical vacuum noise, where the count rate is proportional to intracavity photon number, allowing calibrated thermometry with a precision that is independent of knowledge of the losses in the optical path. Additional knowledge of the optical path losses enables the inference of γ_{OM} from a measurement of $\Gamma_{SB,0}$, which can be used to calibrate the vacuum optomechanical coupling rate. For one representative device as shown in Figure 6 we measure $\Gamma_{SB,0} = 3.263 \times 10^3$ c.p.s. using a measurement photon number of $n_c = 101$, and with κ known from independent measurements we extract $g_0/2\pi = 713$ kHz, consistent with previous measurements on similar devices with similar orientations relative to the Si crystal axes [27].

An important figure of merit for phonon-counting measurements is the sensitivity, expressed as a noise-equivalent phonon number n_{NEP}. This n_{NEP} represents the equivalent occupancy inferred from noise counts only:
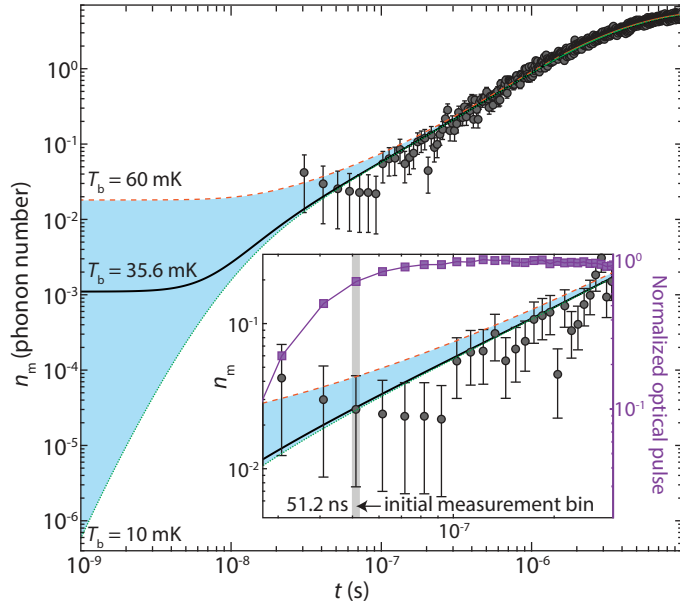
FIG. 7. **Base occupancy measurement and pulse turn-on dynamics.** Mode occupancy during the pulse on-state of the zero-shield device (device A) presented in Fig. 10. The photon number $n_c = 10$ is chosen to be small to minimize parasitic heating during the initial time bins of the pulse (bin size is 10.24 ns.). The model best-fit corresponds to $T_b = 35.6$ mK. Bounding curves to the fit are shown for $T_b = 60$ mK (orange dashed line) and $T_b = 10$ mK (green dotted line). Inset: Overlay plot of the initial time bins of the mode occupancy curve and the input optical pulse (purple squares). Time bins earlier than 51.2 ns occur during the fast rise of the pulse, which occurs at a timescale set by the rise of the EOMs and optical switches. The first measurement bin is chosen at $t = 51.2$ ns, where the input optical pulse has reached $> 70\%$ of its nominal on-state value (here $n_c = 10$). For 10.24 ns binning as shown here, the initial measurement bin is bin #5. For 25.6 ns binning as shown in the Main Text figures, the initial bin is bin #2.

$$n_{\mathrm{NEP}} = \frac{\Gamma_{\mathrm{dark}} + \Gamma_{\mathrm{pump}}}{\Gamma_{\mathrm{SB},0}}. \tag{C3}$$

The acoustic breathing mode thermalizes to a mode temperature T_m which is related to the applied fridge temperature T_f through the thermal conductance G_{th} of the structure. In order to measure the minimum temperature $T_{\mathrm{b,min}}$ to which the mode thermalizes at the lowest fridge temperature $T_f = 10$ mK, we use a low-power ($n_c = 10$) red-detuned pulsed probe and a device having relatively low $Q_m = 3.57 \times 10^5$ (chosen so that the mode thermalizes to its base temperature rapidly between each incident optical pulse). The initial mode occupancy during the pulse then approximately corresponds to the 'off-state' occupancy n_0. However, as the optical probe turns on during the first several time bins of the pulse, the mode is heated such that the initial observed occupancy exceeds n_0. We therefore extract n_0 from a fit to the pulsed heating model which is extrapolated back to $t = 0$.

Figure 7 shows a fit to the pulse occupancy in the on-state, which yields $T_m \approx 35.6$ mK. The lower bound of 10 mK is set by the minimum applied fridge temperature, and the upper bound 60 mK corresponds to the directly observed occupancy value in the initial measurement bin.

4. Summary of Device Parameters

Detailed measurements of several devices are presented in the Main Text and Appendices. For reference, here we provide a look-up table for each of these devices and their measured optical and mechanical properties.

TABLE I. **Measured optical and acoustic device parameters.** ρ_r is the Si device layer resistivity of the SOI wafer from which the device was fabricated (as provided by the manufacturer), λ_c is the optical mode wavelength, κ is the measured total optical linewidth, κ_e is the measured coupling rate between the OMC cavity mode and the on-chip waveguide, g_0 is the measured vacuum optomechanical rate, ω_m is the measured breathing mode frequency, $\gamma_0/2\pi$ is measured breathing mode damping from ringdown measurements at $T_f = 7$ mK.

Device	shield #	ρ_r [Ω-cm]	λ_c [nm]	$\kappa/2\pi$ [GHz]	$\kappa_e/2\pi$ [MHz]	$g_0/2\pi$ [MHz]	$\omega_m/2\pi$ [GHz]	$\gamma_0/2\pi$ (Hz)
A	0	5-20	1541.850	1.507	778	0.713	5.053	14.1×10^3
B	6	5-20	1539.285	1.13	605	~ 0.713	5.013	0.21
C	7	5-20	1538.716	1.21	362	~ 0.713	5.014	0.27
D	8	$> 5 \times 10^3$	1538.971	0.575	131	1.15	5.31	0.108
E	7	5-20	~ 1540	1.244	261	0.833	4.98	0.33

Appendix D: Optical-Absorption-Induced Bath

1. Theoretical Model of the Bath

Optical absorption is found to induce additional parasitic heating and damping of the high-Q acoustic breathing mode of the Si OMC devices at millikelvin temperatures. This absorption heating is thought to proceed through excitation of sub-bandgap electronic defect states at the Si surfaces which undergo phonon-assisted decay, generating a local bath of thermal phonons coupled to the high-Q breathing mode [50]. We may gain some understanding of the optically-induced bath by considering a simple model of phonon-phonon interactions which can couple the optically-induced hot phonon bath to the breathing mode. As we are concerned in this work with the phonon dynamics at low bath temperature ($T_b \lesssim 10$ K), and the acoustic mode of interest is at microwave frequencies, the phonon-phonon interactions leading to heating and damping of the breathing mode can be understood in terms of a Landau-Rumer scattering process [23, 51] (see App. F). In this context, we may consider a simple model in which our mode of interest at frequency ω_m is coupled to higher-frequency bath phonon modes at frequencies ω_1 and ω_2, with $\omega_2 - \omega_1 = \omega_m$. Then we may write the scattering rates into and out of the mode of interest to first order in perturbation theory [23, 50] as $\Gamma_+ = A(n_m + 1)(n_2 + 1)n_1$ and $\Gamma_- = An_m n_2(n_1 + 1)$, respectively, where n_1, n_2, and n_m are the number of phonons in each mode involved in the scattering and A is a constant describing the Si lattice anharmonicity. Then the overall rate of change in the occupancy of the mode of interest is,

$$\dot{n}_m = \Gamma_+ - \Gamma_- = -A(n_1 - n_2)n_m + An_2(n_1 + 1). \tag{D1}$$

This expression has exactly the form of a harmonic oscillator coupled to a thermal bath with rate $\gamma_p = A(n_1 - n_2)$ and effective occupancy $n_p = An_2(n_1 + 1)/\gamma_p$. Assuming thermal occupancies for each of the higher frequency phonon modes of the hot bath, $n_{1,2} = n_B[\hbar\omega_{1,2}/k_B T_p] \equiv 1/(\exp[\hbar\omega_{1,2}/k_B T] - 1)$, and using the identity $n_B[x + x'](n_B[x] + 1) = (n_B[x] - n_B[x + x'])n_B[x']$ [23], one finds that the mode m thermalizes with the hot bath via 3-phonon scattering to an effective occupancy which is $n_p = n_B[\hbar\omega_m/k_B T_p]$. This result holds when the hot bath thermalizes to some temperature independent of the interactions with mode m.

In the real material system of the nanobeam, the local hot phonon bath at elevated temperature T_b is expected to be generated as electronic states at $\sim$ eV energy undergo phonon-assisted relaxation processes, emitting a shower of high-frequency phonons which subsequently decay by a cascade of nonlinear multi-phonon interactions into a bath of GHz phonons. Due to the geometric aspect ratio of the thin-film nanobeam, the local density of phonon states becomes restricted at lower frequency, decreasing the rates of phonon-phonon scattering at low frequency relative to those of a bulk crystal with a 3D Debye density of states. The beam thickness ($t = 220$ nm, width $w \approx 560$ nm, length $l \approx 15$ μm) corresponds to a relatively high cutoff frequency in the vicinity of $\omega_{co}/2\pi \approx v_l/(2t) \approx 20$ GHz, where $v_l = 8.433$ km/s is the longitudinal-phonon velocity in Si. This cutoff frequency imposes an effective phonon *bottleneck* preventing further rapid thermalization to lower-lying modes and a resulting buildup in the bath phonon population above the bottleneck. For phonon frequencies below the cutoff, where the wavelength is large enough to approach the lattice constant of the acoustic bandgap clamping region, the reflectivity of the clamping region increases as ballistic radiation out of the nanobeam is suppressed. The result is a reduced density of phonon states near and below the cutoff, where the nanobeam supports quasi-discrete (and long-lived, especially in the vicinity of the mirror bandgap and acoustic shield bandgap) phonon modes at lower frequency as outlined in Fig. 8. The phenomenological coupling rate γ_p describes the rate at which the lower-lying modes—in particular the breathing mode at 5 GHz—are coupled to the elevated-temperature bath of higher-frequency phonons above the bottleneck.
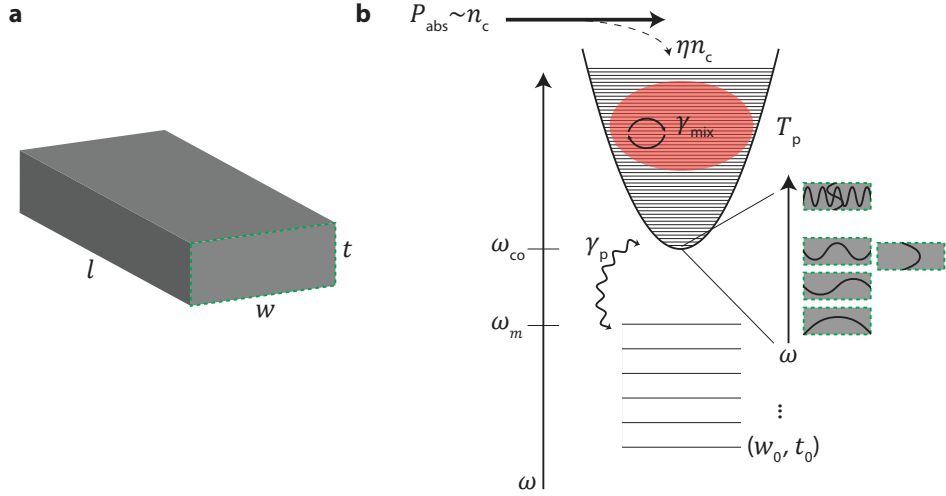
FIG. 8. **Impact of the phonon-bottleneck on the optical-absorption bath. a**, Cross-sectional dimensions of the thin-film nanobeam. **b**, Absorption of sub-Si-bandgap photons gives rise to phonon-assisted decay of THz phonons into a local bath of GHz phonons in the nanobeam. This bath is expected to experience a bottleneck at a cutoff frequency corresponding to the cross-sectional dimensions of the nanobeam, such that a high-frequency phonon bath accumulates and thermalizes among itself to a local temperature T_b at rate γ_{mix}. In the vicinity of the bottleneck frequency the relevant normal modes of the beam are those shown in the inset (black lines are schematics of the local strain in the beam). The lowest-lying discrete mode (w_0, t_0) is a fundamental bowstring mode of the nanobeam at ~ 20 MHz.

In the context of this proposed phonon-bottleneck model, we now consider instead of a discrete pair of modes n_1 and n_2 a quasi-continuum of high-frequency bath modes coupled to the mode of interest via some anharmonicity matrix element $A(\omega; \omega_m)$. We will assume that the thermal phonons populating the bath have sufficient time to thermalize amongst each other before decaying, or in other words, that they couple to each other at a mixing rate γ_{mix} much greater than their coupling rates to the external environment or to the lower-lying phonon modes. Under this assumption, we may define an effective local temperature T_p such that the occupancy of a bath phonon at frequency ω is given by the Bose-Einstein occupation factor

$$n_b[\omega; T_p] \equiv n_B[\hbar(\omega - \omega_{co})/k_B T_p] = \frac{1}{e^{\hbar(\omega-\omega_{co})/k_B T_p} - 1}, \tag{D2}$$

where ω_{co} represents the new effective ground-state frequency due to the phonon bottleneck effect, and $n_B[x] = 1/(\exp[x] - 1)$ is the Bose distribution.

The temperature of the optically-induced hot phonon bath, T_p, can then be related to the absorbed optical power P_{abs} using a model of the lattice thermal conductivity in the low temperature limit. Assuming the optical absorption process is linear, we can write the absorbed optical power as a fraction η of the optical pump power: $P_{abs} = \eta P_{in} = \eta' n_c$. In steady state, the power output into the phonon bath is equal to its input, $P_{out} = P_{abs} \sim n_c$. The lattice thermal conductivity at low temperatures, where phonon transport is ballistic, scales as a power law of the phonon bath temperature [52, 53], $G_{th} \sim (T_p)^\alpha$. The power law exponent α is equal to the effective number of spatial dimensions d of the material/structure under consideration. Effectively, the hot phonon bath radiates energy as a black body, with radiated power scaling as $(T_p)^{\alpha+1}$ via Planck's law. In the case of a structure with 2D phonon density of states, such as the OMC cavity in the frequency range from 10-100 GHz (c.f., Fig. 18), $\alpha = d = 2$ and the hot phonon bath temperature scales as $T_p \sim P_{out}^{1/3} \sim n_c^{1/3}$. This approximate scaling is expected to be valid so long as phonons in the hot phonon bath approximately thermalize each other upon creation from optical absorption events, and then radiate freely (ballistically) into the effective zero temperature substrate. The picture one has then is that the hot bath phonons make multiple passes within the OMC cavity region, scattering with other phonons leading to thermalization, and then eventually radiating into the substrate, i.e., the OMC cavity is still a good cavity for many phonons in the acoustic frequency region above the phononic bandgap.

In analogy with Equation D1, for a phonon bath density of states $\rho(\omega)$ we can calculate the effective coupling rate γ_p between the hot phonon bath and the mode of interest due to 3-phonon scattering:

$$\gamma_{\mathrm{p}} = \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; A[\omega; \omega_{\mathrm{m}}]\rho[\omega]\rho[\omega + \omega_{\mathrm{m}}] \left(n_{\mathrm{b}}[\omega] - n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}]\right), \tag{D3}$$

In a simple continuum elastic model [23, 50], the product of the anharmonicity matrix element $A[\omega; \omega_m]$ and the density of states is taken to obey a polynomial scaling $A[\omega; \omega_m]\rho[\omega]\rho[\omega + \omega_{\mathrm{m}}] = A'(\omega - \omega_{\mathrm{co}})^a$ for some constants A' and a, where we have introduced the cut-off frequency below which we assume the density of states is zero. With this assumption,

$$\gamma_{\mathrm{p}} \cong A' \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; (\omega - \omega_{\mathrm{co}})^a \left(n_{\mathrm{b}}[\omega] - n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}]\right) \tag{D4}$$

$$= A' \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; (\omega - \omega_{\mathrm{co}})^a \left(\frac{n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}](n_{\mathrm{b}}[\omega] + 1)}{n_{\mathrm{B}}[\hbar\omega_{\mathrm{m}}/k_{\mathrm{B}}T_{\mathrm{p}}]}\right) \tag{D5}$$

$$= \frac{A'}{n_{\mathrm{B}}[\hbar\omega_{\mathrm{m}}/k_{\mathrm{B}}T_{\mathrm{p}}] + 1} \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; (\omega - \omega_{\mathrm{co}})^a \left(n_{\mathrm{b}}[\omega](n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}] + 1)\right), \tag{D6}$$

where in the last line we used the identity $n_{\mathrm{B}}[x + x'](n_{\mathrm{B}}[x] + 1)/n_{\mathrm{B}}[x'] = (n_{\mathrm{B}}[x + x'] + 1)n_{\mathrm{B}}[x]/(n_{\mathrm{B}}[x'] + 1)$. Making a change of variables to $x \equiv \hbar(\omega - \omega_{\mathrm{co}})/k_{\mathrm{B}}T_{\mathrm{p}}$ in the integral in Eq. (D6), we have

$$\gamma_{\mathrm{p}} \cong \left(\frac{A'}{n_{\mathrm{B}}[x_{\mathrm{m}}] + 1}\right) \left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar}\right)^{a+1} \int_{0}^{\infty} \mathrm{d}x \; x^a \left(n_{\mathrm{B}}[x](n_{\mathrm{B}}[x + x_{\mathrm{m}}] + 1)\right) \tag{D7}$$

where $x_{\mathrm{m}} = \hbar\omega_{\mathrm{m}}/k_{\mathrm{B}}T_{\mathrm{p}}$. The integral in Equation D7 depends on temperature only through x_{m}, and in the small and large x_{m} limit (corresponding to low and high temperature), is relatively independent of x_{m}. If we assume that the anharmonicity element $A[\omega; \omega_{\mathrm{m}}]$ is approximately frequency independent, and the only frequency dependence in $A'(\omega - \omega_{\mathrm{co}})^a$ comes from the phonon density of states, then $a \approx 2(d-1)$ for a phonon bath of dimension d. We can thus make a general observation about the scaling of the bath-induced damping rate γ_{p} in the low ($x_{\mathrm{m}} \gg 1$) and high ($x_{\mathrm{m}} \ll 1$) temperature regimes:

$$\gamma_{\mathrm{p}} \propto \begin{cases} \left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar}\right)^a \sim n_{\mathrm{c}}^{2(d-1)/(d+1)} & \text{for } T_{\mathrm{p}} \gg \frac{\hbar\omega_m}{k_{\mathrm{B}}}, \\ \left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar}\right)^{a+1} \sim n_{\mathrm{c}}^{(2d-1)/(d+1)} & \text{for } T_{\mathrm{p}} \ll \frac{\hbar\omega_m}{k_{\mathrm{B}}}, \end{cases} \tag{D8}$$

for a generic hot phonon bath of dimension d. In a structure such as the OMC nanobeam cavity we expect the dimensionality of the effective bath density of states to be reduced relative to the Debye 3D density of states for a bulk crystal. Here we will assume - consistent with numerical simulations of the OMC structure in Section F - that the phonon bath has a two-dimensional density of states corresponding to $a = 2$. In this case, we have the following scaling of the damping factor with intra-cavity photon number,

$$\gamma_{\mathrm{p}} \propto \begin{cases} \left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar}\right)^2 \sim n_{\mathrm{c}}^{2/3} & \text{for } T_{\mathrm{p}} \gg \frac{\hbar\omega_{\mathrm{m}}}{k_{\mathrm{B}}}, \\ \left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar}\right)^3 \sim n_{\mathrm{c}} & \text{for } T_{\mathrm{p}} \ll \frac{\hbar\omega_{\mathrm{m}}}{k_{\mathrm{B}}}. \end{cases} \tag{D9}$$

Upon thermalizing with the hot phonon bath, the effective thermal occupancy n_{p} of the high-Q breathing mode of the acoustic cavity can be found from a similar rate equation analysis as considered for the 3-mode scattering in Eq. (D1). Integrating over all the possible 3-phonon scattering events involving the mode of interest at frequency ω_{m} yields,

$$n_{\mathrm{p}} = \frac{1}{\gamma_{\mathrm{p}}} \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; A[\omega; \omega_{\mathrm{m}}]\rho[\omega]\rho[\omega + \omega_{\mathrm{m}}]n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}] \left(n_{\mathrm{b}}[\omega] + 1\right) \tag{D10}$$

$$\cong \frac{n_{\mathrm{b}}[\omega_{\mathrm{co}} + \omega_{\mathrm{m}}]A'}{\gamma_{\mathrm{p}}} \int_{\omega_{\mathrm{co}}}^{\infty} \mathrm{d}\omega \; \omega^a \left(n_{\mathrm{b}}[\omega] - n_{\mathrm{b}}[\omega + \omega_{\mathrm{m}}]\right) \tag{D11}$$

$$= n_{\mathrm{B}}[\hbar\omega_{\mathrm{m}}/k_{\mathrm{B}}T_{\mathrm{p}}]. \tag{D12}$$
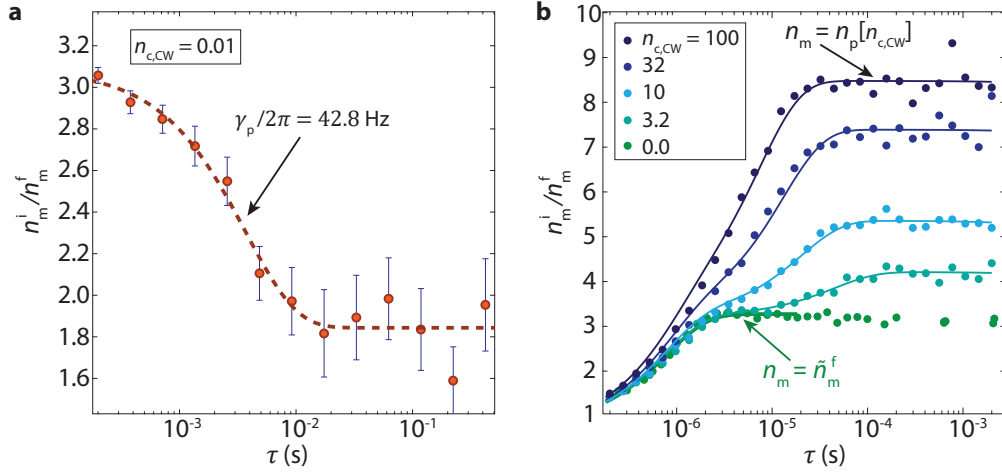
FIG. 9. **Measurement techniques for extracting the optical-bath-induced damping rate γ_{p}. a,** Ringdown measurement in the presence of a continuous-wave pump laser with an average intracavity photon number of $n_{\mathrm{c,CW}} = 10^{-2}$. The total decay rate is $\gamma = \gamma_{\mathrm{p}} + \gamma_0$, and with $\gamma_0/2\pi = 0.21$ Hz known from separate measurements, $\gamma_{\mathrm{p}}/2\pi = 42.8$ Hz is extracted directly from the fitted decay rate. **b,** At larger n_{c}, the bath-heating induced by the pump laser causes net heating in the pulse-off state. Here, the heating is fitted to the phenomenological model of Eq. (D14) to extract γ_{p} due to the CW laser pump. Measurements were performed on a six-acoustic-shield device (device B) with parameters $(\kappa, \kappa_{\mathrm{e}}, g_0, \omega_{\mathrm{m}}, \gamma_0) = 2\pi(1.13$ GHz, 605 MHz, 713 kHz, 5.013 GHz, 0.21 Hz) and with a readout photon number $n_{\mathrm{c,RO}} = 569$.

We therefore have a characteristic scaling behavior for the effective phonon occupancy n_{p} coupled to the cavity mode of interest that is,

$$
n_{\mathrm{p}} \propto
\begin{cases}
\left(\frac{k_{\mathrm{B}}T_{\mathrm{p}}}{\hbar\omega_{\mathrm{m}}}\right) \sim n_{\mathrm{c}}^{1/(d+1)} \overset{d=2}{=} n_{\mathrm{c}}^{1/3} & \text{for } T_{\mathrm{p}} \gg \frac{\hbar\omega_{\mathrm{m}}}{k_{\mathrm{B}}}, \\
\exp\left[-\hbar\omega_{\mathrm{m}}/k_{\mathrm{B}}T_{\mathrm{p}}\right] & \text{for } T_{\mathrm{p}} \ll \frac{\hbar\omega_{\mathrm{m}}}{k_{\mathrm{B}}}.
\end{cases}
\tag{D13}
$$

2. Measurement of optical-absorption-induced damping, γ_{p}

In order to measure the additional bath-induced damping rate γ_{p}, we use a pump-probe technique employing two laser sources. The pump laser is tuned to optical resonance ($\Delta = 0$) to eliminate dynamical back-action effects ($\gamma_{\mathrm{OM}} = 0$), and impinges on the cavity in continuous-wave (CW) operation. The pump laser generates a steady-state intracavity photon population $n_{\mathrm{c,CW}}$ and an absorption-induced bath at elevated temperature in the steady state. A second pulsed laser, the probe laser, is tuned to the red motional sideband of the cavity ($\Delta = +\omega_{\mathrm{m}}$) and is used to periodically read out the phonon occupancy, where the scattering rate of the probe laser at the beginning of the probe pulse provides an estimate of n_{p} due to the CW laser alone. Application of the probe laser not only allows readout of the breathing mode occupancy, but also produces an excess absorption-induced bath above and beyond that of the background CW laser alone. When the readout probe pulse is turned off, the breathing mode initially heats due to the excess hot bath created by the probe pulse (over several microseconds; see Fig. 10(c)), and then after this excess hot bath evaporates away leaving a breathing mode occupancy of n'_{f}, relaxes back to its steady-state occupancy set by the CW laser, $n_{\mathrm{m}}[n_{\mathrm{c,CW}}] = (\gamma_{\mathrm{p}}[n_{\mathrm{c,CW}}]n_{\mathrm{p}}[n_{\mathrm{c,CW}}] + \gamma_0 n_0)/(\gamma_{\mathrm{p}}[n_{\mathrm{c,CW}}] + \gamma_0)$. The rate of relaxation is set by the modified total damping rate of $\gamma_0 + \gamma_{\mathrm{p}}[n_{\mathrm{c,CW}}]$. By observing this modified exponential decay rate we directly extract $\gamma_{\mathrm{p}}[n_{\mathrm{c,CW}}]$, with γ_0 known from independent ringdown measurements in the absence of the CW background laser. For example, in Fig. 9(a) we show the measured ringdown of a high-Q six-shield device (device B; $\gamma_0/2\pi = 0.21$ Hz) for a CW pump laser photon number of $n_{\mathrm{c,CW}} = 10^{-2}$, from which we extract $\gamma_{\mathrm{p}}/2\pi = 42.8$ Hz.

For large $n_{\mathrm{c,CW}}$ ($\gtrsim 1$) the steady-state occupancy of $n_{\mathrm{m}}[n_{\mathrm{c,CW}}]$ becomes larger than the occupancy $\tilde{n}_{\mathrm{m}}^{\mathrm{f}}$ at the end the readout pulse. The readout pulse *should* cool the breathing mode, after all, and it is only the absorption-induced heating caused by the readout pulse itself that manifests as a ring down in absence of heating from the CW laser. For large $n_{\mathrm{c,CW}}$ then, γ_{p} is estimated by observing a ring-*up* in the pulse-off state from the final pulse occupancy $\tilde{n}_{\mathrm{m}}^{\mathrm{f}}$ to the elevated $n_{\mathrm{m}}[n_{\mathrm{c,CW}}]$. Figure 9(b) shows a representative data set for extracting γ_{p} at $n_{\mathrm{c,CW}} > 1$, where an initial fast rise is observed in the mode occupancy in the pulse-off state from $n_{\mathrm{m}}^{\mathrm{f}}$ to $\tilde{n}_{\mathrm{m}}^{\mathrm{f}}$ due to the aforementioned excess bath created by the readout pulse,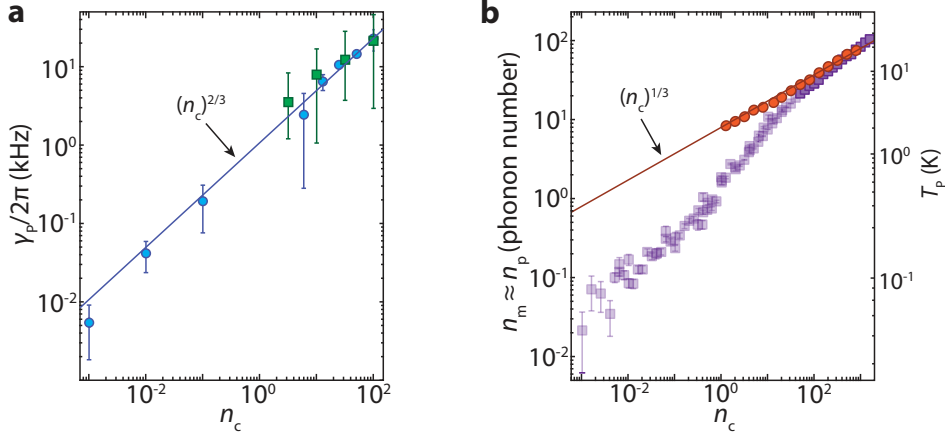 followed by a slower second heating stage from $\tilde{n}_{\mathrm{m}}^{\mathrm{f}}$ to $n_{\mathrm{m}}[n_{\mathrm{c,CW}}]$. As discussed

FIG. 10. **Measured steady-state properties of the optical-absorption-induced bath. a**, Plot of γ_p versus n_c for six-shield (blue circles) and zero-shield (green squares) devices. The solid line is a power-law fit to the six-shield device data: $\gamma_\mathrm{p}/2\pi = (1.07\text{ kHz}) \times n_\mathrm{c}^{2/3}$. The zero-shield device (device A) has parameters $(\kappa, \kappa_\mathrm{e}, g_0, \omega_\mathrm{m}, \gamma_0) = 2\pi(1.507\text{ GHz}, 778\text{ MHz},$ 713 kHz, 5.053 GHz, 14.1 kHz). The six shield device (device B) has parameters $(\kappa, \kappa_\mathrm{e}, g_0, \omega_\mathrm{m}, \gamma_0) = 2\pi(1.13\text{ GHz}, 605\text{ MHz},$ 713 kHz, 5.013 GHz, 0.21 Hz). **b**, Plot of n_p versus n_c for zero-shield (purple symbols) and six-shield (orange circles) devices. Purple squares represent the measured mode occupancy corrected for heating induced by the readout laser tone. The right-hand axis gives the effective bath temperature T_p which corresponds to the measured bath occupancy. Translucent squares show data taken in the regime where the intrinsic decay rate γ_0 is comparable to the bath-induced damping γ_p, indicating that the raw measured occupancy begins to deviate substantially from the inferred occupancy given in the plot. The solid line is a fit to the six-shield data giving $n_\mathrm{p} = (7.94) \times n_\mathrm{c}^{1/3}$. **c**, Normalized phonon occupancy during and after the optical pulse. Squares are data points and the solid line is a best fit to the dynamical model. During the pulse, back-action cooling occurs at a timescale $\gamma_\mathrm{OM}^{-1} \approx 100$ ns. The optical-absorption-induced bath simultaneously heats the mode at a rate $\gamma_\mathrm{p}n_\mathrm{p}$, such that at long T_pulse a steady-state mode occupancy n_f is reached. In the pulse-off state (gray squares), the residual phonon bath heats the mode at a rate $\gamma_\mathrm{p}(t)n_\mathrm{p}(t)$, where the bath damping and effective occupancy are explicitly time-dependent. A full dynamical model of the bath heating is used to generate the fit (dotted line). The purple data point in the off-state plot ($\tau = 200\ \mu$s) corresponds to off-state delay for the measured intra-pulse data shown in the on-state plot.

in more detail in sub-Section D 4, we can fit the ring up curve after the readout pulse is turned off by considering a phenomenological model including decay of the readout-induced hot bath,

$$\dot{n}_\mathrm{m} = -\left\{\gamma_0 + \gamma_\mathrm{p}[n_\mathrm{c,RO}]e^{-\zeta_{\gamma_\mathrm{p}}\tau} + \gamma_\mathrm{p}[n_\mathrm{c,CW}]\right\} n_\mathrm{m}$$
$$+ \left\{\gamma_\mathrm{p}[n_\mathrm{c,RO}]e^{-\zeta_{\gamma_\mathrm{p}}\tau} + \gamma_\mathrm{p}[n_\mathrm{c,CW}]\right\}\left(n_\mathrm{p}[n_\mathrm{c,RO}]e^{-\zeta_{n_\mathrm{p}}\tau} + n_\mathrm{p}[n_\mathrm{c,CW}]\right). \quad \text{(D14)}$$

We first measure the transient readout-induced bath in the absence of the CW laser (dark green curve in Fig. 9b), from which a fit to Eq. (D14) yields $n_\mathrm{p}[n_\mathrm{c,RO}] = 40$ phonons, $\gamma_\mathrm{p}[n_\mathrm{c,RO}]/2\pi = 9.55$ kHz, $\zeta_{\gamma_\mathrm{p}}/2\pi = 143$ kHz, and $\zeta_{n_\mathrm{p}}/2\pi = 15.9$ kHz. With these readout-induced bath values known, Eq. (D14) is numerically integrated to fit the entire heating curve in the pulse-off state to extract the additional CW-pump-induced damping $\gamma_\mathrm{p}[n_\mathrm{c,CW}]$.

The results of the measured optical-absorption-induced damping γ_p versus n_c are summarized in Fig. 10(a) for measurements on both a six-shield (device B) and a zero-shield (device A) nanobeam device. The observed power law scaling fits well to $\gamma_\mathrm{p}/2\pi = (1.07\text{ kHz}) \times n_\mathrm{c}^{2/3}$, in agreement with the scaling predicted in Eq. (D9) for a 2D density of states for the bath phonon population. Note that the much lower γ_0 of the six-shield device allows a much wider range of γ_p (and thus n_c) to be explored.

3. Measurement of optical-absorption-induced bath occupancy, $n_\mathbf{p}$

In order to measure the bath occupancy n_p, again two different methods are used to probe the high- and low-photon-number dependencies of the bath. To measure the bath occupancy at photon numbers $n_\mathrm{c} \gtrsim 1$, a simple readout technique may be used in which a single readout laser is sent to the cavity in continuous-wave operation. The laser is tuned to cavity resonance ($\Delta = 0$) and the resulting sideband scattered photon count rate appearing at either the lower or upper frequency mechanical sideband ($\Delta = \pm\omega_m$) will be
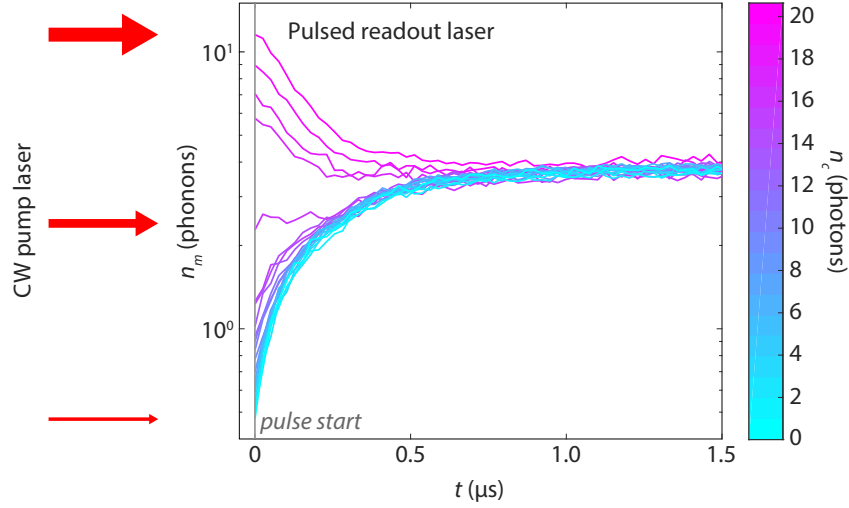
FIG. 11. **Pulsed measurements of the bath occupancy in a low-Q nanobeam.** A continuous-wave background laser (red arrows, detuning $\Delta/2\pi \approx 1$ GHz) is used to generate a constant stead-state absorption bath, while a pulsed readout laser (readout $n_{\mathrm{c,RO}} = 50.6$) is used to probe the resulting bath occupancy for various background laser powers $n_{\mathrm{c,CW}}$. The initial measured occupancy during the pulse is given by $n_{\mathrm{m}}^{\mathrm{i}} \approx (n_{\mathrm{p}}\gamma_{\mathrm{p}} + n_0\gamma_0)/(\gamma_{\mathrm{p}} + \gamma_0) + \tilde{n}_0$, where $\tilde{n}_0$ is residual occupancy due to the finite heating occurring before the first readout time bin. Measurements were performed on the zero-shield device (device A) with parameters $(\kappa, \kappa_{\mathrm{e}}, g_0, \omega_{\mathrm{m}}, \gamma_0) = 2\pi(1.507$ GHz, 778 MHz, 713 kHz, 5.053 GHz, 14.1 kHz$)$.

$$\Gamma = \Gamma_{\mathrm{noise}} + \left(\frac{\kappa}{2\omega_m}\right)^2 \Gamma_{\mathrm{SB},0} n_{\mathrm{m}}. \tag{D15}$$

With the sideband filters aligned to either of the mechanical sidebands of the cavity, the observed count rate is used to extract an equivalent occupancy $n_{\mathrm{m}} = n_{\mathrm{p}}$ at various pump powers n_{c}. The results are shown in Fig. 10b (orange circles) for a six-shield device (device B), exhibiting a power-law scaling of $n_{\mathrm{p}} = (7.94) \times n_{\mathrm{c}}^{1/3}$ in agreement with the model in the limit of high bath temperature $T_{\mathrm{p}} \gg \hbar\omega_m/k_{\mathrm{B}} \approx 200$ mK. The right-hand axis of Fig. 10 gives the effective bath temperature T_{p} corresponding to the measured occupancy, indicating that the measurement regime is indeed well in the high temperature limit.

At lower photon numbers $n_{\mathrm{c}} \lesssim 1$, and corresponding lower $n_{\mathrm{p}} \approx n_{\mathrm{m}}$, the SNR of the counting of photons scattered from cavity resonance into either mechanical sideband begins to drop below 1 due to the large sideband resolution factor $(2\omega_{\mathrm{m}}/\kappa)^2$ of the OMC cavity (c.f., Eq. (D15)). In this regime, an alternative measurement method is employed in which a CW pump laser generates a steady-state optical-absorption bath while a second pulsed readout laser is used to probe the breathing mode occupancy (see Fig. 11). The background pump laser is detuned to $\Delta/2\pi = 1$ GHz from the cavity resonance to minimize back-action as well as bleed-through counts through the sideband filters aligned at $\Delta = 0$. The initial measured occupancy $n_{\mathrm{m}}^{\mathrm{i}}$ during the pulse is a measure of the pump-induced bath occupancy; however, it includes a small residual occupancy $\tilde{n}_0 \approx 0.04$ due to heating caused by the readout laser prior to the first measurement time bin of the pulse-on state. We define a corrected occupancy $n_{\mathrm{m}}^{*} \equiv n_{\mathrm{m}}^{\mathrm{i}} - \tilde{n}_0$ which denotes the measured mode occupancy which is coupled to the fridge bath as well as the absorption-bath induced by the pump laser:

$$n_{\mathrm{m}}^{*} = \frac{n_{\mathrm{p}}\gamma_{\mathrm{p}} + n_0\gamma_0}{\gamma_{\mathrm{p}} + \gamma_0}. \tag{D16}$$

With n_0, γ_0, and the power-dependence of γ_{p} known from independent measurements, we can estimate the equivalent bath occupancy

$$n_{\mathrm{p}}[n_{\mathrm{c}}] = \frac{n_{\mathrm{m}}^{*}\gamma_{\mathrm{p}}[n_{\mathrm{c}}] + (n_{\mathrm{m}}^{*} - n_0)\gamma_0}{\gamma_{\mathrm{p}}[n_{\mathrm{c}}]}. \tag{D17}$$

Using this second method, over a much larger span of n_c, the behavior of the effective bath occupancy n_p for a zero-shield device with intrinsic damping rate $\gamma_0/2\pi = 14.1$ kHz is shown in Fig. 10b as purple squares. Note that measurement of the very high-Q six-shield device (device B) using the pulsed readout scheme is not practical due to the extremely long relaxation times required between readout pulses (we did, however, verify for a few values of n_c that the two schemes give consistent results). For $n_c \gtrsim 1$, again we find good agreement for the zero-shield device with a power-law scaling $n_p \propto n_c^{1/3}$ for $T_b \gg \hbar\omega_m/k_B$. Not only is the scaling of n_p versus n_c the same for both zero-shield and six-shield devices, but so is the absolute value of n_p. For $n_c \lesssim 1$, $\gamma_p(n_c) \approx \gamma_0$ for the zero-shield device and the measured occupancy n_m^* deviates substantially from n_p as the breathing mode thermalizes more strongly with the external substrate temperature set by the fridge ($T_f \approx 10$ mK). In this range we have plotted n_m^* in translucent purple squares to distinguish it from the region of parameter space where n_m^* is expected to faithfully represent n_p.

4. Measurement of optical-absorption-induced bath dynamics

The hot bath created by the application of laser light resonant with the optical mode of the OMC cavity does not instantaneously appear when the laser light is turned on, nor does it instantaneously vanish once the laser is turned off. Rather, there is a somewhat complicated bath dynamics that can be inferred from careful study of the temporal variation of the scattered photon signal from the readout pulse due to excitation of the mechanical breathing mode by the hot phonon bath. Using the measured breathing mode occupancy as a proxy one can infer many subtle features of the bath dynamics.

Figure 12(a) shows the measured scattered photon signal due to a pulsed readout tone applied on the lower motional sideband of the optical cavity ($\Delta = +\omega_m$) of a high mechanical Q-factor six-shield OMC device (device B). Here the readout pulses are $T_{\text{pulse}} = 4$ μs long and a variable delay τ is applied between each successive optical readout pulse. The scattered photons from the readout pulse are filtered by the filter bank resonantly aligned with the optical cavity resonance ($\Delta = 0$), thus yielding a photon count rate throughout the readout pulse which is proportional to the average occupancy of the mechanical breathing mode n_m. This is shown in Fig. 12(a) for a time bin resolution of 10.24 ns, with the first measurement bin occurring at $t = 100$ ns after the pulse-on signal is applied in order to ensure that the optical pulse amplitude has settled and reached its maximum value. In the left panel of Fig. 12(b) we plot time-varying normalized breathing mode occupancy, corresponding to the ratio of the measured signal during the pulse to that at the very end of the pulse. This curve is not a single-shot measurement, but rather averaged over thousands of pulses, for which the normalized signal avoids small, slow drifts in the efficiency of the measurement apparatus. In the right panel we plot the normalized initial measurement bin occupancy (still taken to be at 100 ns after the readout pulse is turned on) as a function of the off-state delay time τ between successive pulses.

Several things are quickly evident from these plots of the measured breathing mode occupancy during and after the applied optical pulse. During the pulse we expect the optomechanical back-action to induce damping and cooling of the breathing mode at a rate $\gamma_{\text{OM}}[n_c]$. Without any parasitic heating effects from the applied optical pulse, the breathing mode should cool down to its equilibrium occupancy, ideally very close to zero at the fridge temperature ($T_f = 10$ mK). This does not occur, but rather the breathing mode occupancy is seen to initially cool to a few phonons over ~ 300 ns, and then slowly heat to a steady-state phonon occupancy at the end of the pulse of $n_m^f = 4.2$ phonons (c.f., Fig. 12(a)). Similarly, once the optical pulse is turned off and light has left the optical cavity, the breathing mode occupancy starts to heat again, levelling off after a few microseconds following a slight overshoot to a modified post-pulse value of $n_m[0] = 27$ phonons ($\rightarrow n_m^i = 13.6$ phonons in the first masurement bin; c.f., Fig. 10(b)). This strange dynamics is a result of the coupling of the breathing mode to the optical-absorption-induced hot bath. The slight undershoot of the cooling and slow heating in the pulse-on state is a result of a slow turn on of the hot bath. Similarly, the transient post-optical-pulse heating results from the slow decay of the hot bath, now without the cooling from optomechanical back-action.

Noticeably, the timescales for the turn on (~ 400 ns) of the bath and the turn off (~ 3 μs) of the bath are different. Less evident from these plots, but nonetheless very clear when attempting to model the hot bath dynamics, is that there seems to be two components to the bath, one whose turn on and turn off transients are very rapid (effectively instantaneous with the optical field), and one with much slower relaxation times. Even more subtle is that to get very good agreement with the measured initial transient in the breathing mode occupancy in the immediate aftermath of turning off the optical pulse, it seems that the hot-bath damping factor, γ_p, should be modeled with a more rapid relaxation rate than that of the hot bath occupancy, n_p. It may be that this is also the case in the transient dynamics during the pulse-on state, however, in the pulse-off state the relaxation rate of the measured breathing mode occupancy is far more sensitive to the value of γ_p as it dominates the total relaxation rate of the breathing mode in the absence of appreciable γ_{OM}.

The model used to fit the data in Fig. 12 consists of a set of coupled differential equations involving the breathing mode occupancy, the hot bath damping factor, and the effective hot bath occupancy. The rate equation for the
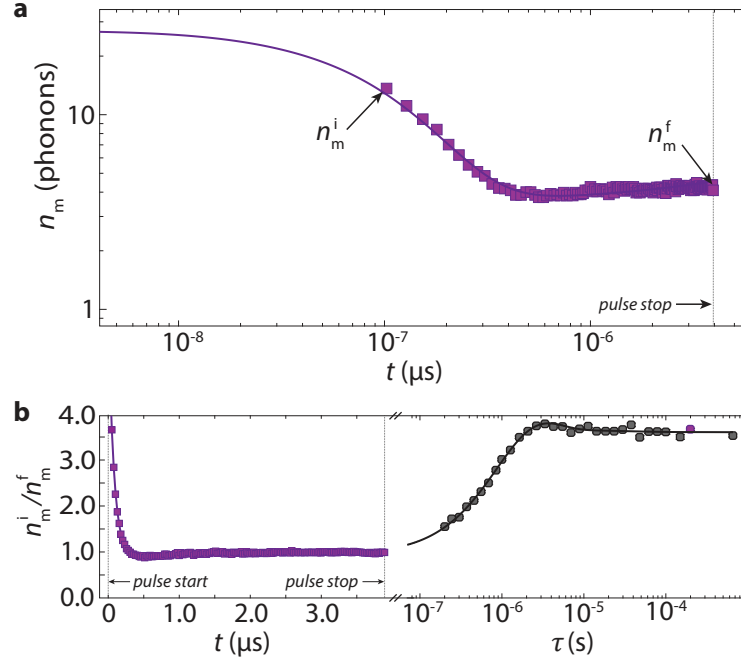
FIG. 12. **Measured pulse dynamics of the breathing mode occupancy and the optical-absorption-induced bath.**
a, Phonon occupancy of the breathing mode as a function of time t during the red-detuned ($\Delta = +\omega_m$) optical excitation pulse. Here, the time delay between successive pulses is $\tau = 654\mu s$. Squares are data points and the solid line is a best fit to the dynamical model. The nanobeam device is device B with six periods of acoustic shielding, and device parameters (κ, κ_e, g_0, ω_m, γ_0) $= 2\pi$ (1.13 GHz, 605 MHz, 713 kHz, 5.013 GHz, 0.21 Hz). During the pulse, back-action cooling occurs at a timescale $\gamma_{\rm OM}^{-1} \approx 100$ ns. Note that the initial mode occupancy $n_{\rm m}[0] = 27$ phonons is determined by extrapolating the model fit back to $t = 0$, while the earliest measurement bin has an occupancy of $n_{\rm m}^{\rm i} = 13.6$ phonons. The optical-absorption-induced bath heats the mode at a rate $\gamma_{\rm p}(t)n_{\rm p}(t)$, such that for long enough $T_{\rm pulse}$ a steady-state mode occupancy $n_{\rm m}^{\rm f}$ is reached. Here $n_{\rm m}^{\rm f} = 4.2$ phonons. The measurement time resolution bin size is 10.24 ns. **b**, Normalized breathing mode phonon occupancy during (left) and after (right) the optical pulse. In the pulse-off state (gray squares), the residual phonon bath heats the mode at a rate $\gamma_{\rm p}(\tau)n_{\rm p}(\tau)$, where the bath damping and effective occupancy are explicitly time-dependent. The purple data point in the off-state plot at $\tau = 200 \ \mu s$ indicates the pulse shown in the on-state plot. For all panels, the measurements were performed using an on-state readout intra-cavity photon number of $n_{\rm c} = 569$, and the solid curves correspond to the phenomenological model including the dynamics of the optical-absorption-induced bath, fit to the data using the parameters shown in Table II.

breathing mode occupancy is given by,

$$\dot{n}_{\rm m} = -(\gamma_{\rm p} + \gamma_{\rm OM} + \gamma_0)n_{\rm m} + \gamma_{\rm p}n_{\rm p} + \gamma_0 n_0 \tag{D18}$$

where in the pulse-on state $\gamma_{\rm OM} = \gamma_{\rm OM}[n_{\rm c}]$ will take on a large value on the order of 1 MHz for a readout pulse amplitude of a few hundred intra-cavity photons, and in the pulse-off state $\gamma_{\rm OM} \approx 0$ due to the large extinction ($\gtrsim 80$ dB) and rapid timescale of the turn-off the optical pulse (20 ns). During the pulse-on state the rate equations for the fast (F) and slow (S) components of the hot bath damping factor and effective occupancy are,

$$(\dot{\gamma_{\rm p}})_{\rm F(S)}(t) = -(\theta_{\gamma_{\rm p}})_{\rm F(S)} \left\{ (\gamma_{\rm p})_{\rm F(S)}(t) - (\delta_{\rm b})_{\rm F(S)}\gamma_{\rm p}[n_{\rm c,RO}] \right\}, \tag{D19}$$

and

$$(\dot{n_{\rm p}})_{\rm F(S)}(t) = -(\theta_{n_{\rm p}})_{\rm F(S)} \left\{ (n_{\rm p})_{\rm F(S)}(t) - (\delta_{\rm b})_{\rm F(S)}n_{\rm p}[n_{\rm c,RO}] \right\}, \tag{D20}$$

where $t = \{0, T_{\rm pulse}\}$ is the time from the start of the pulse to the end of the pulse, $(\theta_{\gamma_{\rm p}})_{\rm F(S)}$ and $(\theta_{\gamma_{\rm p}})_{\rm F(S)}$ are the pulse-on relaxation rate constants for the damping factor and occupancy of the two different bath components, respectively, and $(\delta_{\rm b})_{\rm F(S)}$ is the F(S) fraction of the hot bath. $\gamma_{\rm p}[n_{\rm c,RO}]$ and $n_{\rm p}[n_{\rm c,RO}]$ are the steady-state bath values

reached at the end of the optical readout pulse. The corresponding rate equations for the hot bath in the pulse-off state are,

$$(\dot{\gamma_{\rm p}})_{\rm F(S)}(\tau) = -(\zeta_{\gamma_{\rm p}})_{\rm F(S)} \left\{ (\gamma_{\rm p})_{\rm F(S)}(\tau) - (\delta_{\rm b})_{\rm F(S)} \gamma_{\rm p}[n_{\rm c,RO}] \right\}, \tag{D21}$$

and

$$(\dot{n_{\rm p}})_{\rm F(S)}(\tau) = -(\zeta_{n_{\rm p}})_{\rm F(S)} \left\{ (n_{\rm p})_{\rm F(S)}(\tau) - (\delta_{\rm b})_{\rm F(S)} n_{\rm p}[n_{\rm c,RO}] \right\}. \tag{D22}$$

where τ is the time from the end of the optical pulse, and $(\zeta_{\gamma_{\rm p}})_{\rm F(S)}$ and $(\zeta_{\gamma_{\rm p}})_{\rm F(S)}$ are the pulse-off relaxation rate constants for the damping factor and occupancy of the two different bath components, respectively.

The model parameters used to fit the specific measured data for the six-shield device (device B) presented in Fig. 12 are listed in Table II. Similar bath dynamical parameters are found for all of the measured devices we have studied. Independent of the optical readout pulse power, the fraction of the bath which reacts quickly seems to be consistently close to a value of $(\delta_{\rm b})_F = 0.65$. The fast component of the bath turns on faster than we can resolve ($\gtrsim 50$ MHz), while slow component of the bath turns on with a rate constant of approximately $\theta_S/2\pi = 600$ kHz (for both damping factor and occupancy). The fast component of the bath turns off with an exponential rate constant of $(\zeta_{\gamma_{\rm p}})_F/2\pi = 150$ kHz for $\gamma_{\rm p}$ and $(\zeta_{n_{\rm p}})_F/2\pi = 70$ kHz for $n_{\rm p}$. Even more slowly, the slow component of the bath turns off with a rate constant of $(\zeta_{\gamma_{\rm p}})_S/2\pi = 90$ kHz and $(\zeta_{n_{\rm p}})_S/2\pi = 24$ kHz for the two different bath factors.

Our ability to measure the bare damping rate of the acoustic breathing mode relies on the fact that the hot bath evaporates prior to the actual measurement of the free decay of the breathing mode. This means that the first $\sim 10~\mu$s of the pulse-off state is dead time in which the dynamics of the breathing mode occupancy is still coupled to that of the hot bath. Crucial to the measurement of a ringdown curve using the red-detuned optical pulse as both a readout signal and an excitation source, is that after this dead time there remain a residual, elevated phonon occupancy of the breathing mode from which the mode can decay. This is clearly the case for the data measured in Fig. 12, and is a result of the fact that at this readout power the peak magnitude of $\gamma_{\rm p}$ ($2\pi(85$ kHz)) is still smaller than the fastest decay of the hot bath ($(\zeta_{\gamma_{\rm p}})_F/2\pi = 150$ kHz), so that the breathing mode occupancy cannot follow that of the fast dynamics of the hot bath. This non-adiabatic quenching leaves the breathing mode with an elevated occupancy after the dead time. At readout powers beyond $n_{\rm c} = 1000$ this stops being the case, hence our choice of readout pulse powers $n_{\rm c} \lesssim 600$ in the ringdown measurements.

A few observational comments are warranted. The fact that the hot bath should have faster pulse-on rate constants than pulse-off rate constants might be explained by the fact that there are likely a wide spectrum of phonons which are created by absorption of the optical pulse. This may lead to a hierarchy of phonon baths. Consider for instance a two bath scenario, consisting of a high and a low frequency phonon bath. The high frequency bath is assumed to be directly populated from optical absorption events, while the low frequency bath is predominantly responsible for coupling to the breathing mode of interest. In the high frequency phonon bath, phonons rapidly mix with each other due the large density of states and mode occupancy. The high frequency phonon bath is also well thermalized to the external substrate through acoustic radiation. Phonons in the low frequency bath are fed from the phonon-phonon scattering processes in the high frequency phonon bath, and are less connected via radiation to the substrate. When the optical pulse is on, the high frequency bath is rapidly populated. The high frequency bath not only acts as a source of phonons for the low frequency bath, but through nonlinear phonon mixing also helps bring it into some quasi-equilibrium temperature. When the optical pulse is turned off, the high frequency bath rapidly decays away, leaving the low frequency bath of phonons to more slowly decay away due to the absence of the phonons in the high frequency bath to mix with. This scenario would also explain the difference in the decay of the low frequency bath $\gamma_{\rm p}$ damping rate, which depends on the phonon number density, to that of the effective occupancy $n_{\rm p}$, which is set by the quasi-equilibrium temperature of the bath. The absence of the high frequency bath could greatly slow down the low frequency bath equilibriation rate, and thus the rate of change of the effective bath temperature, while the low frequency bath acoustic coupling to the external substrate will provide a constant decay channel for bath phonons and thus $\gamma_{\rm p}$.

We should further note that the dynamical bath parameters reported in Table II are consistent for devices fabricated from low resistivity SOI. In the case of our high resistivity SOI samples, we have measured devices with much slower post-read-pulse decay of the hot bath. Hot bath decay times as long as tens of milliseconds have been observed. Although requiring further study, we believe that this very slow decay dynamics of the optical-absorption-induced bath indicate that phonons are not the only parties involved in the optically-induced hot bath, but that the hot bath is likely also composed of much longer lifetime two-level system (TLS) defects. The high resistivity SOI seems to harbor much longer-lived TLS states, possibly due to the reduction of electronic relaxation pathways. Our two-phonon-bath scenario above may in fact be a two-bath scenario consisting of one phonon bath coupled to a longer lived TLS bath.

TABLE II. **Dynamical model parameters of the optical-absorption-induced bath.** These parameters were used to fit the measured phonon occupancy shown in Fig. 12 during the entire pulse on period and during the transient initial millisecond after the pulse is turned off. The specific device measured was a six-shield device, device B in Table I, fabricated from low resistivity SOI (ρ = 5-20 Ω-cm).

Parameter	Description	Value	Refs./Notes
$(\delta_{\rm b})_{\rm F(S)}$	fast (slow) fractional component of the optically-induced bath	0.65 (0.35)	from model fit
$(\theta_{n_{\rm p}})_{\rm F(S)}/2\pi$	$n_{\rm p}$ exponential relaxation rate with optical *pulse on*	$\gtrsim$ 50 MHz (600 kHz)	from model fit
$(\theta_{\gamma_{\rm p}})_{\rm F(S)}/2\pi$	$\gamma_{\rm p}$ exponential relaxation rate with optical *pulse on*	$\gtrsim$ 50 MHz (600 kHz)	from model fit
$(\zeta_{n_{\rm p}})_{\rm F(S)}/2\pi$	$n_{\rm p}$ exponential relaxation rate with optical *pulse off*	70 kHz (24 kHz)	from model fit
$(\zeta_{\gamma_{\rm p}})_{\rm F(S)}/2\pi$	$\gamma_{\rm p}$ exponential relaxation rate with optical *pulse off*	150 kHz (90 kHz)	from model fit
$n_{\rm c,RO}$	readout pulse amplitude	569 photons	
$\gamma_{\rm OM}/2\pi$	optomechanical back-action rate during *pulse on*	1.4 MHz	measured ind.
$n_{\rm p}$	steady-state hot bath occupancy for $n_{\rm c,RO}$ intra-cavity photons	73 phonons	measured ind.
$\gamma_{\rm p}$	steady-state hot bath damping rate for $n_{\rm c,RO}$ intra-cavity photons	85 kHz	measured ind.
τ	delay time between successive readout pulses	200 μs	
$T_{\rm pulse}$	readout pulse length	4 μs	
$n_{\rm m}^{\rm f}$	breathing mode occupancy at end ($t = 4\ \mu$s) of optical readout pulse	4.2 phonons	measured
$n_{\rm m}^{\rm i}$	breathing mode occupancy in first measurement bin of readout pulse	13.6 phonons	measured
$n_{\rm m}[0]$	breathing mode occupancy referred back to $t = 0$ of optical readout pulse	27 phonons	from model fit

Other evidence for this interpretation is the high values of measured $\gamma_{\rm p}$ which is more consistent with estimated TLS damping rates (damping due to three-phonon scattering is shown to be too slow, at least for bath temperatures below 1 K, in sub-Section H 1).

Appendix E: Coherent Excitation Methods: Blue-Detuned Pumping and RF-Modulation Amplification

1. Low-Threshold Acoustic Self-Oscillation

Owing to the extremely slow intrinsic damping rate γ_0 observed in the ultra-high-Q nanobeam devices at low temperature, it is possible to drive the mechanics into the regime of self-sustained oscillations with a blue-detuned pumping laser at very low input optical powers, or equivalently, a very low rate of measurement back-action. The total effective damping rate experienced by the mechanics in the presence of a blue-detuned drive laser is $\gamma = \gamma_{\rm i} - \gamma_{\rm OM}$, where the intrinsic damping rate $\gamma_{\rm i} = \gamma_0 + \gamma_{\rm p}$ includes damping γ_0 from both the cold fridge bath (with occupancy $n_0 \approx 10^{-3}$) and from the optical absorption-induced phonon bath at rate $\gamma_{\rm p}$. The usual condition for self-oscillation is that the damping rate is matched by the back-action amplification rate $\gamma_{\rm OM}$:

$$\gamma_{\rm OM} = \gamma_0 + \gamma_{\rm p}. \tag{E1}$$

We observe the onset of mechanical self-oscillation at $T_{\rm f} = 10$ mK, in which a CW blue-detuned pump laser drives the cavity and the sideband filters are aligned to the cavity resonance ($\Delta = 0$). The scattered photon count rate $\Gamma_{\rm SB,0}$ is measured in steady-state. In the setup configuration used for these measurements, an additional VOA is placed in the optical path, elevating the measured SPD dark count rate to 10.8 c.p.s. Sweeping the input power (photon number) $n_{\rm c}$ results in a sharp increase in detected count rate at the self-oscillation threshold $n_{\rm c,thresh} = 2 \times 10^{-3}$ as shown in Figure 13, where we estimate the resulting steady-state phonon occupancy to be of order $n_{\rm m} \sim 5 \times 10^4$. At the threshold $n_{\rm c,thresh}$ we can estimate the back-action amplification rate $\gamma_{\rm OM}/2\pi = 4g_0^2 n_{\rm c,thresh}/\kappa \approx 8$ Hz from the known optical device parameters, indicating that the intrinsic damping $\gamma_{\rm i}$ is dominated by the bath damping rate: $\gamma_{\rm p}(n_{\rm c,thresh}) = \gamma_{\rm OM}(n_{\rm c,thresh}) - \gamma_0 \approx 2\pi(7.9$ Hz$)$, in good quantitative agreement with the trend measured on a similar device in Figure 10a. Upon decreasing the driving power (green data in Figure 13), self-oscillation appears to relax at a decreased threshold of $n_{\rm c} = 1.4 \times 10^{-3}$, indicating a hysteresis in the measured count rates as a function of input power. This apparent hysteresis likely arises from a change in the *true* intracavity photon number as a function of driving power $P_{\rm in}$. We have so far adhered to Equation C1 in determining the $n_{\rm c}$ as a function of $P_{\rm in}$; this expression is used to generate the horizontal axis of Figure 13, and so does not represent the true intracavity photon number. However, in order to unambiguously calculate $n_{\rm c}$ in the presence of large phonon amplitude $n_{\rm m}$, a more thorough calculation is needed which accounts for the effective optical reflection profile in the presence of strong modulation by the mechanical motion.
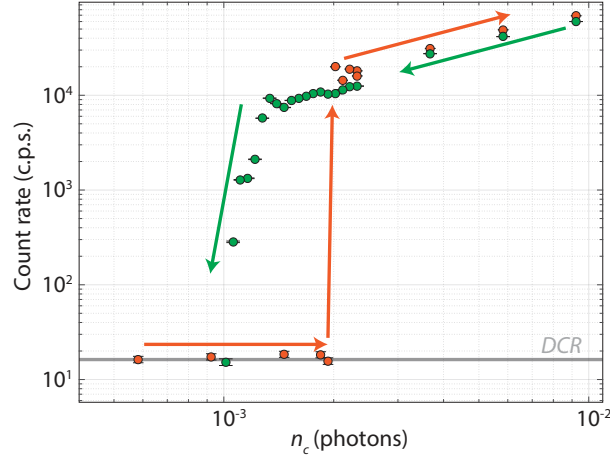
FIG. 13. **Low-temperature measurement of the self-oscillation threshold in a high-Q nanobeam at $T_{\mathrm{f}} = 10$ mK.** Under blue-detuned ($\Delta = -\omega_m$) driving the self-oscillation threshold is reached when $\gamma_{\mathrm{OM}} = \gamma_{\mathrm{i}}$, here for $n_{\mathrm{c,thresh}} = 2 \times 10^{-3}$, or $\gamma_{\mathrm{OM}}/2\pi = 8$ Hz, for increasing optical power (orange points). Measurements were performed on device D.
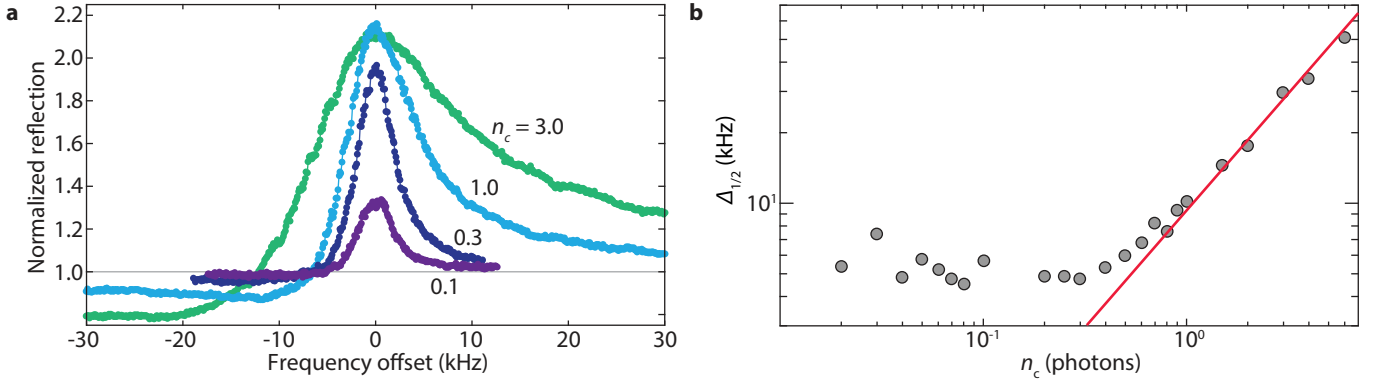

FIG. 14. **Mechanical mode time-averaged linewidth versus probe power. a,** Time-averaged mechanical mode linewidth as a function of optical pump power. Broadening of the linewidth is due to optical back-action. Solid line is a fit to the back-action damping rate $\gamma_{\mathrm{OM}} = 4g_0^2 n_{\mathrm{c}}/\kappa$, from which we extract an estimated $g_0/2\pi = 1.15$ MHz. Measurements are performed on Device D.

2. Electromagnetically Induced Transparency Mechanical Spectroscopy

Electromagnetically induced transparency (EIT) in optomechanical systems allows for a spectral measurement of the mechanical response via observation of a transparency window in the optical reflection spectrum. A pump laser tone at ω_{c} is amplitude modulated to generate a weak probe tone at $\omega_{s,\pm} = \omega_{\mathrm{c}} \pm \Delta_{\mathrm{p}}$. If the pump-cavity detuning is fixed on either the red- or blue-side of the optical cavity ($\Delta = \pm\omega_{\mathrm{m}}$), the optical susceptibility of the cavity strongly suppresses one of the probe sidebands (at $\omega_{s,\mp}$) and only the other probe sideband will have an appreciable intracavity population. For a red-detuned pump, the interaction of the pump tone and mechanics with the probe sideband yields a reflection coefficienct $r(\Delta, \delta)$ for the probe which contains a transparency window having a width on the scale of the mechanical mode linewidth:

$$r(\Delta, \delta) = 1 - \frac{\kappa_{\mathrm{e}}}{\kappa/2 + i(\Delta - (\delta + \omega_{\mathrm{m}})) + \frac{|G|^2}{-i\delta + \gamma_{\mathrm{i}}/2}}, \tag{E2}$$

where we have defined $\delta \equiv \Delta_{\mathrm{p}} - \omega_{\mathrm{m}}$ and $G \equiv g_0\sqrt{n_{\mathrm{c}}}$. We measure the reflection amplitude $R = |r|^2$ by driving an EOM weakly to generate a probe tone and observing the count rates of sideband-scattered probe photons. The pump is locked at $\Delta = +\omega_{\mathrm{m}}$ and the cascaded filter stack is locked to the cavity frequency. The RF modulation power is
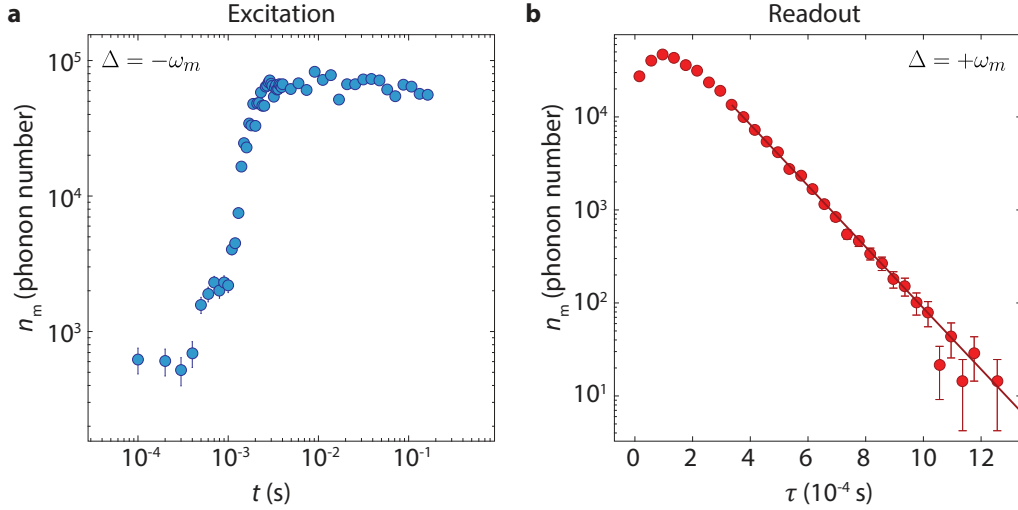
FIG. 15. **Measurement of the mode occupancy during excitation and readout pulses for high-amplitude ringdown. a**, A blue-detuned ($\Delta = -\omega_m$) laser with pulse-on photon number $n_{c,blue} = 0.15$ drives the mechanics into self-oscillation in a timescale of 2 ms, with a saturated phonon occupancy $n_m \approx 5 \times 10^4$. **b**, A red-detuned ($\Delta = +\omega_m$) probe laser with pulse-on photon number $n_{c,red} = 0.30$ serves to reads out the mode occupancy after a variable delay time τ. Measurements are performed on Device D.

chosen to generate a sideband intracavity photon number much smaller than the carrier photon number ($n_{c,+} \ll n_c$) while maintaining a large count rate $\sim 10^5$ c.p.s. at the SPDs to minimize data integration times. This corresponds to modulation indices in the range of $\beta \sim 10^{-3}$ for our system parameters (measurements were performed on device D). The modulation frequency Δ_p is swept over a range of about 1 MHz to map out the transparency window. This range is large enough to include the optomechanically-broadened mechanical linewidth which sets the width of the transparency window, but much narrower than the bandwidth of the FFP filters (≈ 50 MHz), allowing for the filters to be stably locked at a single position in the center of the optical cavity line. Figure 14 shows the normalized reflection level for various optical probe power levels n_c, as well as fits to the data. The extracted total mechanical linewidth $\gamma = \gamma_i + \gamma_p + \gamma_{OM}$ is plotted in Fig. 14. At low probe-power, $\gamma/2\pi$ saturates to a value ≈ 4 kHz, which represents time-averaged broadening of the intrinsic mechanical linewidth due to mechanical frequency jitter as summarized in Fig. 3 of the Main Text. With κ and n_c calibrated, the linear portion of the curve which is dominated by back-action damping is fitted to extract the optomechanical coupling rate $g_0/2\pi = 1.15$ MHz.

3. Blue-Detuned Pumping.

In the limit of high phonon amplitude, we perform ringdown using a pulse sequence consisting of a blue-detuned excitation pulse followed by a red-detuned readout (or *probe*) pulse. Two separate laser sources are used, as shown in Figure 2, for generating the excitation and readout pulses in order to allow a fixed detuning of each laser to avoid instabilities associated with rapid stabilization of the laser frequency on the timescale of the pulse sequencing (tens of μs). Owing to the extremely narrow instantaneous mechanical mode linewidth, a very small back-action amplification rate $\gamma_{OM}/2\pi \lesssim 8$ Hz is required to drive the mechanics into the self-oscillation regime. This enables operation at low driving pulse photon number $n_{c,blue} = 0.15 \gg n_{c,thresh} = 2 \times 10^{-3}$, in order to minimize the effective temperature and coupling rate of the absorption-induced phonon bath. The steady-state phonon amplitude in the presence of the driving pulse is saturated to $n_m \approx 5 \times 10^4$. The driving pulse is turned off and after a variable delay time τ a red-detuned pulse from the readout laser is used to probe the mode occupancy.

The readout photon number $n_{c,red} = 0.30$ is again chosen small to minimize absorption bath effects, as well as to give a total count rate $\Gamma \propto \Gamma_{SB,0} n_m \propto n_c n_m$ within the dynamic range of the single-photon detector, which in our amplifier setup has sensitivity to a maximum count rate of $\sim 2 \times 10^6$ c.p.s. With the present setup efficiencies and device parameters (see caption of Figure 15), the detected count rate is approximately $\Gamma = 31$ c.p.s. per phonon per photon at $\Delta = \pm\omega_m$, and the resulting effective upper bound on $n_{c,red}$ at which the SPD can efficiently detect is 2.2 photons. Now, after the readout pulse is used to probe the mode occupancy, the mode occupancy is cooled via dynamical back-action to near its local bath temperature in preparation for the subsequent series of driving and

readout pulses (effectively 're-setting' the measurement). In practice, a single red-detuned pulse is used for both readout and cooling (re-setting). In Figure 15 we show the phonon occupancy during both the driving pulse and the readout pulse. Note that the excited occupancy saturates to $n_m \approx 5 \times 10^4$ from an initial occupancy $n_m \lesssim 1 \times 10^3$, corresponding to our estimated decay ratio of 50 from one pulse period to the next.

4. Modulated Pump-Probe Excitation and Ringdown.

For finer control of the mechanical mode amplitude during excitation, microwave-frequency modulation of the excitation pulse was used to amplify the mechanics to a fixed phonon amplitude which is tunable by the depth of optical modulation placed on the pump laser tone. This technique allows us to probe the intrinsic energy decay constant γ_0 in the regime lying intermediate between the level of single-phonons and the saturated high-phonon-amplitude limit of self-oscillation. In this measurement scheme, a radio-frequency (RF) signal generator is used to drive an electro-optic intensity modulator (EOM) at the mechanical resonance frequency $\omega_m/2\pi$ to generate the probe sideband. The excitation pulse consists of a red-detuned pump carrier tone which is weakly modulated (RF driving power -4 dBm applied to an EOM with $V_\pi = 4.1$ V, giving a modulation index $\beta = 0.11$) to generate a probe sideband at the cavity resonance frequency. Interference between the pump carrier and probe sideband generates a time-dependent radiation pressure force at the difference frequency $\omega_m/2\pi$, which resonantly excites the acoustic mode. A second pulsed laser source is then used to generate the readout optical pulse, which is a red-detuned pulse of fixed frequency and power. Using this method as shown in Fig. 2, we measure an initial excited mechanical occupation of $n_m = 1.2 \times 10^3$ having a decay rate $\gamma_0/2\pi = 0.108$ Hz at the lowest temperatures $T_f = 10$ mK, corresponding to a mechanical Q-factor of 4.92×10^{10} ($\tau_{\mathrm{coh},0} = 1.47$ s). The decay is observed to be exponential over at least 1.5 orders of magnitude, consistent with measurements in both the high- and low-phonon number limit.

In both the case of blue-detuned driving and RF-modulated driving ring-up techniques, the total repetition rate of the pulse sequence $1/T_{\mathrm{per}}$ is fixed while only the variable delay τ between the driving pulse and readout pulse is varied. This fixing of the overall duty-cycle of the pulse sequence is performed to eliminate systematic variations in the local absorption-induced bath temperature T_b which, in steady-state, is expected to depend on the average power circulating in the cavity (see Fig. 10). We find that the measured ringdown time constant is approximately consistent over more than three orders of magnitude in starting phonon population, from $n_m \lesssim 10$ in the case of thermally-excited ringdown measurements to $n_m > 2 \times 10^4$ in the case of coherently-excited phonon populations.

Appendix F: 3-phonon-scattering damping model.

The anharmonicity of the atomic lattice in solid-state materials leads to frequency mixing of the approximate harmonic modes (the phonons) of the lattice. This frequency mixing - of all different orders - within a continuum of modes leads to different forms of phonon damping depending on the damped phonon frequency (ω_s), wavelength (λ_{q_s}), and the lattice temperature [23, 24]. At low temperatures where phonon relaxation times (τ) are long, and at relatively high phonon frequencies, the dominant source of phonon damping due to the anharmonic lattice potential results from 3-phonon scattering processes in the so-called Landau-Rumer limit ($\omega_s\tau \gg 1$) [22, 54]. In this limit a single-mode relaxation time (SMRT) approximation [23] can be made in which only the damped phonon mode under consideration is disturbed from equilibrium and the other two phonon modes involved in the scattering are assumed to be frozen at their equilibrium occupancies. Using the SMRT approximation one can calculate the 3-phonon-scattering damping rate from second-order perturbation theory of the quantum mechanical model of the anharmonic lattice. At higher temperatures where the thermal phonon relaxation rate ($1/\tau$) is very fast, or for very low frequency phonons, this approximation breaks down and one enters the Akhiezer limit of phonon damping where $\omega_s\tau \ll 1$ [55]. In this limit a phenomenological model is employed in which the strain wave of a phonon mode induces a redistribution of thermal phonons via the lattice anharmoniticity, and damping occurs due to relaxation of the thermal phonons back towards thermal equilibrium. If in addition the phonon wavelength is long relative to the mean free path of thermal phonons (l_{th}), then a local temperature can be defined and damping can also occur via diffusion of thermal phonons. In this limit, $\omega_s\tau \ll 1$ and $l_{th}/\lambda_{q_s} \ll 1$, energy in the acoustic wave is carried away in heat flow due to temperature gradients on the scale of λ_{q_s}, and the resulting relaxation process is called thermoelastic damping [56].

As we are concerned with microwave frequency acoustic waves and sub-Kelvin temperatures, the dominant phonon-phonon scattering damping is expected to arise from 3-phonon scattering under the SMRT approximation. In what follows we present a model of such Landau-Rumer damping utilizing leaky quasi-modes [57]. This quasi-mode picture arises naturally in the context of the OMC cavity structure, in which localized acoustic modes are weakly coupled to the continuum of phonon modes in the surrounding substrate via the peripheral clamping of the Si device layer to the underlying Si dioxide BOX layer (c.f., Fig. 17). We follow closely the derivation of 3-phonon scattering in Ref. [23],

although with slight adjustments to the notation to accommodate the use of quasi-normal modes. The notation developed here will also be used in the analysis of two-level system damping described in the next sub-section.

For a displacement vector field $u_\alpha(\mathbf{r})$, with u the local amplitude of atomic displacement in direction α from equilibrium, the stored potential energy to second and third order in the displacement field can be written as,

$$\mathcal{V}_2 \equiv \text{2nd-order elastic (potential) energy} = \frac{1}{2}\int \mathrm{d}^3 r J^{\alpha\gamma}_{\ \ \beta\delta} \frac{\partial u_\alpha}{\partial r_\beta}\frac{\partial u_\gamma}{\partial r_\delta}, \tag{F1}$$

and

$$\mathcal{V}_3 \equiv \text{third-order elastic (potential) energy} = \frac{1}{3!}\int \mathrm{d}^3 r A^{lmn}_{\ \ ijk}\frac{\partial u_l}{\partial r_i}\frac{\partial u_m}{\partial r_j}\frac{\partial u_n}{\partial r_k}. \tag{F2}$$

Here, $J^{\alpha\gamma}_{\ \ \beta\delta}$ is in general a rank 4 tensor whose coefficients are the 2nd-order elastic coefficients of the material which relate strain to stress and have units of energy density. $A^{lmn}_{\ \ ijk}$ is rank 6 tensor with coefficients arising from the lowest order anharmoniticity of the lattice. $\partial u_{\alpha/\partial r_\beta}$ is a rank 2 tensor representing the local strain created by the displacement vector field $u_\alpha(\mathbf{r})$.

From these expression we can define the total elastic energy density for a classical acoustic wave oscillating harmonically in mode s as,

$$h(\bar{\bar{e}}_s(\mathbf{r})) \equiv \text{(classical) strain field elastic energy density} = \frac{1}{2}J^{\alpha\gamma}_{\ \ \beta\delta}\left(e_s(\mathbf{r})\right)_\alpha^{\ \beta}\left(\left(e_s(\mathbf{r})\right)_\gamma^{\ \delta}\right)^*, \tag{F3}$$

where $\bar{\bar{e}}_s(\mathbf{r})$ is a complex strain tensor field related to the real (physical) strain tensor field by,

$$\frac{\partial\left(u_s(\mathbf{r})\right)_i}{\partial r_j} \equiv \text{(classical) strain tensor field of mode } s \equiv \mathrm{Re}\left(\left(e_s(\mathbf{r})\right)_i^{\ j}\right) = \mathrm{Re}\left(\bar{\bar{e}}_s(\mathbf{r})\right). \tag{F4}$$

Note that we have used the fact that for a harmonic wave the cycle averaged potential and kinetic energies are equal (and thus the total wave energy is twice the potential energy), and $h(\bar{\bar{e}}_s(\mathbf{r}))$ should therefore be strictly considered as the energy density averaged over a single cycle in time and a single wavelength in space. We also define a normalized complex strain field for mode s having a peak strain value of approximately unity (exactly unit for tensor-averaged fields) and a peak energy density of $\bar{J}$,

$$\bar{\bar{\bar{e}}_s(\mathbf{r})} \equiv \text{normalized classical strain field for mode } s = \frac{(\bar{J})^{1/2}\bar{\bar{e}}_s(\mathbf{r})}{(\max[h(\bar{\bar{e}}_s(\mathbf{r}))])^{1/2}}, \tag{F5}$$

where $\bar{J}$ is the tensor-average of the harmonic elastic coefficients,

$$\bar{J} \equiv \text{tensor-averaged 4th-order elastic tensor} = \left\langle \frac{1}{2}\bar{\bar{\bar{J}}}\right\rangle_{\mathrm{t}} = \frac{1}{2\cdot 3^4}\left(\sum_{\alpha,\gamma,\beta,\delta}\left(J^{\alpha\gamma}_{\ \ \beta\delta}\right)^2\right)^{1/2}. \tag{F6}$$

The effective mode volume over which the strain energy of mode s is localized can also be defined as,

$$V_s \equiv \text{effective mode volume of mode } s = \frac{\int h(\bar{\bar{e}}_s(\mathbf{r}))\mathrm{d}^3 r}{\max[h(\bar{\bar{e}}_s(\mathbf{r}))]}. \tag{F7}$$

The peak strain amplitude of mode s containing half a quanta of energy, i.e., the 'vacuum' strain level, is given by,

$$e_{\mathrm{vac},s} \equiv \text{vacuum strain field amplitude for mode } s = \sqrt{\frac{\hbar\omega_s}{2\bar{J}V_s}}, \tag{F8}$$

where $\hbar\omega_s$ is the mode s energy quantum. From the peak strain amplitude of vacuum and the normalized strain field we can define a quantum strain field operator,

$$\hat{\bar{\bar{e}}}_s(\mathbf{r}) \equiv \text{quantum strain tensor field operator for mode } s = (e_{\text{vac},s}) \left[\hat{b}_s \overline{\bar{e}_s(\mathbf{r})} + \hat{b}_s^\dagger \left(\overline{\bar{e}_s(\mathbf{r})} \right)^* \right] \qquad \text{(F9)}$$

where $\hat{b}_s$ and $\hat{b}_s^\dagger$ annihilate and create individual phonon quanta in mode s. The corresponding quantum interaction Hamiltonian for 3-phonon scattering can then be written in terms of triplets of quantum strain field operators directly from the third-order elastic potential energy relation in Eq. (F2),

$$\hat{\mathcal{H}}_{\text{3-ph}} \equiv \text{3-phonon interaction Hamiltonian} = \frac{1}{3!} \sum_{s\,s'\,s''} \int \mathrm{d}^3 r \left(A^{lmn}{}_{ijk} \right) \left(\hat{\bar{\bar{e}}}_s(\mathbf{r}) \right)^i_l \left(\hat{\bar{\bar{e}}}_{s'}(\mathbf{r}) \right)^j_m \left(\hat{\bar{\bar{e}}}_{s''}(\mathbf{r}) \right)^k_n. \qquad \text{(F10)}$$

1. Type-I scattering processes

3-phonon scattering, as it pertains to damping of a particular mode s, can be categorized into two classes of processes [23]. Type-I scattering involves the mode of interest, mode s, as a 'daughter' phonon which combines with another 'sibling' phonon (mode s') to create a higher frequency 'parent' phonon (mode s''). The reverse process is also of type-I. Type-II scattering has the mode s of interest as the high frequency parent phonon. The 3-phonon interaction Hamiltonian for type-I scattering is given in terms of phonon creation and annihilation operators as,

$$\hat{\mathcal{H}}^{\text{3-ph}}_{s+s' \rightleftharpoons s''} \equiv A^{s''}_{ss'} \hat{b}_s \hat{b}_{s'} \hat{b}_{s''}^\dagger + A^{ss'}_{s''} \hat{b}_s^\dagger \hat{b}_{s'}^\dagger \hat{b}_{s''}, \qquad \text{(F11)}$$

where

$$A^{s''}_{ss'} = \left(A^{ss'}_{s''} \right)^* \equiv \left[(e_{\text{vac},s})(e_{\text{vac},s'})(e_{\text{vac},s''}) \right] \int \left(A^{lmn}{}_{ijk} \right) \left(\overline{\bar{e}_s(\mathbf{r})} \right)^i_l \left(\overline{\bar{e}_{s'}(\mathbf{r})} \right)^j_m \left(\left(\overline{\bar{e}_{s''}(\mathbf{r})} \right)^* \right)^k_n \mathrm{d}^3 r. \qquad \text{(F12)}$$

Calculating to 2nd-order in perturbation theory, the energy shift in the phonon Fock state $|n_s, n_{s'}, n_{s''}\rangle$ is given by,

$$\left\langle \left(\delta E_{n_s, n_{s'}, n_{s''}} \right)^{\text{I}}_{\text{3-ph}} \right\rangle = \sum_{s' s''} \left[\frac{\left| \left\langle n_s - 1, n_{s'} - 1, n_{s''} + 1 | A^{s''}_{ss'} \hat{b}_s \hat{b}_{s'} \hat{b}_{s''}^\dagger | n_s, n_{s'}, n_{s''} \right\rangle \right|^2}{\hbar \left((\omega_s + \omega_{s'} - \omega_{s''}) - i(\Gamma_s + \Gamma_{s'} - \Gamma_{s''}) \right)} \right.$$
$$\left. + \frac{\left| \left\langle n_s + 1, n_{s'} + 1, n_{s''} - 1 | A^{ss'}_{s''} \hat{b}_s^\dagger \hat{b}_{s'}^\dagger \hat{b}_{s''} | n_s, n_{s'}, n_{s''} \right\rangle \right|^2}{\hbar \left((\omega_{s''} - \omega_s - \omega_{s'}) - i(\Gamma_{s''} - \Gamma_s - \Gamma_{s'}) \right)} \right]. \qquad \text{(F13)}$$

Note that we have used a modified form of non-Hermitian perturbation theory [57, 58], suitable for leaky quasi-normal modes, in which the finite linewidth of the phonon quasi-modes are included in the denominator of Eq. (F13). Also implicit in our use of 2nd-order perturbation theory is that the phonon-phonon coupling is weak. Collecting terms and specifically identifying mode s as the breathing acoustic mode of the OMC cavity (labeled by m), we have for the complex level shift,

$$\left\langle \left(\delta E_{n_m, n_{s'}, n_{s''}} \right)^{\text{I}}_{\text{3-ph}} \right\rangle = \sum_{s' s''} \left(\frac{\left| A^{s''}_{m\,s'} \right|^2 \left[n_m n_{s'} (n_{s''} + 1) - (n_m + 1)(n_{s'} + 1) n_{s''} \right]}{\hbar \left((\omega_{\text{m}} + \omega_{s'} - \omega_{s''}) - i(\Gamma_m + \Gamma_{s'} - \Gamma_{s''}) \right)} \right). \qquad \text{(F14)}$$

We now invoke the single mode relaxation time approximation, and assume that only mode m is perturbed from equilibrium,

$$\left\langle (\delta E_{m,s',s''})^{\mathrm{I}}_{\text{3-ph}} \right\rangle_{\text{smrt}} = \sum_{s's''} \left(\frac{\left| A^{s''}_{m\,s'} \right|^2 [(\bar{n}_m + \delta n_m)\bar{n}_{s'}(\bar{n}_{s''} + 1) - (\bar{n}_m + \delta n_m + 1)(\bar{n}_{s'} + 1)\bar{n}_{s''}]}{\hbar \left((\omega_{\mathrm{m}} + \omega_{s'} - \omega_{s''}) - i(\Gamma_m + \Gamma_{s'} - \Gamma_{s''}) \right)} \right), \tag{F15}$$

where $\bar{n}$ are the thermal equilibrium mode occupancies and δn_m is the perturbation in phonon number of the breathing mode from equilibrium. Taking the difference between the complex level shift for $\delta n_m + 1$ and δn_m, we find for the single photon energy shift in mode m,

$$\left\langle \hbar\,(\delta\tilde{\omega}_m)^{\mathrm{I}}_{\text{3-ph}} \right\rangle_{\text{smrt}} = \sum_{s's''} \left(\frac{\left| A^{s''}_{m\,s'} \right|^2 [\bar{n}_{s'} - \bar{n}_{s''}]}{\hbar \left((\omega_{\mathrm{m}} + \omega_{s'} - \omega_{s''}) - i(\Gamma_m + \Gamma_{s'} - \Gamma_{s''}) \right)} \right). \tag{F16}$$

Assuming the phonon mode linewidths are energy-damping limited ($\Gamma_s = \gamma_s/2$), we find for the change in the energy damping rate of mode m due to type-I 3-phonon scattering,

$$\left\langle (\delta\gamma_m)^{\mathrm{I}}_{\text{3-ph}} \right\rangle_{\text{smrt}} \equiv \frac{-2\mathrm{Im}\left\langle \hbar\,(\delta\tilde{\omega}_m)^{\mathrm{I}}_{\text{3-ph}} \right\rangle_{\text{smrt}}}{\hbar\delta n_m} = \frac{1}{\hbar^2} \sum_{s's''} \left(\frac{\left| A^{s''}_{m\,s'} \right|^2 (\gamma_{s''} - \gamma_{s'} - \gamma_m)[\bar{n}_{s'} - \bar{n}_{s''}]}{(\omega_{\mathrm{m}} + \omega_{s'} - \omega_{s''})^2 + (\frac{\gamma_m + \gamma_{s'} - \gamma_{s''}}{2})^2} \right)$$

$$\approx \frac{1}{\hbar^2} \sum_{s's''} \left(\frac{\left| A^{s''}_{m\,s'} \right|^2 (\gamma_{s''})[\bar{n}_{s'} - \bar{n}_{s''}]}{(\omega_{\mathrm{m}} + \omega_{s'} - \omega_{s''})^2 + (\frac{\gamma_m + \gamma_{s'} - \gamma_{s''}}{2})^2} \right), \tag{F17}$$

where in the last approximate equality we have neglected the unperturbed damping of the two 'child' phonons (m and s') and only included the quasi-mode damping of the 'parent' phonon (s'') into which the mode m decays in the type-I process.

Equation (F17]) can be approximately evaluated using the relation between the 'mode-averaged' ($\langle \cdot \rangle_{\mathrm{m}}$) third-order elastic constants and the Grüneisen parameter [23],

$$\left\langle \left| A^{lmn}{}_{ijk} \right|^2 \right\rangle_{\mathrm{m}} \equiv \text{mode-averaged three-phonon scattering strength} \approx 4\rho_{\mathrm{Si}}^2 v_{\mathrm{Si}}^4 \gamma_{\mathrm{G}}^2 \tag{F18}$$

where mode averaging is taken over different bulk phonon mode directions and polarizations. This allows us to write for the 3-phonon scattering amplitude,

$$A^{s''}_{ms'} \approx \left(2\rho_{\mathrm{Si}} v_{\mathrm{Si}}^2 \gamma_{\mathrm{G}} \right) [(e_{\mathrm{vac},m})(e_{\mathrm{vac},s'})(e_{\mathrm{vac},s''})] \left(\mathcal{F}^{s''}_{m\,s'} V_m \right), \tag{F19}$$

in which $\mathcal{F}^{s''}_{m\,s'}$ is a mode overlap factor, or equivalently, for plane wave modes, a phase-matching term. The mode overlap factor is less than or equal to unity and depends approximately upon the tensor-averaged normalized strain fields of each of the three modes participating in the scattering process,

$$\mathcal{F}^{s''}_{m\,s'} = \frac{1}{V_m} \int \left(\left\langle \bar{\bar{e}}_m(\mathbf{r}) \right\rangle_{\mathrm{t}} \right) \left(\left\langle \bar{\bar{e}}_{s'}(\mathbf{r}) \right\rangle_{\mathrm{t}} \right) \left(\left\langle \bar{\bar{e}}_{s''}(\mathbf{r}) \right\rangle_{\mathrm{t}} \right)^* \mathrm{d}^3 r \lesssim 1. \tag{F20}$$

As with the elastic constants, we define a tensor-averaged strain field as,

$$\left\langle \bar{\bar{e}}_s(\mathbf{r}) \right\rangle_{\mathrm{t}} \equiv \text{tensor-averaged local strain field amplitude of mode } s = \frac{1}{3^2} \left(\sum_{i,j} \left(e_s(\mathbf{r})_i{}^j \right)^2 \right)^{1/2}. \tag{F21}$$

2. Type-II scattering processes

Following a similar procedure for type-II scattering ($m \rightleftharpoons s' + s''$) yields a complex energy level shift in Fock state $|n_m, n_{s'}, n_{s''}\rangle$,

$$\left\langle (\delta E_{m,s',s''})^{\mathrm{II}}_{\text{3-ph}} \right\rangle = \frac{1}{2} \sum_{s's''} \left(\frac{|A^m_{s's''}|^2 \left[n_{s'} n_{s''} (n_m + 1) - (n_{s'} + 1)(n_{s''} + 1) n_m \right]}{\hbar \left((\omega_{s'} + \omega_{s''} - \omega_{\mathrm{m}}) - i(\Gamma_{s'} + \Gamma_{s''} - \Gamma_m) \right)} \right). \tag{F22}$$

Assuming the SMRT approximation and taking the difference between the energy level shifts for displaced phonon numbers of $\delta n_m + 1$ and δn_m yields the single photon complex energy level shift,

$$\left\langle \hbar \left(\delta \tilde{\omega}_m \right)^{\mathrm{II}}_{\text{3-ph}} \right\rangle_{\text{smrt}} = \frac{1}{2} \sum_{s's''} \left(\frac{|A^m_{s's''}|^2 \left[-(1 + \bar{n}_{s'} + \bar{n}_{s''}) \right]}{\hbar \left((\omega_{s'} + \omega_{s''} - \omega_{\mathrm{m}}) - i(\Gamma_{s'} + \Gamma_{s''} - \Gamma_m) \right)} \right), \tag{F23}$$

and the perturbation in the energy damping rate of mode m due to type-II 3-phonon scattering,

$$\left\langle (\delta \gamma_m)^{\mathrm{II}}_{\text{3-ph}} \right\rangle_{\text{smrt}} = \frac{1}{2\hbar^2} \sum_{s's''} \left(\frac{|A^m_{s's''}|^2 (\gamma_{s'} + \gamma_{s''} - \gamma_m) \left[1 + \bar{n}_{s'} + \bar{n}_{s''} \right]}{(\omega_{s'} + \omega_{s''} - \omega_{\mathrm{m}})^2 + \left(\frac{\gamma_{s'} + \gamma_{s''} - \gamma_m}{2} \right)^2} \right). \tag{F24}$$

In Section H we use Eqs. (F17) and (F24) to numerically evaluate the expected damping of the breathing mode due to 3-phonon scattering with numerically simulated quasi-normal modes of the OMC cavity.

3. 3-phonon scattering in bulk Si

In order to compare the estimated 3-phonon scattering in the restricted geometry of the OMC cavity to that of a bulk material, here we consider a simplified model of 3-phonon scattering in bulk Si in which we treat Si as an isotropic acoustic material. We are primarily interested in Normal ($\mathcal{N}$), type-I scattering processes. $\mathcal{N}$ processes due to the low temperature, and thus low frequency of the acoustic phonons involved in the scattering, and type-I scattering due to the suppression of type-II scattering processes in the effectively one-dimensional OMC cavity for phonon frequencies below that of the breathing mode. As derived in Ref. [23], for such a bulk material system the acoustic damping of mode s under the single mode relaxation time approximation can be written as,

$$(\gamma_s)^{\mathcal{N}-\mathrm{I}}_{\text{3-ph,bulk}} = \frac{\hbar v_p \langle \gamma_{\mathrm{G}} \rangle^2}{4\pi \rho_{\mathrm{Si}} \langle v^2_{\mathrm{Si}} \rangle} \sum_{p'p''} \left(v^2_{p'} v^2_{p''} \right)^{-1} \int_{\mathcal{R}[p',p'']} d\omega_{s'} \, (\omega_{s'})^2 \, (\omega_{s'} + \omega_s)^2 \, (\bar{n}[\hbar \omega_{s'}/k_B T] - \bar{n}[\hbar(\omega_{s'} + \omega_s)/k_B T]), \tag{F25}$$

where p, p', and p'' label the acoustic polarization of the modes s, s', and s'', respectively. The region of integration, $\mathcal{R}[p, p']$, depends upon the acoustic velocity dispersion version frequency and polarization (having assumed an isotropic bulk model there is no directional dispersion). Rewriting in terms of normalized frequencies ($y = \omega(\hbar/k_B T)$), yields the following simplified form of the phonon damping,

$$(\gamma_s)^{\mathcal{N}-\mathrm{I}}_{\text{3-ph,bulk}} = \frac{\hbar v_p \langle \gamma_{\mathrm{G}} \rangle^2 (k_B T/\hbar)^5}{4\pi \rho_{\mathrm{Si}} \langle v^2_{\mathrm{Si}} \rangle} \sum_{p'p''} \left(v^2_{p'} v^2_{p''} \right)^{-1} \left\{ \int_{\mathcal{R}[p',p'',T]} dy_{s'} \, (y_{s'})^2 \, (y_{s'} + y_s)^2 \, (\bar{n}[y_{s'}] - \bar{n}[y_{s'} + y_s]) \right\}. \tag{F26}$$

The integral in curly brackets is unitless and depends on temperature through both the integration range and the constant $y_s = \hbar \omega_s / k_B T$.

Neglecting frequency dispersion, $\mathcal{R}[p', p'']$ takes on a relatively simple form for the various acoustic polarization scenarios. Due to polarization dispersion, the only allowed Normal, type-I scattering processes are: $L_s + L_{s'} \rightleftharpoons L_{s''}$, $L_s + T_{s'} \rightleftharpoons L_{s''}$, $T_s + L_{s'} \rightleftharpoons L_{s''}$, and $T_s + T_{s'} \rightleftharpoons L_{s''}$, where L (T) corresponds to longitudinal (transverse/shear) polarization acoustic waves. The breathing mode is of mixed polarization character, so all four combinations are potentially relevant for comparison to the numerical calculations performed using the quasi-modes of the OMC
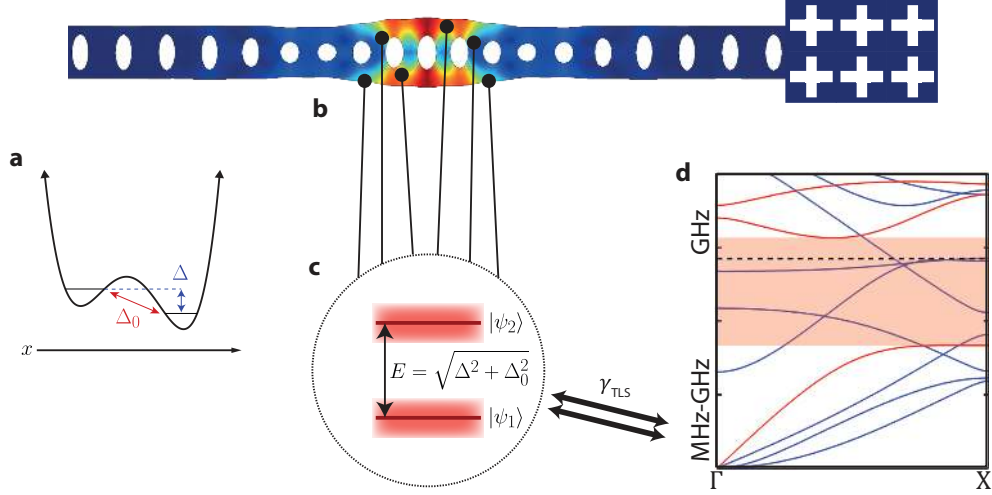
FIG. 16. **Simulation of TLS strain coupling to OMC cavity. a**, Double potential well energy profile of a tunneling-state (TS), showing the asymmetry energy Δ and the tunneling energy Δ_0 between left and right localized potentials. **b** FEM-simulated mode profile of the breathing mode of the OMC cavity, indicating surface-localized TS states which are strain coupled through deformation potential $\bar{\gamma}_{\text{TLS}}$. **c**, TS energy diagram, showing the transition energy, $E = (\Delta^2 + \Delta_0^2)^{1/2}$, between hybridized modes $|\psi_1\rangle$ and $|\psi_2\rangle$ which are mixtures of left and right localized states of the TS double potential well. **d**, Acoustic bandstructure of the OMC nanobeam, with blue (red) bands correspond to even (odd) vector parity acoustic modes with respect to in-plane mirror symmetry. The dashed black curve corresponds to the localized breathing mode frequency. The shaded orange region corresponds to the bandgap of the surrounding acoustic shield.

structure. Defining normalized wavevector magnitudes for the three acoustic waves ($x = |\mathbf{q}_s|/q_D$, $x' = |\mathbf{q}_{s'}|/q_D$, $x'' = |\mathbf{q}_{s''}|/q_D$), the integration range $\mathcal{R}[p,p']$ for x' is given by:

$$L_s + L_{s'} \rightleftharpoons L_{s''}: x' = \{0, 1-x\} \rightarrow \omega_{s'} \simeq \{0, \omega_D\}, \tag{F27}$$

$$T_s + L_{s'} \rightleftharpoons L_{s''}: x' = \{x_1 x, 1-rx\} \rightarrow \omega_{s'} \simeq \{(x_1/r)\omega_s, \omega_D\}, \tag{F28}$$

$$T_s + T_{s'} \rightleftharpoons L_{s''}: x' = \{x_1 x, x/x_1\} \rightarrow \omega_{s'} \simeq \{(x_1/r)\omega_s, (1/x_1 r)\omega_s\}, \tag{F29}$$

$$L_s + T_{s'} \rightleftharpoons L_{s''}: x' = \{0, x/x_3\} \rightarrow \omega_{s'} \simeq \{0, (1/x_3 r)\omega_s\}, \tag{F30}$$

where $q_D = \pi/a$ is the Debye wavevector for an atomic lattice constant a, $\omega_D = q_D \langle v_{\text{Si}} \rangle$, $r = v_{\text{t}}/v_{\text{l}}$ (≈ 0.69 in the [100] direction), $x_1 = (1-r)/(1+r) \approx 0.18$, and $x_3 = (1-r)/2 \approx 0.15$. For a breathing mode at 5 GHz, the corresponding lower frequency cut-off for the integration range of both TLL and TTL scattering would approximately be 1.3 GHz. The upper frequency cut-off of TTL and LTL scattering would be 40 GHz and 47 GHz, respectively. At temperatures below approximately 50 mK these processes would turn off, and they would saturate for temperatures above approximately 2 K. In what follows we consider the LLL scattering combination for comparison as it has effectively unlimited integration range and thus will contribute at both low and high temperatures relative to $T = \hbar\omega_s/k_B$ (≈ 200 mK for the breathing mode).

Appendix G: TLS damping model.

In addition to phonon-phonon scattering, another possible form of damping for the acoustic breathing mode is due to coupling to tunneling states (TS) or two-level systems (TLS). TS (or similar TLS) states correspond to a generic defect state in a solid-state material, typically an amorphous material, in which two local arrangements of atoms are nearly degenerate in energy. The two different atomic arrangements can have both a permanent electric and acoustic dipole associated with them, and atoms can tunnel between the two different arrangements. The TS and TLS models are two different phenomelogical models that are used to describe a wide variety of microscopic situations. Generically, in the TS model one has an asymmetry energy Δ which corresponds to the energy difference between the lowest energy level in each of the local potential energy profiles defining the two independent atomic arrangements, and a tunneling energy Δ_0 related to the energy barrier between the two local atomic arrangements (see Fig. 16(a)). One diagonalizes

the two lowest energy states of the two atomic arrangements into hybridized modes $|\psi_1\rangle$ and $|\psi_2\rangle$, whose energy difference is dependent upon a longitudinal dipole matrix element and which can be coupled via a transition dipole matrix element. In the TS model the ratio of the longitudinal dipole coupling to transition dipole coupling strength depends on the ratio of asymmetry energy to tunneling energy. The TLS model treats the longitudinal and transition dipole couplings as independent.

In the diagonal basis of the TS with asymmetry energy Δ and tunneling energy Δ_0, the interaction between the TS and a stress wave of phonon mode s is,

$$\hat{\mathcal{H}}_{\text{int, TS}-s} \approx \left(\frac{\Delta_0}{E}\hat{\sigma}_x + \frac{\Delta}{E}\hat{\sigma}_z \right) \bar{\gamma}_{\text{TS}}\hat{\bar{e}}_s(\mathbf{r}_0), \tag{G1}$$

where $\mathbf{r}_0$ is the point-like location of the TS and $E = (\Delta^2 + \Delta_0^2)^{1/2}$ is the TS transition energy. Here we treat the stress interaction as approximately scalar, hence, $\bar{\gamma}_{\text{TS}}$ and $\hat{\bar{e}}_s$ are the tensor-averaged deformation potential and stress operator, respectively. The corresponding TLS interaction Hamiltonian is given by,

$$\hat{\mathcal{H}}_{\text{int, TLS}-s} \approx \left(M\hat{\sigma}_x + D\hat{\sigma}_z \right) \hat{\bar{e}}_s(\mathbf{r}_0), \tag{G2}$$

where M is a transverse coupling potential and D is a longitudinal coupling potential. We will follow a TLS model in what follows as it simplifies some of the analysis; however, it is important to note that the TLS model is more constrained than the TS model in that the ratio of transverse and longitudinal coupling is fixed for a given TLS energy. This has the effect of eliminating the wide range of possible excited state decay rates for TS of a fixed energy. In fitting our data with a TS model we found that a model with rather narrow Δ_0 distribution whose mean scales approximately with E fit best, which is effectively a TLS model.

We define corresponding (tensor-averaged) transverse and longitudinal vacuum coupling rates as,

$$\bar{g}_{\text{t},s}(\mathbf{r}_0) = \frac{M}{\hbar}(e_{\text{vac},s})\left\langle \overline{\overline{\bar{e}_s(\mathbf{r}_0)}} \right\rangle_{\text{t}} \tag{G3}$$

and

$$\bar{g}_{\text{l},s}(\mathbf{r}_0) = \frac{D}{\hbar}(e_{\text{vac},s})\left\langle \overline{\overline{\bar{e}_s(\mathbf{r}_0)}} \right\rangle_{\text{t}} \tag{G4}$$

respectively, allowing us to write for the interaction Hamiltonian,

$$\hat{\mathcal{H}}_{\text{int, TLS}-s} \approx \hbar\left[\bar{g}_{\text{t},s}(\mathbf{r}_0)\hat{\sigma}_x + \bar{g}_{\text{l},s}(\mathbf{r}_0)\hat{\sigma}_z\right]\left(\hat{b}_s + \hat{b}_s^\dagger\right). \tag{G5}$$

Including the bare TLS and phonon energy terms, we have for the total Hamiltonian,

$$\hat{\mathcal{H}}_{\text{TLS}-s} \approx \frac{\hbar\omega_{\text{TLS}}}{2}\hat{\sigma}_z + \hbar\omega_s\left(\hat{b}_s^\dagger\hat{b}_s + 1/2\right) + \hbar\left[\bar{g}_{\text{t},s}(\mathbf{r}_0)\hat{\sigma}_x + \bar{g}_{\text{l},s}(\mathbf{r}_0)\hat{\sigma}_z\right]\left(\hat{b}_s + \hat{b}_s^\dagger\right), \tag{G6}$$

where ω_{TLS} is the bare transition frequency of the TLS ($E = \hbar\omega_{\text{TLS}}$). The $\hat{\sigma}_x$ interaction term leads to 'resonant' decay into the phonon bath and a bath-dependent level shift of the TLS which can be treated using 2nd-order perturbation theory. The $\hat{\sigma}_z$ interaction term gives rise to 'relaxation'-type processes of higher-order in perturbation theory.

1. TLS decay into the phonon bath

We first consider a single TLS interacting with the phonons as a dissipative bath (the roles will be reversed when we consider the damping of a given phonon mode). We assume that the TLS decay primarily through resonant $\hat{\sigma}_x$ interactions with the phonon bath, neglecting the $\hat{\sigma}_z$ interaction term. Also, owing to the finite-size of the acoustic cavity structure studied here, we work in a discrete basis of phonon quasi-normal modes. The bare phonon modes of the acoustic cavity have a complex frequency due to coupling to the phonons of the substrate, $\tilde{\omega}_s = \omega_s - i\Gamma_s$, where ω_s is the real angular frequency and $\Gamma_s = \gamma_s/2 + \Gamma_{s,\phi}$ is the phonon *amplitude* decoherence rate given by the sum of

half the energy decay rate (γ_s) and the pure dephasing rate ($\Gamma_{s,\phi}$) of the phonon mode. From 2nd-order perturbation theory [57, 59] we find a complex frequency shift of the TLS transition given by,

$$(\delta\tilde{\omega}_{\mathrm{TLS}})_s \approx \left(2\bar{g}_{\mathrm{t},s}^2 (n_s + 1/2)\right) \left[\frac{1}{\tilde{\Delta}_{\mathrm{TLS},s}} + \frac{1}{\tilde{\Delta}_{\mathrm{TLS},s} + 2\tilde{\omega}_s}\right], \tag{G7}$$

where n_s is the phonon occupancy of mode s and $\tilde{\Delta}_{\mathrm{TLS},s} \equiv \tilde{\omega}_{\mathrm{TLS}} - \tilde{\omega}_s$ is the near-resonant complex detuning. Here we have included the non-resonant term as it contributes non-negligibly to the TLS frequency shift (the real part of Eq. (G7)) when summing over contributions from phonon modes of large detuning. Implicit in our use of 2nd-order perturbation theory is that the TLS-s coupling remain in a small coupling limit ($|\bar{g}_{\mathrm{t},s}|/|\tilde{\Delta}_{\mathrm{TLS},s}| \ll 1$) where non-degenerate perturbation theory is accurate. One can also utilize quasi-degenerate perturbation theory [59] to determine the complex frequency shift without restriction on the strength of the coupling; however, the formulae are more complex and require careful elimination of non-physical solutions. For simplicity of presentation, here we limit ourselves to the small coupling limit. Below, in performing numerical calculations with specific TLS ensembles we found that the small coupling limit is adequate due to the low spectral density of phonon quasi-modes and TLS, and consequently the very unlikely situation where $|\bar{g}_{\mathrm{t},s}|/|\tilde{\Delta}_{\mathrm{TLS},s}| \gtrsim 1$.

One can arrive at a similar result in Hamiltonian form by rotating to a dressed TLS basis that diagonalizes Eq. (G6) in the N-excitation manifold to 2nd-order in the small parameter $\bar{g}_{\mathrm{t},s}/|\tilde{\omega}_{\mathrm{TLS}} - \tilde{\omega}_s|$,

$$\hat{\mathcal{H}}_{\mathrm{TLS}-s}^{\mathrm{eff,res.}} \approx \hbar\left(\tilde{\omega}_{\mathrm{TLS}}\right)\frac{\hat{\sigma}_z}{2} + \hbar\left(\tilde{\omega}_s + \frac{\bar{g}_{\mathrm{t},s}^2}{\tilde{\Delta}_{\mathrm{TLS},s}}\hat{\sigma}_z\right)\left(\hat{b}_s^\dagger \hat{b}_s + 1/2\right), \tag{G8}$$

where we have included only the 'resonant' σ_x interaction for now. Due to our use of quasi-normal modes for both TLS and mode s, $\hat{\mathcal{H}}_{\mathrm{TLS}-s}^{\mathrm{eff,res.}}$ is an effective Hamiltonian with complex energy eigenvalues. We see from this effective Hamiltonian that the resonant σ_x interaction leads both to a dressing of the TLS and the phonon mode: (i) viewed as a Stark-like shift of the TLS, the dressed complex frequency of the TLS is $\tilde{\omega}_{\mathrm{TLS}}' = \tilde{\omega}_{\mathrm{TLS}} + (2\bar{g}_{\mathrm{t},s}^2/\tilde{\Delta}_{\mathrm{TLS},s})\langle\hat{b}_s^\dagger \hat{b}_s + 1/2\rangle$, (ii) viewed as a TLS state-dependent shift of the phonon frequency, the dressed complex frequency of the mode s is $\tilde{\omega}_s' = \tilde{\omega}_s + (\bar{g}_{\mathrm{t},s}^2/\tilde{\Delta}_{\mathrm{TLS},s})\langle\hat{\sigma}_z\rangle$. The absence of the non-resonant term $[(\tilde{\Delta}_{\mathrm{TLS},s} + 2\tilde{\omega}_s)^{-1}]$ in Eq. (G8) is due to our restriction to the N-excitation manifold.

Returning to Eq. G7 and focusing on the damping effect of the phonon mode s on the TLS, we extract the imaginary component of $\delta\tilde{\omega}_{\mathrm{TLS}}$ corresponding to the phonon-induced decoherence rate of the TLS,

$$(\delta\Gamma_{2,\mathrm{TLS}})_s = -\mathrm{Im}\left[(\delta\tilde{\omega}_{\mathrm{TLS}})_s\right] \approx \frac{2\bar{g}_{\mathrm{t},s}^2 \left(\Gamma_s - \Gamma_{2,\mathrm{TLS}}\right)\left(n_s + 1/2\right)}{(\omega_{\mathrm{TLS}} - \omega_s)^2 + (\Gamma_{2,\mathrm{TLS}} - \Gamma_s)^2}. \tag{G9}$$

Assuming the TLS primarily decohere through the phonon bath such that the bare $\Gamma_{2,\mathrm{TLS}} \approx 0$ (neglecting other bath contributions and TLS-TLS dephasing, for instance), and neglecting pure dephasing of the phonon mode ($\Gamma_{s,\phi} = 0$), we can write for the phonon-induced energy decay rate of the TLS due to mode s,

$$(\delta\Gamma_{1,\mathrm{TLS}})_s \approx 2\left(\delta\Gamma_{2,\mathrm{TLS}}\right)_s \approx \left[\frac{\bar{g}_{\mathrm{t},s}^2}{(\omega_{\mathrm{TLS}} - \omega_s)^2 + (\gamma_s/2)^2}\right]\left(\gamma_s(2n_s + 1)\right). \tag{G10}$$

For a phonon bath in thermal equilibrium at temperature T we have that $2n_s + 1 = \coth(\hbar\omega_s/2k_B T)$. Summing over the discrete set of quasi-normal phonon modes allows us to write for the total phonon-induced $\Gamma_{1,\mathrm{TLS}}$ as a function of phonon bath temperature,

$$(\delta\Gamma_{1,\mathrm{TLS}})_{\mathrm{ph}} \approx \sum_s \left[\frac{\bar{g}_{\mathrm{t},s}^2 \gamma_s}{(\omega_{\mathrm{TLS}} - \omega_s)^2 + (\gamma_s/2)^2}\right]\left(\coth(\hbar\omega_s/2k_B T)\right). \tag{G11}$$

One recovers the standard result for a TLS interacting with a continuum phonon bath [25] by integrating Eq. (G11) weighted by the appropriate phonon density of states per unit frequency, $\rho_{\mathrm{ph}}[\omega_s]$,

$$(\delta\Gamma_{1,\mathrm{TLS}})_{\mathrm{ph,cont.}} \approx 2\pi\rho_{\mathrm{ph}}[\omega_{\mathrm{TLS}}](\bar{g}_{\mathrm{t},s}[\omega_{\mathrm{TLS}}])^2 \left(\coth[\hbar\omega_{\mathrm{TLS}}/2k_B T]\right), \tag{G12}$$

For a three-dimensional (3D) bulk material the (polarization-averaged) phonon bath density of states is $\rho_{\rm ph} = (V/(2\pi^2\bar{v}^3))\omega_s^2$, where $\bar{v}$ is an average acoustic velocity in the material. The phonon modes of a homogeneous bulk are plane waves with vacuum strain amplitude $e_{{\rm vac},s} = (\hbar\omega_s/2\bar{J}V)^{1/2}$, where $\bar{J}$ is a (2nd-order in strain) elastic energy density coefficient or bulk modulus of the material. The acoustic velocity and bulk modulus can be related to the bulk material mass density, $\bar{\rho}_{\rm m} = \bar{J}/\bar{v}^2$. Substituting these values into Eq. (G12) yields for an average TLS coupled to a phonon bath in a 3D bulk,

$$(\delta\Gamma_{1,\rm TLS})_{\rm ph,3D} \approx \left(\frac{\bar{M}^2\omega_{\rm TLS}^3}{2\pi\hbar\bar{J}\bar{v}^3}\right)(\coth[\hbar\omega_{\rm TLS}/2k_BT]) = \left(\frac{\bar{M}^2\omega_{\rm TLS}^3}{2\pi\hbar\bar{\rho}_{\rm m}\bar{v}^5}\right)(\coth[\hbar\omega_{\rm TLS}/2k_BT])\,, \qquad (G13)$$

where $\bar{M}$ is an averaged (over TLS orientation and acoustic polarization) transverse coupling potential. For future reference we also quote here the corresponding result for a quasi two-dimensional (2D) material, corresponding to a plate of large area and thickness t smaller than the acoustic wavelength,

$$(\delta\Gamma_{1,\rm TLS})_{\rm ph,2D} \approx \left(\frac{\bar{M}^2\omega_{\rm TLS}^2}{2\hbar\bar{\rho}_{\rm m}\bar{v}^4 t}\right)(\coth[\hbar\omega_{\rm TLS}/2k_BT])\,, \qquad (G14)$$

and a quasi one-dimensional (1D) material, corresponding to a beam of long length and small cross-sectional dimension $\bar{w}$ relative to the acoustic wavelength,

$$(\delta\Gamma_{1,\rm TLS})_{\rm ph,1D} \approx \left(\frac{\bar{M}^2\omega_{\rm TLS}}{2\hbar\bar{\rho}_{\rm m}\bar{v}^3\bar{w}^2}\right)(\coth[\hbar\omega_{\rm TLS}/2k_BT])\,. \qquad (G15)$$

In a prelude to what follows, we note that the frequency scaling with bath dimension will set the temperature scaling of the non-resonant TLS 'relaxation' damping of acoustical modes; hence, a 3D phonon bath will yield a T^3 dependence, a 2D bath a quadratic T^2 dependence, and a 1D bath will result in a linear T dependence.

2. 'Resonant' TLS damping and frequency shift of acoustic cavity quasi-modes

In contrast to the analysis of the prior sub-section, we now reverse roles and consider the TLS to be a bath for the acoustic phonon modes of the structure. In particular, we are interested in the localized high-Q phonon mode which lies within the phononic bandgap of the acoustic radiation shield. As such, this phonon mode should have much smaller intrinsic (radiation) damping to the substrate than the average phonon bath mode considered in the previous sub-section. We first consider the effects of the $\hat{\sigma}_x$ TLS-phonon interaction. Referring to Eq. (G8), the 'resonant' ($\hat{\sigma}_x$) contribution to the TLS state-dependent shift in the complex phonon frequency is given by,

$$(\delta\tilde{\omega}_s)_{\rm TLS,res} \approx \left(\bar{g}_{{\rm t},s}^2[\mathbf{r}_{\rm TLS}]\langle\hat{\sigma}_z\rangle\right)\left[\frac{1}{\tilde{\Delta}_{{\rm TLS},s}} + \frac{1}{\tilde{\Delta}_{{\rm TLS},s} + 2\tilde{\omega}_s}\right]\,, \qquad (G16)$$

where $\mathbf{r}_{\rm TLS}$ is the spatial location of the TLS in the acoustic cavity and we have added in by hand the non-resonant $[(\tilde{\Delta}_{{\rm TLS},s} + 2\tilde{\omega}_s)^{-1}]$ term as found in the perturbation analysis of Eq. (G7). For a TLS in thermal equilibrium at temperature T one has $\langle\hat{\sigma}_z\rangle = -\tanh[\hbar\omega_{\rm TLS}/2k_BT]$. Considering interaction with an ensemble of TLS and summing over this ensemble yields for the TLS state-dependent shift in the real part of the frequency of the phonon mode s,

$$(\delta\omega_s)_{\rm res} \approx \sum_{\rm TLS}{\rm Re}\left[(\delta\tilde{\omega}_s)_{\rm TLS}\right] = -\sum_{\rm TLS}\left(\bar{g}_{{\rm t},s}^2[\mathbf{r}_{\rm TLS}]\tanh[\hbar\omega_{\rm TLS}/2k_BT]\right)$$
$$\times \left[\frac{\omega_{\rm TLS} - \omega_s}{(\omega_{\rm TLS} - \omega_s)^2 + (\Gamma_{2,\rm TLS} - \Gamma_s)^2} + \frac{\omega_{\rm TLS} + \omega_s}{(\omega_{\rm TLS} + \omega_s)^2 + (\Gamma_{2,\rm TLS} + \Gamma_s)^2}\right]\,. \qquad (G17)$$

Substituting for $\Gamma_{2,\rm TLS}$ the estimated energy decay rate due to coupling to the rest of the phonon bath found in Eq. (G10) and a pure dephasing rate ($\Gamma_{\phi,\rm TLS}$), and assuming the phonon mode m of interest has a much smaller decoherence rate than the *dressed* TLS, we have that,

$$(\delta\omega_{\rm m})_{\rm res} \approx -\sum_{\rm TLS}\left(\bar{g}_{{\rm t},m}^2[\mathbf{r}_{\rm TLS}]\tanh[\hbar\omega_{\rm TLS}/2k_BT]\right)$$

$$\times\left[\frac{\omega_{\rm TLS}-\omega_{\rm m}}{(\omega_{\rm TLS}-\omega_{\rm m})^2+((\delta\Gamma_{1,{\rm TLS}})_{\rm ph}/2+\Gamma_{\phi,{\rm TLS}})^2}+\frac{\omega_{\rm TLS}+\omega_{\rm m}}{(\omega_{\rm TLS}+\omega_{\rm m})^2+((\delta\Gamma_{1,{\rm TLS}})_{\rm ph}/2+\Gamma_{\phi,{\rm TLS}})^2}\right]. \quad (G18)$$

Similarly, the energy damping rate of mode m due to resonant interaction processes with the TLS bath is given by,

$$(\delta\gamma_{\rm m})_{\rm res} \approx \sum_{\rm TLS}-2{\rm Im}\left[(\delta\tilde{\omega}_m)_{\rm TLS}\right] \approx \sum_{\rm TLS}\left[\frac{(\bar{g}_{{\rm t},m}^2[\mathbf{r}_{\rm TLS}]\tanh[\hbar\omega_{\rm TLS}/2k_BT])\,(\delta\Gamma_{1,{\rm TLS}})_{\rm ph}}{(\omega_{\rm TLS}-\omega_{\rm m})^2+((\delta\Gamma_{1,{\rm TLS}})_{\rm ph}/2+\Gamma_{\phi,{\rm TLS}})^2}\right], \quad (G19)$$

where we have neglected the non-resonant term $[(\tilde{\Delta}_{{\rm TLS},s}+2\tilde{\omega}_s)^{-1}]$ due to its much weaker contribution to the Lorentzian damping function. Note that we have not included the $\Gamma_{\phi,{\rm TLS}}$ contribution in the numerator of Eq. (G19) as it adds pure dephasing to the phonon mode m.

One recovers the standard result for damping of a phonon interacting with a continuum TLS bath [25] by integrating Eq. (G19) weighted by the TLS density of states per unit *angular* frequency in the acoustic mode volume, $n_{0,m}/2\pi \equiv \hbar n_0\eta_{\rm surf}V_m$,

$$(\delta\gamma_{\rm m})_{\rm res,\,cont.} \approx (2\pi)(\hbar\eta_{\rm surf}n_0V_m)\overline{\bar{g}_{{\rm t},m}^2[\mathbf{r}_m]}\tanh[\hbar\omega_{\rm m}/2k_BT] \approx \left(\frac{\pi\bar{M}^2\omega_{\rm m}}{\bar{\rho}_{\rm m}\bar{v}^2}\right)(\eta_{\rm surf}n_0)\tanh[\hbar\omega_{\rm m}/2k_BT], \quad (G20)$$

where $\eta_{\rm surf}n_0$ is the effective bulk TLS density per unit volume per unit energy, and the average transverse coupling rate for TLS in the acoustic mode volume is approximately, $\overline{\bar{g}_{{\rm t},m}^2[\mathbf{r}_m]} \approx (\bar{M}/\hbar)^2(\hbar\omega_{\rm m}/2\bar{\rho}_{\rm m}\bar{v}^2V_m)$. Following a similar averaging over the cavity mode volume and integration over a TLS density in Eq. (G18), one obtains the corresponding frequency shift of the breathing mode due to resonant interaction with a continuum of TLS,

$$(\delta\omega_{\rm m})_{\rm res,\,cont.} \approx -(\hbar\eta_{\rm surf}n_0V_m)\overline{\bar{g}_{{\rm t},m}^2[\mathbf{r}_m]}$$

$$\times{\rm Re}\left\{\int_0^{\omega_{\rm max}}{\rm d}\omega_{\rm TLS}\tanh[\hbar\omega_{\rm TLS}/2k_BT]\left(\frac{1}{(\omega_{\rm TLS}-\omega_{\rm m})+i\Gamma_{2,{\rm TLS}}}+\frac{1}{(\omega_{\rm TLS}+\omega_{\rm m})-i\Gamma_{2,{\rm TLS}}}\right)\right\}, \quad (G21)$$

where $\omega_{\rm max}$ is the maximum transition frequency of the TLS ensemble. The integral can be evaluated using the digamma function [34], yielding the following simplified result,

$$(\delta\omega_{\rm m})_{\rm res,\,cont.} \approx \left(\frac{\bar{M}^2\omega_{\rm m}}{\bar{\rho}_{\rm m}\bar{v}^2}\right)(\eta_{\rm surf}n_0)\left({\rm Re}\left\{\Psi\left[\frac{1}{2}+i\frac{\hbar\omega_{\rm m}}{2\pi k_BT}\right]\right\}-\ln\left[\frac{\hbar\omega_{\rm max}}{2\pi k_BT}\right]\right). \quad (G22)$$

3. 'Relaxation' TLS damping and frequency shift of acoustic cavity quasi-modes

Relaxation damping of acoustic cavity modes results not from direct energy exchange with nearly-resonant TLS, but rather from the shift in the TLS transition frequencies due to the σ_z interaction of Eq. (G2). This shift, which is linear in the stress amplitude of the acoustic modes and oscillates in time with the frequency of the acoustic mode, displaces the TLS from equilibrium. During this oscillatory displacement out of equilibrium the TLS will relax back towards equilibrium at a rate given by the TLS energy decay rate, stealing away energy from the acoustic mode in the process. Microscopically this is a process involving higher-order perturbation interactions between the TLS and the phonon bath. As such, here we follow the standard semi-classical analysis of relaxation damping by considering the acoustic dipole response of the TLS to a (classical) strain field [25, 34].

The magnitude of the longitudinal acoustic dipole of a TLS is given by $\langle\bar{\bar{p}}_a\rangle_{\rm t} \equiv D\langle\hat{\sigma}_z\rangle$. For a small amplitude, harmonic strain field oscillating in mode s, $\langle\bar{\bar{e}}_s[\mathbf{r};\omega_s]\rangle_{\rm t}$, the harmonically oscillating component of the longitudinal acoustic dipole is linearly related to the applied strain at the site of TLS through a (tensor-averaged) susceptibility, $\langle\bar{\bar{\chi}}_{\rm rel}[\omega_s]\rangle_{\rm t} \equiv \langle\delta\bar{\bar{p}}_a[\omega_s]\rangle_{\rm t}/\langle\bar{\bar{e}}_s[\mathbf{r}_{\rm TLS};\omega_s]\rangle_{\rm t}$, where $\langle\delta\bar{\bar{p}}_a[\omega_s]\rangle_{\rm t} = D(\delta\langle\hat{\sigma}_z[\omega_s]\rangle)$, $\delta\langle\hat{\sigma}_z\rangle = (\langle\hat{\sigma}_z\rangle - \langle\hat{\sigma}_z\rangle_{\rm eq.})$, and $\langle\hat{\sigma}_z\rangle_{\rm eq.} = -\tanh[\hbar\omega_{\rm TLS}/2k_BT]$ is the TLS inversion in thermal equilibrium. Solving the Bloch equations assuming a finite

relaxation rate to equilibrium of $\Gamma_{1,\text{TLS}}$, the displacement of the inversion from equilibrium follows the applied harmonic strain with a phase lag,

$$\delta\langle\hat{\sigma}_z\rangle \approx \frac{\partial\langle\hat{\sigma}_z\rangle_{\text{eq.}}}{\partial\omega_{\text{TLS}}}\frac{\partial\omega_{\text{TLS}}}{\partial\langle\bar{\bar{e}}_s\rangle_{\text{t}}}\left(\frac{1-i\omega_s(\Gamma_{1,\text{TLS}})^{-1}}{1+(\omega_s(\Gamma_{1,\text{TLS}})^{-1})^2}\right)\langle\bar{\bar{e}}_s\rangle_{\text{t}}. \tag{G23}$$

From Eq. (G6) one can show that $\partial\omega_{\text{TLS}}/\partial\langle\bar{\bar{e}}_s\rangle_{\text{t}} = (2D/\hbar)$, and $\partial\langle\hat{\sigma}_z\rangle_{\text{eq.}}/\partial\omega_{\text{TLS}} = (\hbar/2k_BT)\,\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT]$. This yields for the relaxation susceptibility,

$$\langle\bar{\chi}_{\text{rel}}[\omega_s;\omega_{\text{TLS}}]\rangle_{\text{t}} = \left(\frac{D^2}{k_BT}\right)\left(\frac{1-i\omega_s(\Gamma_{1,\text{TLS}})^{-1}}{1+(\omega_s(\Gamma_{1,\text{TLS}})^{-1})^2}\right)\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT]. \tag{G24}$$

The complex energy shift of the acoustic mode due to its interaction with the TLS is given by $(\delta\tilde{E}_s)_{\text{TLS}} \approx \langle\bar{p}_a\rangle_{\text{t}}(\langle\bar{\bar{e}}_s[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^*$. Noting the complex energy shift can be related to a complex frequency shift in the acoustic resonance through the stored phonon number (n_s), $(\delta\tilde{E}_s)_{\text{TLS}} = \hbar(\delta\tilde{\omega}_s n_s)$, and writing the local applied strain amplitude in terms of phonon number, $\langle\bar{\bar{e}}_s[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}} = (e_{\text{vac},s})(\langle\bar{\bar{e}}_s[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})\sqrt{n_s}$, yields for the complex frequency shift in quasi-mode s due to relaxation interactions with a single TLS,

$$(\delta\tilde{\omega}_s)_{\text{TLS,rel}} \approx \frac{(\langle\bar{\bar{e}}_s[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^2(e_{\text{vac},s})^2}{\hbar}\langle\bar{\chi}_{\text{rel}}[\omega_s;\omega_{\text{TLS}}]\rangle_{\text{t}}. \tag{G25}$$

For a given quasi-mode m of interest, interacting with an ensemble of TLS, the corresponding frequency shift due to relaxation processes is given by,

$$\begin{aligned}(\delta\omega_{\text{m}})_{\text{rel}} &= \sum_{\text{TLS}}\text{Re}\left[(\delta\tilde{\omega}_m)_{\text{TLS,rel}}\right] \approx \sum_{\text{TLS}}\frac{\omega_{\text{m}}(\langle\bar{\bar{e}}_m[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^2\text{Re}\left[\langle\bar{\chi}_{\text{rel}}[\omega_{\text{m}};\omega_{\text{TLS}}]\rangle_{\text{t}}\right]}{2\bar{\rho}_{\text{m}}\bar{v}^2V_m}\\ &\approx \sum_{\text{TLS}}\frac{(\langle\bar{\bar{e}}_m[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^2D^2(\Gamma_{1,\text{TLS}})^2}{2\omega_{\text{m}}\bar{\rho}_{\text{m}}\bar{v}^2V_m k_BT}\,\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT]\\ &= \sum_{\text{TLS}}\left(\frac{\bar{g}_{\text{l},m}[\mathbf{r}_{\text{TLS}}]}{\omega_{\text{m}}}\right)^2\left(\frac{\hbar(\Gamma_{1,\text{TLS}})^2}{k_BT}\right)\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT].\end{aligned} \tag{G26}$$

Similarly, the relaxation energy damping rate of mode m is given by,

$$\begin{aligned}(\delta\gamma_{\text{m}})_{\text{rel}} &= -2\sum_{\text{TLS}}\text{Im}\left[(\delta\tilde{\omega}_m)_{\text{TLS,rel}}\right] \approx -\sum_{\text{TLS}}\frac{\omega_{\text{m}}(\langle\bar{\bar{e}}_m[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^2\text{Im}\left[\langle\bar{\chi}_{\text{rel}}[\omega_{\text{m}};\omega_{\text{TLS}}]\rangle_{\text{t}}\right]}{\bar{\rho}_{\text{m}}\bar{v}^2V_m}\\ &\approx \sum_{\text{TLS}}\frac{(\langle\bar{\bar{e}}_m[\mathbf{r}_{\text{TLS}}]\rangle_{\text{t}})^2D^2\Gamma_{1,\text{TLS}}}{\bar{\rho}_{\text{m}}\bar{v}^2V_m k_BT}\,\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT]\\ &= \sum_{\text{TLS}}\left(\frac{2\bar{g}_{\text{l},m}^2[\mathbf{r}_{\text{TLS}}]}{\omega_{\text{m}}}\right)\left(\frac{\hbar\Gamma_{1,\text{TLS}}}{k_BT}\right)\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT].\end{aligned} \tag{G27}$$

Since we are concerned with high frequency, microwave phonon modes and TLS at cryogenic temperatures, we have safely assumed that we are in the non-adiabatic limit, $\omega_{\text{m}}\Gamma_{1,\text{TLS}}^{-1} \gg 1$.

Assuming resonant phonon damping is the dominant decay mechanism for TLS, and substituting $(\delta\Gamma_{1,\text{TLS}})_{\text{ph,cont.}}$ for $\Gamma_{1,\text{TLS}}$ in Eqs. (G26-G27), yields the standard relations for the frequency shift and energy damping for an acoustic mode interacting with a TLS bath in a bulk material which supports a phonon continuum. For a uniform spectral density of states for the TLS, as assumed here, integrating over ω_{TLS} yields a phonon relaxation damping of $(\delta\gamma_{\text{m}})_{\text{rel}} \sim T^d$, where d is the dimension of the phonon bath. This is a result of the fact that $\text{sech}^2[\hbar\omega_{\text{TLS}}/2k_BT]$ effectively limits the TLS frequency integral to frequencies below $\approx k_BT$ (i.e., relaxation damping is limited to thermally occupied TLS), and within this range of frequencies $\coth[\hbar\omega_{\text{TLS}}/2k_BT] \approx 2k_BT/\hbar\omega_{\text{TLS}}$. This observation is especially important for the nanoscale optomechanical structures under study, where at the temperatures and corresponding
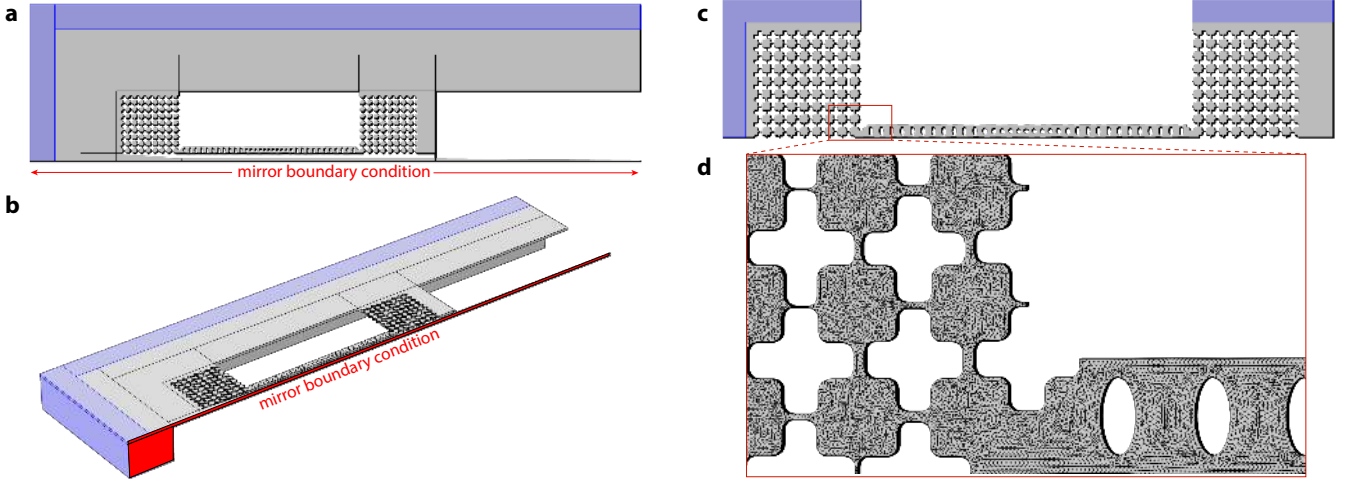
FIG. 17. **FEM simulation layout and mesh. a**, Top view of the FEM-simulated structure for phonon frequencies below 10 GHz. Blue regions correspond to PML radiation boundaries. The structure is mirrored about the center axis of the coupling waveguide (red arrows and label). A maximum mesh size of 20 nm is utilized in the OMC cavity and phononic shield regions. A maximum mesh size of 250 nm is set in the surrounding periphery and PML regions. The mesh resolution is smoothly varied between the two regions. **b**, Isometric view of the same structure in (**a**), showing the underlying BOX clamping layer and the corresponding PML layers. The bottom of the BOX layer has a fixed boundary condition applied to it. **c** Top view of the reduced FEM-simulation volume for phonon frequencies between 10-100 GHz. Again, the blue region is an acoustic PML. All other boundaries other than PML are set to free boundaries. The maximum mesh size is now 20 nm throughout the entire structure, including the PML. **d** Zoom in of the red box region in (**c**), showing the dense meshing of the nanobeam and shield.

frequencies considered, the relevant phonon bath density of states varies between something approximating a 1D bath at temperatures below 100 mK to something approximating a 2.5D bath at temperatures between 100 mK and 1 K. The influence of the geometric patterning of the optomechanical structure also plays a role, modifying the phonon density of states in extreme ways by introducing phononic bandgaps and flat band regions. As we will show below using numerical methods to calculate the full spectrum of quasi-acoustic-modes of the optomechanical structure, the temperature dependence of the relaxation damping rate due to TLS can indeed be significantly modified from that in bulk.

Appendix H: Numerical modeling of TLS interactions and acoustic damping in the OMC cavity

In order to more accurately account for the complex geometry of the optomechanical crystal structure studied and its impact on the phonon mode spectrum and TLS dynamics, we have performed numerical simulations of the acoustic resonances (quasi-modes) of the structure for frequencies below 100 GHz. One half of the simulated structure is shown schematically in Fig. 16, and includes one of the pair of OMC cavities, the optical coupling waveguide, and the phononic shield which make up a single 'device'. In practice, such a device is clamped from below at its periphery through the connection of the top Si device layer to the underlying 3 μm thick SiO_x buried oxide (BOX) layer, which itself is grown on top of a thick (500 μm) Si handle layer. Simulations are performed using the COMSOL finite-element-method solver, using the Amazon AWS cloud computing resources to simulate various parts of the phonon spectrum in parallel. Acoustic perfectly matched layers (PML) at the periphery of the structure are used as a radiation boundary condition.

A series of different meshing schemes are used to cover the acoustic frequency spectrum up to 100 GHz. For all frequencies we utilize a fine mesh with maximum element size of 20 nm in the OMC cavity and phononic shield regions. This yields a meshing resolution of roughly 3-4 points per wavelength even at the highest (100 GHz) frequencies considered. As a result of memory and computing time limitations, we adjust the meshing and structure layout as a function of frequency in the rest of the structure outside of the OMC cavity and shield. For acoustic frequencies below 10 GHz the full structure shown in Fig. 16(b) is simulated, which includes at its periphery a micron thick (in depth and height) BOX layer followed by the PML radiation boundary. For these frequencies a lower mesh density (maximum mesh size of 250 nm) is utilized in the BOX and PML regions, providing a minimum of roughly 4 points per wavelength resolution for modes up to 10 GHz. The simulated quasi-modes of the structure are therefore
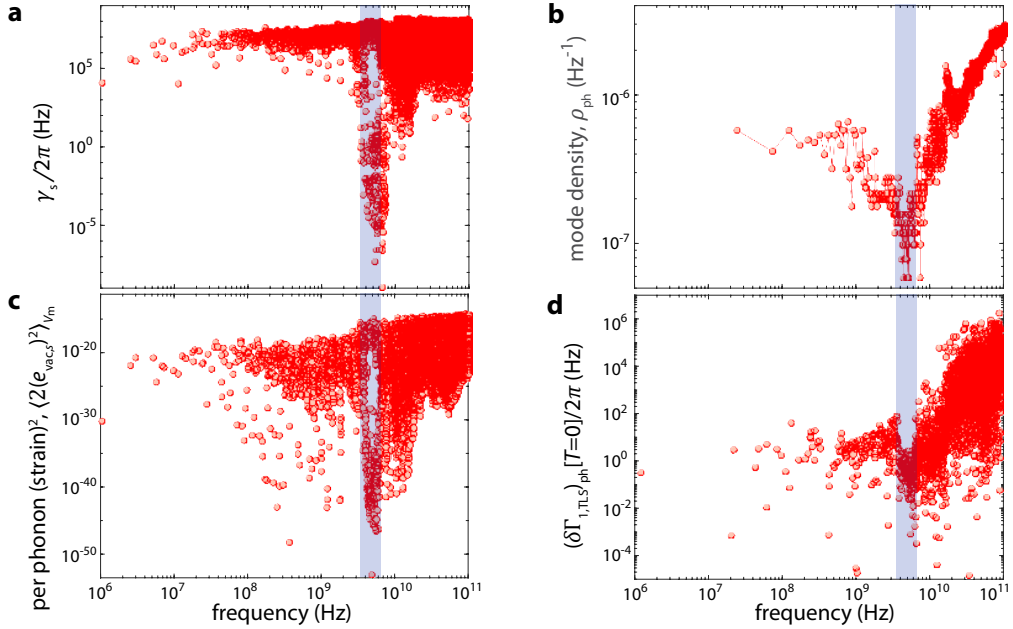
FIG. 18. **Model phonon and TLS properties.** In this model, phonons of the Si optomechanical slab structure are simulated using a FEM numerical solver. The simulation volume consists of the two optomechanical cavities, fiber coupling waveguide, and 10 period phononic shield. As in the fabricated devices, the Si slab is clamped at the periphery of the optomechanical structure to the underlying SiO_x BOX layer of the SOI. An acoustic radiation boundary condition consisting of a perfectly-matched layer allows for radiation to escape into the external Si slab and underlying substrate. All phonon resonances up to 100 GHz frequency are calculated. The lowest phononic bandgap of the shield is shown as a transparent blue band. **a**, Acoustic radiation damping rate for phonon quasi-modes of the optomechanical structure. **b**, Spectral phonon mode density, spectrally averaged over 20 MHz bin size. **c**, Per-phonon squared strain value within mode volume V_m for each of the phonon-quasi modes. Here, the sum of the square of the strain components are averaged over 201 positions within the acoustic mode volume of the breathing mechanical mode. **d**, For an exemplary ensemble instance, zero temperature TLS decay rate of each of the $N_{\text{TLS},m} = 3920$ randomly oriented and distributed within V_m due to resonant ($\hat{\sigma}_x$) coupling to phonon quasi-modes of the optomechanical structure.

damped through their acoustic radiation into the BOX layer, with no further acoustic reflections considered (such as at the Si handle layer). For frequencies above 10 GHz we remove the BOX clamping region and apply the PML layers right at the boundary of the phononic shield, as shown in Fig. 16(c). At these higher frequencies we utilize the same fine mesh in the PML as in the OMC cavity region (maximum mesh size of 20 nm). The thought in doing this is that the BOX clamping region plays less of a role for these short wavelength phonons that can effectively propagate within the thin Si device layer without major reflection at this boundary, hence the removal of the BOX layer and application of the PML region in the Si device layer right at the exit of the acoustic shield. More importantly is that we provide a fine mesh in the PML to avoid unintended reflections. Beyond 100 GHz the resulting memory requirements and computation time are prohibitive, and given the temperature range of interest ($\lesssim 1$ K), 100 GHz frequency is a natural cut-off point. In addition to the above meshing strategy, in order to reduce the memory and computation time we apply a mirror boundary condition along the center of the structure, running down the middle of the coupling waveguide, and double the number of modes recorded in simulation.

For each acoustic resonance found in the simulation we not only record the frequency (ω) and energy damping rate (γ), but also calculate the per phonon strain tensor at 101 locations in the acoustic mode volume V_m of the high-Q breathing mode. These locations were chosen to be in random locations in the Si device layer, but within $\delta w = 15$ nm of the Si-air interface, as this is where we expect TLS to be located due to etch damage. A fixed set of 101 positions are evaluated for all acoustic modes. Plots showing the resulting energy damping rate, effective phonon mode density, and average squared strain in the mode volume V_m as a function of phonon frequency are shown in Figs. 18(a-c). Noteworthy in these plots is the position of the fundamental phononic bandgap of the acoustic shield, which is shown as a semi-transparent blue band from approximately 3.5 to 6.5 GHz. As can be seen, a significant change in the local strain amplitude and mode density occurs around this phononic bandgap. Below the bandgap, the mode density is roughly constant at 1 mode every 5 MHz (a bin size of 50 MHz was used when estimating this spectral quasi-mode density), consistent with that of an effectively 1D system. Within the bandgap of the acoustic

TABLE III. **3-phonon scattering model parameters.**

| Parameter	Description	Value	Refs./Notes
$\left\langle \mathcal{F}^{s''}_{m\,s'} \right\rangle_{\mathrm{m}}$	average 3-phonon mode overlap factor	0.01	$\cong$ phase matching term
$\langle \gamma_{\mathrm{G}} \rangle_{\mathrm{m}}$	mode averaged Grüneisen parameter at cryogenic temperature ($T < 4K$)	0.24	[60]
v_{t}	transverse acoustic phonon velocity in bulk Si ([100] dir., [011] pol.)	8.4×10^3 m/s	[61, 62]
v_{l}	longitudinal acoustic phonon velocity in bulk Si ([100] dir. and pol.)	5.8×10^3 m/s	[61, 62]
$\langle v_{\mathrm{Si}}^2 \rangle$	average square of acoustic phonon velocity in bulk Si ([100] dir.)	$(7.6 \times 10^3$ m/s$)^2$	
ρ_{Si}	Si mass density	2.33×10^3 kg/m^3	[61]
$\langle e_{\mathrm{vac},m}^2 \rangle$	average strain squared of vacuum for mode m within V_m	2.25×10^{-8}	COMSOL sim.
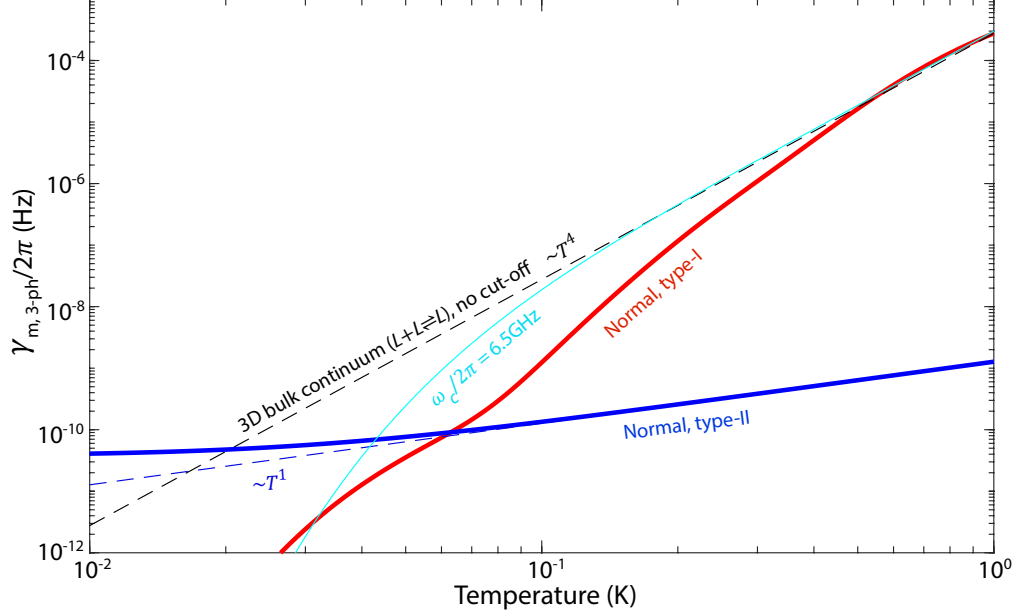
FIG. 19. **3-phonon scattering model.** Simulation of the acoustic damping due to 3-phonon scattering of the localized breathing mechanical mode with other quasi-modes of the OMC cavity structure of Fig. 18. Parameters used in the modeling are listed in Tab. III. Both type I (solid red curve) and type II (solid blue curve) scattering processes involving the breathing mode are modeled. For comparison, we also plot the estimated 3-phonon-scattering damping rate due to type-I processes for longitudinally polarized phonons ($L + L \rightleftharpoons L$) in bulk Si. For the bulk simulation we plot the estimated damping rate without a low-frequency cut-off (dashed black curve), and with a low-frequency cut-off (solid cyan curve) corresponding to the top of the first phononic bandgap ($\omega_c/2\pi = 6.5$ GHz). In all cases we only include Normal scattering processes, and neglect Umklapp scattering, due to the low temperature range considered ($T \lesssim 1$ K).

shield, the mode density drops, and then above the bandgap the mode density rises a little faster than linearly with frequency, corresponding to that of a 2.4-D system. Below the bandgap the per-phonon strain in mode volume V_m is quite small, and then rises within and above the phononic bandgap frequency due to the localization of modes within the shield. The energy damping rate plot shows that a portion of the modes become substantially less damped in and around the bandgap as expected.

1. Numerical simulation of 3-phonon scattering

Utilizing Eqs. (F17), (F24), and (F26), along with the parameters listed in Tab. III and the COMSOL-simulated acoustic modes of the OMC cavity structure (c.f., Fig. 17), in Fig. 19 we plot the calculated acoustic damping of the localized 5 GHz breathing mode due to 3-phonon scattering at temperatures below 1 K. The resulting temperature-dependent damping rate using the numerically computed acoustic quasi-modes of the OMC cavity are shown as a

solid red curve for $\mathcal{N}$-I processes and a solid blue curve for $\mathcal{N}$-II processes. Damping arising from $\mathcal{N}$-II scattering processes is largely suppressed in the OMC cavity structure due to the reduced density of phonon states lying below the breathing mode frequency, a consequence of the effectively reduced dimensionality of the OMC nanobeam at these acoustic wavelengths. At temperatures below $T = \hbar\omega_\mathrm{m}/k_B \approx 200$ mK the $\mathcal{N}$-II damping rate approaches the spontaneous decay rate of the breathing mode, whereas at higher temperatures the damping rate increases linearly with temperature. The $\mathcal{N}$-I damping rate, on the other hand, increases rapidly towards a high temperature scaling of $\sim T^4$. Again, at temperatures below $T = \hbar\omega_\mathrm{m}/k_B \approx 200$ mK the $\mathcal{N}$-I damping rate drops rapidly due to the reduced density of available phonon states in the OMC structure. This is another manifestation of the 'phonon bottleneck' effect discussed above in regards to the optically-induced hot phonon bath. Energy deposited into high frequency phonons decays to lower frequency phonons through processes like the 3-phonon mixing studied here, however, when the phonon wavelengths approach the dimension of the structure the reduced density of phonon states results in a precipitous drop in the nonlinear phonon scattering, effectively trapping the energy in phonons above a certain cut-off frequency.

For comparison purposes we have plotted the estimated damping for an isotropic Si bulk material involving $L+L \rightleftharpoons L$ acoustic scattering. The black dashed curve (solid cyan curve) is the bulk damping rate without (with) a low-frequency cut-off for the range of integration in the $\mathcal{N}$-I process. The cut-off frequency for the solid cyan curve is $\omega_c/2\pi = 6.5$ GHz, chosen to match the top of the first phononic bandgap of the acoustic shield in the OMC cavity structure (c.f., Fig. 18). The temperature scaling for the bulk scattering (above cut-off) is approximately T^4. An average overlap factor $\left\langle \mathcal{F}_{m\,s'}^{s''} \right\rangle = 0.01$ was assumed in the quasi-mode modeling, resulting in a reasonable correspondence with the $\mathcal{N}$-I bulk damping at temperatures above 400 mK.

Although the 3-phonon scattering rapidly rises with temperature ($\sim T^4$), for temperatures below 1 K where the size-scale of the OMC cavity structure comes into play, the magnitude of 3-phonon scattering is estimated to be substantially smaller than that measured in our experiment. In the following subsection we consider a more likely source of the observed damping, two-level systems, which act as an intermediate bath between phonons, greatly increasing the predicted damping rate.

2. Numerical modeling of TLS-phonon interactions in the OMC cavity

The coupling of the high-Q acoustic breathing mode of the OMC cavity to TLS defect states depends on a number factors. First and foremost there is the spatial and spectral density of the TLS, which determine the number of interacting defect states with the breathing mode. In order to constrain the TLS density to a realistic value we consider here that the majority of the TLS are associated with defects in a near-surface layer of the etched Si structure making up the OMC cavity. We assume no TLS defects in the bulk of the crystalline Si layer. The thickness of the defective surface layer of Si depends greatly on its preparation.

In our case, we have used an inductively-coupled reactive ion etch (ICP-RIE) to pattern the 220 nm thick Si device layer. The ICP-RIE etch utilizes an SF_6:C_4F_8 gas chemistry, with low RF power (≈ 30 W) and low DC-bias voltage (≈ 70 V), to reduce optimize the shape of the etched sidewall and to attempt to reduce etch-induced damage on the sidewalls of the etched Si. Nonetheless, it is well known that these etch processes still produce a variety of damage to the exposed near-surface layers of Si in the process. Typically in RIE etching [30, 31, 63, 64], a surface consisting of a super-surface top layer of fluoro-carbons (~ 5 nm) and Si-oxygen (~ 1.5 nm) is followed by a sub-surface heavily damaged layer containing Si-carbon (among other impurities) that can penetrate into the bulk to depths of tens of nanometers. Here we assume an etched sidewall damage layer thickness of $\delta w = 15$ nm. In order to reduce Si oxide growth on the top and bottom surface layers of the released Si device (i.e., those layers that do not see the ICP-RIE etch) we 'flash' the sample with an anhydrous vapor HF etch prior to inserting into the vacuum of the dilution refrigerator (time between flash and vacuum pump down in the cryostat is $\lesssim 45$ minutes). Ideally, this removes surface oxide layers and leaves a hydrogen-terminated Si surface nominally free of oxides. To be conservative, however, we also assume a $\delta t = 0.25$ nm surface oxide layer [65, 66] on the top and bottom surfaces of the released Si device layer. This yields a volume fraction of damaged Si in our devices which is $\eta_\mathrm{surf} = 0.29$.

Assuming a bulk TLS density commensurate with that in vitreous Si dioxide, $n_0 = 1.04$ states/J/m^3 [25], this yields a spectral density of TLS that lie within the acoustic cavity volume of the breathing mode of only $n_{0,m} \sim 20$ states/GHz. For a TLS population that is uniformly distributed spectrally [32], this yields only ~ 2000 TLS with transition frequencies lying below 100 GHz that are in the cavity mode volume. Note that we simulate an ensemble TLS size of $N_\mathrm{TLS,m} = 3920$, taking into account TLS that are within a spatial region of twice that of the cavity mode volume.

The resonant interaction of those TLS that lie spatially in the breathing mode cavity volume V_m and have their transition frequency spectrally nearby the $\omega_\mathrm{m}/2\pi \approx 5$ GHz acoustic resonance frequency is determined by the magni-

TABLE IV. **TLS damping model parameters.**

Parameter	Description	Value	Refs./Notes
D	longitudinal coupling deformation potential	5.6 eV	fit value (see also [25, 32, 67])
$\bar{D}$	angle-averaged longitudinal coupling	3.23 eV	$\equiv D/\sqrt{3}$
M	transverse coupling deformation potential	0.07 eV	fit value (see also [25, 32, 67])
$\bar{M}$	angle-averaged transverse coupling	0.04 eV	$\equiv M/\sqrt{3}$
δw	Si device layer etched sidewall damage layer thickness	15 nm	[30, 31, 63, 64]
δt	Si device layer top and bottom oxide thickness	0.25 nm	[65, 66]
η_{surf}	damaged material/surface oxide volume fraction	0.29	estimated from δt, δw
m	high-Q breathing mode label	N/A	
ω_{m}	mode m frequency	5.3 GHz	COMSOL sim.
V_m	mode m acoustic mode volume	0.11 $(\mu\text{m})^3$	COMSOL sim.
$\langle e_{\text{vac},m}^2 \rangle$	average strain squared of vacuum for mode m within V_m	2.25×10^{-8}	COMSOL sim.
n_0	TLS density per unit volume per unit energy	1.04 states/J/m^3	[25]
$n_{0,m}$	TLS density per unit frequency in V_m $(\equiv (\hbar 2\pi)n_0\eta_{\text{surf}}V_m)$	22.1 states/GHz	calculated
$N_{\text{TLS},m}$	estimated number of TLS in $2 \times V_m$ with $f \leq 100$ GHz	3920	calculated
$\Gamma_{\phi,\text{TLS}}/2\pi$	TLS pure dephasing rate	10 kHz	[25, 68]
Λ_{TLS}	TLS excitation rate (due to optical pumping)	2 Hz	$\sim (2 \times 10^{-3})(\gamma_{\text{p}}n_{\text{p}})$ at $n_{\text{c}} = 0.1$
$\langle \bar{g}_{\text{t},m}/2\pi \rangle$	average transverse vacuum coupling rate to mode m for TLS in V_m	~ 100 kHz	calculated
$\langle \bar{g}_{\text{l},m}/2\pi \rangle$	average longitudinal vacuum coupling rate to mode m for TLS in V_m	~ 8.6 MHz	calculated
$\langle \delta f_{m,\text{max}} \rangle$	average dispersive shift of mode m for nearest resonant TLS	~ 1.5 kHz	calculated $(\cong n_{0,m}(\bar{g}_{\text{t},m}/2\pi)^2)$

tude of the transverse coupling deformation potential, M. These TLS not only act as a bath to damp the breathing mode, but their temporal fluctuations from ground to excited state and back, lead to fluctuations in the acoustic environment of the breathing mode, producing both a frequency jitter of the mechanical mode and an overall shift in the resonance frequency that depends on the TLS temperature through its average excited state population [25, 34]. The acoustic frequency shift due to non-resonant TLS interacting through longitudinal $\hat{\sigma}_z$-coupling is negligible compared to the resonant $\hat{\sigma}_x$-coupling term. As such, the magnitude of the transverse coupling parameter is chosen in our model to be $M = 0.07$ eV, yielding an average transverse vacuum coupling rate to the breathing mode for TLS in V_m of $\langle \bar{g}_{\text{t},m}/2\pi \rangle \sim 100$ kHz. The corresponding dispersive shift due to the nearest resonant TLS (on average) is then approximately $\langle \delta f_{m,\text{max}} \rangle \cong n_{0,m}(\bar{g}_{\text{t},m}/2\pi)^2 \sim 1.5$ kHz. This level of dispersive shift is in line with the both the measured frequency jitter ($\Delta_{1/2} \approx 3.5$ kHz) and the frequency shift around $T \approx \hbar\omega_{\text{m}}/k_{\text{B}}$ for device D shown in Fig. 3(d).

The non-resonant relaxation interactions of TLS with the breathing mode is via a longitudinal coupling deformation potential, D. With the transverse coupling rate set nominally by the measured frequency jitter and temperature-dependent frequency shift of the breathing mode, the longitudinal deformation potential is adjusted to approximately match the measured acoustic damping rate of the breathing mode at the lowest fridge measurement temperatures ($T_{\text{f}} \approx 7$ mK). A value of $D = 5.6$ eV (angle-averaged $\bar{D} = 3.23$ eV) gives a reasonable fit to the measured data of Fig. 3(a). Typical values in the literature for averaged D and M parameters are on the order of 1 eV [25, 32], although these values are hard to distinguish separately from the TLS density [67]. The large value of D and small value of M indicate a set of TLS (or TS) states which have a large asymmetry energy and small tunneling energy. Other parameters and their assumed values in our model are listed, along with references and comments, in Table IV.

The calculated zero-temperature energy decay rate of an ensemble of 3920 TLS, randomly chosen from 101 fixed positions within the breathing mode cavity mode volume, with randomly oriented acoustic dipoles, and with randomly chosen frequency below 100 GHz is displayed Fig. 18(d). This calculation, following Eq. (G11), uses the simulated acoustic strain and radiative decay rate of the localized quasi-mode phonons of the suspended and peripherally-clamped OMC structure whose properties are also displayed in Fig. 18. Several points are worth noting here. The first is that the small number of TLS in the small acoustic mode volume V_m and the small number of localized quasi-normal phonon modes at low frequency means a significant spectral fluctuation in the decay rate of TLS with transition frequency below ~ 1 GHz. Within the phononic bandgap of the OMC cavity ($3.5 - 6.5$ GHz), there is a dramatic reduction in the decay rate of TLS, down to levels on the order of 1 Hz. Above the phononic bandgap, the TLS zero-temperature decay rate rises rapidly, roughly as the cube of the TLS transition frequency, consistent with the approximate 2D phonon quasi-mode density to which the TLS are coupled (c.f., Fig. 18(c)).

Figure 20 displays the resulting damping (c.f., Eqs.(G19,G27)) and frequency shift (c.f., Eqs.(G18,G26)) of the high-Q breathing mode versus temperature due to coupling with the TLS bath in the acoustic mode volume V_m. In these simulations we performed 500 trial runs of random TLS ensembles. Both the average and standard deviation
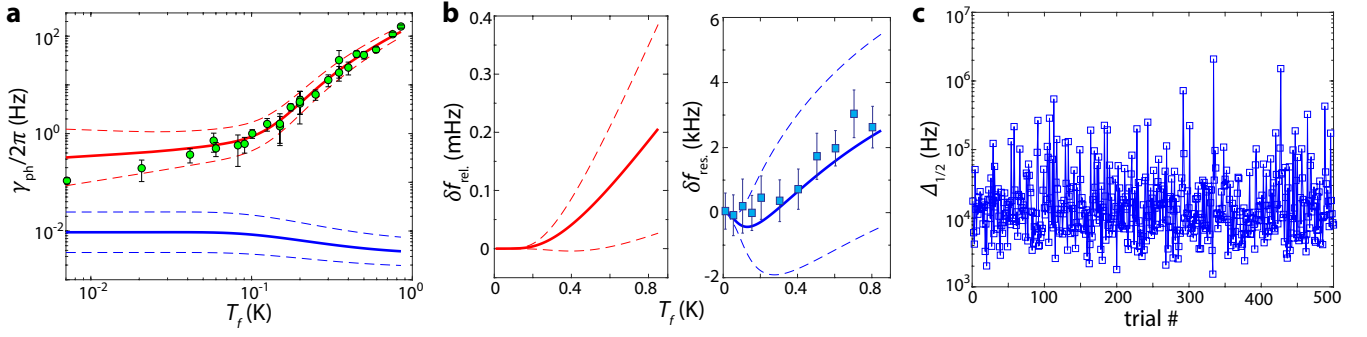
FIG. 20. **Modeled breathing mode interactions with a TLS bath.** Modeling was performed using the acoustic mode and TLS properties found in Tab. IV. Simulations were performed using 100 ensemble instances of $N_{\text{TLS},m} = 3920$ randomly positioned (sampled from a set of 101 fixed positions) and oriented TLS within the breathing mode acoustic volume, V_m. The transition energy of each TLS are also sampled from a random distribution. **a**, Relaxation (red curves) and resonant (blue curves) damping of the acoustic breathing mode versus TLS bath temperature. Solid (dashed) curves represent the average (1-σ standard deviation in log-space) of the 500 simulation trials run. Measured damping of device D is shown as filled green circles. **b**, Breathing mode frequency shift versus temperature due to relaxation (left plot) and resonant (right plot) TLS interactions. The solid cyan curve is a curve resulting from single TLS ensemble trial. Measured frequency shift of mode of device D is shown as filled blue squares. **c**, Full-width half-maximum of the time-averaged frequency jitter of the breathing acoustic mode resulting from resonant interactions with all of the TLS for each of the 500 trial ensembles. Here we assume a $\Lambda_{\text{TLS}} = 2$ Hz excitation rate of each TLS due to weak optical absorption.

of the damping and frequency shift are shown. Also shown are the effects of both 'resonant' $\hat{\sigma}_x$-interactions and 'relaxation' $\hat{\sigma}_z$-interactions with the TLS. As can be clearly seen, the resonant TLS damping of the breathing mode at the lowest temperatures $T_{\text{f}} \lesssim 100$ mK is predicted to be roughly an order of magnitude smaller than the relaxation damping. In addition, as the temperature is increased above that of $\hbar\omega_{\text{m}}/k_{\text{B}} \approx 200$ mK, the resonant damping term begins to saturate due to the thermal excitation of TLS. The overall suppression of the resonant TLS damping is due to the presence of the acoustic bandgap, which dramatically reduces the decay rate of TLS nearly-resonant with the breathing mode due to a lack of localized quasi-normal phonon modes in the gap. Instead, the typically weaker relaxation damping from non-resonant TLS outside the acoustic bandgap dominates the simulated breathing mode damping.

The correspondence of the simulated TLS relaxation damping of the breathing mode with the measured breathing mode damping is striking not only in the overall magnitude of the predicted damping but also in its temperature dependence. At temperatures below $T_{\text{f}} \approx 100$ mK, where the thermally excited TLS that contribute to relaxation damping have transition frequencies below the acoustic bandgap and interact with a quasi-1D phonon bath, the damping is seen to have a reduced, linear to sub-linear dependence with temperature. Above $T_{\text{f}} \approx 100$ mK, thermal excitation of TLS with transition frequencies above that of the acoustic bandgap begin to contribute to the damping. In this spectral range the phonon density of states in the OMC cavity structure is approximately linear with frequency, resulting in TLS decay times which scale quadratically with transition frequency. This sets the temperature scaling of the breathing mode relaxation damping, which is also seen to scale approximately quadratically with temperature above $T_{\text{f}} \approx 100$ mK.

Another feature of the simulated TLS damping of the breathing mode is the large fluctuation from trial-to-trial in the low temperature ($T_{\text{f}} \lesssim 100$ mK) portion of the damping versus temperature curve. This shows up as a large variance in magnitude and temperature trend of the relaxation damping at these low temperatures in simulation. The dashed curves in Fig.20(a) represent the range of temperature curves within one standard deviation of the mean curve *on a log scale*. In order to better appreciate what the level of fluctuations in the predicted low temperature damping are for both resonant and relaxation TLS damping of the breathing mode, we plot in Fig. 21 the trial-to-trial variations of the simulated damping factors at $T_{\text{f}} = 7$ mK. The variations are substantial, with a standard deviation in log-space of a little over an order of magnitude. This is consistent with our measured observations, as can be gleaned from the plot of measured breathing mode energy damping rates versus phononic bandgap shield number in Fig. 2(d). For acoustic shield periods greater than about 6, where radiation damping is estimated to be minimal, the measured damping rates also range over a little over an order of magnitude. This variance is explained in the TLS model by the very small number (handful) of TLS with transition frequency below 100 MHz within the breathing mode volume, which owing to their random positioning and orientation, results in a large variation in breathing mode damping.

The simulated frequency shift versus temperature, for both resonant and relaxation interactions with the TLS
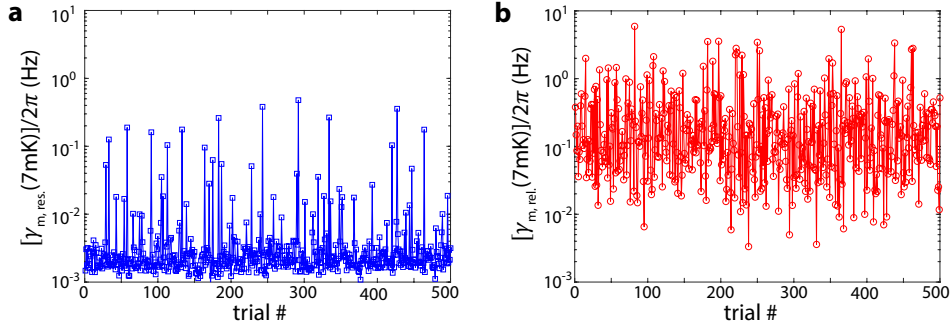
FIG. 21. **Fluctuation from trial-to-trial in the simulated low temperature damping at 7mK. a,** Simulated TLS-resonant-interaction damping at temperature $T = 7$ mK of breathing mode as a function of different TLS ensemble trials. **b,** Simulated TLS-relaxation damping at temperature $T = 7$ mK of breathing mode as a function of different TLS ensemble trials.
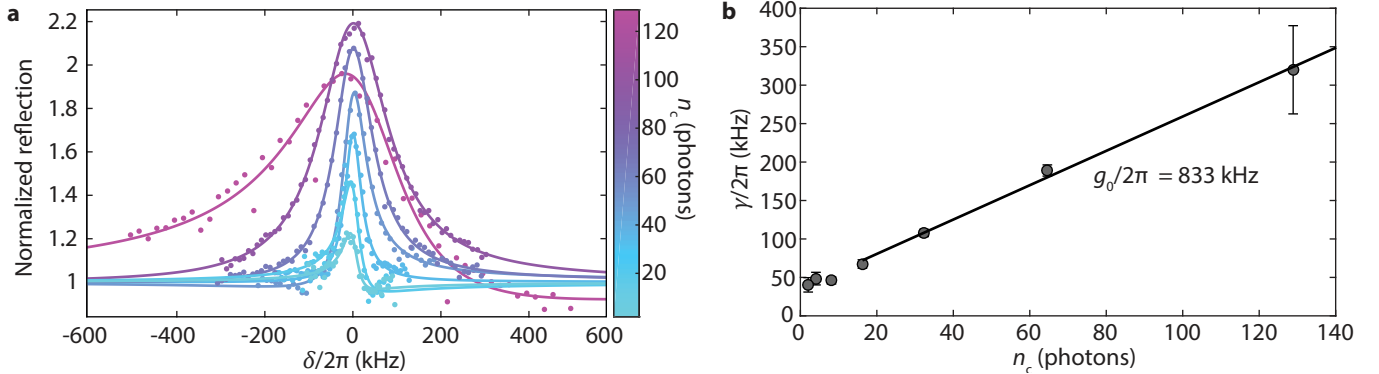

FIG. 22. **EIT linewidth measurements of Device E at $T_f = 10$ mK. a,** Normalized reflection amplitude for probe photons as a function of pump photon number. The reflection peak represents an EIT-like transparency window approximately centered within the bare optical cavity line (~ 1 GHz wide). Asymmetry in the trace at $n_c = 129$ can be attributed to an effective detuning shift, likely caused by thermal shifting of the cavity at high input power. These EIT measurements were performed on a device E having seven acoustic shield periods and mechanical $Q = 1.5 \times 10^{10}$ measured via ringdown. **b,** Plot of the fit mechanical linewidth versus pump photon number from the EIT curves of (**a**). At $n_c \lesssim 10$, the time-averaged mechanical linewidth saturates to 40 kHz due to mechanical frequency jitter. At higher n_c, the mechanical mode is broadened by optomechanical back-action, the slope of which yields $g_0/2\pi = 833$ kHz.

ensemble, is shown in Fig. 20(b). In these plots we have referenced the frequency shift to that at the lowest measured temperature ($T_f \approx 7$ mK). The relaxation component of the frequency shift (left panel) is estimated to be in the milli-Hz range, far below the kHz-scale frequency shift due to resonant interactions with the TLS (right panel). The resonant TLS frequency shift of the breathing mode versus temperature has a simulated mean curve averaged over the 500 ensemble trials that roughly follows the digamma function response of Eq. (G18), with the breathing mode frequency initially shifting lower and then begin to increase around $\hbar\omega_m/k_B \approx 200$ mK, followed by a monotonic (logarithmic) increase in frequency for higher temperatures. The simulated frequency shift versus temperature curve has a large variance however (range of curves within standard deviation of mean curve are bounded by dashed curves), depending sensitively on the magnitude and sign of the detuning of the TLS closest to resonance with the breathing mode. The measured frequency shift with temperature of device D is plotted alongside the simulated resonant component of the TLS-induced frequency shift in Fig. 20(b), showing good correspondence with the mean curve.

As discussed in the main text, and shown in Fig. 3(d), the measured frequency jitter when averaged over minutes has a spectral full-width at half-maximum (FWHM) of $\Delta_{1/2} \approx 3.5$ kHz, approximately independent of temperature ($T_f = 7 - 850$ mK) and optical probing power ($n_c = 0.3 - 0.02$). Over $\lesssim 0.1$ s timescales one can resolve this frequency jitter in the time domain (c.f., main text Fig. 3(c)). In order to estimate the amount of frequency jitter that the acoustic breathing mode might incur due to interactions with TLS we have assumed in our simulations that all TLS are being excited at a rate faster than the measurement averaging time of a few minutes. Assuming all TLS are fluctuating independently, we can then write using Eq. (G18) without the temperature dependence and assuming that

$$\langle \hat{\sigma_z}^2 \rangle = 1,$$

$$\Delta_{1/2} = 2(2\log[2])\left(\sum_{\text{TLS}} \left(\bar{g}_{\text{t},m}^2[\mathbf{r}_{\text{TLS}}] \right)^2 \right.$$

$$\left. \times \left\{ \frac{\omega_{\text{TLS}} - \omega_{\text{m}}}{(\omega_{\text{TLS}} - \omega_{\text{m}})^2 + ((\delta\Gamma_{1,\text{TLS}})_{\text{ph}}/2 + \Gamma_{\phi,\text{TLS}})^2} + \frac{\omega_{\text{TLS}} + \omega_{\text{m}}}{(\omega_{\text{TLS}} + \omega_{\text{m}})^2 + ((\delta\Gamma_{1,\text{TLS}})_{\text{ph}}/2 + \Gamma_{\phi,\text{TLS}})^2} \right\}^2 \right)^{1/2}, \quad \text{(H1)}$$

where the prefactor $2(2\log[2])$ accounts for the conversion between the standard deviation and the FWHM of a normal distribution.

TLS excitation occurs naturally through thermal excitation, although the rate of excitation in that case depends strongly on temperature, which we do not observe in our measurements. However, we also know that the optical probing of the mechanics can lead to optical-absorption-induced excitation of a hot bath that damps the mechanics. It seems reasonable then to assume that the same optical absorption would also excite a broad spectrum of TLS, thus leading to their contribution to the breathing mode frequency jitter. In the model we have taken the optical-absorption-induced pumping rate of the TLS to be $\Lambda_{\text{TLS}} = 2$ Hz, which for context represents 0.2% of the hot bath heating rate $\gamma_{\text{p}} n_{\text{p}}$ of the breathing mode at $n_{\text{c}} = 0.1$. So even if the TLS are driven much more weakly than the breathing mode due to optical absorption effects, one would need to probe at lower optical probe powers than currently accessible in our experiments ($n_{\text{c}} = 0.02$) to see a reduction in the time-averaged frequency jitter of the mechanics, where time averaging is performed over a second or longer. Note that pumping-induced saturation effects of the TLS have also been included in the simulation although their effects on both the damping and the overall frequency shift is minor.

The resulting simulated frequency jitter FWHM of the breathing mode, $\Delta_{1/2}$, is plotted in Fig. 20(d) for each of the 500 trial TLS ensembles. The variation from trial-to-trial of the frequency jitter is substantial, with the jitter ranging from kHz to (in rare cases) MHz. Although the measured frequency jitter of device D, comprehensively studied in the main text, lies on the low end of this spectrum at $\Delta_{1/2} \approx 3.5$ kHz, this device is also on the low end of the range of measured values in our experience. As an example, the measured linewidth of another high-Q device (device E in Table I) is shown in Fig. 22. In this case, the linewidth at low optical probe power is found to saturate to a FWHM of ~ 40 kHz, closer to the mean simulated value. The large fluctuation in the measured frequency jitter from device-to-device, consistent with the model, is again an indication of the sensitivity of the breathing mode to a select few TLS in the near-resonant regime.