I. OPTICAL AND MECHANICAL DESIGN

The geometry of the optomechanical crystal (OMC) studied in this work is numerically optimized for optical and mechanical quality, as well as optomechanical coupling, via finite-element method (FEM) simulation in COMSOL Multiphysics [1]. In Fig. S-1a, a top view of the OMC shows the center defect in the 600 nm wide by 220 nm thick Si nanobeam. The larger holes on the ends of the nanobeam support simultaneous bandgaps by 220 nm thick Si nanobeam. While the smaller holes in the center of the nanobeam perturb the bandgaps such that optical and mechanical modes are co-localized [2]. The fundamental optical mode has a nominal wavelength of 1535 nm (Fig. S-1b), and the acoustic breathing mode has a nominal resonance frequency of 3.85 GHz (Fig. S-1c). Through simulation of the dependence of the effective index of refraction of the structure on the mechanical mode through moving-boundary and photo-elastic effects, we determine the nominal optomechanical vacuum coupling rate $g_0$ to be 870 kHz. This constant represents the shift in optical resonance frequency due to zero-point fluctuations of the mechanical resonator. In fabrication, arrays of the nominal design in Fig. S-1 are scaled by $\pm$ 2% to account for geometrical imperfections, leading to a range of realized optical and mechanical resonance frequencies. For the particular device studied here, the optical wavelength is 1545 nm, the mechanical frequency is 3.6 GHz, and $g_0 = 840$ kHz, as discussed in the main text.

II. FABRICATION

The devices are fabricated from a silicon-on-insulator (SOI) wafer (SOITEC, 220 nm device layer, 3 µm buried oxide) using electron beam lithography followed by reactive ion etching (RIE/ICP). The Si device layer is then masked using ProTEK PSB photoresist to define a mesa region of the chip to which a tapered lensed fiber can access. Outside of the protected mesa region, the buried oxide is removed with a plasma etch and a trench is formed in the underlying silicon substrate using tetramethylammonium hydroxide (TMAH). The devices are then released in hydrofluoric acid (49% aqueous HF solution) and cleaned in a piranha solution (3-to-1 H$_2$SO$_4$:H$_2$O$_2$) before a final hydrogen termination in diluted HF.

III. EXPERIMENTAL SETUP

The full experimental setup for heterodyne spectroscopy and mechanical thermometry is shown in Fig. S-2. A fiber-coupled, wavelength-tunable external cavity diode laser is used as the light source, and a small percentage of the laser output is sent to a wavemeter ($\lambda$-meter) for frequency stabilization. The remaining laser power is split into a high-power ($\approx$0.7 mW) local oscillator (LO) path and a low-power ($\approx$20 µW) signal path. The signal beam is sent through an electro-optic modulator (a-m) to stabilize the signal intensity and a variable optical attenuator to allow control of the probe power sent to the cavity. The signal is sent into an optical circulator which directs the signal beam to the dilution refrigerator in which the fiber terminates with a lensed tip for end-fire coupling to the device. The cavity reflection then circulates to a variable coupler (VC) and mixes with the LO before being detected on a pair of balanced photodiodes (BPD). The difference photocurrent is then amplified and its noise power spectral density (NPSD) is measured on a real-time spectrum analyzer (RSA). The LO is first sent through an electro-optic phase modulator ($\varphi$-m) which produces optical sidebands at $\pm(\omega_m/2\pi - 50$ MHz) for the purpose of mixing the mechanically induced signal modulation down to within the 100 MHz bandwidth of the BPD circuit. An erbium-doped fiber amplifier (EDFA) and a variable optical attenuator (VOA) are used to set the power of the LO sidebands, and the appropriate sideband is selected by a high-finesse tunable Fabry-Pérot filter before recombining it with the signal.

Fig. S-2b shows a scanning electron micrograph (SEM) of a typical device, consisting of a central Si waveguide...
side-coupled to two nanobeam OMCs, all of which is surrounded by an acoustic radiation shield phononic crystal. Light coupled into the central waveguide either reflects from a photonic crystal mirror or, when resonant, evanescently couples to one of two side-coupled OMCs [3,4], where the coupling rate $\kappa_e$ is controlled by the design of the gap size between the waveguide and the cavity (Fig. S-2c). Some cavity light decays through parasitic scattering at a rate $\kappa_i$, while the remainder couples back into the central waveguide and is collected from the central waveguide tip (Fig. S-2d) by the lensed fiber to be guided to the photodetector in Fig. S-2b. By tuning the laser over the cavity mode, we observe in the reflected signal on the output of the dilution refrigerator (Fig. S-2e) a fiber collection efficiency of $\eta_{cpl} = 34.7\%$ and an optical linewidth of $\kappa/2\pi = 529$ MHz, composed of $\kappa_e/2\pi = 153$ MHz and $\kappa_i/2\pi = 376$ MHz.
FIG. S-3: Sample mounting and fiber coupling in the dilution refrigerator. a. Photo of the attocube nanopositioning stages mounted to the bottom of the dilution refrigerator mixing plate. b. Photo of the lensed fiber tip mounted on the attocube positioners, illustrating alignment to a Si test chip. c. Diagram showing the lensed fiber (not-to-scale) coupling to a silicon OMC shown in the SEM image, with the superimposed optical mode FEM simulation of optical intensity.

IV. FIBER COUPLING IN THE DILUTION REFRIGERATOR

The microchip sample is mounted to the mixing chamber of the dilution refrigerator on a copper bracket assembly to allow for in-situ fiber coupling. The bracket is located 10 cm from a calibrated ruthenium oxide resistive temperature sensor. The temperature readings from this sensor are used for comparison and validation of our calibrated optomechanical thermometry measurements, with accuracies of 50 mK near temperatures of 4 K, and 0.06 mK near 10 mK.

We utilize an end-fire coupling scheme to probe individual devices with an anti-reflection-coated tapered lensed fiber. The lensed fiber tip is clamped down on a position encoded piezo xyz-stage inside the dilution refrigerator (Fig. S-3b), which allows nanopositioning with respect to the sample. When mounting the fiber and sample, the fiber is only roughly aligned to within a few millimeters (Fig. S-3b). After cooling the experiment from room temperature to 4 K, we monitor the reflected optical power on a slow photodetector as we carefully lower the fiber tip to match the height of the device layer. The distinct reflection of the device layer allows us to iteratively adjust the fiber position and optimize the coupling to each device (Fig. S-3c).

The fiber-tip launches the light to free-space and focuses it to a beam waist of 2.5 μm at a focal distance of 14 μm. We position the fiber such that the beam waist aligns to a silicon waveguide tip (Fig. S-2b,d) that matches the convergence of the optical field, which becomes a guided mode in the waveguide [5–7].

The design of the tapered waveguide coupler is similar to that presented in Ref. [7], where the tip of the waveguide is mode matched to an input Gaussian field of the appropriate width and adiabatically tapered up to the full width of the photonic crystal mirror section. The major distinction is the use of end-fire coupling utilizing a lensed fiber rather than butt-coupling with a cleaved single-mode fiber. In contrast to the high-stress Si₃N₄ utilized in Ref. [7] a Si waveguide will begin to sag if the waveguide is made too long, even in the presence of a supporting tether, which will lead to misalignment of the waveguide tip and bending losses that inhibit the coupling efficiency. As such, it is necessary to substantially reduce the length of the waveguide taper, which necessitates a smaller overall change in waveguide width to maintain the adiabaticity of the taper. Consequently we are forced to use a larger waveguide width at the tip and a correspondingly smaller mode field diameter such as can be obtained with a lensed fiber.

The broadband reflection spectrum of the coupler, calculated using finite-difference-time-domain simulation [3], is shown for an ideal coupler in Fig. S-4a. The presence of fringes in the reflection spectrum are consistent with the estimated weak reflectivity of the waveguide–air interface ($R \approx 0.5\%$), which forms a low-finesse Fabry–Pérot cavity with the high reflectivity photonic crystal mirror at the other end of the waveguide. From the fringe visibility we can back out the expected single-pass coupling efficiency $\eta_{\text{cpl}}$ shown in Fig. S-4b [7].

The calculated reflection spectrum and single-pass efficiency for the device used in this work (using dimensions determined from SEM images) is also shown in Fig. S-4a,b and is expected to be $\eta_{\text{cpl}} \approx 60\%$ at the cavity resonance. The actual single-pass efficiency of the measured device is found to be $\eta_{\text{cpl}} \approx 50\%$. The difference from simulations is attributed to the difficulty of measuring the exact dimensions of the sample using the SEM, as a small difference of even 10 nm in the width of the waveguide tip can have a significant effect on the mode matching and thus on the overall coupling efficiency.
V. CALIBRATION OF THE OPTICAL TRANSDUCTION OF MECHANICAL MOTION

To perform accurate thermometry of the mechanical mode, it is necessary to calibrate the detection efficiency of the setup. We first measure the efficiency of transmission in the circulator from port 1 to port 2 ($\eta_{12} = 88\%$), and from port 2 to port 3 ($\eta_{23} = 84\%$). These values are measured once when the optical components are connected and do not change. These calibrations are used to determine the reflection efficiency of the device and the overall detection efficiency of the heterodyne setup.

To measure device efficiency, the laser is tuned off-resonance from the optical mode (where the device should act as a near-perfect mirror) and a continuous-wave signal of input power $P_{\text{in}}$ is sent into port 1 of the circulator, leading to a power $\eta_{12} P_{\text{in}}$ exiting port 2 of the circulator. The optical losses incurred in the path from port 2 to the device-under-test are accumulated into an efficiency factor $\eta_{\text{cpl}}$, which includes signal loss in the fiber path through the fridge, mode-mismatch between the lensed-fiber to tapered-waveguide tip, and mode-scattering from the waveguide tip to the photonic-crystal mirror (Fig. S-3). These losses are incurred twice in reflection back to the circulator, so a power of $\eta_{\text{cpl}}^2 \eta_{12} P_{\text{in}}$ propagates into port 2 and $\eta_{23} \eta_{\text{cpl}}^2 \eta_{12} P_{\text{in}}$ propagates out of port 3 of the circulator. An optical switch (SW3) is used to send this signal to a power meter (PM), and thus the coupling efficiency is determined as

$$\eta_{\text{cpl}} = \sqrt{\frac{P_{\text{PM}}}{\eta_{23} \eta_{12} P_{\text{in}}}} = 34\%. \quad (S-1)$$

To calibrate the overall detection efficiency, we must also determine the efficiency of the heterodyne detector itself, which includes the intrinsic quantum efficiency of the BPD, the alignment of the polarization between the LO and the signal, and the degree to which the LO power overcomes the electronic noise of the detector. This is accomplished by using the amplitude modulator to create optical sidebands detuned from the signal by the mechanical frequency while the laser is tuned off-resonance from the optical mode. The optical switches SW1 and SW2 are used to route the signal through a tunable filter to select a single sideband which is sent through the device and onto the BPD. The power $P_{\text{cal}}$ in this sideband is directly measured on the PM at SW3, and the photocurrent NPSD ($S_{\text{II}}(\omega)$) as transduced on the RSA is given by

$$S_{\text{II}}(\omega) = S_{\text{dark}} + \frac{G_e^2}{P_{\text{LO}}} S_{\text{SN}}^2 \left(1 + \frac{\eta_{\text{VC}} \eta_{\text{det}} S_{\text{cal}}(\omega)}{h\omega_o}\right), \quad (S-2)$$

where $S_{\text{dark}}(\omega)$ is the electronic NPSD of the detector, $S_{\text{SN}} = \sqrt{2h\omega_o P_{\text{LO}}}$ is the optical shot-noise NPSD arising from $P_{\text{LO}}$ of LO optical power at optical frequency $\omega_o$, which lies an order of magnitude above the electronic noise, and $S_{\text{cal}}$ is the NPSD of the signal, where $\int_{-\infty}^{\infty} S_{\text{cal}}(\omega) d\omega = P_{\text{cal}}$. The gain factor $G_e$ represents the conversion from optical power to voltage while $R_{\text{L}}$ is the input impedance of the RSA. The total noise floor $S_{\text{noise}} = \frac{G_e^2 S_{\text{SN}}^2}{P_{\text{LO}}} + S_{\text{dark}}$ is measured with the signal beam blocked, while $S_{\text{dark}}$ is measured independently with both signal and LO beams blocked. When referenced back to the PM, the calibration tone (with NPSD $S_{\text{cal}}(\omega)$) picks up losses in the VC and BPD, as parametrized in $\eta_{\text{VC}}$ and $\eta_{\text{det}}$, respectively. The efficiency of the heterodyne receiver is extracted as

$$\eta_{\text{VC}} \eta_{\text{det}} = \frac{h\omega_o}{P_{\text{cal}}} \int_{-\infty}^{\infty} S_{\text{II}}(\omega) - S_{\text{noise}} d\omega \frac{S_{\text{noise}}}{S_{\text{dark}}} 2\pi = 72\%. \quad (S-3)$$
This combined with the measured device coupling efficiency, yields the overall measurement efficiency $\eta$ used for calibrated mechanical thermometry as

$$\eta = \eta_{cpl} \eta_{23} \eta_{VC} \eta_{det} = 20\%.$$  \hspace{1cm} (S-4)

VI. HEATING AND DAMPING VIA THREE-PHONON SCATTERING PROCESSES

Though a detailed microscopic calculation of the additional heating and damping due to optical absorption is beyond the scope of this work, some qualitative insight into the nature of the locally heated mechanical bath and its coupling to the mechanical mode of interest can be obtained from consideration of a simplified model of the phonon-phonon interactions. At the temperatures considered in this work ($T < 10$ K) the mean free path of the thermal phonons is expected to be much larger than the wavelength of the mechanical mode of interest. Consequently, the damping should be described by the Landau-Rumer theory, where losses occur primarily due to three-phonon mixing with the local thermal environment due to anharmonicity in the Si lattice [9][10].

Consider first a toy model where two high-frequency modes, with frequencies $\omega_1$ and $\omega_2$ respectively, are coupled with the mode of interest at frequency $\omega_m$ ($\omega_1 - \omega_2 = \omega_m$). To first order in perturbation theory, the scattering rates into and out of the mechanical cavity mode due to the lowest order anharmonic interactions can be given by [10] $\Gamma_+ = A(n_m + 1)(n_2 + 1)n_1$ and $\Gamma_- = A(n_1 + 1)n_m n_2$, where $A$ is a constant which depends on the matrix element of the anharmonic potential, and $n_1$, $n_2$ and $n_m$ are the number of quanta in each of the three mechanical modes. Thus, in the absence of other dissipative processes a simple rate equation for the population of the mechanical cavity mode is given by

$$\dot{n}_m = \Gamma_+ - \Gamma_- = -A(n_2 - n_1)n_m + A(n_2 + 1)n_1,$$  \hspace{1cm} (S-5)

which has the same form as the equation for a harmonic oscillator interacting with a bath of occupation $n_p$ with a coupling rate $\gamma_p$, where

$$n_p = \frac{n_2(n_2 + 1)}{n_2 - n_1}, \quad \gamma_p = A(n_2 - n_1).$$  \hspace{1cm} (S-6)

If the two high-frequency modes are both in equilibrium with each other at some elevated temperature $T_p$, such that $n_i = (\exp\left(\frac{h\omega_i}{k_B T_p}\right) - 1)^{-1}$, it is easy to show that $n_p$ is simply given by the Bose-Einstein occupation factor for the mechanical mode at temperature $T_p$ ($n_p = (\exp\left(\frac{h\omega_m}{k_B T_p}\right) - 1)^{-1}$). We can also see that the scattering rate $\gamma_p$ will depend on $T_p$ through the temperature dependence of the population difference $n_2 - n_1$. This dependence can be approximately linear or exponential in $T_p$, depending on the value of $(h\omega_2)/(k_BT_p)$.

More realistically, the optical absorption process will populate a range of high-frequency phonon modes above some cutoff frequency ($\omega_c$), which can contribute to the heating. On the assumption that these modes come into equilibrium with each other at some elevated temperature $T_p$, we can easily show that the expression for $n_p$ is unchanged, and the effective bath occupancy $n_p$ will be given by the above expression. The scattering rate will now be given by

$$\gamma_p = \int_{\omega_c}^{\infty} d\omega \rho(\omega)A(\omega, \omega_m)(n(\omega, T_p) - n(\omega + \omega_m, T_p)),$$  \hspace{1cm} (S-7)

where $\rho(\omega)$ is the density of modes at frequency $\omega$, the matrix element $A$ now can depend explicitly on frequency, and $n(\omega, T)$ is just the Bose-Einstein occupation at frequency $\omega$ and temperature $T$. In general, we can obtain limited analytical results if we assume that the product of the density of states and the matrix element obeys some power law as a function of frequency, $\rho(\omega) A(\omega, \omega_m) \propto \omega^a$. For example, under a simple continuum elastic model the product would be given by $\rho(\omega) A(\omega, \omega_m) \propto \omega(\omega_m + \omega)^b$ [10]. Making a simple change of variables $x = \frac{h\omega}{k_B T_p}$, and assuming $\omega_m \ll \omega_c$, we then arrive at the approximate relation

$$\gamma_p \propto \omega_m T_p^{a+1} \int_{x_c}^{\infty} dx x^a e^{-x^2}; \quad x_c = \frac{h\omega_c}{k_B T_p},$$  \hspace{1cm} (S-8)

Considering first the limiting case $x_c \gg 1$, we note that the integral can be approximated by

$$\int_{x_c}^{\infty} dx x^a e^{-x^2} \approx \int_{x_c}^{\infty} dx x^a e^{-1};$$  \hspace{1cm} (S-9)

where $\Gamma(a, z)$ is the upper incomplete Gamma function. For real values of $z$, this function has the asymptotic behavior $\Gamma(a, z) \rightarrow z^{a-1} e^{-z}$ as $|z| \rightarrow \infty$. This leads to the approximate scaling law

$$\gamma_p \propto T_p^{a+1} x_c e^{-x_c} = T_p e^{-\left(\frac{h\omega_c}{k_B T_p}\right)};$$  \hspace{1cm} (S-10)

for $T_p \ll \frac{h\omega_c}{k_B}$. In the limiting case $x_c \ll 1$, the lower limit of the integral can be extended to 0, and a simple integration by parts shows that, for $a > 1$,

$$\int_0^{\infty} dx x^a e^{-x^2} = a \int_0^{\infty} dx x^{a-1} e^{-1} = a \Gamma(a) \zeta(a),$$  \hspace{1cm} (S-11)

where $\Gamma(a)$ is the gamma function and $\zeta(a)$ is the Riemann zeta function. As this is a simple constant, we get the scaling law $\gamma_p \propto T_p^{a+1}$, when $T_p \gg \frac{h\omega_c}{k_B}$ and $a > 1$. 

In Figure S-5b, the low temperature end of the inferred $\gamma_p(T_p)$ data is seen to fit well to a curve $\propto T_p \exp (-T_c/T_p)$, where the cut-off frequency is $T_c \approx 2$ K ($\omega_c/2\pi \approx 35$ GHz). Although this simple model does not capture all of the features of the measured $\gamma_p(T_p)$ curve (in particular, the kink at $T_p \approx 3$ K), it does show that the damping rate due to 3-phonon mixing can vary substantially as a function of $T_p$, particularly at low temperatures in the vicinity of $T_p \lesssim \hbar \omega_c/2k_B$.

VII. TEMPERATURE DEPENDENCE OF FREQUENCY NOISE

At sub-kelvin fridge temperatures, frequency noise of the 'breathing' mode resonance is seen to dominate the time-averaged measured linewidth for low optical probe powers (Fig. S-5a). In the case of the red-detuned probe, frequency noise is observed for $n_c < 1$ before optomechanical damping kicks in, while the on-resonance measurement shows frequency noise for $n_c < 10$ before the intrinsic energy damping ($\gamma_p$) begins to dominate. Combining the sections of both data sets which are dominated by frequency noise, a power law fit shows linewidth scaling as $n_c^{-0.23}$. By comparing to the on-resonance measured mode occupancy (Fig. 4a in main text), the frequency noise can also be plotted with respect to the optical absorption driven temperature $T_p$ as shown in Fig. S-5b. A power law fit of the time-averaged linewidth in the frequency jitter dominated regime, shows a frequency jitter scaling $\propto T_p^{-0.9}$.

Similar inverse temperature scaling of frequency noise has been observed in superconducting microwave resonators [1], with substantial evidence indicating the source to be fluctuations in two-level tunneling states (TLS) of near-field amorphous materials [12, 13]. These TLS also couple to phonons, contributing to not only the dielectric properties of the material, but also the elastic properties. As such, the microwave mechanical modes of the devices studied here can be expected to couple to TLS in a similar fashion as for the microwave electromagnetic resonators. The decrease in frequency noise with temperature can then be attributed to the thermal excitation and saturation of the TLS. The presence of TLS in the Si devices of this work likely stems from the formation of a native oxide on the Si surfaces of the patterned nanobeam. Careful measures, involving a final HF etch during which time a native oxide would at least partially reform on the Si surfaces of the patterned nanobeam. Careful measures, involving a final wet etch in HF acid [14], were made to passivate the Si surfaces of the nanobeam in order to reduce the optical absorption from sub-bandgap surface electronic states. This surface cleaning should have been effective in removing the surface oxide as well, however, a one hour procedure was required to load the sample into the dilution refrigerator, during which time a native oxide would at least partially reform on the Si surfaces of the device. This native oxide is likely to be the source of the TLS. Future work will look to substantially reduce the device loading time after HF surface cleaning, hopefully reducing both the optical absorption and TLS density.


