

State preparation and tomography of a nanomechanical resonator with fast light pulses

Juha T. Muhonen,^{1,2,*} Giada R. La Gala,¹ Rick Leijssen,¹ and Ewold Verhagen^{1,†}

¹*Center for Nanophotonics, AMOLF, Science Park 104, 1098 XG, Amsterdam, The Netherlands*

²*Department of Physics and Nanoscience Center, University of Jyväskylä,
P.O. Box 35, FI-40014 University of Jyväskylä, Finland*

Pulsed optomechanical measurements enable squeezing, non-classical state creation and backaction-free sensing. We demonstrate pulsed measurement of a cryogenic nanomechanical resonator with record precision close to the quantum regime. We use these to prepare thermally squeezed and purified conditional mechanical states, and to perform full state tomography. These demonstrations exploit large photon-phonon coupling in a nanophotonic cavity to reach a single-pulse imprecision of 9 times the mechanical zero-point amplitude x_{zpf} . We study the effect of other mechanical modes which limit the conditional state width to $58x_{\text{zpf}}$, and show how decoherence causes the state to grow in time.

Measurement and control of mechanical motion at the quantum level is of wide interest because of the quantum technologies it would enable and the possibility to probe decoherence in massive quantum systems. Cavity optomechanical demonstrations [1] of quantum control over mechanical resonators included ground state cooling [2–4], quantum squeezing [5–7], entanglement [8, 9], and exchanging individual quanta between the mechanical oscillator and qubits [2, 10].

Measurement and control are intimately linked. At a basic level, any measurement process is a competition between information gain (measurement rate), decoherence processes, and noise added to the system by the measurement (backaction). If the measurement rate overcomes those detrimental effects, control over the system can be achieved, allowing for mechanical oscillators for example feedback cooling to the ground state [11–13]. A continuous measurement of a mechanical resonator’s displacement is subject to the so-called standard quantum limit (SQL). The SQL is a manifestation of Heisenberg’s uncertainty principle as these measurements simultaneously observe the two non-commuting motional quadratures, hence giving a lower limit, equal to the zero-point fluctuation amplitude x_{zpf} , to the noise added by the measurement [14, 15]. Methods to evade this backaction limit by moving away from simple continuous displacement measurements were demonstrated in various quantum systems [16–21], including in sideband-resolved opto- or electromechanical cavities probed by two-tone fields [22–24], where the measurement is only sensitive to one motional quadrature.

Another method for backaction-evading measurements that has been put forward is that of pulsed measurements, where mechanical motion can be neglected during a short interaction [25, 26]. A single ‘snapshot’ measurement of a harmonic oscillator’s position measures one of the quadratures with potentially unlimited precision, as all backaction is introduced to the orthogonal quadrature with no effect to the future evolution of the mea-

sured quadrature. With sufficient precision the mechanical oscillator is then prepared to a squeezed state, conditioned on the measurement. Importantly, the ability to probe either quadrature precisely allows full quantum state tomography [27]. Combined with nonlinear displacement measurement or non-Gaussian states of light, pulsed interactions can also induce other nonclassical states [28, 29]. Moreover, proposals recently suggested to exploit pulsed measurement for swap operations between mechanics and light [30] and creation of macroscopic superpositions [29].

Despite these extensive theoretical proposals, so far only a single proof-of-principle experimental demonstration of pulsed optomechanical measurements has been reported [31], at elevated temperature and without cavity enhancement. In order to reach quantum-level accuracy with a single pulsed measurement one needs to fulfill the challenging requirement $8\eta\sqrt{N_P}g_0/\kappa \gtrsim 1$, where g_0 is the cavity frequency shift for a displacement x_{zpf} (the photon-phonon coupling rate), η the coupling efficiency of light to the cavity, and N_P the number of photons in the pulse [26], while at the same time offering sufficient cavity linewidth κ to accommodate a pulse of duration $\tau_P \ll 2\pi/\omega_m$, where ω_m is the mechanical oscillation frequency.

In this work, we address these challenges using a cavity optomechanical system based on a sliced photonic crystal nanobeam, allowing large photon-phonon coupling rates [32, 33]. We demonstrate pulsed optomechanical measurements close to the quantum regime, achieving a record-low shot-noise limited single-pulse measurement imprecision of $9x_{\text{zpf}}$, constrained mainly by the optical detection efficiency. We prepare both thermally squeezed and purified (cooled) conditional mechanical states, and perform full state tomography on these states. We study how additional mechanical modes affect the conditional state, limiting its width in this experiment to $58x_{\text{zpf}}$. We also demonstrate how thermal dephasing can be tracked by recording the evolution of the state at longer

timescales. In addition, we show how post-selection allows maximizing measurement sensitivity even though the large optomechanical interaction strength pushes our system deep into the regime of nonlinear optomechanical interaction where sensitivity is reduced [33].

Figure 1(a) shows a diagram of the sliced silicon nanobeam. The device, presented in Ref. 33, hosts a photonic crystal nanocavity mode whose resonance frequency (204 THz) depends very sensitively on the flexural movement of the two beam halves. These move roughly independent of each other with two mechanical mode frequencies around $\omega_m/(2\pi) \approx 3$ MHz, separated by ~ 120 kHz. The photon-phonon coupling rate $g_0/(2\pi) \approx 25$ MHz is approximately equal for both modes. The optical cavity linewidth $\kappa/(2\pi) \approx 20.4$ GHz enables practically instantaneous measurements of mechanical position while still achieving $g_0/\kappa > 10^{-3}$. All reported measurements are performed on the same sample at a temperature of 3.2 K. The exact mechanical frequencies and damping rates drift with time for reasons not fully understood, and are mentioned in the captions for each dataset.

The sample is incorporated in a homodyne interferometer (Fig. 1(b)). An electro-optic modulator driven by a waveform generator produces optical pulses using light from a continuous-wave tunable narrowband diode laser, which are sent into both interferometer arms. Light is focused on the sample through an NA ≈ 0.55 lens, coupling to the nanocavity with efficiency $\eta = \sqrt{\eta_{\text{in}}\eta_{\text{out}}} \approx 0.01$ [33]. Here $\eta_{\text{in,out}}$ are the efficiencies with which light is coupled from the incident laser beam to the cavity and from the cavity to the detectors, respectively. Each incident pulse carries $\sim 2 \times 10^6$ photons in a duration of $\tau_P = 20$ ns, such that the estimate maximum number of simultaneous intracavity photons is ~ 60 . The same lens collects emitted cavity radiation, whose phase quadrature is measured by recording the output of a balanced detector after interference with the local oscillator pulse. The resultant detector voltage thus reflects optical phase and mechanical displacement x .

Figure 1(c) shows examples of recorded pulse traces. Between traces we wait for a time (30 ms) larger than the mechanical damping time $2\pi/\Gamma$. It can be directly recognized that the recorded pulse heights are correlated when they are separated by a full oscillation period (two last pulses) whereas the pulses separated by half a period (e.g., the third and fourth pulse) are anti-correlated (around a non-zero offset voltage). Figure 1(d) depicts histograms of the difference of the recorded pulses (integrating the voltage over the pulse duration), demonstrating this correlation and anti-correlation behaviour. This is direct indication that the thermal nanobeam motion is imprinted on the detected pulses.

To understand the histogram shapes, we need to consider the nonlinearity of the transduction between mechanical position and optical phase. At the phase-sensitive operation point of the homodyne interferometer

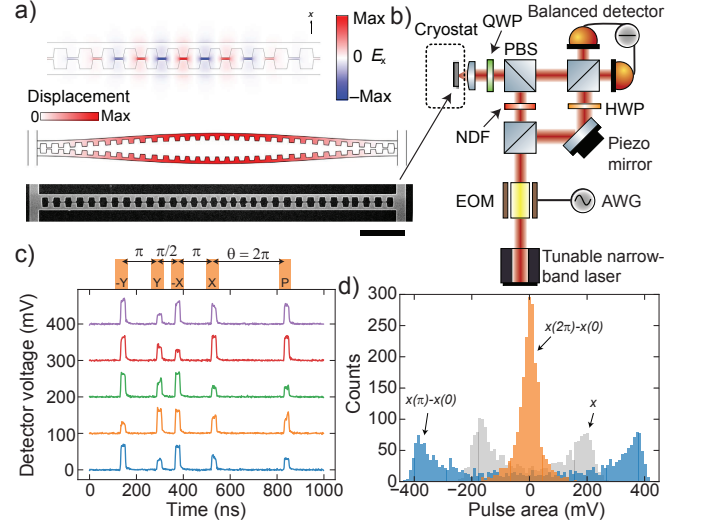


FIG. 1. (a) (Bottom) Electron microscope image of the free-standing, 250 nm thick sliced silicon nanobeam. Scale bar is $2 \mu\text{m}$. (Middle) Simulated displacement profile of the flexural vibrations of the two beam halves. (Top) Simulated transverse electric field of the optical cavity mode. (b) Schematic diagram of the employed balanced homodyne interferometer set-up. Pulses are created by an electro-optical-modulator (EOM) controlled by an arbitrary waveform generator (AWG). Quarter-wave-plate (QWP), half-wave plate (HWP), neutral density filter (NDF) and polarising beam-splitter (PBS) are also shown. (c) Example measurement traces, and schematic of the used pulsing sequence (separations ωt indicated). The different measurement traces are offset for clarity. (d) Examples of extracted histograms: thermal state [random sampling, mean subtracted] (grey), difference of two pulses a half-period apart (blue) and difference of two pulses a full period apart (orange).

and on resonance with the optical cavity the balanced detector output is [34]

$$H = 4|a_{\text{in}}||a_{\text{lo}}|\eta \frac{\beta x_n}{\beta^2 x_n^2 + 1}, \quad (1)$$

with $\beta = 2g_0/\kappa$, $x_n = x/x_{\text{zpf}}$ the normalized displacement and a_{in} and a_{lo} the optical field amplitudes towards the sample and in the local oscillator, respectively. This homodyne signal (Fig. 2(a)) depends linearly on displacement only when $\beta x_n \ll 1$. Outside this regime, the relationship between signal and displacement is nonlinear and even multivalued. Therefore, our value of $\beta \approx 2.5 \times 10^{-3}$ causes the thermal Gaussian displacement due to both modes with width (square-root of variance) $\sigma_{\text{th}} \approx 290x_{\text{zpf}}$ to be distorted into the double-peaked probability histogram in Fig. 1(d). To calibrate the homodyne signal we fit this thermal histogram to an analytical model using the known sample temperature [34]. The fit allows converting the measurement signal to normalized homodyne signal $H_{\text{norm}} = H/4|a_{\text{s}}||a_{\text{lo}}|\eta$.

The pulsing sequence in Fig. 1(c) is used for conditional

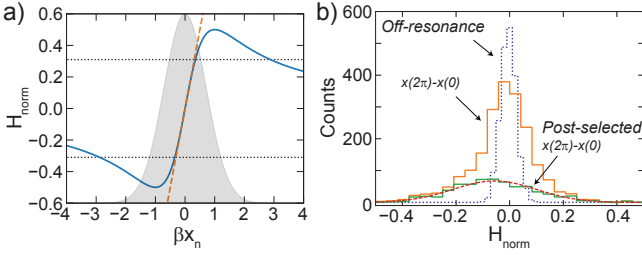


FIG. 2. **(a)** Full transduction function (Eq. 1, solid line), showing the linear approximation (dashed line) and the post-selection threshold (dotted lines). Gray area depicts a Gaussian with variance $\beta^2 \sigma_{\text{th}}^2$. **(b)** The histogram of the difference of two pulses a full period apart (as in Fig. 1(d)), before and after post-selection (solid lines). After post-selection we recover a Gaussian shape (dashed line shows fit). Also shown is a reference measurement off-resonant with the optical cavity (dotted line), from which the measurement noise floor is extracted.

state preparation and tomography. The pulses are separated by an angle $\theta \equiv \omega_m t$, where ω_m is the frequency of the mechanical mode of interest. (We address the impact of multiple modes below.) The mechanical resonator motion can be written as $x(\theta) = X \cos(\theta) + Y \sin(\theta)$, using the quadrature amplitudes X, Y which vary slowly within pulse trains but are randomized between pulse trains with zero mean and variance $\text{Var}(X) = \text{Var}(Y) = 2n_{\text{th}}x_{\text{zpf}}^2$, where $n_{\text{th}} = k_B T / (\hbar \omega_m)$ is the number of thermal phonons. The four first pulses of the sequence measure the instantaneous value of the two quadrature amplitudes (state preparation) and the last pulse quantifies the difference between the expected mechanical position and the actual position (tomography). By varying the waiting time between the state preparation and the tomography pulse, we can map out this difference in all quadrature angles and perform full state tomography of the conditional state. Note that we measure both $-X$ and X using two pulses to cancel measurement offsets caused by low-frequency drifts [34].

To maximize sensitivity and allow single-valued estimation, we post-select the data so that the measured value for the quadrature amplitude of interest falls in the linear transduction regime. The dotted lines in Fig. 2(a) show the chosen post-selection thresholds. Figure 2(b) shows the effect of post-selection on the histogram of the difference of two pulses separated by $\theta = 2\pi$. The original histogram had non-Gaussian shape with variance dominated by shot-noise due to a large contribution from parts of the transduction function with reduced sensitivity at $|\beta x_n| \approx 1$. The post-selected histogram has a Gaussian shape with larger variance. Hence, the post-selection protocol allowed recovering the linear operating regime.

To understand the principles of state preparation, consider a measurement at $t = 0$, measuring X . With just this measurement, the prediction for the oscillator's po-

sition at later times is $\cos(\theta)X$, differing from the actual position by $x(\theta) - \cos(\theta)X = \sin(\theta)Y$ (with variance $\sin^2(\theta)\text{Var}(Y) = 2n_{\text{th}}x_{\text{zpf}}^2 \sin^2(\theta)$ for a thermal state). This notably goes to zero at $\theta = \{\pi, 2\pi\}$, demonstrating that the knowledge of one quadrature allows ideally predicting the mechanical position exactly every half-period. In a phase-space defined by the two quadrature amplitudes this is a squeezed state. Combining two ideal measurements (a quarter period apart) then allows to measure both quadratures and fully predict the oscillator's evolution. In other words, we have then prepared a pure state with no classical uncertainty in x . This pure state will then decay towards the thermal equilibrium distribution with the time constant $2\pi/\Gamma$.

In the above idealized case, the pulsed measurements are assumed to be infinitely accurate and backaction-free. In practice, finite measurement imprecision leads to a Gaussian probability distribution for the measured quadrature amplitude. In addition, unavoidably, any measurement disturbs the other quadrature. These intuitions are formalized with a measurement operator given by Caves and Milburn for a free particle [35] and Vanner *et al.* for harmonic motion [26]

$$\hat{M} = \frac{1}{\sqrt[4]{\pi}} \exp \left(i\Omega \hat{\mathcal{X}} - \frac{(\hat{\mathcal{X}} - \mathcal{M})^2}{2/\chi^2} \right), \quad (2)$$

where $\mathcal{M} = x_n^{\text{meas}}/\sqrt{2}$ and x_n^{meas} is the dimensionless measurement result (normalized by x_{zpf}) and $\hat{\mathcal{X}} = (\hat{b}^\dagger + \hat{b})/\sqrt{2} = \hat{x}/(\sqrt{2}x_{\text{zpf}})$ is the quadrature operator with \hat{b} the phonon annihilation operator. The parameter $\chi = 8\sqrt{\eta_{\text{in}}\eta_{\text{out}}}\bar{N}_P g_0/\kappa$ [26, 29] characterizes the pulsed measurement strength and hence the conditional state variance. Physical insight into χ is provided by noting that the signal-to-noise-ratio of a shot-noise limited measurement is χx_n [34], meaning that with $\chi = 1$ one measures a displacement of x_{zpf} with unity signal-to-noise ratio when comparing two pulses.

Performing a measurement (transforming the state with $\rho_f = \hat{M}\rho_i\hat{M}^\dagger$) transforms one quadrature of an arbitrary initial state (ρ_i) into a Gaussian with width $\sigma_m = x_{\text{zpf}}/\chi$ and mean given by the random measurement result (which follows statistics determined by ρ_i), while adding $\Omega\sqrt{2}x_{\text{zpf}}$ to the other quadrature. That quadrature will also gain uncertainty $\sigma_{\text{ba}} = \sqrt{\eta_{\text{in}}/\eta_{\text{out}}}\chi x_{\text{zpf}}$, hence forcing for a state prepared by two sequential measurements quarter-period apart $\sqrt{\sigma_m(\sigma_m + \sigma_{\text{ba}})} \geq x_{\text{zpf}}$. In Fig. 2(b) we show a histogram measured away from optical resonance to determine the noise floor of the measurement. From this we can extract a single-pulse measurement imprecision of $\sigma_m \approx 9x_{\text{zpf}}$ [34], corresponding to $\chi \approx 0.11$, some three orders of magnitude higher than previous state-of-the-art [31]. This value is consistent with measured sample parameters including $\eta_{\text{in}} \approx 1.3\%$ and for $\eta_{\text{out}} \approx 0.35\eta_{\text{in}}$, close to our previous independent estimation [33].

Figure 3 presents experimental results of conditional state preparation and tomography. Scanning the delay between preparation pulses and the tomography pulse (i.e., varying θ) allows mapping the mechanical marginals and reconstructing the phase-space Wigner function of the conditional state via an inverse Radon transform from the histograms [26]. As we subtract the measured (random) values for the quadrature amplitudes, the plots depict the conditional state shifted to origin. The upper panels of Fig. 3(a) show this Radon transform for only the tomography pulse, without (left) and with (right) post-selection. These should be circularly symmetric and without the nonlinearity would simply depict the thermal Gaussian distribution. The post-selected data now depicts a small central part of that distribution. The lower left panel shows the thermally squeezed state prepared with one quadrature conditionalization. The state is extracted with $P - \cos(\theta)X$, with P the result of the tomography pulse and X the measured quadrature amplitude from the previous pulses. Finally, in the lower right panel of Fig. 3(a), we plot the Wigner function for the state conditionalized in both quadratures $P - \cos(\theta)X - \sin(\theta)Y$, which would have the same area as the ground state if the measurements would be ideal. In our case the state has an average width that corresponds to a one mode thermal state at a temperature of 380 mK, reflecting purification from the original temperature of 3.2 K.

The width of the state conditionalized on both quadratures is plotted in Fig. 3(b) as a function of the tomography angle. In the measured data the minimum width reached is $58x_{\text{zpf}}$, significantly exceeding the shot noise floor, and is maximized at $\theta = 3\pi/2$. These features are explained through the existence of a second mechanical mode. As in this device the two mechanical modes couple equally strongly to the cavity, the contribution of the second mechanical mode is captured with $x(\theta) = X_1 \cos(\theta) + Y_1 \sin(\theta) + X_2 \cos(r\theta) + Y_2 \sin(r\theta)$, with r the ratio of the mechanical frequencies and subscripts refer to the two modes. The resulting uncertainty caused by the second mode [34] as a function of θ is plotted in Fig. 3(b), matching the data well. There are no fitting parameters here as frequency and temperature of both modes are known. The non-monotonic shape is caused by measuring the Y quadrature before the X quadrature, causing it to have a larger contribution as it has more time to evolve out-of-sync with the mode of interest [34].

Assuming that $r \approx 1$ and that the two modes have equal n_{th} and x_{zpf} , the expected contribution to the conditional state width from the second mode can be approximated as $\sqrt{4n_{\text{th}} [1 - \cos(r\theta) \cos(\theta) - \sin(r\theta) \sin(\theta)]} x_{\text{zpf}}$ [34]. This is expected to vanish when $\cos(\theta) \cos(r\theta) = 1$ and $\sin(\theta) \sin(r\theta) = 0$, or vice versa. Although this cannot be fulfilled exactly unless r is a rational number, it is approximated when $\theta = 2n\pi$, where $n \approx \omega_m / (r\omega_m - \omega_m)$. In Fig. 4(a-b) we compare the one-pulse conditional state

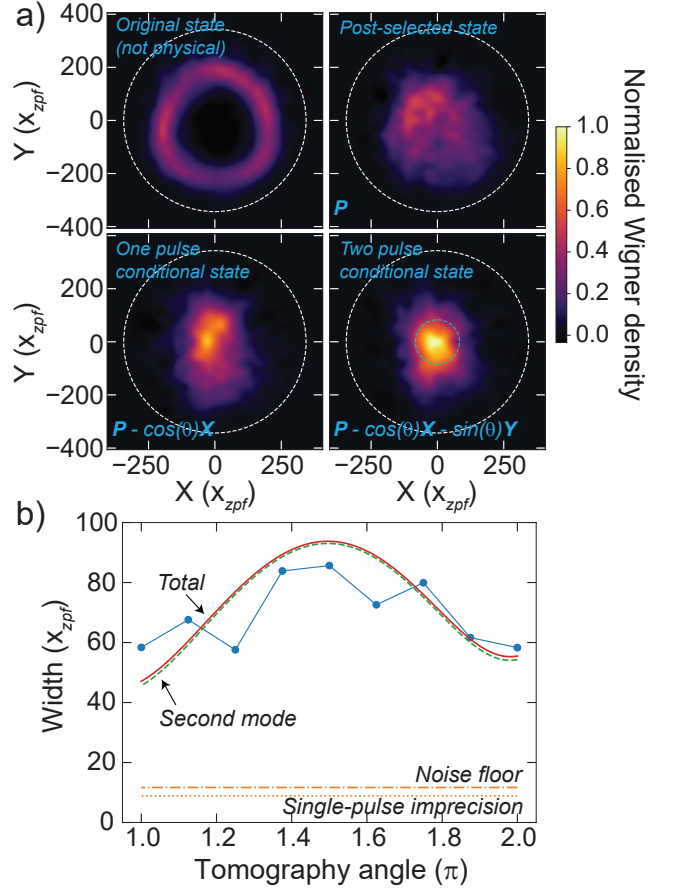


FIG. 3. (a) Reconstructed Wigner densities, for non-post-selected data (top left), post-selected data (top right), one-pulse conditional data (bottom left) and two-pulse conditional data (bottom right). A single normalization is applied on all panels. White dashed line shows the calculated full-width-half-maximum (FWHM) of the original thermal state at 3.2 K, and green dashed line the measured FWHM of the conditional state. Each panel is reconstructed from nine different measurement angles θ each combining 2000 acquired traces. (b) Width of different mechanical marginals for the two-pulse conditional data. Also shown are the noise floor of the measurement (dash-dotted line), the expected contribution from the second mechanical mode (dashed line) and the squared sum of these (solid line). Note that the noise floor differs slightly from the single-pulse imprecision (dotted line) as multiple pulses are used in the measurement and shot noise of the tomography pulse is not subtracted [34]. The two mechanical frequencies were $\omega = 2\pi \times 3.1081$ MHz and $r\omega = 2\pi \times 3.2280$ MHz.

width after one mechanical period ($\theta = 2\pi$) and where this condition is fulfilled ($n \approx 54, 56$). A slightly lower conditional state width is measured at $\theta = 54\pi, 56\pi$ than at $\theta = 2\pi$.

The measured conditional state widths at $\theta = 54\pi, 56\pi$ differ strongly from expectation based on the formula above. This is because we neglected thermal decoherence, which will cause the minimum conditional state

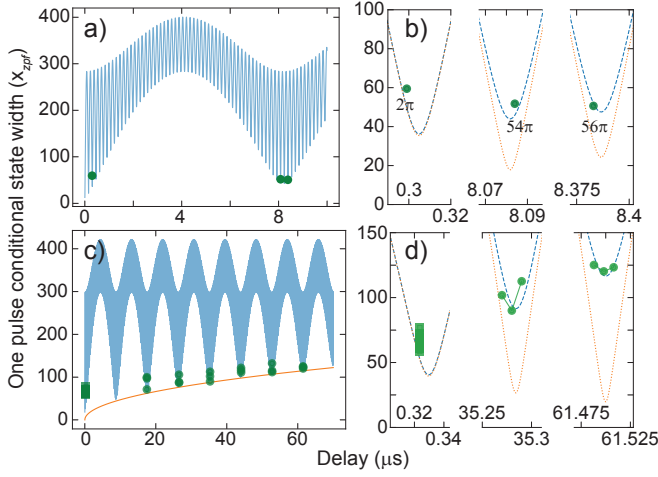


FIG. 4. **(a-b)** Common mode measurement, showing that smaller conditional state width is achieved at $\theta = 54\pi, 56\pi$ than at $\theta = 2\pi$ due to minimization of the second mode contribution at these times. Panel (b) shows close-ups of (a). Blue line shows the expected behaviour with fitted $\Gamma/(2\pi) = 400$ Hz, yellow dashed line [only shown in (b)] without any decoherence. Circles are measured data. Frequencies $\omega = 2\pi \times 3.340$ MHz and $r\omega = 2\pi \times 3.218$ MHz. **(c-d)** Thermal decoherence measurement demonstrating the growing conditional state width. Panel (d) shows close-ups of (c). Yellow line in (c) is the envelope function $\sqrt{[1 - \exp(-t\Gamma/2)]8n_{th}x_{zpf}}$, and the blue line the full expected curve, both with fitted $\Gamma/(2\pi) = 400$ Hz, yellow dashed line in (d) without decoherence. Each pulse sequence here has two tomography pulses, one at $\theta = 2\pi$ (shown as squares) and other at variable distance (circles). Also shown is data at ± 10 ns around $\theta = 2n\pi$ points. Frequencies $\omega = 2\pi \times 3.090$ MHz and $r\omega = 2\pi \times 2.976$ MHz. All datasets contained 1000 samples before post-selection.

width to increase in time as $\sqrt{8n_{th}[1 - \exp(-t\Gamma/2)]x_{zpf}}$, in the case of two mechanical modes, assuming the modes have identical n_{th} , x_{zpf} and Γ [34]. Figure 4(c-d) shows measurements around the times where the second mode contribution should vanish on a longer timescale, tracking the loss of coherence due to thermalization quantitatively in time-domain.

These results show that nano-optomechanical systems can bring quantum-level mechanical measurements with single nanosecond pulses within reach. Notably, achieving $\eta \gtrsim 8\%$ without changing any other parameters would bring the uncertainty in one quadrature below x_{zpf} , allowing squeezed state preparation and observing quantum backaction [15]. Preparing a pure state of a single resonator would require reducing the second mode contribution, through (opto)mechanical design (coupling the two beam halves more strongly to create a single optically bright mode), or by further cooling (cryogenic or feedback). Alternatively, one could exploit the fact that with quantum-level precision, a single pulse would entangle the quadratures of the two mechanical modes,

providing a new path to explore many-mode quantum optomechanics in the time domain.

Indeed, our experiments demonstrate how pulsed measurements yield interesting possibilities for measurement and control of mechanical motion, complementing the conventional frequency domain analyses. This “time-domain optomechanics” may give rise to new protocols for quantum sensing that exploit the fast backaction-free determination of a mechanical quadrature, as well as new paradigms to create quantum states of motion and mechanical entanglement.

Acknowledgements. The authors thank Hugo Doleman for critical reading of the manuscript. This work is part of the research programme of the Netherlands Organisation for Scientific Research (NWO), and supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 732894 (FET Proactive HOT). E.V. gratefully acknowledges an NWO-Vidi grant for financial support. J.T.M. thankfully acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 707364.

* juha.t.muhonen@jyu.fi

† verhagen@amolf.nl

- [1] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Rev. Mod. Phys.* **86**, 1391 (2014).
- [2] A. D. O’Connell, M. Hofheinz, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, J. M. Martinis, and A. N. Cleland, *Nature* **464**, 697 (2010).
- [3] J. Chan, T. P. M. Alegre, A. H. Safavi-Naeini, J. T. Hill, A. Krause, S. Gröblacher, M. Aspelmeyer, and O. Painter, *Nature* **478**, 89 (2011).
- [4] J. D. Teufel, T. Donner, D. Li, J. W. Harlow, M. S. Allman, K. Cicak, A. J. Sirois, J. D. Whittaker, K. W. Lehnert, and R. W. Simmonds, *Nature* **475**, 359 (2011).
- [5] E. E. Wollman, C. U. Lei, A. J. Weinstein, J. Suh, A. Kronwald, F. Marquardt, A. A. Clerk, and K. C. Schwab, *Science* **349**, 952 (2015).
- [6] J.-M. Pirkkalainen, E. Damskägg, M. Brandt, F. Massel, and M. A. Sillanpää, *Phys. Rev. Lett.* **115**, 243601 (2015).
- [7] F. Lecocq, J. B. Clark, R. W. Simmonds, J. Aumentado, and J. D. Teufel, *Phys. Rev. X* **5**, 041037 (2015).
- [8] R. Riedinger, A. Wallucks, I. Marinkovi, C. Lössnauer, M. Aspelmeyer, S. Hong, and S. Gröblacher, *Nature* **556**, 473 (2018).
- [9] C. F. Ockeloen-Korppi, E. Damskägg, J.-M. Pirkkalainen, M. Asjad, A. A. Clerk, F. Massel, M. J. Woolley, and M. A. Sillanpää, *Nature* **556**, 478 (2018).
- [10] Y. Chu, P. Kharel, W. H. Renninger, L. D. Burkhardt, L. Frunzio, P. T. Rakich, and R. J. Schoelkopf, *Science* **358**, 199 (2017).
- [11] P. F. Cohadon, A. Heidmann, and M. Pinard, *Phys. Rev.*

- Lett. **83**, 3174 (1999).
- [12] D. J. Wilson, V. Sudhir, N. Piro, R. Schilling, A. Ghadimi, and T. J. Kippenberg, *Nature* **524**, 325 (2015).
 - [13] M. Rossi, D. Mason, J. Chen, Y. Tsaturyan, and A. Schliesser, *Nature* **563**, 53 (2018).
 - [14] K. W. Murch, K. L. Moore, S. Gupta, and D. M. Stamper-Kurn, *Nat. Phys.* **4**, 561 (2008).
 - [15] T. P. Purdy, R. W. Peterson, and C. A. Regal, *Science* **339**, 801 (2013).
 - [16] G. Vasilakis, H. Shen, K. Jensen, M. Balabas, D. Salart, B. Chen, and E. S. Polzik, *Nat. Phys.* **11**, 389 (2015).
 - [17] C. F. Ockeloen-Korppi, E. Damskägg, J.-M. Pirkkalainen, A. A. Clerk, M. J. Woolley, and M. A. Sillanpää, *Phys. Rev. Lett.* **117**, 140401 (2016).
 - [18] C. B. Møller, R. A. Thomas, G. Vasilakis, E. Zeuthen, Y. Tsaturyan, M. Balabas, K. Jensen, A. Schliesser, K. Hammerer, and E. S. Polzik, *Nature* **547**, 191 (2017).
 - [19] N. S. Kampel, R. W. Peterson, R. Fischer, P.-L. Yu, K. Cicak, R. W. Simmonds, K. W. Lehnert, and C. A. Regal, *Phys. Rev. X* **7**, 021008 (2017).
 - [20] V. Sudhir, R. Schilling, S. A. Fedorov, H. Schütz, D. J. Wilson, and T. J. Kippenberg, *Phys. Rev. X* **7**, 031055 (2017).
 - [21] D. Mason, J. Chen, M. Rossi, Y. Tsaturyan, and A. Schliesser, *ArXiv e-prints* (2018), arXiv:1809.10629.
 - [22] J. B. Hertzberg, T. Rocheleau, T. Ndukum, M. Savva, A. A. Clerk, and K. C. Schwab, *Nat. Phys.* **6**, 213 (2009).
 - [23] J. Suh, A. J. Weinstein, C. U. Lei, E. E. Wollman, S. K. Steinke, P. Meystre, A. A. Clerk, and K. C. Schwab, *Science* **344**, 1262 (2014).
 - [24] I. Shomroni, L. Qiu, D. Malz, A. Nunnenkamp, and T. J. Kippenberg, *ArXiv e-prints* (2018), arXiv:1809.01007.
 - [25] V. B. Braginsky, Y. I. Vorontsov, and F. Y. Khalili, *JETP Lett.* **27**, 276 (1978).
 - [26] M. R. Vanner, I. Pikovski, G. D. Cole, M. S. Kim, Č. Brukner, K. Hammerer, G. J. Milburn, and M. Aspelmeyer, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16182 (2011).
 - [27] M. R. Vanner, I. Pikovski, and M. S. Kim, *Ann. Phys.* **527**, 15 (2014).
 - [28] M. R. Vanner, *Phys. Rev. X* **1**, 021011 (2011).
 - [29] U. B. Hoff, J. Kollath-Bönig, J. S. Neergaard-Nielsen, and U. L. Andersen, *Phys. Rev. Lett.* **117**, 143601 (2016).
 - [30] J. S. Bennett, K. Khosla, L. S. Madsen, M. R. Vanner, H. Rubinsztein-Dunlop, and P. Warwick, *New J. Phys.* **18**, 053030 (2016).
 - [31] M. R. Vanner, J. Hofer, G. D. Cole, and M. Aspelmeyer, *Nat. Commun.* **4**, 2295 (2013).
 - [32] R. Leijssen and E. Verhagen, *Sci. Rep.* **5**, 15974 (2015).
 - [33] R. Leijssen, G. R. La Gala, L. Freisem, J. T. Muhonen, and E. Verhagen, *Nat. Commun.* **8**, 16024 (2017).
 - [34] See Supplemental Material at [URL will be inserted by publisher].
 - [35] C. M. Caves and G. J. Milburn, *Phys. Rev. A* **36**, 5543 (1987).

SUPPLEMENTARY INFORMATION

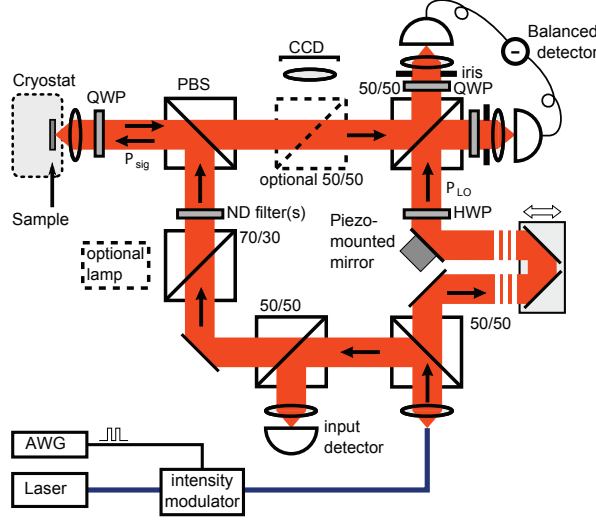


FIG. S1. **Detailed measurement setup.** The lamp and the CCD are used to align the laser beam with our sample using a moving lens outside the cryostat window. The extra quarter wave plates (QWP) in front of the detector are for preventing detector reflections of the local oscillator from reaching the sample as they will be reflected by the polarising beam splitter. Irises are used to balance the two detectors. Piezo mounted mirror is used to adjust the local oscillator phase with respect to the signal phase. P_{lo} and P_{sig} refer to the power in the local oscillator arm and signal arm, respectively.

Details of the pulsing sequences and single-pulse imprecision extraction

All pulse trains we apply are separated from each other by a time that is large compared to the mechanical damping time in order to 'reset' the mechanical state (~ 30 ms). For reasons not fully understood (but most likely deriving from the fact that our homodyne phase is not actively locked during measurements and/or technical laser noise) our measured voltages have an offset that is constant inside a pulse train but varies between different pulse trains. In order to cancel out this offset we always take two pulses per quadrature for the state preparation, as well as subtract this offset from the tomography pulse.

The pulse train for two pulse conditionalization is shown in Fig. 1(c). Assuming the signal is given by $s(\omega t) = X \cos(\omega t) + Y \sin(\omega t) + V_{\text{off}} + \delta$, where V_{off} is the offset and δ random noise, and setting $t = 0$ at the final pulse before the tomography pulse, the four preparation pulses read

$$P_0 = s(-5\pi/2) = -Y + V_{\text{off}} + \delta_0 \quad (\text{S1})$$

$$P_1 = s(-3\pi/2) = Y + V_{\text{off}} + \delta_1 \quad (\text{S2})$$

$$P_2 = s(-\pi) = -X + V_{\text{off}} + \delta_2 \quad (\text{S3})$$

$$P_3 = s(0) = X + V_{\text{off}} + \delta_3. \quad (\text{S4})$$

Hence, to extract the quadrature amplitudes we use $Y = (P_1 - P_0)/2$ and $X = (P_3 - P_2)/2$. Then the full two pulse conditional state reads

$$s_{\text{cond}}(\omega t) = s(\omega t) - \frac{P_3 - P_2}{2} \cos(\omega t) - \frac{P_1 - P_0}{2} \sin(\omega t) - \frac{P_0 + P_1 + P_2 + P_3}{4}, \quad (\text{S5})$$

where the last term is used to cancel the offset in the final tomography pulse $s(\omega t)$. Note that replacing that term with either $(P_0 + P_1)/2$ or $(P_3 + P_2)/2$ cancels out some of the noise terms in one quadrature at the expense of increasing the noise in the other quadrature. For the one-pulse conditional state in Fig. 3 we use the same formula without the sine term, and for the non-conditional state we also remove the cosine term (but keep the offset subtraction term).

The noise floor presented in the main text Fig.3(b) was extracted by running the same pulse sequence and analysis for measurement data gathered with laser tuned off-resonance with the optical cavity. The resulting conditional state

histogram with $\theta = 2\pi$ is presented in main Fig. 2(b), and has the width $11.67x_{\text{zpf}}$, when converted to displacement. As the off-resonance data will presumably have no signal ($X, Y=0$) we can write

$$s_{\text{cond}}^{\text{noise}}(2\pi) = s(2\pi) - \frac{P_3 - P_2}{2} \cos(2\pi) - \frac{P_0 + P_1 + P_2 + P_3}{4} \quad (\text{S6})$$

$$= V_{\text{off}} + \delta_4 - \frac{V_{\text{off}} + \delta_3 - V_{\text{off}} - \delta_2}{2} - \frac{4V_{\text{off}} + \delta_0 + \delta_1 + \delta_2 + \delta_3}{4} \quad (\text{S7})$$

$$= \delta_4 - \frac{\delta_0 + \delta_1}{4} - \frac{3\delta_3}{4} + \frac{\delta_2}{4}. \quad (\text{S8})$$

Hence, we have for the variance (we assume all noise terms are independent and hence have no mutual covariance)

$$\text{Var}[s_{\text{cond}}^{\text{noise}}(2\pi)] = \text{Var}(\delta_4) + \frac{\text{Var}(\delta_0) + \text{Var}(\delta_1) + \text{Var}(\delta_2) + 9\text{Var}(\delta_3)}{16} \quad (\text{S9})$$

$$= (1 + \frac{3}{4})\text{Var}(\delta_{\text{sn}}), \quad (\text{S10})$$

where for the last line we have assumed that the variance of noise in each pulse is the same and marked it $\text{Var}(\delta_{\text{sn}})$. The noise floor variance is the sum of the imprecision in the measurement of X [$\text{Var}(\delta_{\text{sn}})/2$ as two pulses are used], the imprecision in the tomography pulse [$\text{Var}(\delta_{\text{sn}})$] and the added variance due to the offset correction [$\text{Var}(\delta_{\text{sn}})/4$]. Hence, in order to extract the single-pulse imprecision $\sqrt{\text{Var}(\delta_{\text{sn}})}$ we need to divide the variance we extract by $7/4$, meaning we need to divide the width we extracted $11.67x_{\text{zpf}}$ with square root of $7/4$, giving $8.8x_{\text{zpf}}$.

We can make an easy check on this, as we can also extract the variance of the "non-conditional state" from the off-resonance data

$$s_{\text{non-cond}}^{\text{noise}}(2\pi) = s(\omega t) - \frac{P_0 + P_1 + P_2 + P_3}{4} \quad (\text{S11})$$

$$= \delta_4 - \frac{\delta_0 + \delta_1 + \delta_2 + \delta_3}{4}, \quad (\text{S12})$$

from which using similar arguments as before we can extract $\text{Var}[s_{\text{non-cond}}^{\text{noise}}(2\pi)] = (1 + 1/4)\text{Var}(\delta_{\text{sn}})$. From the data (not shown) we can extract a non-conditional width of $9.92x_{\text{zpf}}$ which would again give a single pulse shot-noise of $8.8x_{\text{zpf}}$.

Note that the actual imprecision in the measurement of X (without the tomography pulse) is half the single-pulse imprecision, due to the fact that we use two pulses to measure it.

Homodyne signal calibration

As mentioned in the main text (and is derived below) the homodyne signal as a function of the normalized displacement can be written as

$$H = A \frac{\beta x_n}{\beta^2 x_n^2 + 1} \equiv A \frac{\Delta}{\Delta^2 + 1}, \quad (\text{S13})$$

where we have made explicit the relative change in optical cavity frequency due to the mechanical motion $\Delta = \beta x_n$ and used a constant A to absorb all the normalization terms. Inverting this equation gives

$$\Delta^{\pm} = \frac{1 \pm \sqrt{1 - 4H'^2}}{2H'}, \quad (\text{S14})$$

where we have defined $H' \equiv H/A$. As expected, this is multivalued and hence we cannot generally reliably deduce the mechanical position from a single measurement. This is only possible in the linearised regime $\beta^2 x_n^2 \ll 1$, where $H \approx A\beta x_n$. Hence, in the main manuscript we only use data where this approximation is valid (post-selection).

We can still, however, reliably predict ensemble histograms and this is used for calibrating the measurement signal. At thermal equilibrium x_n will have a Gaussian probability density with variance $\text{Var}(x_n) = (2k_B T)/(\hbar\omega_m) = 2n_{\text{th}}$, where k_B is the Boltzmann constant, T is temperature, \hbar the reduced Planck constant and ω_m the mechanical oscillation frequency. For multiple independent mechanical modes, we need to sum the variances. The parameter β

is approximately equal for both modes. It follows that $\text{Var}(\Delta) = \beta^2 \text{Var}(x_n) = 8(g_0/\kappa)^2(n_{\text{th},1} + n_{\text{th},2}) \equiv \sigma_\Delta^2$, where we have marked $n_{\text{th},i}$ as the mean thermal phonon number for mode i . From probability calculus

$$\mathcal{P}(H') = \mathcal{P}(\Delta(H')) \left| \frac{d\Delta}{dH'} \right| = \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp\left(-\frac{\Delta^2}{2\sigma_\Delta^2}\right) \left| \frac{d\Delta}{dH'} \right|, \quad (\text{S15})$$

where we use \mathcal{P} to mark the probability density. In order to deal with the multivaluedness we need to sum over both branches

$$\mathcal{P}(H') = \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp\left(-\frac{(\Delta^+)^2}{2\sigma_\Delta^2}\right) \left| \frac{d\Delta^+}{dH'} \right| + \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp\left(-\frac{(\Delta^-)^2}{2\sigma_\Delta^2}\right) \left| \frac{d\Delta^-}{dH'} \right|. \quad (\text{S16})$$

Inserting Eq. S14 and its derivative

$$\left| \frac{d\Delta^\pm}{dH'} \right| = -\frac{1}{2H'} \left(1 \mp \frac{1}{\sqrt{1-4H'^2}} \right), \quad (\text{S17})$$

will give an algebraic form for the expected probability density, that only depends on the parameter σ_Δ^2 . Hence, measuring the thermal equilibrium distribution of our pulse outputs and fitting it, allows us to relate the pulse outputs to the parameter H' , and further, in the linear regime, to $x_n = H'/\beta$. This fit is plotted together with data in Fig. S2.

In order to avoid any errors caused by low frequency drifts the thermal histogram points are extracted by taking the difference of two pulses half an oscillation period apart and dividing by two. For one mode, this would give exactly the same histogram as random sampling. Although the fact that we have two modes with differing frequencies produces an error to this procedure, the two frequencies are close enough that the error in σ_Δ^2 is less than 0.5% in all cases considered in this paper. (The error is the difference between a factor of 8 and factor $(6 - 2\cos(r\pi))$. See equation (S21) and compare it to two times equation (S17), taking $\theta = \pi$.)

It should also be emphasized that with our parameters, the expected distribution will always be doubly peaked around $H' = \pm 0.5$. Any small uncertainty in parameter σ_Δ^2 will not change the position of the peaks. Hence, our calibration is relatively insensitive to any imprecision in σ_Δ^2 . On the other hand, the vertical scale here is fixed (with proper normalization) and hence the correctness of the height of the center flat part of the histogram will show the accuracy of our σ_Δ^2 extraction. As can be seen in Fig. S2 (and this was true generally) the line does lie somewhat above the measured histogram (this error would correspond to $\sim 10\%$ higher σ_Δ^2). We note that if this error would be in the β parameter, this would make the extracted imprecisions in the main manuscript lower (meaning, better) by $\sim 5\%$.

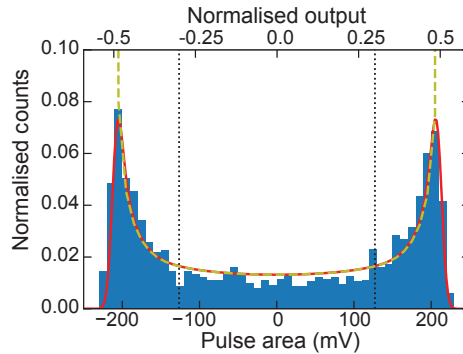


FIG. S2. Example of a thermal probability histogram of measurement outputs and fit to Eq. S16 (dashed line). From this we can relate the measured voltages (bottom x-axis) to the normalized output (H' , top x-axis). Solid line shows a numerical convolution of the analytical solution (that diverges at $H' = \pm 0.5$) and a Gaussian with the width given by our noise floor. The thermal histogram points are extracted by taking the difference of two pulses half an oscillation period apart and dividing by two. This eliminates errors from a constant offset drift (see previous section). Time between individual measurements is roughly three times the mechanical damping time. Dotted lines show our post-selection thresholds at $H' = \pm 0.31$.

Finally, we note that we obviously cannot use this method for the off-resonance data we extract the noise floor and single-pulse imprecision from. For that data we first gathered data on-resonance to run the above analysis and then without changing anything except laser frequency run the off-resonance measurement and used the on-resonance data for calibration.

Post-selection threshold dependance

As mentioned in the main manuscript, all the data presented has been post-selected so that the instantaneous fluctuation amplitude during the measurements is in the linear regime of optomechanical response. In Fig. S3 we show how the conditional state width of data in Fig. 3(b) depends on the post-selection point. As expected, the end result is independent of the post-selection point, as long as the point is in the linear regime. If the post-selection point is set higher, the measured width starts to decrease as the measured signal does not correspond linearly to the mechanical displacement.

It is also important to note that the threshold has to be chosen so that there is only a small probability for the higher branch of the transduction function (above $\beta x_n = 1$) to produce the result. For our chosen post-selection threshold the transduction function crosses the threshold first at $0.48\sigma_{\text{th}}$ and then again at $4\sigma_{\text{th}}$. Concretely, this means that we retain roughly one third of the acquired data (per quadrature that is post-selected) and on average only one point in 5000 in that post-selected dataset is from the wrong branch.

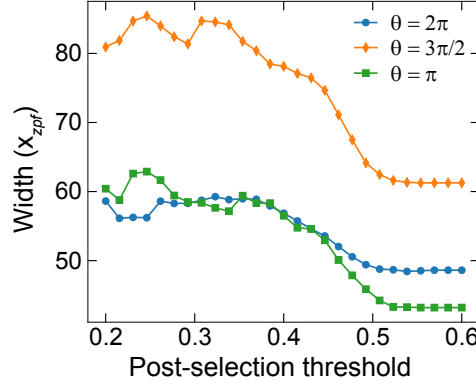


FIG. S3. Width of different mechanical marginals as a function of the post-selection threshold. The data shown are the three points in Fig. 3, corresponding to tomography angles 2π , $3\pi/2$, and π .

Homodyne signal and measurement strength

To derive Eq. (1) of the main manuscript we consider the interference of two beams each described by the complex field amplitudes: a_s and a_{lo} , where the subscripts stand for signal and local oscillator respectively. This happens at the final 50:50 beam splitter of the interferometer and produces the output beams $a_+ = 1/\sqrt{2}(a_s + ia_{\text{lo}})$ and $a_- = 1/\sqrt{2}(ia_s + a_{\text{lo}})$. Each beam is then detected by a photo-detector sensitive to power and the two signals are electronically subtracted. Hence the balanced detector gives as output voltage a signal that is proportional to the difference of the optical powers of the beams $H = |a_+|^2 - |a_-|^2$. Substituting a_+ and a_- one obtains

$$H = i(a_s^* a_{\text{lo}} - a_{\text{lo}}^* a_s), \quad (\text{S18})$$

where the “*” indicates the complex conjugate.

The expression for our signal beam can be obtained from input-output theory. We consider that the signal beam is the output field of the interaction between an incident light beam and the opto-mechanical cavity $a_s = \sqrt{\kappa_{\text{out}}}a$, where a is the field amplitude inside the cavity. The equation of motion for the field amplitude is

$$\frac{da(t)}{dt} = -i\Delta(t) + \frac{\kappa}{2}a(t) + \sqrt{\kappa_{\text{in}}}a_{\text{in}} \quad (\text{S19})$$

where κ_{in} is the coupling rate of the incident light with the cavity mode, κ is the total cavity energy decay rate and a_{in} is the complex amplitude describing the incident beam. An important quantity in this equation is the detuning $\Delta(t) = \Delta_0 + Gx(t)$, where $\Delta_0 = \omega - \omega_c$ is the static detuning of the laser frequency ω from the cavity frequency ω_c . In all measurements in the paper we use $\Delta_0 = 0$. The mechanical system enters the description through the position dependent frequency displacement $Gx(t) = g_0 x_n(t)$. Since in our system the cavity linewidth is large compared to

the mechanical frequency $\kappa \gg \omega_m$, the intracavity field adiabatically follows the evolution of the mechanics. Hence the equation of motion can be adiabatically eliminated leading to

$$a(t) = \frac{\sqrt{\kappa_{\text{in}}} a_{\text{in}}}{-i\Delta(t) + \frac{\kappa}{2}}. \quad (\text{S20})$$

As the pulse duration in the measurements is also much shorter than the mechanical oscillation period $\tau_P \ll 1/\omega_m$ we consider the mechanical displacement static during the short measurement interaction and hence $a(t) \rightarrow a$ as $\Delta(t) = \Delta = g_0 x_n$ with x_n the static normalized mechanical displacement during the measurement.

Now we can return to Eq. S18 and substitute $a_s = \sqrt{\kappa_{\text{out}}} a$, and $a_{\text{lo}} = |a_{\text{lo}}| e^{i\phi}$, where ϕ is the phase difference between the a_{lo} and a_{in} (hence we can also substitute $a_{\text{in}} = |a_{\text{in}}|$). We control this parameter with a mirror attached to a piezo stage. We obtain

$$H = |a_{\text{lo}}| \frac{4\sqrt{\kappa_{\text{in}}\kappa_{\text{out}}}}{\kappa} \frac{|a_{\text{in}}|}{1 + \left(\frac{2\Delta}{\kappa}\right)^2} \left(\cos \phi + \frac{2\Delta}{\kappa} \sin \phi \right), \quad (\text{S21})$$

which reduces to Eq. (1) when $\phi = \pi/2$, meaning at the phase sensitive operation point of the homodyne interferometer. Note how the signal is a product of three parts: the amplitude of the local oscillator, the amplitude of the signal beam after interaction with the cavity and the homodyne angle sensitive part. At $\phi = 0$ the signal is proportional to only the amplitudes of the two fields, whereas at $\phi = \pi/2$ it becomes linearly sensitive to the detuning of the cavity, meaning in our case the mechanical displacement.

It is instructive to notice that assuming the linear transduction regime in which $4\Delta^2 \ll \kappa^2$, and $\phi = \pi/2$ we have

$$H \simeq 4 |a_{\text{lo}}| |a_{\text{in}}| \eta \beta x_n, \quad (\text{S22})$$

which we can integrate over the pulse duration

$$\int_0^{\tau_P} H dt \approx \langle H \rangle \tau_P = 4 \langle |a_{\text{lo}}| \rangle \sqrt{\tau_P} \langle |a_{\text{in}}| \rangle \sqrt{\tau_P} \eta \beta x_n = 4 \sqrt{N_{\text{lo}}} \sqrt{N_P} \eta \beta x_n = \sqrt{N_{\text{lo}}} \chi x_n, \quad (\text{S23})$$

using the definition of χ given in main text and in references [26, 29]. This gives an intuitive feeling for the parameter χ which is simply the transduction parameter between normalized displacement and the homodyne output. In addition the output is amplified by the strength of the local oscillator as would be expected. As usual, the purpose of this amplification is to bring the signal strength above all electric noise signals. Note that although we have implicitly disregarded some technical variables like the interferometer overlap and the aperture balancing the two detector sides, these can be included in parameter κ_{out} (and hence η). Note also that assuming our noise floor is given by optical shot noise and that the power of local oscillator is much larger than the signal power, the standard deviation of the measurement is given by $\sqrt{N_{\text{lo}}}$ and hence the signal-to-noise ratio (mean over standard deviation) is simply χx_n .

Conditional state variances

To illustrate the basic principle of calculating the conditional state variances, consider first the variance of the difference between two pulses separated by angle $\theta = \omega t$, assuming the motion is given by $x(\theta) = X \cos \theta + Y \sin \theta$

$$\begin{aligned} \text{Var}[x(\theta) - x(0)]_{\text{mode}} &= \text{Var}[X \cos \theta + Y \sin \theta - X] = \text{Var}[X(\cos \theta - 1) + Y \sin \theta] \\ &= (\cos \theta - 1)^2 \text{Var}(X) + \sin^2 \theta \text{Var}(Y) = (2 - 2 \cos \theta) \text{Var}(Q), \end{aligned} \quad (\text{S24})$$

where we have marked $\text{Var}(Q) = \text{Var}(X) = \text{Var}(Y)$ as the variance of the quadrature amplitudes, which for a thermal state is given by $\text{Var}(Q) = 2n_{\text{th}} x_{\text{zpf}}^2$. Note that for the difference of two uncorrelated pulses we would expect the variance to be $2\text{Var}(Q)$.

Any dephasing process can be added to this formula by requiring that the covariance of the quadrature amplitude with itself goes down with time as $\text{Cov}[X(t), X(0)] = \exp(-\gamma t) \text{Var}[X(0)]$, where $\gamma = \Gamma/2$ for the mechanical damping. Note that this produces the correct limits $\text{Cov}[X(0), X(0)] = \text{Var}[X(0)]$ and $\text{Cov}[X(\infty), X(0)] = 0$. Then using the formula $\text{Var}[aX + bX'] = a^2 \text{Var}[X] + b^2 \text{Var}[X'] + 2ab \text{Cov}[X, X']$ we have

$$\begin{aligned} \text{Var}[X' \cos \theta - X + Y \sin \theta]_{\text{mode}} &= \cos^2 \theta \text{Var}(X') + \text{Var}(X) - 2 \cos \theta \exp(-\gamma t) \text{Var}(X) + \sin^2 \theta \text{Var}(Y) \\ &= [2 - 2 \exp(-\gamma t) \cos \theta] \text{Var}(Q), \end{aligned} \quad (\text{S25})$$

which produces the expected limits. Note that we assume $\text{Cov}(X, Y) = 0$ at all times. Going through similar algebra for the one and two pulse conditional state will produce

$$\text{Var}[x(\theta) - \cos \theta x(0)]_{1\text{mode}} = \{1 + \cos^2 \theta [1 - 2 \exp(-\gamma t)]\} \text{Var}(Q) \quad (\text{S26})$$

$$\text{Var}[x(\theta) - \cos \theta x(0) - \sin \theta x(\pi/2)]_{1\text{mode}} = [2 - 2 \exp(-\gamma t)] \text{Var}(Q). \quad (\text{S27})$$

Now, we want to derive the same formulae in the case of two mechanical modes, $x(\theta) = X_1 \cos \theta + Y_1 \sin \theta + X_2 \cos r\theta + Y_2 \sin r\theta$, where r is the ratio of the two mechanical frequencies. For generality we can assume they can have different n_{th} and/or x_{zpf} , and hence $\text{Var}(Q_i) = \text{Var}(X_i) = \text{Var}(Y_i)$, where $i = 1, 2$. Note that although we do not consider the coupling from the modes to the light field explicitly here, ultimately differing coupling constants would affect the results similarly as differing x_{zpf} . Going through similar algebra as above will then produce

$$\text{Var}[x(\theta) - x(0)]_{2\text{modes}} = [2 - 2 \exp(-\gamma_1 t) \cos \theta] \text{Var}(Q_1) + [2 - 2 \exp(-\gamma_2 t) \cos r\theta] \text{Var}(Q_2) \quad (\text{S28})$$

$$\begin{aligned} \text{Var}[x(\theta) - \cos \theta x(0)]_{2\text{modes}} = & \{1 + \cos^2 \theta [1 - 2 \exp(-\gamma_1 t)]\} \text{Var}(Q_1) \\ & + \{1 + \cos^2 \theta [1 - 2 \frac{\cos r\theta}{\cos \theta} \exp(-\gamma_2 t)]\} \text{Var}(Q_2) \end{aligned} \quad (\text{S29})$$

$$\begin{aligned} \text{Var}[x(\theta) - \cos \theta x(0) - \sin \theta x(\pi/2)]_{2\text{modes}} = & [(2 - 2 \exp(-\gamma_1 t)) \text{Var}(Q_1) \\ & + \{2 - 2 \exp(-\gamma_2 t) \\ & \times [\cos \theta \cos r\theta + \sin \theta \sin r\theta \sin r\pi/2 + \cos r\pi/2 \sin \theta (\cos \theta + \cos r\theta)]\} \text{Var}(Q_2) \\ \approx & [(2 - 2 \exp(-\gamma_1 t)) \text{Var}(Q_1) \\ & + \{2 - 2 \exp(-\gamma_2 t) [\cos \theta \cos r\theta + \sin \theta \sin r\theta]\} \text{Var}(Q_2)], \end{aligned} \quad (\text{S30})$$

where the last approximation ($\sin r\pi/2 = 1, \cos r\pi/2 = 0$), assumes r is sufficiently close to one.

As our pulsing sequence consists actually of more pulses than what is assumed above, the exact analytical formulae are more complicated but the above derivations capture the essential physics. For completeness we quote here the full formula that is plotted with data in Fig. 3(b), which is for the pulse sequence given by Eq. S5

$$\begin{aligned} \sigma^2 = & \left[\frac{1}{8} \left(2 \cos(\theta) + (1 - 2 \cos(\theta)) \cos(\pi r) - 4 \cos(\theta r) + (2 \sin(\theta) + 1) \cos\left(\frac{3\pi r}{2}\right) + (1 - 2 \sin(\theta)) \cos\left(\frac{5\pi r}{2}\right) + 1 \right)^2 \right. \\ & \left. + \frac{1}{8} \left((2 \sin(\theta) + 1) \sin\left(\frac{3\pi r}{2}\right) + (1 - 2 \sin(\theta)) \sin\left(\frac{5\pi r}{2}\right) + 4 \sin(\theta r) + (1 - 2 \cos(\theta)) \sin(\pi r) \right)^2 \right] \text{Var}(Q_2) \end{aligned} \quad (\text{S31})$$

This is without thermal dephasing (meaning $\gamma_i = 0$), which means there is no contribution from $\text{Var}(Q_1)$.