Supplemental Information

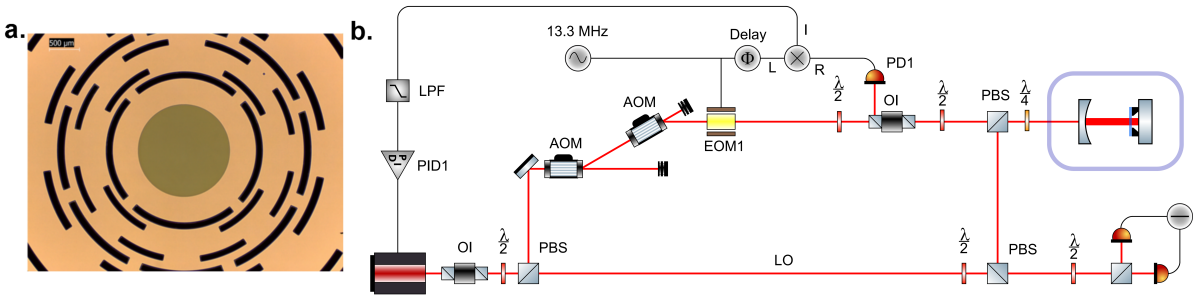

FIG. 1: **a.** Optical microscope image of the circular membrane, radius 0.75 mm and thickness 97 nm. **b.** Scheme of the experimental apparatus. Optical Isolator (OI); acousto-optic modulator (AOM); electro-optic modulator (EOM); half-wave plate ($\lambda/2$); quarter-wave plate ($\lambda/4$); Polarizing beam-splitter (PBS); photodiode (PD); Low pass filter (LPF); servo-loop electronics (PID).

EXPERIMENTAL SETUP

The experimental setup is sketched in Fig.1 b. A Nd:YAG laser at $\lambda = 1064$ nm is split into two beams, with a polarizing beam-splitter (PBS), after a 40 dB optical isolator. The first beam is frequency shifted with two acousto-optic modulators (AOM) operating at opposite orders. A subsequent resonant electro-optic modulator (EOM) provides a phase modulation at 13.3 MHz with a depth $\simeq 1$ rad to implement a Pound-Drever-Hall (PDH) locking scheme. After going through an optical isolator the beam is mode-matched ($> 90\%$) to the optical cavity after having gone through an additional PBS. The cavity reflected beam is collected by the input polarizer of the optical isolator and sent on a fast photodiode for the PDH signal detection. A quarter-wave plate, placed between the cavity and the last PBS allows to pick up a fraction of the reflected beam($\simeq 70\%$). The second beam is used as Local Oscillator for the r-heterodyne detection. The two beams are mode-matched and superimposed on a PBS and sent to a balance detection.

A major source of uncertainty arises because θ in our experiment was not actively stabilized. The LO phase time series $\theta(t) = \theta(0) + \delta\theta(t)$ (including any errors in θ) was estimated by implementing a numerical lock-in amplifier to demodulate the 10 kHz beat note. Fig.2 illustrates a 25 Hz environmental noise modulation which frequently exceeded 2π amplitude. In order to mitigate this potential source of error, about two dozen stretches of data with a drift $|\delta\theta(t)| < 1$ were identified and selected as shown in the figure.

POWER SPECTRAL DENSITY (PSD)

Experimental investigation of detected cavity output time-traces of the photocurrent $i_h(t)$, for analysis of narrowband features, generally proceeds via analysis in frequency space. For generality (and with a view to future experiments approaching quantum ground-state cooling) we consider general quantum operators in the expressions below, noting that measured currents $\hat{i}_h^\dagger = \hat{i}_h \equiv i_h$ are real.

As the time traces are of long but finite duration T we employ the usual gated Fourier transform:

$$\hat{A}(\omega) = \lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_0^T \hat{A}(t) e^{i\omega t} dt \tag{1}$$

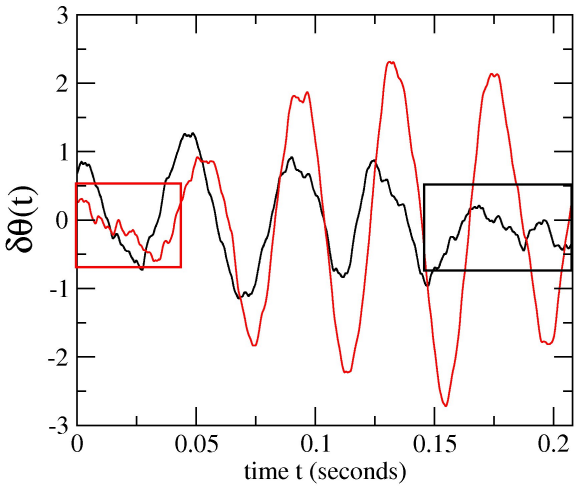

FIG. 2: Drift in phase θ for typical experimental time traces. Smaller stretches with lower amplitude phase errors were selected for the PSD calculations.

for an arbitrary stochastic observable $\hat{A}$. We may also define a cross-spectrum between two such observables:

$$S_{\hat{A}\hat{B}} = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt' \int_0^T dt \langle \hat{A}(t)\hat{B}(t')\rangle e^{i\omega(t'-t)} \tag{2}$$

where the power spectral density (PSD) represents a special instance:

$$S_{i_h i_h}(\omega) \equiv \langle |i_h(\omega)|^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt' \int_0^T dt \langle i_h(t) i_h(t')\rangle e^{i\omega(t'-t)} \tag{3}$$

The measured current $\equiv i_h(t_k)$, for $k = 1, 2, 3..N_T$, where typically here $N_T \sim 10^5 - 10^6$, is obtained as a discrete time series so conversion to frequency space is via a numerical discrete Fourier transform. For practical calculation, it is assumed that $T \to \infty$ simply denotes a long enough trace.

For a steady-state observable Q, which is also *stationary* and time-translation invariant, an important result is the Wiener-Kinchin theorem, which simplifies calculation of the PSD to a Fourier transform of its autocorrelation function:

$$\langle |\hat{Q}(\omega)|^2 \rangle \equiv S_{\hat{Q}^\dagger \hat{Q}}(\omega) = \int_0^T ds \langle \hat{Q}^\dagger(0)\hat{Q}(s)\rangle e^{i\omega s} \tag{4}$$

In other words, the double integral in Eq.3 reduces to a single integral, since then the autocorrelator depends only on the time-difference $s = t - t'$. Note the absence of the $1/T$ prefactor which is eliminated since $1/T \int_0^T dt' = 1$.

This approximation enables the derivation of analytical forms for optomechanical noise spectra in the quantum regime. We cannot use it for i_h as the correlations in heterodyne spectra are non-stationary. However the heterodyne current:

$$i_h(t) = \hat{a}^\dagger e^{i(\Omega t+\theta)} + \hat{a} e^{-i(\Omega t+\theta)} \tag{5}$$

is given in terms of field operators $\hat{a}$ and $\hat{a}^\dagger$ which are stationary; for these we define correlators:

$$S_{\hat{A}\hat{B}} = \int_0^T ds \langle \hat{A}(0)\hat{B}(s)\rangle e^{i\omega s} \tag{6}$$

where $\hat{A}, \hat{B}$ can be any combination of $\hat{a}$ and $\hat{a}^\dagger$, e.g. $S_{a\hat{a}^\dagger}(\omega) = \int_0^T \langle \hat{a}(-s/2)\hat{a}^\dagger(s/2)\rangle e^{i\omega s} ds \equiv \int_0^T \langle \hat{a}(0)\hat{a}^\dagger(s)\rangle e^{i\omega s} ds$.

As the limit $T \to \infty$ is implied throughout (i.e. a sufficiently long trace) for the gated FTs, the limits of integration are not significant provided the normalisation prefactor is adjusted appropriately.

THE EFFECT OF FILTER FUNCTIONS

If we now calculate the PSD of the measured current from the experimental time-series with a filter function, $\mathcal{F}(\bar{t})$ for example, we calculate:

$$S^{rhet}_{i_h i_h}(\omega) = \frac{1}{T}\int_0^T \mathcal{F}(\bar{t})d\bar{t} \;\; \int_0^T \langle i_h(\bar{t}-s/2)i_h(\bar{t}+s/2)\rangle e^{i\omega s}ds \tag{7}$$

where $\mathcal{F}(\bar{t}) = 1$ gives the ordinary PSD as in Eq.3. For the numerics, we simple evaluate Eq.7 using the experimental traces.

However, to understand the effect of the filters we can obtain approximate analytical forms; we first decompose the r-heterodyne 'PSD' in terms of field operators to obtain:

$$\begin{aligned}
T \quad & S^{rhet}_{i_h i_h}(\omega) = \\
\int_0^T \; & \mathcal{F}(\bar{t})d\bar{t}\int_0^T \langle \hat{a}^\dagger(\bar{t}-s/2)\hat{a}(\bar{t}+s/2)\rangle e^{i(\omega-\Omega)s}ds \\
\int_0^T \; & \mathcal{F}(\bar{t})d\bar{t}\int_0^T \langle \hat{a}(\bar{t}-s/2)\hat{a}^\dagger(\bar{t}+s/2)\rangle e^{i(\omega+\Omega)s}ds \\
\int_0^T \; & \mathcal{F}(\bar{t})e^{2i(\Omega\bar{t}+\theta)}d\bar{t}\int_0^T \langle \hat{a}^\dagger(\bar{t}-s/2)\hat{a}^\dagger(\bar{t}+s/2)\rangle e^{i\omega s}ds \\
\int_0^T \; & \mathcal{F}(\bar{t})e^{-2i(\Omega\bar{t}+\theta)}d\bar{t}\int_0^T \langle \hat{a}(\bar{t}-s/2)\hat{a}(\bar{t}+s/2)\rangle e^{i\omega s}ds
\end{aligned} \tag{8}$$

Clearly $\mathcal{F}(\bar{t}) = 1$ is the usual heterodyne spectrum. We then simply note that all the integrals over ds involve time translation-invariant integrals hence we can use the usual Wiener-Kinchin approximation and we then replace e.g. the first integral by $S_{\hat{a}^\dagger a}(\omega-\Omega)$ and so forth.

The integrals are then separable; the left-hand side integrals are replaced by the appropriate FTs of $\mathcal{F}(t)$. Hence we obtain:

$$\begin{aligned}
S^{rhet}_{i_h i_h}(\omega) \simeq & \frac{1}{T}\tilde{\mathcal{F}}(0)[S_{\hat{a}^\dagger\hat{a}}(\omega-\Omega) + S_{\hat{a}\hat{a}^\dagger}(\omega+\Omega)] \\
& + \frac{1}{T}\tilde{\mathcal{F}}(2\Omega)[e^{-2i\theta}S_{\hat{a}\hat{a}}(\omega) + S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega)e^{2i\theta}]
\end{aligned} \tag{9}$$

Noting that, for convenience, $\tilde{\mathcal{F}}(\omega) = FT[\mathcal{F}] = \int_0^T \mathcal{F}(t)e^{i\omega t}dt$ does not include the $1/\sqrt{T}$ prefactor.

We then seek appropriate filter functions for our purposes. To obtain hybrid homodyne-heterodyne spectra combining both stationary terms plus the correlations, we require $\tilde{\mathcal{F}}(0)/T \simeq \tilde{\mathcal{F}}(2\Omega)/T \simeq 1$, while to isolate correlations, we require $\tilde{\mathcal{F}}(0) = 0$. In the latter case, the spectra can be approximated by:

$$\begin{aligned}
S^{rhet}_{i_h i_h}(\omega) &\simeq \frac{1}{T}\tilde{\mathcal{F}}(2\Omega)[e^{-2i\theta}S_{\hat{a}\hat{a}}(\omega) + S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega)e^{2i\theta}] \\
&\simeq -[S_{\hat{a}\hat{a}}(\omega) + S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega)]
\end{aligned} \tag{10}$$

for the case $\theta = \pi/2$. These represent correlation spectra (exemplified by Fig.2a(i) in the main text) with stronger sidebands than the heterodyne case, but still less strong than the homodyne case. However, one advantage is that correlation spectra have no imprecision noise floor, even for the case of a shot-noise limited laser, so offer possibilities for enhancing displacement sensing, in scenarios where the sideband area is evaluated (a standard procedure for thermometry). The absence of this floor, could remove a source of error. The correlation spectra $S_{\hat{a}\hat{a}}(\omega)e^{-2i\theta} + S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega)e^{2i\theta}$ spectra are nonetheless real.

For the case of initial time, t_0-sampled spectra, we begin from Eq.1 once again, but use the substitution $t = t_0$ as well as $t - t' = s$. In that case we obtain an expression very similar to Eqs.8 and 9 except that we have:

$$\begin{aligned}
S_{\hat{a}\hat{a}}(\omega) \;&\rightarrow\; S_{\hat{a}\hat{a}}(\omega-\Omega) \\
S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega) \;&\rightarrow\; S_{\hat{a}^\dagger\hat{a}^\dagger}(\omega+\Omega)
\end{aligned} \tag{11}$$

so that the restored correlations coincide in frequency with the main heterodyne peaks. Hence, the $\tilde{\mathcal{F}}(0)/T \simeq \tilde{\mathcal{F}}(2\Omega)/T \simeq 1$ case (exemplified by Fig.3 in the main text) represents the closest to the homodyne case as it yields spectra of comparable strength to homodyne, but which still retain Stokes/anti-Stokes asymmetries in the quantum regime.

For computing spectra, expression Eq.7 is not practicable, since it would involve a sum over $N_T \sim 10^6$ FTs, combined with 100 realisations for the ensemble average and 800 values of θ for the correlations maps; instead we invert the order of the integrals and evaluate the equivalent expression:

$$S^{rhet}_{i_h i_h}(\omega) = \frac{1}{T}\int e^{i\omega s} ds \int \langle i(\bar{t} - s/2) i(\bar{t} + s/2)\rangle \mathcal{F}(\bar{t}) d\bar{t} \equiv \ \mathrm{FT}[\mathcal{F} \otimes \mathcal{A}_{\bar{t}}] \tag{12}$$

as in the main text, noting that we do not necessarily use the full $0 : T$ range (and this may be optimised with skilful coding) but adjust normalisation factors in the FT accordingly. As the method is most useful for experiments with long $T \to \infty$, this is not a significant factor.

We also note that above, $\mathcal{F} \otimes \mathcal{A}$ does not denote the standard convolution, but rather a modified form where:

$$\mathcal{F} \otimes \mathcal{A}_{\bar{t}} = \frac{1}{T}\int \langle i(\bar{t} - s/2) i(\bar{t} + s/2)\rangle \mathcal{F}(\bar{t}) d\bar{t} \tag{13}$$

where $\langle .. \rangle$ is the usual average over different realisations which is necessary to obtain a smooth PSD. Equivalently,

$$\mathcal{F} \otimes \mathcal{A}_{t_0} = \frac{1}{T}\int \langle i(t_0) i(t_0 + s)\rangle \mathcal{F}(t_0) dt_0. \tag{14}$$

As this is not a standard convolution, the PSD does not straightforwardly reduce to the product of the corresponding Fourier transforms. However, as shown in Eqs.9 and 10, the resulting spectra are constructed from individual products of Fourier transforms.

COMPARISON WITH CROSS-CORRELATION SPECTRA

The t_0 spectra shown in Fig.3 of the main text, which in future experiments in quantum regimes, with stable θ, may offer the "best of both worlds" scenario of strong sidebands, detecting both mechanical quantisation (via asymmetry) and squeezing regimes of the optical mode, are not obtainable by cross-correlation methods.

However, the isolated correlations, shown in Fig.2a(i) of the main text, correspond to the cross-correlation between the red and the blue heterodyne sideband and thus the spectrum of $\langle i_h^*(\omega + \Omega) i_h(\omega - \Omega)\rangle$. As this is complex in general, we show its real component. In general, this requires spectral demodulation, usually electronically, which has an amplitude cost as half the signal is lost. The reason for this is that shifting the red and blue sidebands onto each other requires that $\Omega = j\Delta_k$ where Δ_k is the frequency resolution of the Discrete Fourier Transform. This is a function of the time sampling and N_T. For our spectra, with $\Delta_k = 40$ Hz, and with the shifted sidebands out by only 20 Hz, the cross-correlation spectra were almost an order of magnitude weaker and much noisier as shown in Fig.3.

Demodulation can overcome this problem and such cross-correlation spectra were very recently measured (T.P. Purdy, K.E. Grutter, K. Srinivasan and J.M. Taylor, Science vol.356, 1265 (2017)) in a sophisticated experiment with electronic demodulation, which could even identify the quantum back-action component at room temperature. As the present experimental set-up has mechanical frequencies $\omega_M/(2\pi) = 0.39$ MHz which are $\sim 10^4$ smaller than in the above, its room -temperature phonon occupation is considerably higher.

However, with a 10^4 factor cooling from room-temperature, and proper phase stabilisation, such effects may be mapped directly from heterodyne traces with no further experimental requirement.

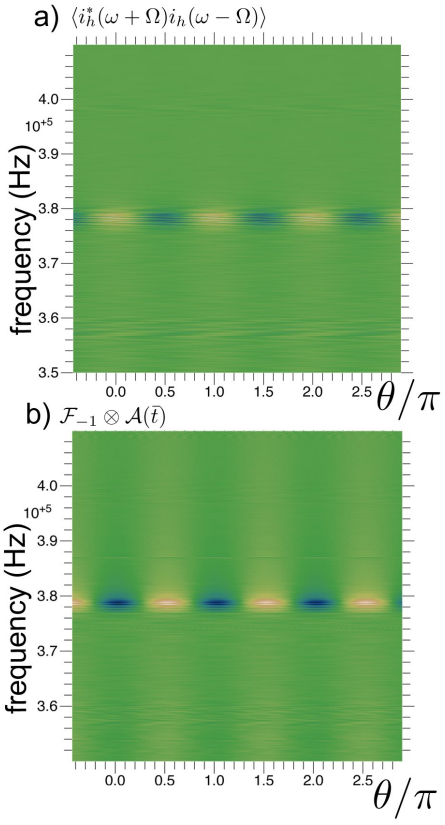

FIG. 3: **a.** Comparison of data analysed via cross-correlation obtained from $S(\omega)\langle i_h^*(\omega+\Omega)i_h(\omega-\Omega)\rangle + C.C.$ compared with the similar spectrum obtained **b.** with a filter function. The π shift is not significant it just means $-S(\omega)$ was calculated.