nature

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41586-018-0036-z

In the format provided by the authors and unedited.

Remote quantum entanglement between two micromechanical oscillators

Ralf Riedinger[1,3], Andreas Wallucks[2,3], Igor Marinković[2,3], Clemens Löschnauer[1], Markus Aspelmeyer[1], Sungkun Hong[1]* & Simon Gröblacher[2]*

[1]Vienna Center for Quantum Science and Technology, Faculty of Physics, University of Vienna, Vienna, Austria. [2]Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands. [3]These authors contributed equally: Ralf Riedinger, Andreas Wallucks, Igor Marinkovic´. *e-mail: sungkun.hong@univie.ac.at; s.groeblacher@tudelft.nl

Ralf Riedinger,[1,*] Andreas Wallucks,[2,*] Igor Marinković,[2,*] Clemens Löschnauer,[1] Markus Aspelmeyer,[1] Sungkun Hong,[1,†] and Simon Gröblacher[2,‡]

[1] *Vienna Center for Quantum Science and Technology (VCQ), Faculty of Physics, University of Vienna, A-1090 Vienna, Austria*

[2] *Kavli Institute of Nanoscience, Delft University of Technology, 2628CJ Delft, The Netherlands*

Supplementary Information
Remote quantum entanglement between two micromechanical oscillators

Device fabrication and characterization

The devices in the main part are fabricated as described in reference [1]. The most crucial steps for generating two identical chips are the electron beam lithography and the inductively coupled plasma reactive ion etching. We beamwrite and etch on a single proto-chip containing two sets of devices. This chip is then diced into two halves, each with several hundred nominally identical resonators. The structures are subsequently released in 40% hydrofluoric acid and cleaned with the RCA method, followed by a dip in 2% hydrofluoric acid. When characterizing the two chips, we find the center wavelengths to be 1552.4 nm on chip A and 1550.0 nm on chip B (see Figure S1). The standard deviation on the spread of the optical resonances is around 2 nm on both chips. For the experiments in the main text, we search for resonances that overlap to within 10% of their linewidth, which is equal to around 100 MHz. We find a total of 5 pairs fulfilling this requirement within 234 devices tested per chip.

In order to verify that finding identical devices is not just lucky coincidence and that this can even be done with a smaller sample size per chip, one can estimate the number of devices needed for a birthday paradox type approach. Therefore, we assume a pair of chips with 234 devices each that are centered at the same target wavelength. Taking similar parameters as found in our actual chips, we use a spread in resonance wavelength of 2 nm and we define resonances to be identical if they match to within 100 MHz. While the probability of obtaining a single device exactly at the center wavelength is only 0.03%, the probability of finding two matching devices at any wavelength within this distribution is 99.9996%. This is reduced if an offset in the mean wavelength of the two chips is introduced. For an offset of 2.5 nm, the probability is 99.98%, and for 5 nm, it is still 92.7%. By extending this approach to, for example, four chips, with the same parameters as above, no offset in the center wavelength and 500 devices per chip, we calculate the probability of finding four identical resonances to be 51.6%.

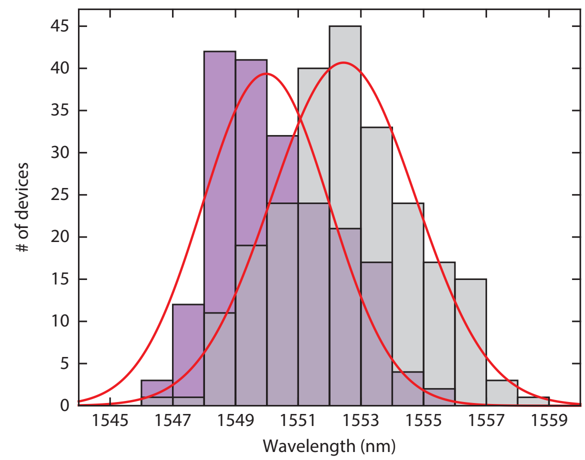

Figure S1: **Distribution of optical wavelengths.** We plot a histogram (bin size 1 nm) of the optical resonance wavelengths for a large set of devices on each of the chips containing devices A (gray) and B (magenta). The fits of a Gaussian distribution to the data sets (red solid lines) give a standard deviation of 2.3 nm and 2.0 nm, respectively. The large overlap of the optical resonance frequencies highlights the feasibility of extending the entanglement to even more optomechanical devices in the future.

Such a quartet would directly enable experiments on entanglement swapping and tests of a Bell inequality, as proposed by DLCZ. Further improvements could include post-fabrication wavelength tuning, as has recently been demonstrated for similar devices [2]. This could significantly improve the prospect of scalability of our approach, as it would allow to fabricate identical devices more deterministically.

In addition, the mechanical resonances are also susceptible to fabrication errors and vary by up to 50 MHz for our devices. To overcome this mismatch, we use serrodyne frequency shifting (see sections below).

Experimental setup

A detailed drawing of the experimental setup is shown in Figure S2. The light sources for our pump and read beams are two New Focus 6728 CW lasers, tuned and stabilized on their respective sideband of the opti-

*These authors contributed equally to this work.
†Electronic address: sungkun.hong@univie.ac.at
‡Electronic address: s.groeblacher@tudelft.nl

cal resonance. The beams are filtered by MicronOptics FFP-TF2 tunable optical filters in order to reduce the laser phase noise in the GHz regime. We then proceed to generate the actual pump and read pulses by driving acousto-optic modulators (Gooch&Housego T-M110-0.2C2J-3-F2S) with an arbitrary function generator (Tektronic AFG3152C). These pulses are then combined on a variable ratio coupler (Newport F-CPL-1550-N-FA). The combined optical mode is subsequently split by another variable ratio coupler and fed into the Mach Zehnder interferometer. The coupling ratio is adjusted to primarily compensate for a small difference in total losses between two paths. The power in the interferometer arms can additionally be balanced by an electrically driven variable optical attenuator (Sercalo VP1). We reflect the pulses from the two devices via optical circulators and recombine on a 50/50 coupler (measured deviation of 0.6%, see below). The strong pump pulses are filtered with two MicronOptics FFP-TF2 fiber filters per detection arm, tuned to transmit only the scattered (anti-) Stokes photons (bandwidth 50 MHz). We detect the resonant photons with superconducting nanowire single photon detectors (Photonspot) and register their arrival times on a TimeHarp 260 NANO correlation board.

Serrodyne frequency shifting

In our experiment, the mechanical frequency of device B ($\Omega_{\mathrm{m,B}}$) is greater than of device A ($\Omega_{\mathrm{m,A}}$) by $\Delta\Omega_{\mathrm{m}} = 2\pi \cdot 45$ MHz. If we were to send pump pulses with exactly the same frequency ω_{pump} to both of the devices, they would produce scattered photons with frequencies $\omega_{\mathrm{o,A}} = \omega_{\mathrm{pump}} - \Omega_{\mathrm{m,A}}$ and $\omega_{\mathrm{o,B}} = \omega_{\mathrm{pump}} - \Omega_{\mathrm{m,B}} = \omega_{\mathrm{o,A}} - \Delta\Omega_{\mathrm{m}}$. This frequency mismatch of scattered photons from the two devices would make them distinguishable, therefore preventing the entangled state. A simple solution is to shift the frequency of the pump pulse going to the device A by $\Delta\Omega_{\mathrm{m}}$, i.e. $\omega_{\mathrm{pump,A}} = \omega_{\mathrm{pump}} - \Delta\Omega_{\mathrm{m}}$. We experimentally realize this by electrically driving the electro-optic phase modulator on the path to device A, with a sawtooth waveform. This so-called serrodyne modulation with frequency $\omega_{\mathrm{s}} = \Delta\Omega_{\mathrm{m}}$ and peak-to-peak phase amplitude of 2π results in an optical frequency shift of ω_{s} [3, 4]. We use an arbitrary waveform generator (Agilent 81180A, bandwidth DC to 600 MHz) to generate the sawtooth voltage signal, amplify it with a broadband amplifier (Minicircuits TVA-R5-13A+, bandwidth 0.5 to 1000 MHz) and apply it to the optical phase modulator (Photline MPZ-LN-10-P-P-FA-FA-P, bandwidth DC to 12 GHz). We also apply an additional DC-bias to the serrodyne signal in order to generate a fixed phase offset $\Delta\phi$ in the interferometer arms. Due to the high analog bandwidth of the AWG compared to the frequency shift of 45 MHz, higher order sidebands were negligible and were not observed in the experiment.

Phase stabilization of the interferometer

For stabilizing the phase of the interferometer, we use an additional laser pulse $\sim$5 μs after the read pulse. To produce these auxiliary pulses, we also use the red-detuned laser that generates the read pulses and send them along the same beam paths. After being reflected from the optical filters in the detection line, the pulses are re-routed by optical circulators and picked up by a balanced detector (see Figure S2). These signals are then sent to a PID controller, which regulates a fiber stretcher to stabilize the relative path length and therefore locking the phase of the interferometer on a slow timescale (i.e. with the experiment repetition period of 50 μs).

In principle, the read or pump pulses that are reflected off the filter cavities could also be used for the phase locking. However, the serrodyne modulation during pump and read results in a beat signal of the pulses behind the beam splitter. This beating requires more sophisticated signal processing, which we avoid by using the auxiliary pulses, during which the serrodyne modulation is off. We note that the auxiliary pulses also induce some absorption heating of the devices. However, the 50 μs repetition period is sufficiently long compared to the decay times of the devices for the extra heating not to influence our experimental result.

Entanglement witness and Systematic Errors

The entanglement witness [5]

$$R(\tau, j) = \frac{\left\langle \hat{m}_{\mathrm{A}}^{\dagger}(\tau)\hat{m}_{\mathrm{A}}(\tau)\hat{m}_{\mathrm{B}}^{\dagger}(\tau)\hat{m}_{\mathrm{B}}(\tau)\right\rangle_j}{\left|\left\langle \hat{m}_{\mathrm{A}}^{\dagger}(\tau)\hat{m}_{\mathrm{B}}(\tau)\right\rangle_j\right|^2} \tag{S1}$$

derived in reference [6] is based on the concurrence [7] of the bipartite mechanical system. Here, the conditional average $\langle \hat{o}\rangle_j = \langle \hat{o}\, \hat{p}_j^{\dagger}\hat{p}_j\rangle / \langle \hat{p}_j^{\dagger}\hat{p}_j\rangle$ for an operator $\hat{o}$ is its expectation value of the state heralded by a Stokes photon detected by detector j. The $\hat{m}_i$ ($\hat{m}_i^{\dagger}$) are the mechanical annihilation (creation) operators of device $i = \mathrm{A}, \mathrm{B}$. While R is experimentally not directly accessible, the upper bound to this witness R_{m}, see Eq. (3), is a measurable quantity in our interferometry setup. The derivation of the inequality $R_{\mathrm{m}} \geq R$, as described in reference [6] and its supplementary material, is based on several of assumptions: The unheralded state must be Gaussian at all times, and the interference on the combining beamsplitter must be symmetric. In this section, we would like to discuss the validity of each of these assumptions in more detail.

To obtain R_{m}, threefold coincidence measurements are re-expressed as twofold coincidences, which can be done for Gaussian states. Note that as the degrees of second order coherence in Eq. (3) are measured between the

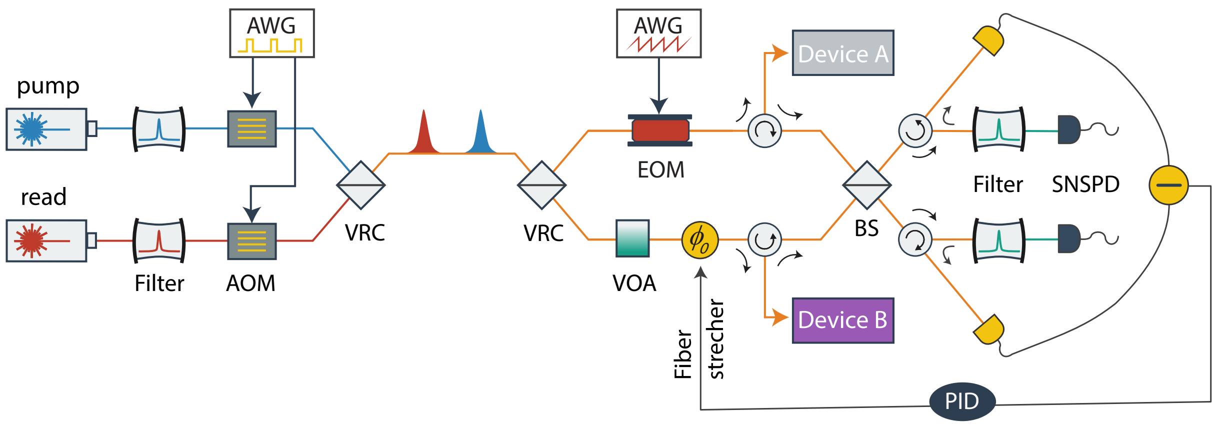

Figure S2: **Experimental setup.** A detailed schematics of our setup is shown here and described in the text. AOM are the acousto-optic modulators, AWG the arbitrary waveform generators, VRC the variable ratio couplers, EOM the electro-optic modulator, VOA the variable optical attenuator, BS the 50/50 beamsplitter and SNSPD the superconducting nanowire single-photon detectors.

pump and the read pulse, they are applied to this Gaussian state, not to the heralded, non-Gaussian entangled state $|\Psi\rangle \sim |A\rangle + e^{i\theta}|B\rangle$. We ensure that the mechanical states at the beginning of our protocol are Gaussian by allowing sufficiently long thermalization times (7x the mechanical decay time) prior to any optical manipulations. Consequently, the initial state is thermal, which for a bosonic system [1, 8] implies Gaussian quadrature statistics [9]. Next, all optomechanical interactions involved in our protocol are linear [10], therefore conserving the Gaussianity of the state [9]. Specifically, the Stokes process is described by the linear interaction Hamiltonian $\hat{H}_S \propto g_0 \hat{m}_i^\dagger \hat{o}_i^\dagger + h.c.$ and the anti-Stokes process by $\hat{H}_{AS} \propto g_0 \hat{m}_i \hat{o}_i^\dagger + h.c.$ for device $i = \mathrm{A}, \mathrm{B}$ with the annihilation (creation) operator of optical resonance $\hat{o}_i$ ($\hat{o}_i^\dagger$).

Unintentional interactions, like absorption heating, happen probabilistically and in a remote frequency regime, such that it effectively acts as a Gaussian thermal bath [1, 11]. Though not strictly contributing to the mechanical state, we also consider false positive detection events: drive photons leaking through the filter stages can be described by a coherent state (with Gaussian intensity fluctuations). Detection of stray photons and electrically caused false positive events are rare (~0.3% of the total count rate in the detection window) and uncorrelated (autocorrelation $g^{(2)}(0) = 1.05 \pm 0.09$), such that it is reasonable to model them as a Gaussian process as well. More specifically, we observe an average added noise of ~0.14 phonons during the measurement, from which the individual contributions of leaked drive photons, background detection events, and optical absorption heating can be estimated with additional measurements (for details see the section on "Second Order Coherence and Entanglement"). With the latter being a thermal process and therefore yielding Gaussian quadrature statistics, and an estimation of the initial thermal occupation from the nominal cryostat temperature, this leaves $\sim 4^{+2.8}_{-2.7} \cdot 10^{-2}$ phonons in device A and $\sim 0^{+1.4}_{-0} \cdot 10^{-2}$ in device B of unidentified origin. These are likely thermal phonons stemming from the non-ideal thermalization of the chips with their environment [1, 8].

For the modes leaving the interferometer $\hat{r}_\mathrm{j}$, $\hat{p}_\mathrm{j}$, reference [6] assumes an ideal 50/50 beamsplitter with equal powers on each input. Small experimental deviations from this idealized scenario result in quadratic corrections to the witness. For example, when the ratio of read photons at the beamsplitter originating from devices A and B is $1+\delta$, the measurable upper bound changes to $R_\mathrm{m}(1+\delta^2/2) \geq R$ for small δ. In our experimental setup, we choose a fused fiber beamsplitter with a measured deviation of 0.6%, leading to a relative correction on the order of 10^{-5}. Experimentally, we cannot achieve power balancing at the input of the beamsplitter for photons scattered from the pump and the read pulses at the same time because the devices have slightly different thermal occupation. We choose to match the detection rates of heralding photons, i.e. photons scattered by the pump pulse. This preserves the unknown origin of the heralding photons. Differences in the optomechanical coupling strength and optical losses on the path from the device to the combining beamsplitter are compensated by adjusting the drive power in each path. After the balancing procedure, the relative difference in count rates of heralding photons is below 2%, limited by the measurement precision during the balancing run and laser power fluctuations during the measurements. This leads to a relative correction of the witness below 10^{-3}. The slightly different heating dynamics between devices A and B result in a measured flux ratio deviation of the scattered readout photons of $\delta \sim 5-10\%$. It can easily be seen from Equation (3) that an increased heating will increase the witness R_{m}, and therefore the heating induced imbalance does not limit the validity of the witness. Yet, neglecting the thermal origin of the imbalancing, employing the correction $R_{\mathrm{m}}(1+\delta^2/2) \geq R$, we obtain a relative systematic correction of 0.5% to our result in the worst case scenario. Adding all systematic errors, we obtain a conservative upper bound of all relative systematic corrections of $\sim$0.5%. Consequently, we use a reduced classicality bound of $R_m \geq 0.995$ instead of 1, reducing the confidence level slightly from 99.84% to 99.82%.

Statistical Analysis

Results in the text and figures are given as maximum likelihood values and, where applicable, with a confidence interval of $\pm 34\%$ around this value. For the second order coherences $g^{(2)}_{ri,pj}$, $(i,j = 1,2)$, we apply binomial statistics based on the number of counted two-fold coincidences, which dominates the statistical uncertainty [8]. The entanglement witness $R_{\mathrm{m}}(\theta, j)$ in Eq. (3) is a non-trivial function of multiple such $g^{(2)}_{ri,pj}$, expressed here as $R_{\mathrm{m}}(\theta, j) \equiv \mathcal{R}_{\mathrm{m}}\left(g^{(2)}_{r1,pj}(\theta), g^{(2)}_{r2,pj}(\theta)\right)$. To estimate its confidence intervals, we discretize the probability density function of the second order coherences $P(g^{(2)}_{ri,pj} \in [a; a+\delta a])$ at equidistant $a = n\delta a, n \in \mathbb{N}$. The probabilities for finding R_{m} in an interval $[f, f+\delta f]$ is then given by $P(R_{\mathrm{m}} \in [f, f+\delta f]) = \sum_{(a,b)\in\mathcal{M}} P(g^{(2)}_{r1,pj} \in [a, a+\delta a]) P(g^{(2)}_{r2,pj} \in [b, b+\delta b])$ on the set $\mathcal{M}$ for which $\mathcal{R}_{\mathrm{m}}(a,b) \in [f, f+\delta f] \forall (a,b) \in \mathcal{M}$. For the optimal read phase θ_{opt} we obtain as maximum likelihood values for the witness bounds $R_{\mathrm{m}}(\theta_{\mathrm{opt}}, 1) = 0.612^{+0.152}_{-0.057}$ and $R_{\mathrm{m}}(\theta_{\mathrm{opt}}, 2) = 0.846^{+0.210}_{-0.090}$. We obtain the symmetrized witness by treating the experimentally observed $R_{\mathrm{m}}(\theta_{\mathrm{opt}}, 1)$ and $R_{\mathrm{m}}(\theta_{\mathrm{opt}}, 2)$ as two independent measurements of the expectation value $R_{\mathrm{m,\,sym}}(\theta) \equiv \langle R_{\mathrm{m}}(\theta, 1)\rangle \overset{!}{=} \langle R_{\mathrm{m}}(\theta, 2)\rangle$ of the optomechanical state in a symmetric setup with no detector noise. For a realistic setup with detectors exhibiting different noise properties, without loss of generality, let detector j have more false positive detection events than detector i. We then find the symmetrized witness $\langle R_{\mathrm{m}}(\theta_{\mathrm{opt}}, i)\rangle \leq R_{\mathrm{m,\,sym}} \leq \langle R_{\mathrm{m}}(\theta_{\mathrm{opt}}, j)\rangle$ to be an upper bound to the entanglement witness of the state heralded by the better detector i. Consequently $1 > R_{\mathrm{m,\,sym}} \geq R_{\mathrm{m}}(\theta_{\mathrm{opt}}, i) \geq R$ implies entanglement between the two remote mechanical oscillators. The observed values yield a confidence level for $R_{\mathrm{m,\,sym}} < 1$ of 98.4%. When correcting for the conservative upper bound of systematic errors, see above, the confidence level for having observed entanglement remains at 98.3%.

The complete counting statistics of the witness measurement at θ_{opt} are accumulated over $N = 1.114 \cdot 10^9$ trials (i.e. sets of pulses). We obtain $C(p_1) = 111134$ and $C(p_2) = 184114$ counts for photons scattered by the pump pulse on detectors $i = 1,2$ and $C(r_1) = 108723$ and $C(r_2) = 167427$ counts from the read pulse. This yields the coincidence counts $C_{ri,pj}$, $(i,j = 1,2)$ for counts on detector i, heralded by detector j, $C_{r1,p1} = 9$, $C_{r2,p1} = 130$, $C_{r1,p2} = 129$, $C_{r2,p2} = 37$.

For non-optimal phases $\theta \neq \theta_{\mathrm{opt}}$, $R_{\mathrm{m}}(\theta, i) \geq R_{\mathrm{m}}(\theta_{\mathrm{opt}}, i)$. Consequently, by adding the photon counts of measurements from an interval $[\theta_1, \theta_2]$, the resulting $R_{\mathrm{m}}([\theta_1, \theta_2], i) \geq R_{\mathrm{m}}(\theta \in [\theta_1, \theta_2], i)$ serves as an upper bound to any phase θ within that interval. Including the measurements at the non-optimal phase $\Delta\phi = 0$, we obtain $R_{\mathrm{m}}([\theta_{\mathrm{opt}} - 0.2\pi, \theta_{\mathrm{opt}}], 1) = 0.66^{+0.114}_{-0.055}$ and $R_{\mathrm{m}}([\theta_{\mathrm{opt}} - 0.2\pi, \theta_{\mathrm{opt}}], 2) = 0.806^{+0.129}_{-0.071}$, resulting in $R_{\mathrm{m,\,sym}}([\theta_{\mathrm{opt}} - 0.2\pi, \theta_{\mathrm{opt}}]) = 0.74^{+0.08}_{-0.05} \geq R_{\mathrm{m,\,sym}}(\theta_{\mathrm{opt}})$. The confidence level for $R_{\mathrm{m,\,sym}} < 1$ is 99.84%, dropping to 99.82% when correcting for the conservative upper bound for systematic errors. Note that for states heralded only with the more efficient detector 1, we already have a confidence level for entanglement between the two mechanical oscillators of 99.50%, including corrections for systematic errors.

The statistics of the witness within the extended phase region are obtained from $N = 1.949 \cdot 10^9$ experimental trials. We get $C(p_1) = 196080$ and $C(p_2) = 322608$ counts for photons scattered by the pump pulse on detectors $i = 1,2$ and $C(r_1) = 194023$ and $C(r_2) = 300373$ counts from the read pulse. This yields the coincidence counts $C_{ri,pj}$, $(i,j = 1,2)$ for counts on detector i, heralded by detector j, $C_{r1,p1} = 16$, $C_{r2,p1} = 223$, $C_{r1,p2} = 242$, $C_{r2,p2} = 67$.

Second Order Coherence and Entanglement

The second order coherence between the scattered photons from the pump pulse and signal phonons transfered by the read pulse after a delay t, $g^{(2)}_{i,rp}(t) = \langle \hat{r}^\dagger_{\mathrm{j}} \hat{p}^\dagger_{\mathrm{j}} \hat{r}_{\mathrm{j}} \hat{p}_{\mathrm{j}}\rangle / \langle \hat{r}^\dagger_{\mathrm{j}} \hat{r}_{\mathrm{j}}\rangle \langle \hat{p}^\dagger_{\mathrm{j}} \hat{p}_{\mathrm{j}}\rangle$, of the individual devices $i = A, B$, allows to us quantify the total noise contribution limiting the interference visibility. The measurements are performed the same way as described in the main text, however with the optical path to the other device blocked (see Fig. 2). Though there is no fundamental difference between the detectors $j = 1,2$, we only use detector $j = 2$ for the single device measurements. Following [12] and starting from a thermal state for the mechanical system ρ_m and vacuum $|0\rangle\langle 0|_o$ in the optical sidebands we obtain in the low temperature limit

$$g^{(2)}_{i,rp}(t) \approx 1 + \frac{e^{-\Gamma_i t}}{n_{i,\mathrm{th}}(t) + p_{\mathrm{pump},i} \cdot e^{-\Gamma_i t} + n_{\mathrm{leak}} + n_{\mathrm{bg}}}, \qquad \text{(S2)}$$

where $n_{i,\mathrm{th}} \ll 1$ is the mean phonon occupation of the device, $p_{\mathrm{pump},i} \ll n_{\mathrm{th}}$ is the Stokes excitation probability, $n_{\rm leak} \ll n_{\rm th}$ is the average number of leaked pump photons per transfered phonon, and $n_{\rm bg} \ll n_{\rm th}$ is the average number of background counts per detected phonon. At a delay of $\tau = 123$ ns $\ll 1/\Gamma_i$ we measure $g^{(2)}_{\rm A,rp} = 7.1^{+1.2}_{-0.9}$ and $g^{(2)}_{\rm B,rp} = 9.6^{+1.1}_{-0.9}$. With calibrated rates of $n_{\rm bg} = 3\cdot 10^{-3}$, $p_{\rm pump,A} = 0.56\cdot 10^{-2}$, $p_{\rm pump,B} = 0.80\cdot 10^{-2}$ and leaks at detector 1 $n_{\rm leak,1} = 4.2\cdot 10^{-2}$ and detector 2 $n_{\rm leak,2} = 3.2\cdot 10^{-2}$ we can estimate the number of incoherent phonons to be $n_{\rm th,A} = 11.9^{+2.8}_{-2.7}\cdot 10^{-2}$ for device A and $n_{\rm th,B} = 6.9^{+1.4}_{-1.3}\cdot 10^{-2}$ for device B. In a more detailed analysis, including the mechanical decay measurement (see below), we can estimate that absorption by the pump pulse contributes $n_{\rm pump} \sim 3\cdot 10^{-2}$ phonons and absorption from the read pulse contributes $n_{\rm read} \sim 5\cdot 10^{-2}$ phonons. This suggest, that the performance of device A is limited by imperfect thermalization with the cryostat at a temperature of 60 mK.

The interference contrast of the entangled state, i.e. $C_{\rm e}(t) = \max(g^{(2)}_{ri,pj}(t,\theta)) - \min(g^{(2)}_{ri,pj}(t,\theta))$ is bound by the cross correlation of the individual devices $C_{\rm e,max}(t) \approx \min(g^{(2)}_{{\rm A},rp}(t), g^{(2)}_{{\rm B},rp}(t)) - 1$ [13]. A pump-probe measurement of the thermal response of the devices [8] allows us to predict the cross correlations and thus an upper bound to the interference contrast. We excite the devices with a blue detuned pump pulse with $p_{\rm pump,A} = 2.8\%$ ($p_{\rm pump,B} = 4.0\%$) and vary the delay of a read (or probe) pulse with 5.6% (8.0%) state swap fidelity for device A (B). Note that for this experiment, we deliberately choose pulse energies higher than for the interference experiments in order to reduce the measurement time. As is done for sideband asymmetry [1, 8], the scattering rate of the probe pulse can be converted into the number of phonons at time t, when using the rate of the pump pulse (after scaling the signal according to the pump-to-probe power ratio) as a reference signal of single-phonon strength. This is only valid when $n_{\rm th} \ll 1$ at the time of pump pulse, which we find is the case later in the section. The results of these measurements are shown in Figure S3. We observe an initial rise in phonon occupation after the pump pulse, followed by a decay to an equilibrium state. The delayed heating can be understood by the presence of long lived high frequency phonons, which weakly couple to the 5 GHz modes under investigation. We model these high frequency phonons by an effective thermal bath, exponentially decaying with rate γ_i. This results in the rate equation $\dot{n}_i(t) = -\Gamma_i n_i(t) + k_i e^{-\gamma_i t} + \Gamma_i n_{i,\rm init}$ of the mean occupation number n_i of device $i =$ A,B. The additional bath couples with strength k_i, and the thermal environment of the chip has an equilibrium temperature of $n_{i,\rm init}$. The detected signal of the probe pulse $d_i(t) = n_{i,\rm pump}(t) + n_{i,\rm final}$ contains the average phonon number $n_{i,\rm pump}(t) = n_i(t) - n_{i,\rm init}$ induced by the pump beam, as well as a constant offset $n_{i,\rm final} = n_{i,\rm init} + n_{\rm i,probe} + n_{\rm leak} + n_{\rm bg}$ given by the thermal environment $n_{i,\rm init}$, the false positive events $n_{\rm leak}$ and $n_{\rm bg}$ and the heating during the probe pulse itself $n_{\rm i,probe}$. Consequently, we fit the pump-probe data with the general solution of the above differential equation $d_i(t) = a_i\cdot e^{-\Gamma_i t} - b_i\cdot e^{-\gamma_i t} + n_{i,\rm final}$, with a_i, b_i fitting parameters and including the offset $n_{i,\rm final}$ of the probe pulse detection rate. We obtain $1/\Gamma_A = 4.0\ \mu$s, $1/\Gamma_B = 5.8\ \mu$s, $1/\gamma_A = 0.5\ \mu$s and $1/\gamma_B = 0.5\ \mu$s. Using the calibrated false positive detection rates and estimating the equilibrium occupation from the thermal environment $n_{i,\rm init} \sim 1/(e^{\hbar\Omega/k_B T} - 1)$, with the Boltzmann constant k_B and cryostat temperature T, we can obtain the individual contributions to the absorption heating by the pump ($n_{i,\rm pump}$) and the probe pulse ($n_{i,\rm probe}$) for any given delay t. Adapting the number of phonons added by the optical drive pulses $n_{i,\rm pump}$ and $n_{i,\rm probe}$ for the lower energies, under the assumption of linear absorption and optomechanical processes, we obtain an estimate of the thermal occupation $n_{i,\rm th}(t) = n_i(t) + n_{i,\rm probe} - p_{{\rm pump},i} e^{-\Gamma_i t}$ during the entanglement experiment. Using the calibrated rates and equation (S2), we can obtain an upper bound to the interference contrast $C_{\rm e,max}$ (see above) and therefore also for the visibility $V_{\rm max} = C_{\rm e,max}/(C_{\rm e,max} + 2)$, which is shown in Figure 4.

Rates and Extrapolation of Results

In order to highlight the scalability of the mechanical entanglement, we recapitulate the detection rates in the entanglement witness measurement. The experiment is repeated every 50 μs, limited by the thermalization time of the mechanical modes. Using the counting statistics of the measurement at the optimal phase $\theta_{\rm opt}$, see above, we have a probability of $p_{\rm herald} \approx 2.7\cdot 10^{-4}$ and the unconditional probability to also detect the anti-Stokes photon from the readout pulse of $p_{\rm read} \approx 2.8\cdot 10^{-7}$. These probabilities contain the optomechanically generated photons, leaked pump photons and background counts. In the current setup these rates are limited by losses in the filtering setup and low optomechanical scattering rates to retain the effects of absorption heating. When the two mechanical devices are placed in two separate refrigerators and placed in different locations, additional fiber would be inserted, causing additional losses. While the detection rate of leaked pump photons reduces in the same manner as the of the optomechanical photons, the rate of background counts stays the same. Consequently, the signal-to-noise ratio reduces, lowering the normalized cross-correlation between Stokes and anti-Stokes photons. In the present measurements, the inverse signal-to-background ratio is $n_{\rm bg,1} = 2.9\cdot 10^{-3}$ ($n_{\rm bg,2} = 3.2\cdot 10^{-3}$) for detectors 1 (2). The most reliable estimate of the second order coherence of device A, which has worse properties, can be extracted from the observed interference contrast, resulting in $g^{(2)}_{A,\rm rp} \sim 7.5 \pm 0.35$, as it has better statistics than the direct measurement. For device B, we

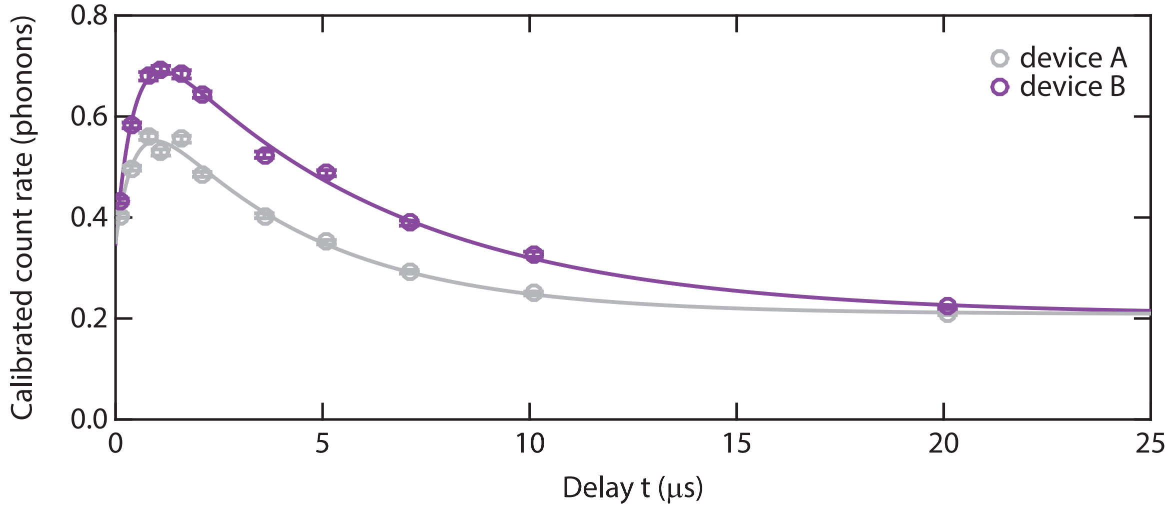

Figure S3: **Pump-probe experiment.** The experiment reveals the response of the devices mechanical modes to the initial optical pump. See text for details. Error bars are s.d.

use the direct measurement $g^{(2)}_{B,\mathrm{rp}} = 9.6^{+1.1}_{-0.9}$. To maintain an interference contrast of 95% of the current level, the second order coherence of both devices is allowed to decrease to ∼7.1. With equation (S2) we can estimate that an additional loss of 5.4 dB (10.6 dB) loss in the optical path from device A (B) to the combining beamsplitter decreases the signal-to-background ratio, such that $g^{(2)}_{A,rp} \approx g^{(2)}_{B,rp} \approx 7.1$. Using the nominal attenuation of commercial low-loss telecom fiber (Corning SMF-28 ULL ∼0.17 dB/km), we can estimate that an additional fiber length of 94^{+8}_{-12} km could be inserted between the devices. The projected entanglement witness in this configuration is $R_m \sim 0.76$ and would need roughly the same number of coincidence events to clear the classicality bound of 1 by 3 standard deviations. Including the reduced scattering rates from matching both paths, this would require ∼170 days of integration time, including a 15% overhead time for stabilization of filters and interferometer as well as data management. Reducing the separation distance to 75 km, i.e. an additional fiber length of 32 km (43 km) between device A (B) and the beamsplitter, requires 38 days of continuous measurement time for a statistically significant demonstration of remote entanglement. While our cryostat currently only allows for a few weeks of measurements at a time due to technical limitations, much longer times should in principle be easily achievable.

[1] S. Hong, R. Riedinger, I. Marinković, A. Wallucks, S. G. Hofer, R. A. Norte, M. Aspelmeyer, and S. Gröblacher, Science **358**, 203 (2017).

[2] K. Fang, J. Luo, A. Metelmann, M. H. Matheny, F. Marquardt, A. A. Clerk, and O. Painter, Nature Phys. **13**, 465 (2017).

[3] R. C. Cumming, Proc. IRE **45**, 175 (1957).

[4] K. K. Wong, R. M. D. L. Rue, and S. Wright, Opt. Lett. **7**, 546 (1982).

[5] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[6] K. Børkje, A. Nunnenkamp, and S. M. Girvin, Phys. Rev. Lett. **107**, 123601 (2011).

[7] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).

[8] R. Riedinger, S. Hong, R. A. Norte, J. A. Slater, J. Shang, A. G. Krause, V. Anant, M. Aspelmeyer, and S. Gröblacher, Nature **530**, 313 (2016).

[9] W. Wieczorek, S. G. Hofer, J. Hoelscher-Obermaier, R. Riedinger, K. Hammerer, and M. Aspelmeyer, Phys. Rev. Lett. **114**, 223601 (2015).

[10] S. G. Hofer, W. Wieczorek, M. Aspelmeyer, and K. Hammerer, Phys. Rev. A **84**, 52327 (2011).

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

[12] A. Kuzmich, W. P. Bowen, A. D. Boozer, A. Boca, C. W. Chou, L.-M. Duan, and H. J. Kimble, Nature **423**, 731 (2003).

[13] K. C. Lee, Ph.D. thesis, University of Oxford (2012).