Non-classical correlations between single photons and phonons from a mechanical oscillator*

Ralf Riedinger,[1,†] Sungkun Hong,[1,†] Richard A. Norte,[2] Joshua A. Slater,[1] Juying Shang,[3] Alexander G. Krause,[2] Vikas Anant,[3] Markus Aspelmeyer,[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*

[3] *Photon Spot Inc., Monrovia, CA 91016, USA*

Interfacing a single photon with another quantum system is a key capability in modern quantum information science. It allows quantum states of matter, such as spin states of atoms [1, 2], atomic ensembles [3, 4] or solids [5], to be prepared and manipulated by photon counting and, in particular, to be distributed over long distances. Such light-matter interfaces have become crucial to fundamental tests of quantum physics [6] and realizations of quantum networks [7]. Here we report non-classical correlations between single photons and phonons – the quanta of mechanical motion – from a nanomechanical resonator. We implement a full quantu protocol involving initialization of the resonator in its quantum ground state of motion and subsequent generation and read-out of correlated photon-phonon pairs. The observed violation of a Cauchy-Schwarz inequality is clear evidence for the non-classical nature of the mechanical state generated. Our results demonstrate the availability of on-chip solid-state mechanical resonators as light-matter quantum interfaces. The performance we achieved will enable studies of macroscopic quantum phenomena [8] as well as applications in quantum communication [9], as quantum memories [10] and as quantum transducers [11, 12].

arXiv:1512.05360v2 [quant-ph] 23 Feb 2016

Over the past few years, nanomechanical devices have been discussed as possible building blocks for quantum information architectures [9, 13]. Their unique feature is that they combine an engineerable solid-state platform on the nanoscale with the possibility to coherently interact with a variety of physical quantum systems including electronic or nuclear spins, single charges, and photons [14, 15]. This feature enables mechanics-based hybrid quantum systems that interconnect different, independent physical qubits through mechanical modes.

A successful implementation of such quantum transducers requires the ability to create and control quantum states of mechanical motion. The first step – the initialization of micro- and nanomechanical systems in their quantum ground state of motion – has been realized in various mechanical systems either through direct cryogenic cooling [16, 17] or laser cooling using microwave [18] and optical cavity fields [19]. Further progress in quantum state control has mainly been limited to the domain of electromechanical devices, in which mechanical motion couples to superconducting circuits in the form of qubits and microwave cavities [15]. Recent achievements include single-phonon control of a micromechanical resonator by a superconducting flux qubit [16], the generation of quantum entanglement between quadratures of a microwave cavity field and micromechanical motion [20], and the preparation of quantum squeezed micromechanical states [21–23].

Interfacing mechanics with optical photons in the quantum regime is highly desirable because it adds important features such as the ability to transfer mechanical excitations over long distances [9, 24]. In addition, the available toolbox of single-photon generation and detection allows for remote quantum state control [7]. However, micro- and nano-mechanical quantum control through single optical photons has not yet been demonstrated. One of the outstanding challenges is to achieve single-particle coupling rates that are sufficiently large to alleviate effects of optical and mechanical decoherence in the system, that is, single-photon strong co-operativity. Some of the largest optomechanical couplings have been reported in nanomechanical photonic crystal cavities [25], but are still two orders of magnitude short of that regime. Although low coupling rates can be overcome in principle by a strong and detuned coherent drive field [15], such measures typically result in unwanted heating of the mechanical device (see Methods).

Here we take a different approach that allows us to circumvent these problems and to realize quantum control of single phonons through single optical photons. We use a probabilistic scheme based on the well-known DLCZ protocol (Duan, Lukin, Cirac and Zoller) [26], which, in its original form, uses Raman scattering for efficient generation and read-out of collective spin states of atomic ensembles. In essence, the scheme generates entanglement through single-photon interference and post-selection, which does not require strong coupling [27]. In the context of mechanical quanta, this protocol has

*This work was published in Nature **530**, 313–316 (2016).

†These authors contributed equally to this work.

‡Electronic address: markus.aspelmeyer@univie.ac.at

§Electronic address: s.groeblacher@tudelft.nl

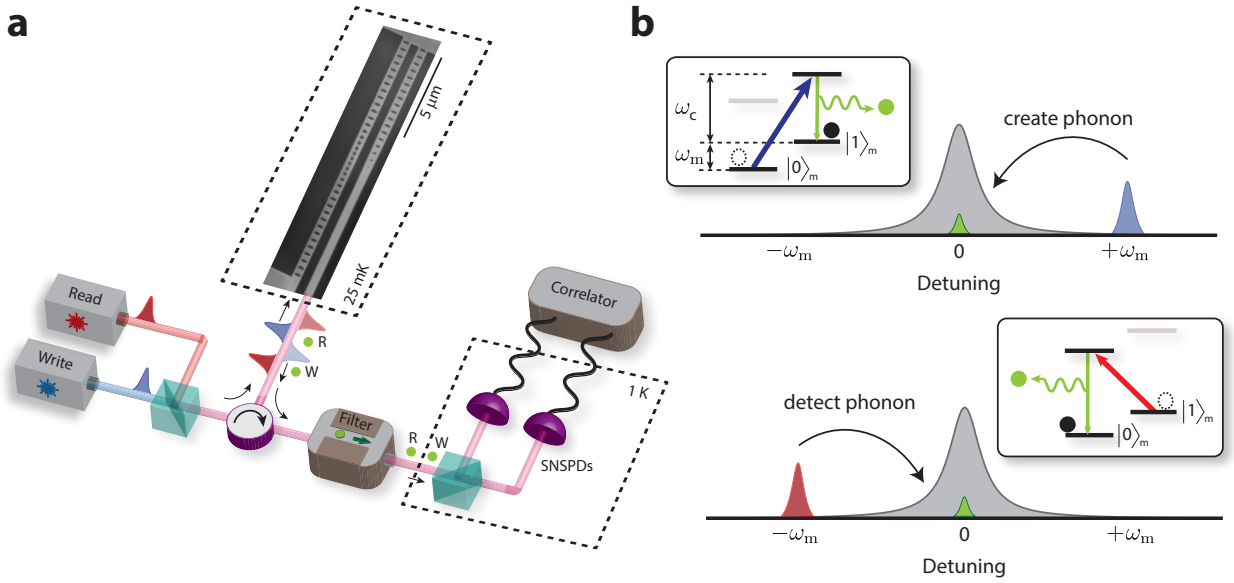


Figure 1: **Generation and read-out of photon-phonon pairs. a**, Schematic of the experiment. Two independent lasers (stabilized to a wave-meter) are used to generate a sequence of 'write' and 'read' pulses with tunable time delay δt. They are sent through a circulator and drive a nanomechanical photonic crystal cavity (a scanning electron microscope image of which is shown in the inset) that is mounted inside a dilution refrigerator at a base temperature of 25 mK, which prepares the device in its quantum ground state of motion. For each pulse, Stokes and anti-Stokes Raman scattering creates single photons (green dots) from the write (W) and the read (R) pulse, respectively, that are emitted at a frequency ω_c. The detuned pump fields are strongly suppressed by optical filtering and only the Raman scattered photons are measured by two superconducting nanowire single-photon detectors (SNSPDs) in the output ports of a 50/50 beam-splitter. The time of each photon detection event is recorded and is then correlated in post-processing to obtain both auto- and cross-correlations of the emitted photons. A more detailed explanation of the experimental set-up is provided in Methods. **b**, Pulsed optomechanical interactions in frequency space. A blue-detuned write pulse realizes a two-mode squeezing interaction (blue and green pulses; see text). Cavity-enhanced Stokes Raman scattering generates a single phonon, stored as an excitation on the mechanical resonator, and a single (W) photon, which is emitted from the cavity on resonance (upper panel). Reading out of the phonon utilizes a red-detuned read pulse, which swaps the mechanical excitation onto the optical cavity field, hence creating a single (R) photon (lower panel). The insets depict the relevant energy level diagrams for the two processes, reminiscent of the Λ-schemes in atomic Raman scattering. The grey bars indicate the energy levels that are not involved in the depicted process (Stokes or anti-Stokes), but in the other one.

been used in an experiment to entangle high-frequency (40 THz) optical phonons of two bulk diamond lattices [28]. However, the small interaction and coherence times of such phonons are incompatible with their use in quantum transduction and storage, and so it is necessary to take this approach to the level of chip-scale optomechanical systems. In addition, we minimize absorption heating by using short optical pulses in a cryogenic environment [17]. The combination of these techniques allows us to overcome the previous limitations and realize a photon-phonon quantum interface.

Our experiment complements previous work on single- and two-mode (opto-)mechanical squeezing in microwave circuits [20–23]. Although these experiments were based on the same underlying interactions, they involved homodyne or heterodyne detection of light to access continuous-variable degrees of freedom of a quantum state – specifically, quadrature fluctuations in the mechanical and optical canonical variables. In contrast, the DLCZ scheme uses photon counting, which allows access to discrete quantum variables – here, in form of energy eigenstates (phonons) of the mechanical motion – and thereby enables realistic architectures for entanglement distribution and quantum networking [7].

The mechanical system studied here is a microfabricated silicon photonic crystal nanobeam structure (Fig. 1a). Such optomechanical crystals co-localize optical and mechanical modes and couple them via a combination of radiation pressure and photostriction [15]. Our device exhibits an optical cavity resonance at wavelength $\lambda_c = 1{,}556$ nm and a mechanical breathing mode at frequency $\omega_m/2\pi = 5.3$ GHz. The cavity decay rate (full-width at half-maximum, FWHM) is $\kappa_c/2\pi = 1.3$ GHz and the mechanical quality factor at cryogenic temperature is $Q_m = 1.1 \cdot 10^6$ (see Methods). Pulsed optical driving at laser frequency $\omega_L = \omega_c \pm \omega_m$ (in which $\omega_c = 2\pi c/\lambda_c$ is the cavity frequency and c is the vacuum speed of light) allows to realize two different types of interactions on

the basis of cavity-enhanced Stokes (+) and anti-Stokes (−) Raman scattering (Fig. 1b). A blue-detuned pulse ($\omega_L = \omega_c + \omega_m$) results in two-mode squeezing with interaction Hamiltonian $H_{tms} \propto \hbar g_0 \left(\hat{a}_m^\dagger \hat{a}_o^\dagger + \hat{a}_m \hat{a}_o\right)$, in which $\hat{a}_m^{(\dagger)}$ and $\hat{a}_o^{(\dagger)}$ are the creation (annihilation) operators of the mechanical and optical mode, respectively, g_0 is the effective optomechanical coupling rate (here, $g_0/2\pi = 825$ kHz; see Methods) and $\hbar$ is the reduced Planck constant. This interaction generates photon-phonon pairs in close analogy to the photon-photon pairs generated in parametric down-conversion [29]. A red-detuned pulse ($\omega_L = \omega_c - \omega_m$) allows read-out of the mechanical state through the optomechanical beam-splitter interaction $H_{bs} \propto \hbar g_0 \left(\hat{a}_m \hat{a}_o^\dagger + \hat{a}_m^\dagger \hat{a}_o\right)$, in which an anti-Stokes scattering event realizes a state swap between the mechanical and optical cavity mode.

Our protocol consists of three distinctive steps. First, we initialize the mechanical system in its quantum ground state of motion by cryogenic cooling. Second, a short blue pulse creates a photon-phonon pair and leaves the originally empty mechanical and optical modes $|0\rangle_m$ and $|0\rangle_o$ at frequencies ω_m and ω_c, respectively, in the state $|\Phi\rangle_{om} = |00\rangle + \sqrt{p}|11\rangle + p|22\rangle + \mathcal{O}(p^{3/2})$. Here p is the probability for a single Stokes scattering event to take place. Residual heating through optical absorption introduces additional noise to the state (see Methods). Finally, a strong red pulse is used to read out the phonon state via emission of an anti-Stokes scattered photon [30]. We confirm the non-classical photon-phonon correlations on the basis of an observed violation of a Cauchy-Schwarz inequality for the cross-correlation of the coincidence measurements between the Stokes and anti-Stokes photons [3].

Precooling of the nanomechanical device is performed using a dilution refrigerator that operates at a base temperature of approximately 25 mK. If the mechanical system is in its quantum ground state of motion, then anti-Stokes processes cannot occur because no additional phonons can be extracted to support the scattering. This is in contrast to Stokes processes, which deposit mechanical energy and hence can always occur. As a consequence, the asymmetry in the scattering rates of these two processes is a direct measurement of the mean thermal phonon occupancy n_{th}. Using such photon-counting based sideband thermometry [17], we find $n_{th} \lesssim 0.025$ (see Fig. 2).

We create the desired photon-phonon pairs using a blue-detuned 'write' pulse that is sufficiently weak to minimize the effects of residual absorption heating (FWHM, 28.4 ns; energy, 40 fJ). We find the relevant probability to generate a Stokes scattered photon on cavity resonance to be $p \approx 3.0\%$. Subsequently, a red-detuned 'read' pulse (effective length, 55 ns; energy of approximately 50 fJ) is injected at a time delay δt (see Fig. 3a), resulting in a phonon-to-photon conversion efficiency of approximately

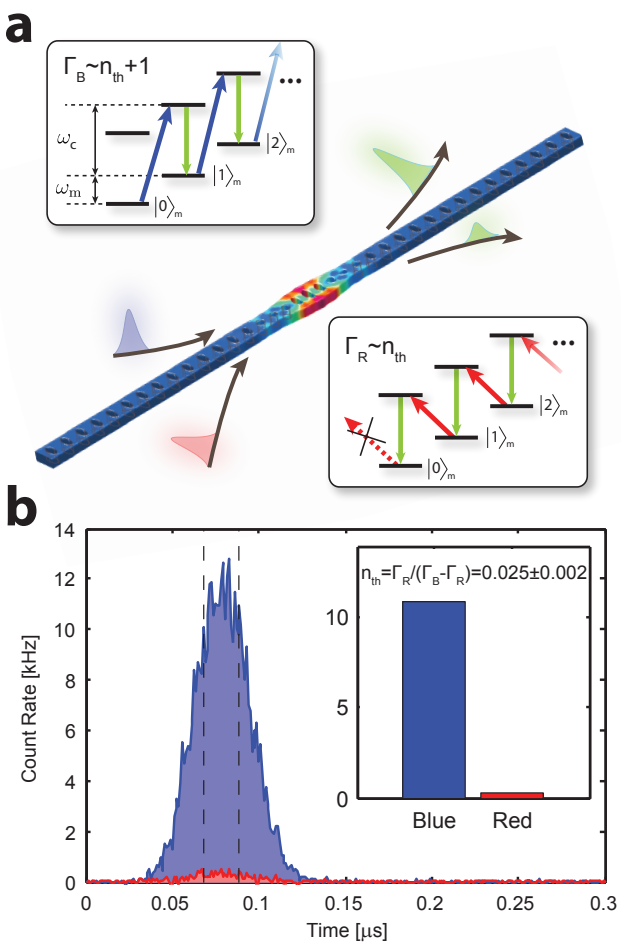


Figure 2: **Mechanical quantum ground state preparation. a**, Principle of sideband thermometry. The finite element method simulation depicted in the main panel shows the structure of the mechanical breathing mode under investigation. The upper (lower) inset shows the energy level scheme in case of blue- (red-) detuned pumping and the resultant cavity-enhanced Stokes (anti-Stokes) scattering. The corresponding scattering rates Γ_R and Γ_B are proportional to thermal occupation of the mechanics n_{th} and $n_{th} + 1$, respectively, and hence show a strong asymmetry when the mechanics are close to the quantum ground state. **b**, Sideband asymmetry. The optomechanical device is pumped with a sequence of alternating blue- and red-detuned optical pulses at frequency $\omega_c \pm \omega_m$ (optical energy per pulse $E_{opt} = 33$ fJ; FWHM of 28.4 ns; 500 μs separation of pulse sequences). Shown are the count rates recorded by the SNSPDs as a function of the arrival time of the scattered photons (blue, blue-detuned pulse; red, red-detuned pulse). This data has been corrected for leakage of pump photons through the optical filters, which was independently measured and subtracted from our data (see Methods). The inset shows a histogram of the total counts that are obtained when averaging over a 20 ns window centred on the peak (within the dashed lines). The pronounced asymmetry in the rates (of more than a factor of 40) corresponds to a thermal occupancy of $n_{th} = \Gamma_R/(\Gamma_B - \Gamma_R) = 0.025 \pm 0.002$ and to a mode temperature of 69 mK.

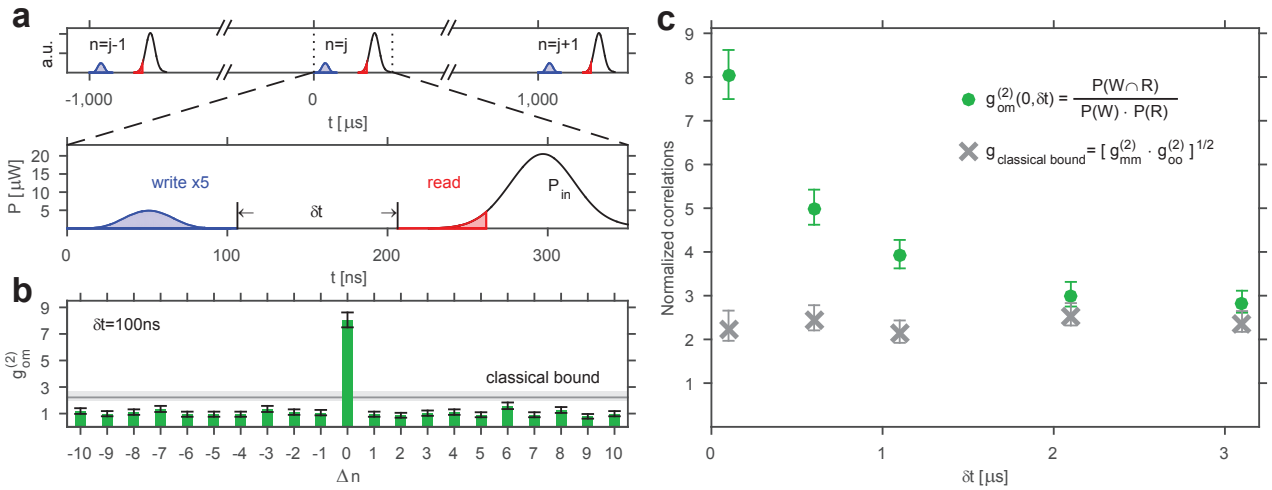


Figure 3: **Non-classical photon-phonon correlations. a**, Driving pulse sequence. A pair of one write (blue) and one read (red) pulse is sent to the device every 1 ms. The long idle phase between pulse pairs ensures the ground-state initialization by cryogenic cooling. Each pulse sequence is labelled with a number (n). The read pulse is delayed by δt with respect to the write pulse, and only the first 55 ns, equivalent to a read-pulse power of about 50 fJ, are used for the data evaluation. This reduces the influence of absorption heating while maintaining reasonable state swap fidelity. **b**, Violating a Cauchy-Schwarz inequality. Shown is the cross-correlation (green bars) between the mechanical (read pulse) and optical state (write pulse) for $\delta t = 100$ ns, as well as the classical (Cauchy-Schwarz) bound obtained from the autocorrelations at $\Delta n = 0$ (grey horizontal line, shading indicates a 68% confidence interval; see text). For photon-phonon pairs that emerge from different pulse sequences ($\Delta n \neq 0$) the Cauchy-Schwarz inequality is fulfilled, $\langle g^{(2)}_{\rm om}(\Delta n \neq 0, 100\ {\rm ns})\rangle = 1.04 \pm 0.04$, consistent with statistical independence. For pulses from the same pair, the cross-correlation $g^{(2)}_{\rm om}(0, 100\ {\rm ns})$ clearly exceeds the classical bound. $g^{(2)}_{\rm om}$ can be interpreted as the ratio of heralded phonons $n_{\rm h}$ to unheralded (thermal) phonons $n_{\rm th}$ at the time of the read pulse. **c**, Storage of non-classical correlations. Shown is the dependence of the cross-correlation on the time delay δt between the write and read pulses. For each data point, the classical bound is measured independently through the normalized autocorrelation functions of the write (W) and read (R) photons. For increasing δt, the photon-phonon cross-correlations decrease, but stay above the classical limit even beyond 1 μs. The main contribution to the loss of correlation is heating by absorption of the write pulse (see Methods). All error bars represent a 68% confidence interval.

3.7% (see Methods).

We correlate the measured Stokes- and anti-Stokes photons via the cross-correlation function $g^{(2)}_{\rm om}(\Delta n, \delta t) = P(W \cap R)/\left[P(R)P(W)\right]$, which is computed for read and write pulses originating from pulse sequences from different trials separated by Δn iterations (see Fig. 3). $P(W \cap R)$ is the probability for a joint detection of both a Stokes (W, 'write') and an anti-Stokes (R, 'read') photon from these pulses, and $P(W)$ and $P(R)$ are the unconditional probabilities to detect either of the two photons. For all pair correlations of classical origin, the value of $g^{(2)}_{\rm om}$ is bounded by a Cauchy-Schwarz inequality of the form [3] $g^{(2)}_{\rm om}(0, \delta t) \leq \left[g^{(2)}_{{\rm oo},\delta t}(0) g^{(2)}_{{\rm mm},\delta t}(0)\right]^{1/2}$, in which $g^{(2)}_{{\rm oo},\delta t}(0)$ and $g^{(2)}_{{\rm mm},\delta t}(0)$ are the autocorrelation functions for the optical and mechanical mode, respectively (see Methods). A violation of this inequality [3, 31, 32] is an unambiguous measure for the non-classicality of the generated photon-phonon state. The Cauchy-Schwarz inequality for coincidence detection marks a well-defined border between the quantum and classical domain. It is based on the fact that the Glauber-Sudarshan phase-space function, or P-function, is positive definite for every classical field. This places a fundamental limit on the relative strength of measurable cross-correlations versus autocorrelations between classical fields. Previous applications of this limit include the distinction between the classical and quantum field theoretical predictions for the photoelectric effect [31], and the storage and retrieval of non-classical states in the collective emission from an atomic ensemble [3]. A detailed derivation of the Cauchy-Schwarz inequality for the case of non-stationary fields, as are being used here, is provided in ref. [3].

We find a clear violation for an extended regime of time delays. Figure 3b shows the value of $g_{\rm om}$ at a time delay of 100 ns. For pairs emitted from the same pulse sequence ($\Delta n = 0$) we find that $\left\{g^{(2)}_{\rm om}(0, 100\ {\rm ns}) = 8.0 + 0.6 - 0.5\right\} \not\leq \left\{\left[g^{(2)}_{{\rm oo},100\ {\rm ns}}(0) g^{(2)}_{{\rm mm},100\ {\rm ns}}(0)\right]^{1/2} = 2.09 + 0.23 - 0.16\right\}$, which obviously violates the classical bound. As expected, pairs emitted from different pulse sequences ($\Delta n \neq 0$) are uncorrelated and hence fulfil the inequality. Upon increasing the time delay further, we find a violation even beyond $\delta t = 1$ μs (see Fig. 3c), which demonstrates that we can store and retrieve non-classical

states for an extended time interval. Nevertheless, the lifetime of these non-classical correlations is still much shorter than the lifetime of the mechanical excitations, $Q/\omega_{\rm m} \approx 34$ μs. We attribute this to the fact that the dynamics are dominated by heating caused by absorption of pump photons, which after some onset time drives the mechanical system towards a thermal state (see Methods). As a consequence, reducing the energy of the write pulse further should allow non-classical correlations to be maintained for much longer times. In addition, upon further reduction of the absorption heating of the read pulse, even higher values for the cross-correlation are obtained.

The cross-correlation is also linked to the autocorrelation of the heralded mechanical state. If one considers two-mode optomechanical squeezing acting on an initial mechanical thermal state, and if $g^{(2)}_{\rm om} \gg 1$ – as is the case in our experiment – then one obtains $g^{(2)}_{\rm mm,heralded} \approx 4/\left(g^{(2)}_{\rm om}-1\right)$. The largest value for $g^{(2)}_{\rm om}$ observed in our experiment was $g^{(2)}_{\rm om}(0, 100\text{ ns}) = 19.6 - 2.8 + 3.9$ (using an energy of 1.7 fJ in the first 30 ns of the read pulse; see Methods). In other words, our system should allow for a Hanbury Brown and Twiss experiment with phonons yielding $g^{(2)}_{\rm mm,heralded} \approx 0.22$. A direct measurement of this value with the current experimental parameters is difficult without a prohibitively large number of pulse sequences.

In summary, we have demonstrated non-classical correlations between single photons and phonons from a nanomechanical resonator. This is a crucial step towards on-chip photon-phonon quantum interfaces, which are relevant for future solid-state based quantum information and communication architectures. For example, the observed photon-phonon correlation of $g^{(2)}_{\rm om} = 19.6$ suggests that conditional mechanical Fock-state preparation should be possible with fidelities exceeding 85% (see Methods). The ability to store and retrieve non-classical states over extended storage times that we reported also shows that nano-optomechanical resonators are a promising candidate for quantum memories. The performance of the system we have demonstrated constitutes an improvement of almost two orders of magnitude on previous lifetimes of stored non-classical single-phonon states [16]. Finally, photon-phonon conversion on the single particle level is required to extend the ongoing efforts on mechanically transduced conversion between microwave and optical fields [12] into the quantum domain [11].

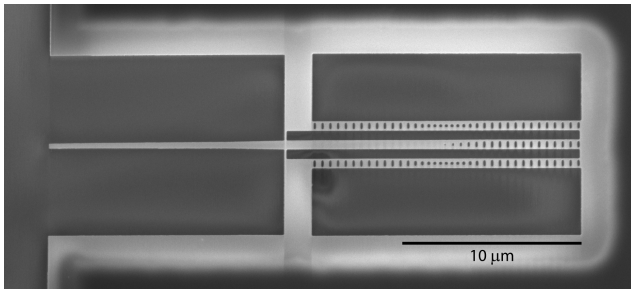
Acknowledgments Acknowledgements We thank K. Hammerer and S. Hofer for discussions, and T. Graziosi, J. Hill, J. Hoelscher-Obermaier, Y. Liu, L. Procopio, A. Safavi-Naeini, E. Schafler, G. Steele and W. Wieczorek for experimental support. We acknowledge assistance from the Kavli Nanolab Delft, in particular from M. Zuiddam and F. Dirne. This project was supported by the European Commission (cQOM, SIQS, IQUOEMS), a Foundation for Fundamental Research on Matter (FOM) Projectruimte grant (15PR3210), the Vienna Science and Technology Fund WWTF (ICT12-049), the European Research Council (ERC CoG QLev4G), and the Austrian Science Fund (FWF) under projects F40 (SFB FOQUS) and P28172. R. R. is supported by the FWF under project W1210 (CoQuS) and is a recipient of a DOC fellowship of the Austrian Academy of Sciences at the University of Vienna.

[1] T. Wilk, S. C. Webster, A. Kuhn, and G. Rempe, Science **317**, 488 (2007).
[2] A. Stute, B. Casabone, B. Brandstätter, K. Friebe, T. E. Northup, and R. Blatt, Nature Photon. **7**, 219 (2013).
[3] A. Kuzmich, W. P. Bowen, A. D. Boozer, A. Boca, C. W. Chou, L.-M. Duan, and H. J. Kimble, Nature **423**, 731 (2003).
[4] C. H. van der Wal, M. D. Eisaman, A. André, R. L. Walsworth, D. F. Phillips, A. S. Zibrov, and M. D. Lukin, Science **301**, 196 (2003).
[5] S. T. Yilmaz, P. Fallahi, and A. Imamoğlu, Phys. Rev. Lett. **105**, 033601 (2010).
[6] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Nature **526**, 682 (2015).
[7] H. J. Kimble, Nature **453**, 1023 (2008).
[8] O. Romero-Isart, Phys. Rev. A **84**, 52121 (2011).
[9] K. Stannigel, P. Rabl, A. S. Sørensen, P. Zoller, and M. D. Lukin, Phys. Rev. Lett. **105**, 220501 (2010).
[10] D. Chang, A. H. Safavi-Naeini, M. Hafezi, and O. Painter, New J. Phys. **13**, 023003 (2011).
[11] S. Barzanjeh, M. Abdi, G. J. Milburn, P. Tombesi, and D. Vitali, Phys. Rev. Lett. **109**, 130503 (2012).
[12] J. Bochmann, A. Vainsencher, D. D. Awschalom, and A. N. Cleland, Nature Phys. **9**, 712 (2013).
[13] M. Wallquist, K. Hammerer, P. Rabl, M. Lukin, and P. Zoller, Phys. Scr. **T137**, 014001 (2009).
[14] M. Poot and H. S. J. van der Zant, Phys. Rep. **511**, 273 (2012).
[15] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, Rev. Mod. Phys. **86**, 1391 (2014).
[16] 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).
[17] S. M. Meenehan, J. D. Cohen, G. S. MacCabe, F. Marsili, M. D. Shaw, and O. Painter, Phys. Rev. X **5**, 041002 (2015).
[18] 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).
[19] 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).
[20] T. Palomaki, J. Teufel, R. Simmonds, and K. Lehnert,

Science **342**, 710 (2013).
[21] 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).
[22] J.-M. Pirkkalainen, E. Damskägg, M. Brandt, F. Massel, and M. A. Sillanpää, Phys. Rev. Lett. **115**, 243601 (2015).
[23] F. Lecocq, J. B. Clark, R. W. Simmonds, J. Aumentado, and J. D. Teufel, Phys. Rev. X **5**, 041037 (2015).
[24] A. H. Safavi-Naeini and O. Painter, New J. Phys. **13**, 013017 (2011).
[25] A. H. Safavi-Naeini, T. P. M. Alegre, M. Winger, and O. Painter, Appl. Phys. Lett. **97**, 181106 (2010).
[26] L. M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Nature **414**, 413 (2001).
[27] C. Cabrillo, J. I. Cirac, P. Garcia-Fernandez, and P. Zoller, Phys. Rev. A **59**, 1025 (1998).
[28] K. C. Lee, M. R. Sprague, B. J. Sussman, J. Nunn, N. K. Langford, X.-M. Jin, T. Champion, P. Michelberger, K. F. Reim, D. England, D. Jaksch, and I. A. Walmsley, Science **334**, 1253 (2011).
[29] L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, Phys. Rev. Lett. **57**, 2520 (1986).
[30] J. D. Cohen, S. M. Meenehan, G. S. MacCabe, S. Gröblacher, A. H. Safavi-Naeini, F. Marsili, M. D. Shaw, and O. Painter, Nature **520**, 522 (2015).
[31] J. F. Clauser, Phys. Rev. D **9**, 853 (1974).
[32] M. Förtsch, J. U. Fürst, C. Wittmann, D. Strekalov, A. Aiello, M. V. Chekhova, C. Silberhorn, G. Leuchs, and C. Marquardt, Nature Commun. **4**, 1818 (2013).
[33] J. Chan, A. H. Safavi-Naeini, J. T. Hill, S. Meenehan, and O. Painter, App. Phys. Lett. **101**, 081115 (2012).
[34] S. M. Meenehan, J. D. Cohen, S. Gröblacher, J. T. Hill, A. H. Safavi-Naeini, M. Aspelmeyer, and O. Painter, Phys. Rev. A **90**, 011803 (2014).
[35] V. Fiore, Y. Yang, M. C. Kuzyk, R. Barbour, L. Tian, and H. Wang, Phys. Rev. Lett. **107**, 133601 (2011).
[36] E. Verhagen, S. Deléglise, S. Weis, A. Schliesser, and T. J. Kippenberg, Nature **482**, 63 (2012).
[37] R. W. Andrews, R. W. Peterson, T. P. Purdy, K. Cicak, R. W. Simmonds, C. A. Regal, and K. W. Lehnert, Nature Phys. **10**, 321 (2014).
[38] T. Bagci, A. Simonsen, S. Schmid, L. G. Villanueva, E. Zeuthen, J. Appel, J. M. Taylor, A. Sørensen, K. Usami, A. Schliesser, and E. S. Polzik, Nature **507**, 81 (2014).
[39] U. Akram, N. Kiesel, M. Aspelmeyer, and G. J. Milburn, New J. Phys. **12**, 083030 (2010).
[40] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, Supercond. Sci. Technol. **25**, 063001 (2012).
[41] S. G. Hofer, W. Wieczorek, M. Aspelmeyer, and K. Hammerer, Phys. Rev. A **84**, 52327 (2011).
[42] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).
[43] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, Rev. Mod. Phys. **83**, 33 (2011).
[44] P. Zoller, J. I. Cirac, L. Duan, and J. J. García-Ripoll, in *Quantum Entanglement and Information Processing, École d'été de Physique des Houches Session LXXIX*, edited by D. Estève, J.-M. Raimond, and J. Dalibard (2004).
[45] M. R. Vanner, M. Aspelmeyer, and M. S. Kim, Phys. Rev. Lett. **110**, 10504 (2013).
[46] C. Galland, N. Sangouard, N. Piro, N. Gisin, and T. J. Kippenberg, Phys. Rev. Lett. **112**, 143602 (2014).
[47] B. Zhao, Y.-A. Chen, X.-H. Bao, T. Strassel, C.-S. Chuu, X.-M. Jin, J. Schmiedmayer, Z.-S. Yuan, S. Chen, and J.-W. Pan, Nature Phys. **5**, 95 (2009).

METHODS

Device fabrication and characterization



Methods Figure 1: **Optomechanical device.** Shown is a scanning electron microscope image of a set of nanobeams, which are fabricated in silicon, as described in the text. Light is coupled into the central, adiabatically tapered waveguide through a lensed optical fiber (not shown) from the left of the image. The field then evanescently couples to each nanobeam (top and bottom). The two devices have slightly different resonance frequency, which makes it possible to distinguish them.

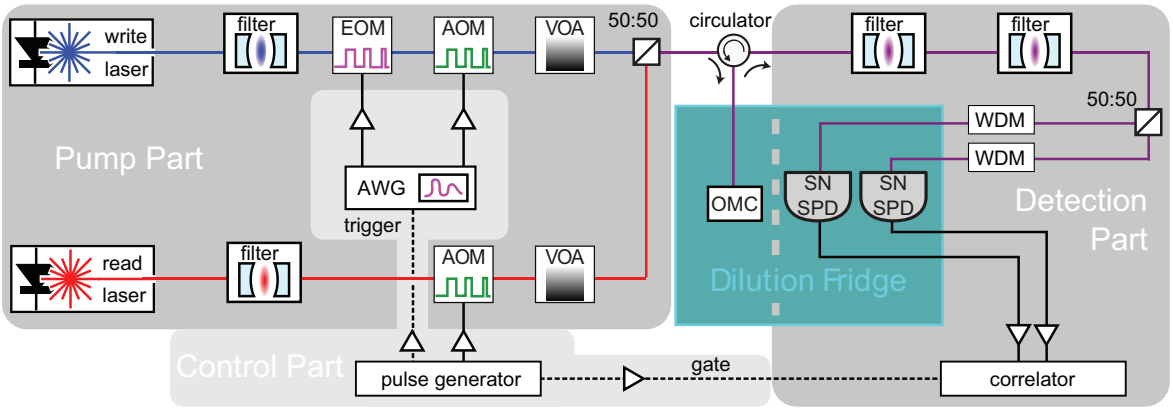
The optomechanical device used for this experiment (see Methods Figure 1) is fabricated from a silicon-on-insulator wafer, with a device layer thickness of 250 nm and 3 μm of buried oxide. The structures are patterned using an electron beam writer and are then transferred into the top silicon layer in an SF_6/O_2 atmosphere using a reactive ion-etcher. The devices are finally released and undercut using concentrated hydrofluoric acid. We design the nanobeams such that the fundamental mechanical breathing mode is at 5.3 GHz (cf. Figure 2a, main text) and the optical resonance is around 1550 nm (the measured wavelength for the device used here is 1556 nm) [33]. The optical and the mechanical modes are co-localized in the center of the beam, where we create a defect region of the photonic- and phononic-bandgap, allowing for an optomechanical coupling rate $g_0/2\pi = 825$ kHz. In order to minimize the thermalization time to the surrounding bath we opted, unlike previous designs, to not use any additional phononic shielding. As a consequence, the mechanical quality factors at base temperature are found to be around $1.1 \cdot 10^6$ (see section *Mechanical response to optical pulses*), compared to values above 10^7 with a phononic shield [17]. The laser pulses are coupled directly into a tapered waveguide through an optical fiber with a lensed tip [34], achieving efficencies of about 60%. The optical mode of the nanobeam is evanescently coupled to the waveguide, which is terminated with a periodic array of holes, acting as a mirror, allowing us to collect the light in reflection. For this experiment we chose a critically coupled device (internal losses equal external losses) with an optical linewidth $\kappa_c/2\pi$ of approximately 1.3 GHz. This places us well within the so-called resolved-sideband regime ($\omega_m > \kappa_c$).

Setup

In this section, we provide a detailed description of the experimental setup. It consists of a 'pump part', 'detection part', and the 'electronic control part' (cf. Methods Figure 2).

Pump part. We use two identical, tunable continuous-wave (CW) lasers (New Focus 6728) as our light sources. The lasers are detuned and stabilized to the blue and red side respectively of the devices cavity resonance (1556.21 nm). The detuning is set to be the mechanical frequency (5.307 GHz). The two lasers separately pass through voltage-controlled tunable optical filters (MicronOptics FFP-TF2, free spectral range ∼18 GHz, bandwidth ∼50 MHz) to suppress any potential background emissions dispersed in frequency space. In order to create short optical pulses we modulate the filtered CW fields using acousto-optic modulators (AOM; IntraAction) and an additional electro-optic amplitude modulator (EOM; EOSpace). We employ variable optical attenuators (VOA; Sercalo) on each path to control the pulse power. The pulses are combined on a variable optical coupler and then sent to the device in the dilution refrigerator (Vericold E21) via an optical circulator. At the device (OMC; optomechanical crystal), the optomechanical interaction with the blue (red) detuned pulses generates down-(up-) converted photons, whose frequency is on resonance with the device's optical cavity frequency. The scattered photons are reflected back from the OMC into the optical fiber and routed to the detection part through the output port of the circulator.

Detection part. Two voltage-controlled optical filters (MicronOptics FFP-TF2, specification as above) are installed in series at the beginning of the detection path. These filters are tuned on resonance with the OMC cavity frequency such that they only allow (anti-) Stokes scattered photons to be transmitted, while strong off-resonant pump photons are rejected (suppression of about 84 dB). After the filters, a 50:50 beam splitter divides the path. Each output is additionally filtered by broadband wavelength-division multiplexors (WDM), and fiber-coupled to two superconducting nanowire single photon detectors (SNSPD; PhotonSpot, detection efficiency ∼90%, dark count rate <10 Hz). The SNSPDs are mounted on the 1 K plate inside the dilution refrigerator. Upon receiving a photon the SNSPD generates a brief voltage spike, which is then electrically registered by a time-correlated single photon counting module (TCSPC; PicoQuant TimeHarp 260 NANO). The overall efficiency of detecting a photon leaving the OMC is ∼2.7% (see below).



Methods Figure 2: **Detailed experimental setup.** See the Methods text for a detailed description.

Control part. In order to generate programmable optical pulses and to detect photons synchronously, we use a digital pulse generator (DPG; Highland Technology P400) and an arbitrary waveform generator (AWG; Agilent Technologies 81180A). We first program the DPG to generate a TTL gate voltage signal for the AOM on the read (red) path and to trigger the TCSPC synchronously. The DPG additionally triggers the AWG, which then generates a TTL gate voltage for the AOM and a voltage pulse for the EOM on the write (blue) path.

Mechanical response to optical pulses

Over the past few years, several experiments have demonstrated precise control over optical and mechanical states through continuous optomechanical driving, including coherent state transfer [20, 35, 36] and microwave-to-optics conversion [12, 37, 38]. Due to the unavailability of the regime of single-photon strong cooperativity, strong drive fields have to be used in order to achieve the wanted coupling strength [39]. This leads to unwanted heating effects, in particular in the optical domain. Since the mechanism of optical absorption couples only indirectly to the mechanical mode of interest [34], using short optical pulses as nonstationary drive fields can substantially suppress the heating on short time scales – in particular at low temperatures [17].

Here, we probe the thermal response of the mechanical mode by pump-probe type measurements; we first send a short blue-detuned pump pulse onto the OMC cavity to intentionally heat the mode, and subsequently inject a red-detuned probe pulse to read out the modes phonon occupancy. By repeating the experiment with varying time delay between the pump and the probe pulses, we monitor time-dependent evolution of the modes phonon occupancy with a fixed initial impulse heating. The time delay δt is defined as the delay between the end of the pump (blue) pulse and the start of the probe (red) pulse detection window, as indicated in Figure 3a in the main text. In that way the probing is performed after the optical absorption of the pump photons is completed. In order to ensure that the mechanical mode fully re-thermalizes to the bath, we set the duty cycle of sending another blue pulse after the red pulse to be one millisecond. For an improved signal to noise ratio, the pulse energies of the blue (200 fJ) and red pulses (2 pJ) used here are substantially larger than in the cross-correlation measurements.

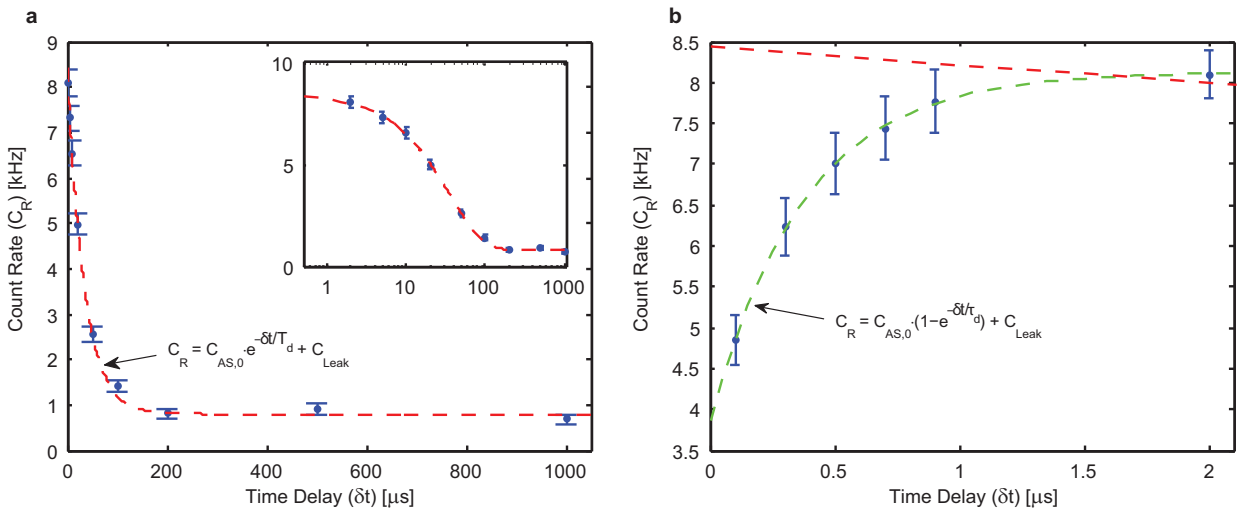
The effective mode temperature is inferred from the average count rate observed after sending the red pulse (C_R). C_R can be decomposed into three terms: (1) the rate proportional to the (on-resonance) anti-Stokes Raman scattered pump photons (C_AS), (2) the term corresponding to pump photons leaked through the optical filters (C_Leak), and (3) the additional anti-Stokes scattering term due to heating (ref. [17]) of the mode during the readout pulse (C_Heat). We minimize C_Heat by only taking into account the first 30 ns of the red pulse as 'logical' red pulse. C_Leak gives a constant offset to the signal. C_AS directly reflects the mode's effective temperature, as the anti-Stokes scattering rate is proportional to the average number of phonons (n_m) in the mode (see Figure 2a in the main text). To that end, we deduce the following equation

$$C_\mathrm{R}(\delta t) = C_\mathrm{AS}(\delta t) + C_\mathrm{Leak} = \alpha \cdot n_\mathrm{m}(\delta t) + C_\mathrm{Leak},$$

in which α is the constant of proportionality.

The long-term response of the mechanical mode to the initial blue pump pulse is shown in Methods Figure 3a. It exhibits an exponential decay with a time constant of $T_\mathrm{d} = 34.4\ \mu$s, which is interpreted as the mechanical damping time. The corresponding mechanical quality factor is then $Q = \omega_\mathrm{m} \cdot T_\mathrm{d} \approx 1.1 \cdot 10^6$.

In addition, we probe the short-term response of the mechanics within one microsecond after the blue pulse in more detail (Methods Figure 3b). We observe an increase of C_R with a time constant of 0.37 μs (fit to a simple



Methods Figure 3: **Pump-probe measurement of the mechanical response.** We send in a brief, intense blue detuned optical pulse (pump) and measure the mechanical response via red detuned optical probe pulse as a function of pump-probe time delay (δt). **a** Long-term mechanical response. The result fits well with a simple exponential decay (red dashed line; see the equation in the plot) with a damping time constant (T_d) of 34.4 μs. The inset shows the same data/fit with a logarithmic scale on the x axis. $C_{\mathrm{AS},0}$ is the extrapolated $C_\mathrm{AS}(\delta t = 0)$. **b** Short-term mechanical response. The data is fitted to a simple exponential curve (green dashed line; see the equation in the plot). The fitted time constant (τ_d) is 0.37 μs. The fit results of long-term response (red dashed line) projected to 0 μs delay is also shown for comparison. As the pump pulse had 5 times stronger energies than the write pulses in the correlation experiment, it is expected that the delayed heating occurs on longer time scale, due to the temperature dependence of the thermal conductivity of silicon [17]. Error bars in **a** and **b** represent a 68% confidence interval.

exponential curve). This data reveals slow turn-on dynamics of pulse-induced heating, as previously studied in reference [17]. This time constant is even shorter than the decay of the cross-correlations (see Figure 3c in the main text), which we attribute to the increased thermal conductivity of silicon at higher temperatures, caused by absorption of increased optical pump energies.

Characterization of the detection scheme

Detection Efficiency. We first calibrate the fiber-to-chip coupling efficiency (η_fc) by sending in light far off-resonant from the OMC cavity and then measure the reflected power (η_fc = 60.3% one-way). The device impedance ratio (η_c), i.e. the ratio of external coupling losses κ_ext to total losses κ_c, is measured through the depth and the linewidth of the optical resonance, which we find to be $\eta_\mathrm{c} = \kappa_\mathrm{ext}/\kappa_\mathrm{c} = 0.5$. The detection efficiency of scattered photons for each detector (η_i; i=1,2) consists of η_fc, η_c, the total losses of the remaining detection paths ($\eta_{\mathrm{path,i}}$), and the SNSPDs' quantum efficiencies ($\eta_{\mathrm{QE,i}}$). To measure η_i, pulses with calibrated energy are sent off-resonantly to the OMC (P_in), and the reflected photons transmitted through the optical filters are detected by the SNSPDs (P_out). $P_\mathrm{in}/P_\mathrm{out}$ corresponds to $\eta_\mathrm{fc}\cdot\eta_\mathrm{fc}\cdot\eta_{\mathrm{path,i}}\cdot\eta_{\mathrm{QE,i}}$, which we measure to be 0.013 for SNSPD1 and 0.019 for SNSPD2. Therefore, we deduce $\eta_\mathrm{i} = \eta_\mathrm{c}\cdot\eta_\mathrm{fc}\cdot\eta_{\mathrm{path,i}}\cdot\eta_{\mathrm{QE,i}}$ to be

$$\eta_1 = 1.1\%$$
$$\eta_2 = 1.6\%.$$

The detection efficiency of SNSPD1 (characterized quantum efficiency $\eta_{\mathrm{QE},1} = 65\%$) is lower than SNSPD2 (characterized quantum efficiency $\eta_{\mathrm{QE},2} = 90\%$), as we needed to reduce the bias current to prevent the detector from latching [40]. This latching is probably caused by a nearby heater of the dilution refrigerator. It also results in a slow drift in the quantum efficiency of SNSPD1. We note that the deduced $\eta_{\mathrm{path,i}}$ come from the various optical elements in the beam path of the detection part and are in good agreement with their specified insertion losses.

Scattering rates and optomechanical coupling rate. With the total detection efficiency of resonantly generated cavity photons, we can estimate the pair generation probability per write pulse (optical energy $E_\mathrm{opt}\sim$40 fJ) to be $p\sim$3.0%, including the effects of a finite starting temperature and leaked pump photons. The latter is calibrated by sending detuned optical pulses ($E_\mathrm{opt}\sim$40 fJ, $\omega_\mathrm{L} = \omega_\mathrm{c} - \omega_\mathrm{m} - 2\pi\cdot 200$ MHz) to the device. The generated optomechanical sidebands are now blocked by the filters and only leaked pump photons are detected. We measure a suppression of the pump pulse by 84 dB compared to an

on- resonance transmission. Thus, approximately 1 out of 25 photons detected during the write pulse is a leaked pump photon. Knowing the scattering rate and the energy of the detuned pump pulse, we can determine the single-photon coupling rate of our OMC to be

$$g_0 = \frac{\partial \omega_\mathrm{c}}{\partial x}\sqrt{\frac{\hbar}{2m\omega_\mathrm{m}}} = 2\pi \cdot 825 \text{ kHz}.$$

With this coupling rate, we can estimate the state-transfer efficiency of the red-detuned optical readout pulse of $E_\mathrm{opt} = 50$ fJ to be $\varepsilon_\mathrm{R} = 3.7\%$, where

$$\hat{a}_\mathrm{opt,out} \approx \sqrt{1-\varepsilon_\mathrm{R}}\hat{a}_\mathrm{opt,in} + e^{i\phi}\sqrt{\varepsilon_\mathrm{R}}\hat{a}_\mathrm{mech,in}.$$

Here, $\hat{a}_\mathrm{opt,in(out)}$ are the annihilation operators of the temporal optical input (output) mode of the cavity resonance, $\hat{a}_\mathrm{mech,in}$ the mechanical mode before the interaction, and ϕ an arbitrary but fixed phase between the inputs [41].

Definition and properties of the second order correlation function

We define the normalized second-order correlation function for two, not necessarily different modes α and β, and the respective annihilation operators $\hat{a}_\alpha$ and $\hat{a}_\beta$, to be (references [3, 42] and references therein)

$$g^{(2)}_{\alpha\beta} = \frac{\langle : \hat{a}^\dagger_\alpha \hat{a}_\alpha \hat{a}^\dagger_\beta \hat{a}_\beta : \rangle}{\langle \hat{a}^\dagger_\alpha \hat{a}_\alpha \rangle \langle \hat{a}^\dagger_\beta \hat{a}_\beta \rangle},$$

where $: \hat{O} :$ denotes normal ordering of the operators. For the autocorrelation of the optical field (photons scattered by the write pulse), $\alpha = \beta = \mathrm{o}$, for the mechanical field $\alpha = \beta = \mathrm{m}$ and for the cross-correlation $\alpha = \mathrm{o}$, $\beta = \mathrm{m}$. By introducing effective modes γ, δ it can be seen that this correlation-function is independent of losses in the detection. Assuming the loss angles φ_α and φ_β for detection of modes α, β, we define the annihilation operators of the effectively detected modes γ, δ

$$\hat{a}_{\gamma/\delta} = \cos(\varphi_{\alpha/\beta})\hat{a}_{\alpha/\beta} + \sin(\varphi_{\alpha/\beta})\hat{l}_{\alpha/\beta}$$

by coupling the original modes α, β to modes l_α and l_β represented by the annihilation operator $\hat{l}_{\alpha/\beta}$. As the detected modes γ, δ have frequencies in the optical domain, we can assume the in-coupled modes l_α, l_β to be in their respective ground state. Tracing over l_α, l_β, we find that

$$g^{(2)}_{\gamma\delta} = g^{(2)}_{\alpha\beta}$$

i.e. the second order correlation function is independent of losses or, in the case of the mechanical mode, of "ineffective" partial state-transfer to the cavity mode. Thus, e.g. $g^{(2)}_\mathrm{mm}$ is equivalent to the autocorrelation of the photons scattered by the read pulse.

For autocorrelation measurements, we use a Hanbury Brown and Twiss setup, by splitting the mode on a symmetric beamsplitter and sending it to a pair of detectors. We define the modes detected by the individual detectors d_1, d_2 with their annihilation operators

$$\hat{a}_{1/2} = \cos(\theta)\hat{a}_\alpha \pm \sin(\theta)\hat{l}_\mathrm{d}$$

with the splitting angle θ of the beam splitter and the annihilation operator $\hat{l}_\mathrm{d}$ of the second input of the beamsplitter. The input state can as before be approximated to be in its vacuum state. We find that the autocorrelation of mode α equals the cross-correlation of the two detectors:

$$g^{(2)}_{\alpha\alpha} = g^{(2)}_{12}.$$

For a definition in terms of probabilities, see below.

Statistical Analysis

Due to the low detection probability, the uncertainty in the estimation of the second-order correlation functions is completely dominated by the estimation of the coincidence rate $\langle : \hat{a}^\dagger_\alpha \hat{a}_\alpha \hat{a}^\dagger_\beta \hat{a}_\beta : \rangle$ of the two modes α, β. As the absolute number of coincidences is low in some measurements, Gaussian statistics cannot be used for estimating uncertainties. Instead, we use the likelihood function based on the binomial distribution for estimating the probability p of the underlying process, i.e. to obtain N counts in T tries

$$L(p,N,T) = \frac{1}{K}p^N(1-p)^{T-N}.$$

The normalization K is chosen such that $\int_0^1 L(p,N,T)dp = 1$. The upper and lower uncertainty σ_+ and σ_- are chosen numerically, such that they cover a 68% confidence interval around the maximum likelihood estimator $p_\mathrm{ML} = N/T$, i.e. $\int_0^{p_\mathrm{ML}-\sigma_-} L(p,N,T)dp = 0.16$, $\int_{p_\mathrm{ML}+\sigma_+}^1 L(p,N,T)dp = 0.16$.

For the classical bound of the cross-correlation, $g^{(2)}_\mathrm{cb} = \sqrt{g^{(2)}_\mathrm{mm} \cdot g^{(2)}_\mathrm{oo}}$, the likelihood functions of the individual autocorrelations are convoluted. Due to their asymmetry, the maximum likelihood estimator of the classical bound is slightly lower than when using the individual maximum likelihood estimators $g^{(2)}_\mathrm{cb,ML} \leq \sqrt{g^{(2)}_\mathrm{mm,ML} \cdot g^{(2)}_\mathrm{oo,ML}}$.

As estimators for the cross-correlation function, the probabilities P of a coincidence- or single detection event during the read (R) and write pulse (W) were used, with $g^{(2)}_\mathrm{om} = P(W \cap R)/P(R)P(W)$. This is valid for low event probabilities $P \ll 1$. Autocorrelations were estimated by probabilities of coincidence- and single detection events on individual SNSPDs (1,2), $g^{(2)}_\mathrm{yy} = P(X_1 \cap X_2)/P(X_1)P(X_2)$ during the evaluation periods of the

δt	0.1 µs	0.6 µs	1.1 µs	2.1 µs	3.1 µs	total	0.1 µs*
N(R∩W)	202	153	144	113	127		34
$N(R_1 \cap R_2)$	13	24	23	36	37		0
$N(W_1 \cap W_2)$	16	15	17	12	20	80	16
$N(R_1)$	13,172	16,523	18,751	18,316	23,629		966
$N(R_2)$	17,490	22,278	25,394	26,122	31,892		1145
$N(W_1)$	12,471	12,061	13,051	11,870	15,032	64,485	12,471
$N(W_2)$	19,409	18,616	20,176	18,329	23,601	100,131	19,409
T	38,806,017	38,829,923	39,958,216	35,712,159	47,964,927	201,271,242	38,806,017

Methods Table 1: **Counts of the cross-correlation measurements.** The row label 'N(event)' represents the number of counts for a certain event, e.g. detection of a photon during the measurement window of read pulse on detector 1/2 ($R_{1/2}$), or the coincidence of a detection event of a subsequent write and read pulse on either detector 1 or 2, $R \cap W = (R_1 \cup R_2) \cap (W_1 \cup W_2)$. T denotes the total number of pulse pairs sent to the optomechanical device. For the calculation of the autocorrelation function of the read pulse, only counts from the delay setting δt are used, as the delayed heating of the blue pulse (cf. Figure 3) influences the mechanical state. For the autocorrelation function of the write pulse, counts from all delay settings are summed, as the mechanical state is reinitialized by cryogenic cooling before measurement, independent of the delay δt. The numbers for this are summarized in the column labeled 'total'. The highest reported cross-correlation value was obtained by reducing the measurement window of the read pulse from 55 ns to 30 ns, with a delay of $\delta t = 100$ ns between the write and the read pulse. The counts for this evaluation window are presented in the column marked with *. The underlying dataset is the same as for the standard evaluation period of 55 ns, i.e. the first column.

write (y = o, $X = W$) and read (y = m, $X = R$) pulse, respectively. The statistics of the cross-correlation measurements are summarized in Methods Table 1.

For the read pulse, only the first 55 ns of the pulse were evaluated (cf. Figure 3, main text). A further reduction of the evaluation period to $t_{\mathrm{eval}} = 30$ ns (R^*) has the advantage of reducing the influence of optical absorption of pump photons from the read pulse, while still obtaining solid statistics for the cross-correlation, $g^{(2)}_{\mathrm{om}}(\Delta n = 0, \delta t = 100\ \mathrm{ns}, t_{\mathrm{eval}} = 30\ \mathrm{ns}) = 19.6 - 2.8 + 3.9$. However, this reduction also results in a much lower state transfer efficiency $\varepsilon^*_{\mathrm{R}} \approx 0.1\%$ (compared to $\varepsilon_{\mathrm{R}} = 3.7\%$ above; see Methods section *Characterization of the detection scheme*). As a consequence we cannot obtain independent statistics on the autocorrelation function of the read pulse, as the number of pulse sequences is too low to observe coincidences during the reduced read pulse $N(R^*_1 \cap R^*_2) = 0$. Thus, no independent classical bound g^*_{cb} can be obtained for this case. As the measurement is identical to the one with longer evaluation window in the first column of Methods Table 1, it is reasonable to assume the same autocorrelation of the mechanical state and thus the same classical limit.

We note that slight differences in the polarization of the two input lasers and the optimal axis of the SNSPDs can lead to different detection rates of leaked pump photons between the read and the write pulse. While this does not influence the cross-correlation measurement, it is important to use the same laser source for the sideband asymmetry measurements.

Interpretation of the cross-correlation measurements

Classical bound. The classical bound $g_{\mathrm{cb,ML}}$ is found to be slightly above 2, the value expected for a thermal state of the mechanical system (cf. Figure 3c, main text). Although this increase of the autocorrelation is not significant in our measurements, a behavior like this would be expected in the case of mixed thermal states of different temperatures, caused e.g. by fluctuations in the absorbed power. Effects that usually decrease the measured autocorrelation function of a thermal state, such as dark counts of the detectors and instantaneous heating by the read pulse, do not play a major role in our experiment due to the choice of pulse parameters. In conclusion, the classical bound in the present experiment is slightly elevated compared to cross-correlation experiments in atomic physics or non-linear optics, where the classical bound is usually assumed to be [43] below 2.

Decay of cross-correlations due to delayed heating. The cross-correlation can be interpreted as

$$g^{(2)}_{\mathrm{om}} \sim \frac{\langle n_{\mathrm{m}} \rangle_{\mathrm{h}}}{\langle n_{\mathrm{m}} \rangle}$$

where $\langle n_{\mathrm{m}} \rangle_{\mathrm{h}}$ is the average number of mechanical excitations in the state heralded on a detection event of the write pulse (indicating the presence of an anti-Stokes scattered photon), and $\langle n_{\mathrm{m}} \rangle$ the average number of unheralded events (essentially probing the thermal excitation of the system when $p \ll 1$). In case of a delayed heating, the thermal occupation of the system is a function of the

delay δt after the write pulse $\langle n_{\mathrm{m,th}}\rangle = \langle n_{\mathrm{m,th}}\rangle(\delta t)$. Assuming our cross-correlation is dominated by the thermal occupation, we obtain for $\delta t \ll T_{\mathrm{d}}$,

$$g^{(2)}_{\mathrm{om}}(\delta t) \sim \frac{1+\langle n_{\mathrm{m,th}}\rangle(\delta t)}{\langle n_{\mathrm{m,th}}\rangle(\delta t)},$$

which clearly decays in the case of substantial delayed heating as observed here (cf. Methods Figure 3). Theoretical models of the complex thermodynamic non-equilibrium processes contain many device dependent parameters [17], which will be subject of further studies.

Estimation of the heralded single-phonon fidelity. In general, the toolbox of quantum optics provides unique means for quantum state control of various systems [44]. As an example we discuss the application of single-photon detection for the heralded generation of single-phonon Fock states of our mechanical resonator [41, 45, 46]. To estimate the fidelity of the single-phonon state directly after heralding on the detection of a resonant photon generated by the write pulse, we need to know all contributions to the diagonal of the density matrix, which are not a single phonon. These contributions can either be higher excitations by thermal contribution, multi-pair generation, or vacuum states by false positive heralding events. Higher excitations can be estimated by the auto-correlation function of the heralded state, which is related to the cross-correlations function [47]. As the target is to estimate the state immediately after heralding it, we reduce the evaluation window of the read pulse as much as possible, while maintaining reasonable statistics on the cross-correlation (cf. Methods Table 1). From measured $g^{(2)}_{\mathrm{om}}(\Delta n = 0, \delta t = 100\ \mathrm{ns}, t_{\mathrm{eval}} = 30\ \mathrm{ns}) = 19.6 - 2.8 + 3.9$, we infer an auto-correlation function for the heralded mechanical state of $g^{(2)}_{\mathrm{mm,heralded}} \approx 0.22 \pm 0.04$, which approximately relates to the ratio of probabilities of higher excitations $p_{n>1}$ to single phonon excitations $p_{n=1}$, $2 \cdot p_{n>1} \approx g^{(2)}_{\mathrm{mm,heralded}} \cdot p^2_{n=1}$. In the meantime, the main contribution for non-zero $p_{n=0}$ (i.e. the probability of the heralded mechanical state being the ground state) is false positive heralding events, i.e. dark counts and leaked pump photons. With the known ratio of true positive to false positive heralding events (cf. section *Characterization of the detection scheme*), we obtain an estimate of $p_{n=0} \sim 1/25$. With these conservative estimates, we obtain a heralded Fock-state fidelity of $p_{n=1} = 87.7 \pm 1.2\%$ on the basis of the standard system Hamiltonian of the optomechanical device [41].