

Bidirectional microwave-optical transduction based on integration of high-overtone bulk acoustic resonators and photonic circuits

Terence Blésin,^{1,2} Wil Kao,^{1,2} Anat Siddharth,^{1,2} Rui N. Wang,^{1,2} Alaina Attanasio,³ Hao Tian,³ Sunil A. Bhavé,^{3,*} and Tobias J. Kippenberg^{1,2,†}

¹*Institute of Physics, Swiss Federal Institute of Technology Lausanne (EPFL), CH-1015 Lausanne, Switzerland*

²*Center of Quantum Science and Engineering (EPFL), CH-1015 Lausanne, Switzerland*

³*OxideMEMS lab, Purdue University, West Lafayette, IN, USA*

Appendix A: Theory

1. Coupled mode theory for optical modes hybridization

Using a pair of optical resonances increases the intracavity photon number for a given input optical power, resulting in an enhanced optomechanical interaction. Such an optical doublet is implemented here by evanescently coupling two micro-rings, forming a photonic molecule. Consider two bare optical modes with resonance frequencies ω_l and ω_r that correspond to the left and right micro-rings, respectively. The two modes couple with a rate J , which yields the Hamiltonian [1]

$$\hat{\mathcal{H}} = \hbar\omega_l\hat{a}_l^\dagger\hat{a}_l + \hbar\omega_r\hat{a}_r^\dagger\hat{a}_r - \hbar J \left(\hat{a}_l\hat{a}_r^\dagger + \hat{a}_l^\dagger\hat{a}_r \right) \quad (\text{A1})$$

and the equations of motion

$$\begin{aligned} \frac{d}{dt}\hat{a}_l(t) &= \left(-i\omega_l - \frac{\kappa_l}{2} \right) \hat{a}_l(t) + iJ\hat{a}_r(t), \\ \frac{d}{dt}\hat{a}_r(t) &= \left(-i\omega_r - \frac{\kappa_r}{2} \right) \hat{a}_r(t) + iJ\hat{a}_l(t). \end{aligned} \quad (\text{A2})$$

With Laplace transform, they can be rewritten in frequency domain as

$$\begin{pmatrix} \chi_l^{-1}[s] & -iJ \\ -iJ & \chi_r^{-1}[s] \end{pmatrix} \begin{pmatrix} \hat{a}_l[s] \\ \hat{a}_r[s] \end{pmatrix} = \begin{pmatrix} \hat{a}_l(0) \\ \hat{a}_r(0) \end{pmatrix}, \quad (\text{A3})$$

where $\chi_o[s] = (s + i\omega_o + \kappa_o/2)^{-1}$, $o = l, r$ denote the susceptibilities. Diagonalizing the matrix yields the eigenvalues

$$\begin{aligned} \lambda_{\pm} &= \frac{\chi_l^{-1} + \chi_r^{-1}}{2} \pm \frac{1}{2} \sqrt{(\chi_l^{-1} - \chi_r^{-1})^2 - 4J^2} \\ &= (s + i\bar{\omega} + \bar{\kappa}/2) \pm \frac{1}{2} \sqrt{\frac{\mu^2}{4} - \delta^2 + i\mu\delta - 4J^2}. \end{aligned} \quad (\text{A4})$$

Here we define the average optical frequency $\bar{\omega} = (\omega_l + \omega_r)/2$, average linewidth $\bar{\kappa} = (\kappa_l + \kappa_r)/2$, relative detuning of the micro-rings $\delta = \omega_l - \omega_r$, and linewidth difference $\mu = \kappa_l - \kappa_r$. The difference in resonance frequency and linewidth of the two supermodes can be computed as

$$\Delta\omega = \frac{1}{\sqrt{2}} \sqrt{-\frac{\mu^2}{4} + \delta^2 + 4J^2 + \sqrt{\left[\left(2J - \frac{\mu}{2} \right)^2 + \delta^2 \right] \left[\left(2J + \frac{\mu}{2} \right)^2 + \delta^2 \right]}}, \quad (\text{A5})$$

$$\Delta\kappa = \sqrt{2} \sqrt{\frac{\mu^2}{4} - \delta^2 - 4J^2 + \sqrt{\left[\left(2J - \frac{\mu}{2} \right)^2 + \delta^2 \right] \left[\left(2J + \frac{\mu}{2} \right)^2 + \delta^2 \right]}}, \quad (\text{A6})$$

* bhavé@purdue.edu

† tobias.kippenberg@epfl.ch

where $\lambda_+ - \lambda_- = \Delta\kappa/2 + i\Delta\omega$. In the strong coupling regime where $2J \gg \mu/2$, we simply have $\Delta\omega \approx \sqrt{4J^2 + \delta^2}$ and $\Delta\kappa \approx \mu\sqrt{1 - 4J^2/(4J^2 + \delta^2)}$. The two supermodes are thus centered at $\bar{\omega}$ with a splitting of $\Delta\omega$, resulting in susceptibilities

$$\chi_{\pm}[s] = \left[s + i \left(\bar{\omega} \pm \frac{\Delta\omega}{2} \right) + \frac{1}{2} \left(\bar{\kappa} \pm \frac{\Delta\kappa}{2} \right) \right]^{-1}. \quad (\text{A7})$$

Solving for the eigenvectors gives the participation ratios of the bare cavity modes in the supermodes

$$\begin{aligned} u_{\pm} &= \frac{2J}{\sqrt{4J^2 + |\mu \mp \Delta\kappa/2 + i(\delta \mp \Delta\omega)|^2}} \exp(i\phi_{\pm}), \\ v_{\pm} &= \frac{|\mu \mp \Delta\kappa/2 + i(\delta \mp \Delta\omega)|}{\sqrt{4J^2 + |\mu \mp \Delta\kappa/2 + i(\delta \mp \Delta\omega)|^2}} \exp \left[i\phi_{\pm} + i \arg \frac{\mu \mp \Delta\kappa/2 + i(\delta \mp \Delta\omega)}{2iJ} \right], \end{aligned} \quad (\text{A8})$$

where $|u_{\pm}|^2 + |v_{\pm}|^2 = 1$ and $\phi_{\pm}$ is a global phase factor. In particular, the supermode $\hat{a}_- = u_- \hat{a}_l + v_- \hat{a}_r$ has a resonance frequency $\omega_- = \bar{\omega} - \Delta\omega/2$ and linewidth $\kappa_- = \bar{\kappa} - \Delta\kappa/2$. It is referred to as the symmetric supermode. The rationale behind the nomenclature can be seen by considering the limit of zero detuning $\delta = 0$, where we simply have $\hat{a}_- = (\hat{a}_l + \hat{a}_r)/\sqrt{2}$ with $\omega_- = \bar{\omega} - J$ and $\kappa_- = \bar{\kappa}$. On the other hand, the antisymmetric supermode $\hat{a}_+ = u_+ \hat{a}_l + v_+ \hat{a}_r$ has a resonance frequency $\omega_+ = \bar{\omega} + \Delta\omega/2$ and linewidth $\kappa_+ = \bar{\kappa} + \Delta\kappa/2$. When $\delta = 0$, this reduces to $\hat{a}_+ = (\hat{a}_l - \hat{a}_r)/\sqrt{2}$ with $\omega_+ = \bar{\omega} + J$ and $\kappa_+ = \bar{\kappa}$.

2. Transducer Hamiltonian and input-output relations

In terms of bare micro-ring modes, the transducer Hamiltonian is given by

$$\hat{\mathcal{H}} = \hbar\omega_m \hat{b}^\dagger \hat{b} + \hbar\omega_l \hat{a}_l^\dagger \hat{a}_l + \hbar\omega_r \hat{a}_r^\dagger \hat{a}_r - \hbar J (\hat{a}_l^\dagger \hat{a}_r + \hat{a}_r^\dagger \hat{a}_l) - \hbar g_0 \hat{a}_l^\dagger \hat{a}_l (\hat{b} + \hat{b}^\dagger) + \hat{\mathcal{H}}_{\text{drive}}. \quad (\text{A9})$$

Note that the piezo-coupled acoustic mode interacts with only one of the micro-rings. To compute quantities directly related to the experiments, we rewrite the Hamiltonian in terms of the supermodes (Appendix A 1) as

$$\hat{\mathcal{H}} = \hbar\omega_m \hat{b}^\dagger \hat{b} + \hbar\omega_- \hat{a}_-^\dagger \hat{a}_- + \hbar\omega_+ \hat{a}_+^\dagger \hat{a}_+ - \hbar g_0 (|x|^2 \hat{a}_-^\dagger \hat{a}_- + |y|^2 \hat{a}_+^\dagger \hat{a}_+ + xy^* \hat{a}_+^\dagger \hat{a}_- + x^* y \hat{a}_-^\dagger \hat{a}_+) (\hat{b} + \hat{b}^\dagger) + \hat{\mathcal{H}}_{\text{drive}} \quad (\text{A10})$$

such that $\hat{a}_l = x\hat{a}_- + y\hat{a}_+$. The driving term is given by

$$\begin{aligned} \hat{\mathcal{H}}_{\text{drive}} &= i\hbar\sqrt{\kappa_{\text{ex},-}} \left(\hat{a}_{\text{in}} \hat{a}_-^\dagger e^{-i\omega_L t} - \hat{a}_{\text{in}}^\dagger \hat{a}_- e^{i\omega_L t} \right) + i\hbar\sqrt{\kappa_{\text{ex},+}} \left(\hat{a}_{\text{in}} \hat{a}_+^\dagger e^{-i\omega_L t} - \hat{a}_{\text{in}}^\dagger \hat{a}_+ e^{i\omega_L t} \right) \\ &\quad + i\hbar\sqrt{\kappa_{\text{ex},m}} \left(\hat{c}_{\text{in}} \hat{b}^\dagger e^{i\omega_m t} - \hat{c}_{\text{in}}^\dagger \hat{b} e^{-i\omega_m t} \right) \end{aligned} \quad (\text{A11})$$

for input optical and microwave fields $\hat{a}_{\text{in}}$ and $\hat{c}_{\text{in}}$. By going into the frames rotating at the laser drive frequency ω_L , we obtain

$$\begin{aligned} \hat{\mathcal{H}} &= \hbar\omega_m \hat{b}^\dagger \hat{b} - \hbar\Delta_- \hat{a}_-^\dagger \hat{a}_- - \hbar\Delta_+ \hat{a}_+^\dagger \hat{a}_+ \\ &\quad - \hbar g_0 (|x|^2 \hat{a}_-^\dagger \hat{a}_- + |y|^2 \hat{a}_+^\dagger \hat{a}_+) (\hat{b} + \hat{b}^\dagger) - \hbar g_0 (xy^* \hat{a}_+^\dagger \hat{a}_- + x^* y \hat{a}_-^\dagger \hat{a}_+) (\hat{b} + \hat{b}^\dagger) \\ &\quad + i\hbar\sqrt{\kappa_{\text{ex},-}} (\hat{a}_{\text{in}} \hat{a}_-^\dagger - \hat{a}_{\text{in}}^\dagger \hat{a}_-) + i\hbar\sqrt{\kappa_{\text{ex},+}} (\hat{a}_{\text{in}} \hat{a}_+^\dagger - \hat{a}_{\text{in}}^\dagger \hat{a}_+) + i\hbar\sqrt{\kappa_{\text{ex},m}} (\hat{c}_{\text{in}} \hat{b}^\dagger e^{i\omega_m t} - \hat{c}_{\text{in}}^\dagger \hat{b} e^{-i\omega_m t}), \end{aligned} \quad (\text{A12})$$

where $\Delta_{\pm} = \omega_L - \omega_{\pm}$. Considering small deviations of the supermodes around their steady-state amplitudes $\alpha_{\pm} \approx [\sqrt{\kappa_{\text{ex},\pm}}/(\kappa_{\pm}/2 - i\Delta_{\pm})] \sqrt{\hat{n}_{\text{in}}^{\pm}} e^{i\phi_{\pm}}$ and letting the input field operators denote just the Langevin force, the Hamiltonian

$$\begin{aligned}
\hat{\mathcal{H}} = & \hbar\omega_m \hat{b}^{\dagger} \hat{b} - \hbar\Delta_- \hat{a}_{-}^{\dagger} \hat{a}_{-} - \hbar\Delta_+ \hat{a}_{+}^{\dagger} \hat{a}_{+} \\
& - \hbar g_0 |x|^2 \left(|\alpha_-|^2 + \alpha_-^* \hat{a}_{-} + \hat{a}_{-}^{\dagger} \alpha_- + \hat{a}_{-}^{\dagger} \hat{a}_{-} \right) (\hat{b} + \hat{b}^{\dagger}) \\
& - \hbar g_0 |y|^2 \left(|\alpha_+|^2 + \alpha_+^* \hat{a}_{+} + \hat{a}_{+}^{\dagger} \alpha_+ + \hat{a}_{+}^{\dagger} \hat{a}_{+} \right) (\hat{b} + \hat{b}^{\dagger}) \\
& - \hbar g_0 x y^* \left(\alpha_+^* \alpha_- + \alpha_+^* \hat{a}_{-} + \hat{a}_{+}^{\dagger} \alpha_- + \hat{a}_{+}^{\dagger} \hat{a}_{-} \right) (\hat{b} + \hat{b}^{\dagger}) \\
& - \hbar g_0 x^* y \left(\alpha_-^* \alpha_+ + \alpha_-^* \hat{a}_{+} + \hat{a}_{-}^{\dagger} \alpha_+ + \hat{a}_{-}^{\dagger} \hat{a}_{+} \right) (\hat{b} + \hat{b}^{\dagger}) \\
& + i\hbar\sqrt{\kappa_{\text{ex},-}} \left(\hat{a}_{\text{in}} \hat{a}_{-}^{\dagger} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{-} \right) + i\hbar\sqrt{\kappa_{\text{ex},+}} \left(\hat{a}_{\text{in}} \hat{a}_{+}^{\dagger} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{+} \right) \\
& + i\hbar\sqrt{\kappa_{\text{ex},m}} \left(\hat{c}_{\text{in}} \hat{b}^{\dagger} e^{i\omega_m t} - \hat{c}_{\text{in}}^{\dagger} \hat{b} e^{-i\omega_m t} \right)
\end{aligned} \tag{A13}$$

can be further simplified. In the large intracavity photon number limit, the $\mathcal{O}(\hat{a}^2)$ terms are omitted as their contribution to the system dynamics is hidden by the $\mathcal{O}(\hat{a})$ terms. The $\mathcal{O}(\alpha^2)$ constant terms are removed by shifting the position origin by the steady-state displacement

$$x_{\text{ss}} = \frac{2g_0}{\omega_m} (|x|^2 |\alpha_-|^2 + |y|^2 |\alpha_+|^2 + x y^* \alpha_+^* \alpha_- + x^* y \alpha_-^* \alpha_+) x_{\text{ZPF}} \tag{A14}$$

using the translation operator $\hat{T}_{\hat{x}}(x_{\text{ss}}) = \exp(x_{\text{ss}} \hat{p}/(i\hbar))$, $\hat{p}$ being the momentum operator associated with $\hat{x} = x_{\text{ZPF}} (\hat{b} + \hat{b}^{\dagger})$. The resulting effective Hamiltonian is then given by

$$\begin{aligned}
\hat{\mathcal{H}} = & \hbar\omega_m \hat{b}^{\dagger} \hat{b} - \hbar\Delta_- \hat{a}_{-}^{\dagger} \hat{a}_{-} - \hbar\Delta_+ \hat{a}_{+}^{\dagger} \hat{a}_{+} \\
& - \hbar g_0 (|x|^2 \alpha_-^* + x y^* \alpha_+^*) \hat{a}_{-} (\hat{b} + \hat{b}^{\dagger}) - \hbar g_0 (|x|^2 \alpha_- + x^* y \alpha_+) \hat{a}_{-}^{\dagger} (\hat{b} + \hat{b}^{\dagger}) \\
& - \hbar g_0 (|y|^2 \alpha_+^* + x^* y \alpha_-^*) \hat{a}_{+} (\hat{b} + \hat{b}^{\dagger}) - \hbar g_0 (|y|^2 \alpha_+ + x y^* \alpha_-) \hat{a}_{+}^{\dagger} (\hat{b} + \hat{b}^{\dagger}) \\
& + i\hbar\sqrt{\kappa_{\text{ex},-}} \left(\hat{a}_{\text{in}} \hat{a}_{-}^{\dagger} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{-} \right) + i\hbar\sqrt{\kappa_{\text{ex},+}} \left(\hat{a}_{\text{in}} \hat{a}_{+}^{\dagger} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{+} \right) + i\hbar\sqrt{\kappa_{\text{ex},m}} \left(\hat{c}_{\text{in}} \hat{b}^{\dagger} e^{i\omega_m t} - \hat{c}_{\text{in}}^{\dagger} \hat{b} e^{-i\omega_m t} \right).
\end{aligned} \tag{A15}$$

Finally, by defining effective optomechanical coupling rates

$$\begin{aligned}
g_- &= g_0 (|x|^2 \alpha_- + x^* y \alpha_+), \\
g_+ &= g_0 (|y|^2 \alpha_+ + x y^* \alpha_-),
\end{aligned} \tag{A16}$$

and going into a rotating frame defined by the operator

$$\hat{U} = \exp\left(\frac{1}{i\hbar} \hbar\Delta_- \hat{a}_{-}^{\dagger} \hat{a}_{-} t\right) \exp\left(\frac{1}{i\hbar} \hbar\Delta_+ \hat{a}_{+}^{\dagger} \hat{a}_{+} t\right) \exp\left(\frac{-1}{i\hbar} \hbar\omega_m \hat{b}^{\dagger} \hat{b} t\right), \tag{A17}$$

we have

$$\begin{aligned}
\hat{\mathcal{H}} = & -\hbar \left(g_-^* \hat{a}_{-} e^{i\Delta_- t} + g_- \hat{a}_{-}^{\dagger} e^{-i\Delta_- t} \right) \left(\hat{b} e^{-i\omega_m t} + \hat{b}^{\dagger} e^{i\omega_m t} \right) \\
& -\hbar \left(g_+^* \hat{a}_{+} e^{i\Delta_+ t} + g_+ \hat{a}_{+}^{\dagger} e^{-i\Delta_+ t} \right) \left(\hat{b} e^{-i\omega_m t} + \hat{b}^{\dagger} e^{i\omega_m t} \right) \\
& + i\hbar\sqrt{\kappa_{\text{ex},-}} \left(\hat{a}_{\text{in}} \hat{a}_{-}^{\dagger} e^{-i\Delta_- t} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{-} e^{i\Delta_- t} \right) + i\hbar\sqrt{\kappa_{\text{ex},+}} \left(\hat{a}_{\text{in}} \hat{a}_{+}^{\dagger} e^{-i\Delta_+ t} - \hat{a}_{\text{in}}^{\dagger} \hat{a}_{+} e^{i\Delta_+ t} \right) \\
& + i\hbar\sqrt{\kappa_{\text{ex},m}} \left(\hat{c}_{\text{in}} \hat{b}^{\dagger} - \hat{c}_{\text{in}}^{\dagger} \hat{b} \right)
\end{aligned} \tag{A18}$$

in the interaction picture. The time evolution of the bosonic modes can be computed using the Heisenberg-Langevin equations given by

$$\begin{aligned}
\frac{d}{dt}\hat{a}_-(t) &= \frac{-\kappa_-}{2}\hat{a}_-(t) + ig_- \left[\hat{b}(t)e^{-i(\omega_m+\Delta_-)t} + \hat{b}^\dagger(t)e^{i(\omega_m-\Delta_-)t} \right] + \sqrt{\kappa_{\text{ex},-}}\hat{a}_{\text{in}}(t)e^{-i\Delta_-t}, \\
\frac{d}{dt}\hat{a}_+(t) &= \frac{-\kappa_+}{2}\hat{a}_+(t) + ig_+ \left[\hat{b}(t)e^{-i(\omega_m+\Delta_+)t} + \hat{b}^\dagger(t)e^{i(\omega_m-\Delta_+)t} \right] + \sqrt{\kappa_{\text{ex},+}}\hat{a}_{\text{in}}(t)e^{-i\Delta_+t}, \\
\frac{d}{dt}\hat{b}(t) &= \frac{-\kappa_m}{2}\hat{b}(t) + ig_-^*\hat{a}_-(t)e^{i(\omega_m+\Delta_-)t} + ig_- \hat{a}_-^\dagger(t)e^{i(\omega_m-\Delta_-)t} \\
&\quad + ig_+^*\hat{a}_+(t)e^{i(\omega_m+\Delta_+)t} + ig_+ \hat{a}_+^\dagger(t)e^{i(\omega_m-\Delta_+)t} + \sqrt{\kappa_{\text{ex},m}}\hat{c}_{\text{in}}(t).
\end{aligned} \tag{A19}$$

We can move to the frequency domain by using the Fourier transform

$$\hat{a}[\omega] = \mathcal{F}(\hat{a}(t))[\omega] = \int_{-\infty}^{\infty} \hat{a}(t)e^{i\omega t} dt \tag{A20}$$

along with the properties

$$\begin{aligned}
\mathcal{F}\left(\frac{d}{dt}\hat{a}(t)\right)[\omega] &= -i\omega\hat{a}[\omega], \\
\mathcal{F}(\hat{a}(t)e^{-i\Delta t})[\omega] &= \hat{a}[\omega - \Delta], \\
\hat{a}^\dagger[\omega] &\equiv \mathcal{F}(\hat{a}^\dagger(t))[\omega] = (\hat{a}[-\omega])^\dagger.
\end{aligned} \tag{A21}$$

We then have

$$\begin{aligned}
\hat{a}_-[\omega] &= ig_- \chi_-[\omega] \left(\hat{b}[\omega - (\omega_m + \Delta_-)] + \hat{b}^\dagger[\omega + (\omega_m - \Delta_-)] \right) + \sqrt{\kappa_{\text{ex},-}}\chi_-[\omega]\hat{a}_{\text{in}}[\omega - \Delta_-], \\
\hat{a}_+[\omega] &= ig_+ \chi_+[\omega] \left(\hat{b}[\omega - (\omega_m + \Delta_+)] + \hat{b}^\dagger[\omega + (\omega_m - \Delta_+)] \right) + \sqrt{\kappa_{\text{ex},+}}\chi_+[\omega]\hat{a}_{\text{in}}[\omega - \Delta_+], \\
\hat{b}[\omega] &= ig_-^* \chi_m[\omega]\hat{a}_-[\omega + (\omega_m + \Delta_-)] + ig_- \chi_m[\omega]\hat{a}_-^\dagger[\omega + (\omega_m - \Delta_-)] \\
&\quad + ig_+^* \chi_m[\omega]\hat{a}_+[\omega + (\omega_m + \Delta_+)] + ig_+ \chi_m[\omega]\hat{a}_+^\dagger[\omega + (\omega_m - \Delta_+)] + \sqrt{\kappa_{\text{ex},m}}\chi_m[\omega]\hat{c}_{\text{in}}[\omega],
\end{aligned} \tag{A22}$$

with the susceptibilities $\chi_-[\omega] = (-i\omega + \kappa_-/2)^{-1}$, $\chi_+[\omega] = (-i\omega + \kappa_+/2)^{-1}$ and $\chi_m[\omega] = (-i\omega + \kappa_m/2)^{-1}$. The bath operators are linked through the input-output relations

$$\begin{aligned}
\hat{a}_{\text{out}}(t)e^{-i\omega_L t} &= \hat{a}_{\text{in}}(t)e^{-i\omega_L t} - \sqrt{\kappa_{\text{ex},-}}\hat{a}_-(t)e^{-i\omega_- t} - \sqrt{\kappa_{\text{ex},+}}\hat{a}_+(t)e^{-i\omega_+ t}, \\
\hat{c}_{\text{out}}(t)e^{-i\omega_m t} &= -\hat{c}_{\text{in}}(t)e^{-i\omega_m t} + \sqrt{\kappa_{\text{ex},m}}\hat{b}(t)e^{-i\omega_m t}
\end{aligned} \tag{A23}$$

or

$$\begin{aligned}
\hat{a}_{\text{out}}[\omega] &= \hat{a}_{\text{in}}[\omega] - \sqrt{\kappa_{\text{ex},-}}\hat{a}_-[\omega + \Delta_-] - \sqrt{\kappa_{\text{ex},+}}\hat{a}_+[\omega + \Delta_+], \\
\hat{c}_{\text{out}}[\omega] &= -\hat{c}_{\text{in}}[\omega] + \sqrt{\kappa_{\text{ex},m}}\hat{b}[\omega]
\end{aligned} \tag{A24}$$

in frequency domain. They satisfy the commutation relations $[\hat{a}_{\text{in}}[\omega], \hat{a}_{\text{in}}^\dagger[\omega']] = \delta[\omega - \omega']$ and $[\hat{c}_{\text{in}}[\omega], \hat{c}_{\text{in}}^\dagger[\omega']] = \delta[\omega - \omega']$.

3. Conversion efficiency and added noise

a. Anti-Stokes configuration

For the following analysis, we consider the triply resonant configuration where $\omega_+ - \omega_- = \omega_m$ and the optical pump is on resonance with one of the optical supermodes. First, we consider the anti-Stokes configuration where $\Delta_- = 0$ and $\Delta_+ = -\omega_m$. From Eq. A24, we have the input-output relation

$$\hat{a}_{\text{out}}[\omega] = \hat{a}_{\text{in}}[\omega] - \sqrt{\kappa_{\text{ex},-}}\hat{a}_-[\omega] - \sqrt{\kappa_{\text{ex},+}}\hat{a}_+[\omega - \omega_m]. \tag{A25}$$

The equations of motion is given by

$$\begin{aligned}
\hat{a}_-[\omega] &= ig_- \chi_-[\omega] \left(\hat{b}[\omega - \omega_m] + \hat{b}^\dagger[\omega + \omega_m] \right) + \sqrt{\kappa_{\text{ex},-}} \chi_-[\omega] \hat{a}_{\text{in}}[\omega] \\
&\approx \sqrt{\kappa_{\text{ex},-}} \chi_-[\omega] \hat{a}_{\text{in}}[\omega], \\
\hat{a}_+[\omega] &= ig_+ \chi_+[\omega] \left(\hat{b}[\omega] + \hat{b}^\dagger[\omega + 2\omega_m] \right) + \sqrt{\kappa_{\text{ex},+}} \chi_+[\omega] \hat{a}_{\text{in}}[\omega + \omega_m] \\
&\approx ig_+ \chi_+[\omega] \hat{b}[\omega] + \sqrt{\kappa_{\text{ex},+}} \chi_+[\omega] \hat{a}_{\text{in}}[\omega + \omega_m], \\
\hat{b}[\omega] &= ig_-^* \chi_m[\omega] \hat{a}_-[\omega + \omega_m] + ig_- \chi_m[\omega] \hat{a}_-^\dagger[\omega + \omega_m] + ig_+^* \chi_m[\omega] \hat{a}_+[\omega] + ig_+ \chi_m[\omega] \hat{a}_+^\dagger[\omega + 2\omega_m] + \sqrt{\kappa_{\text{ex},m}} \chi_m[\omega] \hat{c}_{\text{in}}[\omega] \\
&\approx ig_+^* \chi_m[\omega] \hat{a}_+[\omega] + \sqrt{\kappa_{\text{ex},m}} \chi_m[\omega] \hat{c}_{\text{in}}[\omega].
\end{aligned} \tag{A26}$$

Here we drop the terms that contain susceptibilities and ladder operators with an offset frequency dependence (with the exception of the bath operators), since they necessarily imply an attenuation of at least $\kappa_+ / (2\omega_m)$ compared to the dominant terms. This approximation is valid in the resolved-sideband regime $\kappa_-, \kappa_+ \ll \omega_m$. Supplementary Figure 1a shows the signal flow graph representing the equations of motion, from which we deduced the transfer functions between the microwave field at ω_m and the optical field at $\omega_L + \omega_m$ using Mason's rule [2]

$$\begin{aligned}
S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]} &= S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]} = S_{\hat{a}_{\text{out}}^\dagger[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}^\dagger[\omega]} = S_{\hat{c}_{\text{out}}^\dagger[\omega] \leftarrow \hat{a}_{\text{in}}^\dagger[\omega + \omega_m]} \\
&= \frac{-i\sqrt{\kappa_{\text{ex},+}}\sqrt{\kappa_{\text{ex},m}}g_+\chi_+[\omega]\chi_m[\omega]}{1 + |g_+|^2\chi_+[\omega]\chi_m[\omega]}.
\end{aligned} \tag{A27}$$

The optical-optical and microwave-microwave transfer functions can also be derived as

$$S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]} = 1 - \frac{\kappa_{\text{ex},+}\chi_+[\omega]}{1 + |g_+|^2\chi_+[\omega]\chi_m[\omega]} - \kappa_{\text{ex},-}\chi_-[\omega + \omega_m], \tag{A28}$$

$$S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{c}_{\text{in}}[\omega]} = -1 + \frac{\kappa_{\text{ex},m}\chi_m[\omega]}{1 + |g_+|^2\chi_+[\omega]\chi_m[\omega]}. \tag{A29}$$

The other transfer functions (e.g., $S_{\hat{a}_{\text{out}}^\dagger[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]}$) are expected to be small given the minuscule weight of the corresponding edges in the signal flow graph. The photon number flux at the optical output of the transducer is given by

$$\begin{aligned}
\hat{a}_{\text{out}}^\dagger[\omega + \omega_m] \hat{a}_{\text{out}}[\omega + \omega_m] &\approx \left(S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]} \hat{a}_{\text{in}}^\dagger[\omega + \omega_m] + S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]} \hat{c}_{\text{in}}^\dagger[\omega] \right) \\
&\quad \left(S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]} \hat{a}_{\text{in}}[\omega + \omega_m] + S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]} \hat{c}_{\text{in}}[\omega] \right).
\end{aligned} \tag{A30}$$

To capture the noise properties of the optical output, one has to measure its symmetrized power spectral density

$$\begin{aligned}
\mathcal{S}_{\hat{a}_{\text{out}}}[\omega + \omega_m] &= \frac{1}{2} \left\langle \left\{ \hat{a}_{\text{out}}[\omega + \omega_m], \hat{a}_{\text{out}}^\dagger[\omega + \omega_m] \right\} \right\rangle = \frac{1}{2} + \left\langle \hat{a}_{\text{out}}^\dagger[\omega + \omega_m] \hat{a}_{\text{out}}[\omega + \omega_m] \right\rangle \\
&= \eta_+[\omega] \left(\mathcal{S}_{\hat{c}_{\text{in}}}[\omega] + \mathcal{S}_{\text{added},+}^\uparrow[\omega] \right),
\end{aligned} \tag{A31}$$

where $\mathcal{S}_{\hat{c}_{\text{in}}}[\omega] = \frac{1}{2} + \left\langle \hat{c}_{\text{in}}^\dagger[\omega] \hat{c}_{\text{in}}[\omega] \right\rangle$ is the power spectral density of the microwave photon flux at the input port. We define the on-chip photon number transduction efficiency as

$$\begin{aligned}
\eta_+[\omega] &= |S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]}|^2 = |S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]}|^2 \\
&= \frac{\kappa_{\text{ex},+}\kappa_{\text{ex},m}|g_+|^2|\chi_+[\omega]|^2|\chi_m[\omega]|^2}{|1 + |g_+|^2\chi_+[\omega]\chi_m[\omega]|^2} \approx \frac{\kappa_{\text{ex},+}}{\kappa_+} \frac{\kappa_{\text{ex},m}}{\kappa_m} \frac{4C_+}{(1 + C_+)^2}
\end{aligned} \tag{A32}$$

and the added noise during up-conversion as

$$\mathcal{S}_{\text{added},+}^\uparrow[\omega] = \frac{1}{2\eta_+[\omega]} \left(1 - \eta_+[\omega] - |S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]}|^2 \right) + \frac{|S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]}|^2}{\eta_+[\omega]} \mathcal{S}_{\hat{a}_{\text{in}}}[\omega + \omega_m]. \tag{A33}$$

We used the fact that the noise from the two input ports are not correlated, i.e. $\left\langle \hat{c}_{\text{in}}^\dagger[\omega] \hat{a}_{\text{in}}[\omega + \omega_m] \right\rangle = \left\langle \hat{c}_{\text{in}}[\omega] \hat{a}_{\text{in}}^\dagger[\omega + \omega_m] \right\rangle = 0$. An additional noise of $\frac{1}{2\eta_+[\omega]} \left(1 - \eta_+[\omega] - |S_{\hat{a}_{\text{out}}[\omega + \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega + \omega_m]}|^2 \right)$ is introduced so that

the output channel still respect the bosonic commutation relations [3]. The last approximation of Eq. A32 gives the steady-state efficiency of the transducer ($\omega \rightarrow 0$), which depends on anti-Stokes cooperativity $C_+ = 4g_+^2/(\kappa_+ \kappa_m)$ as well as extraction efficiencies of the antisymmetric supermode $\kappa_{\text{ex},+}/\kappa_+$ and the acoustic mode $\kappa_{\text{ex},m}/\kappa_m$. Similarly, the symmetrized power spectral density of the output photon number flux at the microwave port is

$$\mathcal{S}_{\hat{c}_{\text{out}}}[\omega] = \frac{1}{2} \left\langle \left\{ \hat{c}_{\text{out}}[\omega], \hat{c}_{\text{out}}^\dagger[\omega] \right\} \right\rangle = \frac{1}{2} + \left\langle \hat{c}_{\text{out}}^\dagger[\omega] \hat{c}_{\text{out}}[\omega] \right\rangle = \eta_+[\omega] \left(\mathcal{S}_{\hat{a}_{\text{in}}}[\omega + \omega_m] + \mathcal{S}_{\text{added},+}^\dagger[\omega] \right). \quad (\text{A34})$$

While the expressions for the conversion efficiency are identical for both up- and down-conversion, the added noise for down-conversion is instead

$$\mathcal{S}_{\text{added},+}^\dagger[\omega] = \frac{1}{2\eta_+[\omega]} \left(1 - \eta_+[\omega] - |S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{c}_{\text{in}}[\omega]}|^2 \right) + \frac{|S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{c}_{\text{in}}[\omega]}|^2}{\eta_+[\omega]} \mathcal{S}_{\hat{c}_{\text{in}}}[\omega]. \quad (\text{A35})$$

b. Stokes configuration

We follow the same recipe to derive the conversion efficiency and added noise for the Stokes configuration where $\Delta_- = \omega_m$ and $\Delta_+ = 0$. We have the optical input-output relation

$$\hat{a}_{\text{out}}[\omega] = \hat{a}_{\text{in}}[\omega] - \sqrt{\kappa_{\text{ex},-}} \hat{a}_-[\omega + \omega_m] - \sqrt{\kappa_{\text{ex},+}} \hat{a}_+[\omega] \quad (\text{A36})$$

and the equations of motion

$$\begin{aligned} \hat{a}_-[\omega] &= ig_- \chi_-[\omega] \left(\hat{b}[\omega - 2\omega_m] + \hat{b}^\dagger[\omega] \right) + \sqrt{\kappa_{\text{ex},-}} \chi_-[\omega] \hat{a}_{\text{in}}[\omega - \omega_m] \\ &\approx ig_- \chi_-[\omega] \hat{b}^\dagger[\omega] + \sqrt{\kappa_{\text{ex},-}} \chi_-[\omega] \hat{a}_{\text{in}}[\omega - \omega_m], \\ \hat{a}_+[\omega] &= ig_+ \chi_+[\omega] \left(\hat{b}[\omega - \omega_m] + \hat{b}^\dagger[\omega + \omega_m] \right) + \sqrt{\kappa_{\text{ex},+}} \chi_+[\omega] \hat{a}_{\text{in}}[\omega] \\ &\approx \sqrt{\kappa_{\text{ex},+}} \chi_+[\omega] \hat{a}_{\text{in}}[\omega], \\ \hat{b}[\omega] &= ig_-^* \chi_m[\omega] \hat{a}_-[\omega + 2\omega_m] + ig_- \chi_m[\omega] \hat{a}_-^\dagger[\omega] + ig_+^* \chi_m[\omega] \hat{a}_+[\omega + \omega_m] + ig_+ \chi_m[\omega] \hat{a}_+^\dagger[\omega + \omega_m] + \sqrt{\kappa_{\text{ex},m}} \chi_m[\omega] \hat{c}_{\text{in}}[\omega] \\ &\approx ig_- \chi_m[\omega] \hat{a}_-^\dagger[\omega] + \sqrt{\kappa_{\text{ex},m}} \chi_m[\omega] \hat{c}_{\text{in}}[\omega]. \end{aligned} \quad (\text{A37})$$

Using the signal flow graph shown in Supplementary Fig. 1b, the transfer functions between the microwave field at ω_m and the optical field at $\omega_L - \omega_m$ can once again be derived as

$$S_{\hat{a}_{\text{out}}[\omega - \omega_m] \leftarrow \hat{c}_{\text{in}}^\dagger[\omega]} = S_{\hat{a}_{\text{out}}^\dagger[\omega - \omega_m] \leftarrow \hat{c}_{\text{in}}[\omega]}^* = \frac{-i\sqrt{\kappa_{\text{ex},-}} \sqrt{\kappa_{\text{ex},m}} g_- \chi_-[\omega] \chi_m^*[\omega]}{1 - |g_-|^2 \chi_-^*[\omega] \chi_m[\omega]}, \quad (\text{A38})$$

$$S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{a}_{\text{in}}^\dagger[\omega - \omega_m]} = S_{\hat{c}_{\text{out}}^\dagger[\omega] \leftarrow \hat{a}_{\text{in}}[\omega - \omega_m]}^* = \frac{i\sqrt{\kappa_{\text{ex},-}} \sqrt{\kappa_{\text{ex},m}} g_- \chi_-^*[\omega] \chi_m[\omega]}{1 - |g_-|^2 \chi_-[\omega] \chi_m^*[\omega]}. \quad (\text{A39})$$

While the microwave-microwave transfer function is identical to Eq. A29, the optical-optical transfer function is

$$S_{\hat{a}_{\text{out}}[\omega - \omega_m] \leftarrow \hat{a}_{\text{in}}[\omega - \omega_m]} = 1 - \frac{\kappa_{\text{ex},-} \chi_-[\omega]}{1 - |g_-|^2 \chi_-^*[\omega] \chi_m[\omega]} - \kappa_{\text{ex},+} \chi_+[\omega - \omega_m]. \quad (\text{A40})$$

The photon number flux at the optical output of the transducer has a symmetrized power spectral density

$$\begin{aligned} \mathcal{S}_{\hat{a}_{\text{out}}}[\omega - \omega_m] &= \frac{1}{2} \left\langle \left\{ \hat{a}_{\text{out}}[\omega - \omega_m], \hat{a}_{\text{out}}^\dagger[\omega - \omega_m] \right\} \right\rangle = \frac{1}{2} + \left\langle \hat{a}_{\text{out}}^\dagger[\omega - \omega_m] \hat{a}_{\text{out}}[\omega - \omega_m] \right\rangle \\ &= \eta_-[\omega] \left(\mathcal{S}_{\hat{c}_{\text{in}}}[\omega] + \mathcal{S}_{\text{added},-}^\dagger[\omega] \right). \end{aligned} \quad (\text{A41})$$

We define the on-chip photon number conversion efficiency for the Stokes configuration as

$$\begin{aligned} \eta_-[\omega] &= |S_{\hat{a}_{\text{out}}[\omega - \omega_m] \leftarrow \hat{c}_{\text{in}}^\dagger[\omega]}|^2 = |S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{a}_{\text{in}}^\dagger[\omega - \omega_m]}|^2 \\ &= \frac{\kappa_{\text{ex},-} \kappa_{\text{ex},m} |g_-|^2 |\chi_-[\omega]|^2 |\chi_m[\omega]|^2}{|1 + |g_-|^2 \chi_-[\omega] \chi_m[\omega]|^2} \approx \frac{\kappa_{\text{ex},-}}{\kappa_-} \frac{\kappa_{\text{ex},m}}{\kappa_m} \frac{4C_-}{(1 - C_-)^2} \end{aligned} \quad (\text{A42})$$

as well as the added noise during up-conversion

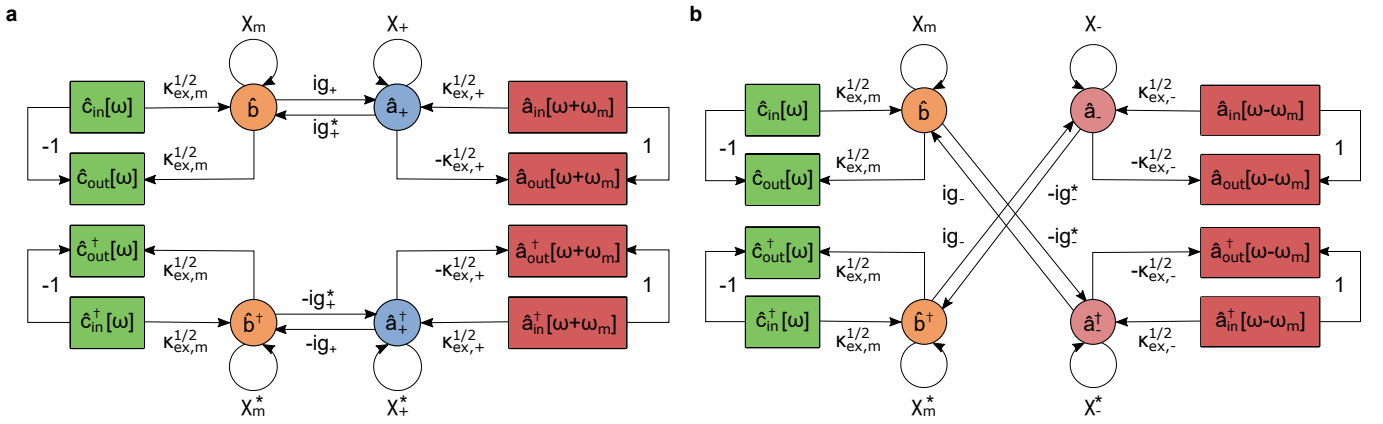
$$\mathcal{S}_{\text{added},-}^{\uparrow}[\omega] = \frac{1}{2\eta_-[\omega]} (1 + \eta_-[\omega] - |S_{\hat{a}_{\text{out}}[\omega-\omega_m] \leftarrow \hat{a}_{\text{in}}[\omega-\omega_m]}|^2) + \frac{|S_{\hat{a}_{\text{out}}[\omega-\omega_m] \leftarrow \hat{a}_{\text{in}}[\omega-\omega_m]}|^2}{\eta_-[\omega]} \mathcal{S}_{\hat{a}_{\text{in}}}[\omega - \omega_m] \quad (\text{A43})$$

where $C_- = 4g_-^2/(\kappa_- \kappa_m)$ is the Stokes cooperativity. The output photon number at the microwave port for the Stokes configuration is given by

$$\mathcal{S}_{\hat{c}_{\text{out}}}[\omega] = \frac{1}{2} \langle \{ \hat{c}_{\text{out}}[\omega], \hat{c}_{\text{out}}^{\dagger}[\omega] \} \rangle = \frac{1}{2} + \langle \hat{c}_{\text{out}}^{\dagger}[\omega] \hat{c}_{\text{out}}[\omega] \rangle = \eta_-[\omega] (\mathcal{S}_{\hat{a}_{\text{in}}}[\omega - \omega_m] + \mathcal{S}_{\text{added},-}^{\downarrow}[\omega]). \quad (\text{A44})$$

The photon number on-chip conversion efficiency is again the same for both up- and down-conversion, but the added noise for the down-conversion process is instead

$$\mathcal{S}_{\text{added},-}^{\downarrow}[\omega] = \frac{1}{2\eta_-[\omega]} (1 + \eta_-[\omega] - |S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{c}_{\text{in}}[\omega]}|^2) + \frac{|S_{\hat{c}_{\text{out}}[\omega] \leftarrow \hat{c}_{\text{in}}[\omega]}|^2}{\eta_-[\omega]} \mathcal{S}_{\hat{c}_{\text{in}}}[\omega]. \quad (\text{A45})$$



Supplementary Fig. 1. **Signal flow graphs of the frequency conversion processes.** **a** Effective beam-splitter process. **b** Effective two-mode-squeezing process. The bath operators are indicated by rectangles, the internal modes of the transducers by circles and the coupling rates by arrows.

4. Estimation of correlated microwave-optical pair generation rates

Equations A41 and A44 suggest that entangled pairs of microwave and optical photons can be generated in the Stokes configuration since it gives rise to an effective two-mode squeezing Hamiltonian [4–6]. In order to verify that the HBAR transducer can generate entangled pairs of microwave and optical photons, we compute the on-chip pair generation rate, expected to be equal to the output photon flux from the lower frequency optical mode as in standard spontaneous parametric down-conversion (SPDC) scheme. Integrating over the transduction bandwidth, we have

$$R = \int_{-\infty}^{\infty} \langle \hat{a}_{\text{out}}^{\dagger}[\omega] \hat{a}_{\text{out}}[\omega] \rangle d\omega \approx \int_{-\infty}^{\infty} \eta_-[\omega] d\omega \approx \frac{\pi}{2} (\kappa_-^{-1} + \kappa_m^{-1})^{-1} \eta_-[0]. \quad (\text{A46})$$

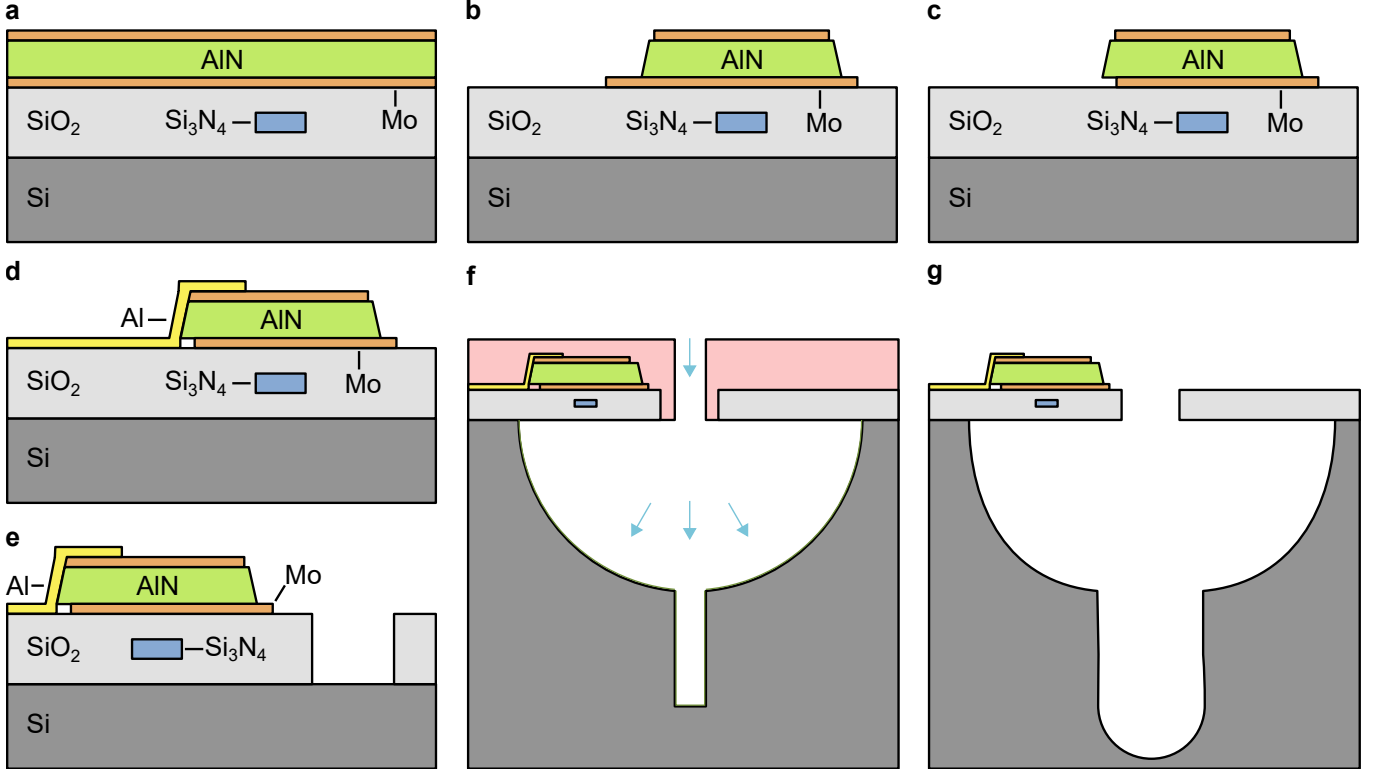
In practice, the pump laser should be filtered for the optical output to be dominated by heralding photons, and other losses from the measurement setup can further reduce the effective rate. For $\eta^{\text{oc}} = 7.9 \times 10^{-5}$, $\kappa_- = 2\pi \times 166$ MHz and $\kappa_m = 2\pi \times 13$ MHz, this would correspond to an on-chip pair generation rate of $R \approx 2\pi \times 1.5$ kHz. Losses in the measurement chain have to be accounted for to estimate the count rate at the detector. In order to measure entanglement between the microwave and optical fields, the on-chip rate must be higher than the thermal decoherence rate of the acoustic mode $\Gamma^{\text{dec}} = \kappa_m n_{\text{th},m}(\omega_m)$, where the thermal occupancy is given by the Bose-Einstein distribution $n_{\text{th},m}(\omega) = [\exp(\hbar\omega/(k_B T)) - 1]^{-1}$ given a temperature T and Boltzmann's constant k_B . For $\omega_m = 2\pi \times 3.5$ GHz and $\kappa_m = 2\pi \times 13$ MHz, the decoherence rate $\Gamma^{\text{dec}}/(2\pi)$ at 800 mK is 56 MHz and 0.7 Hz at 10 mK. Therefore, the transducer needs to be operated in the mixing chamber of the dilution refrigerator, with a pulsed pump to reduce heat load. The total rate of optical heralding is hence gated by the pump duty cycle.

To evaluate the non-classical correlation required for the DLCZ protocol [6, 7], one can compute the second-order cross-correlation function. Since no input signal needs to be sent to the transducer for SPDC, $\langle \hat{a}_{\text{in}} \rangle = \langle \hat{a}_{\text{in}}^\dagger \rangle = \langle \hat{a}_{\text{in}}^\dagger \hat{a}_{\text{in}} \rangle = \langle \hat{a}_{\text{in}} \hat{a}_{\text{in}} \rangle = \langle \hat{a}_{\text{in}}^\dagger \hat{a}_{\text{in}}^\dagger \rangle = 0$ and $\langle \hat{c}_{\text{in}}^\dagger[\omega] \hat{c}_{\text{in}}[\omega] \rangle = n_{\text{th},m}(\omega)$, leading to

$$g_{ac}^{(2)} = \frac{\langle \hat{c}_{\text{out}}^\dagger \hat{a}_{\text{out}}^\dagger \hat{a}_{\text{out}} \hat{c}_{\text{out}} \rangle}{\langle \hat{c}_{\text{out}}^\dagger \hat{c}_{\text{out}} \rangle \langle \hat{a}_{\text{out}}^\dagger \hat{a}_{\text{out}} \rangle} \approx \frac{(|S_{\hat{a}_{\text{out}} \leftarrow \hat{a}_{\text{in}}}|^2 + \eta_-)}{(\eta_- + |S_{\hat{c}_{\text{out}} \leftarrow \hat{c}_{\text{in}}}|^2 n_{\text{th},m})(1 + n_{\text{th},m})} + \frac{n_{\text{th},m}}{1 + n_{\text{th},m}} + \frac{\left(S_{\hat{a}_{\text{out}} \leftarrow \hat{a}_{\text{in}}}^* S_{\hat{a}_{\text{out}} \leftarrow \hat{c}_{\text{in}}} S_{\hat{c}_{\text{out}} \leftarrow \hat{a}_{\text{in}}}^* S_{\hat{c}_{\text{out}} \leftarrow \hat{c}_{\text{in}}} + S_{\hat{a}_{\text{out}} \leftarrow \hat{a}_{\text{in}}} S_{\hat{a}_{\text{out}} \leftarrow \hat{c}_{\text{in}}}^* S_{\hat{c}_{\text{out}} \leftarrow \hat{a}_{\text{in}}} S_{\hat{c}_{\text{out}} \leftarrow \hat{c}_{\text{in}}}^* \right) n_{\text{th},m}}{\eta_- (\eta_- + |S_{\hat{c}_{\text{out}} \leftarrow \hat{c}_{\text{in}}}|^2 n_{\text{th},m})(1 + n_{\text{th},m})} \quad (\text{A47})$$

where the frequency dependences are omitted for readability. Non-classicality is expected to lead to violation of the Cauchy-Schwarz inequality, given by $g_{ac}^{(2)} \leq \sqrt{g_{aa}^{(2)} g_{cc}^{(2)}}$. Therefore, it is desirable to minimize both microwave thermal occupancy (cooling) and optical transmission (attaining critical coupling).

Appendix B: Fabrication process flow



Supplementary Fig. 2. **Process flow for the fabrication of piezoelectric actuators on top of the photonic circuitry.** **a** Sputtering of the metal and piezoelectric layers. **b** Patterning of the top metal, piezoelectric and bottom metal layers. **c** XeF_2 etching of the bottom metal to avoid shorting the electrodes. **d** Sputtering of a second top metal layer to connect the top of the actuators to a feedline. **e** Dry etching of the oxide cladding to define the chip facets and the actuator holes. **f** Dry etching of the silicon substrate to suspend the actuators. **g** Deep reactive ion etching of the substrate to facilitate chip release.

The piezoelectric actuators are monolithically integrated on Si_3N_4 waveguides, fabricated using the photonic Damascene process [8, 9]. The 850-nm-thick Si_3N_4 film is deposited using low-pressure chemical vapor deposition into

the Damascene preform, with 2.8 μm thermal oxide below the waveguides. After annealing to drive hydrogen impurities out of the Si_3N_4 layer, 2.1- μm -thick TEOS oxide and 1.0- μm -thick low-temperature oxide top claddings are deposited and subsequently annealed. The metallic and piezoelectric films—95 nm of Mo, 1.0 μm of AlN and another 95 nm of Mo—are sputtered through foundry services provided by Plasma-Therm [10]. They are then patterned using chlorine-based reactive ion etching to form the actuators, ground planes and integrated heaters. We employ the same Band-Aid process as in Ref. [11] to localize the electrodes atop the suspended cladding. In particular, when connecting the Al feedline to the top electrode by a lift-off process, the bottom electrode is etched back using XeF_2 to avoid short-circuiting the electrodes. The process to suspend the cladding in the vicinity of the actuator is combined with the chip singulation steps of the Damascene process. First, a hole is opened in the middle of the arc-shaped actuator. The chip facets are simultaneously defined where trenches in the SiO_2 are etched by C_4F_8 deep reactive ion etching (DRIE) between neighboring chips. A second photolithography is performed to protect the facets while leaving the etched holes exposed for further processing, resulting in narrower trenches between chips. The Si substrate is then isotropically etched with SF_6 until the SiO_2 cladding below the actuator is suspended [12]. Following cladding suspension, the Si substrate is etched anisotropically until the desired chip thickness of 250 μm is reached. A final isotropic Si etch removes the parts of the substrate protruding from the SiO_2 facet, providing proper access to the bus waveguide nanotapers with lensed fibers, or even using the butt-coupling scheme. To facilitate this last step, we employ ion-beam etching at a 20° tilt angle to remove the passivation layer formed on the hole sidewalls during the Si DRIE. The chips are ultimately released by grinding the back side of the wafer. This convoluted process, summarized in Supplementary Fig. 2, ensures that all the actuators on the wafer are fully suspended, while the chip facets do not become exceedingly fragile from the undercut.

Appendix C: Optical characterization

1. Resonance linewidths

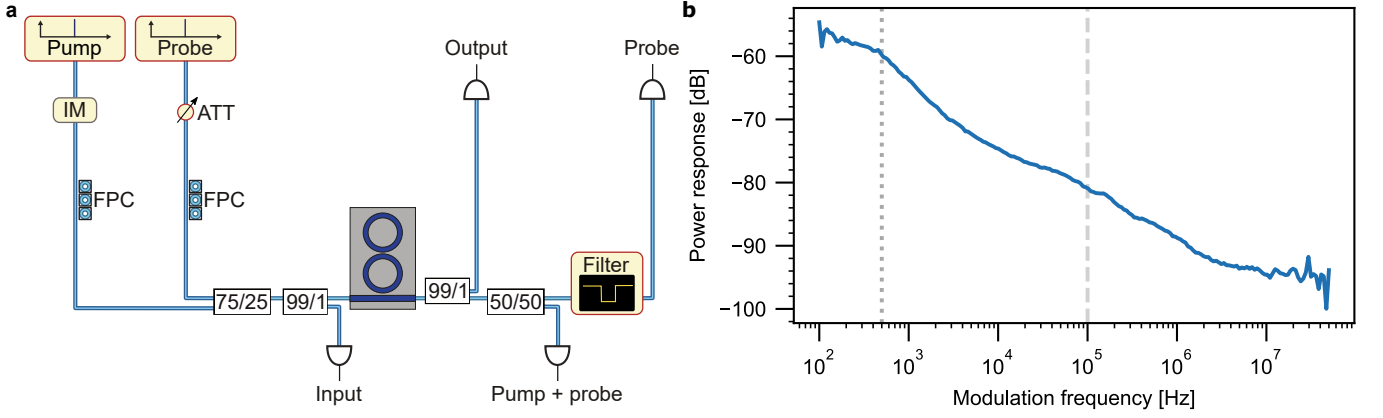
Fitting the hybridized transmission spectrum in Fig. 2c to the coupled mode theory presented in Appendix A 1 and in Ref. [2] yields $\kappa_r/(2\pi) = 154 \text{ MHz}$ and $\kappa_l/(2\pi) = 190 \text{ MHz}$, with $\kappa_{\text{ex},l}/(2\pi) = \kappa_{\text{ex},r}/(2\pi) = 60 \text{ MHz}$.

2. Hybridization of optical resonances via integrated thermo-optic heaters

For the purpose of rapid characterization of the photonic molecules at room temperature, the bottom electrode of the piezoelectric actuator also serves as an integrated heater to control the relative detuning between the micro-rings by thermo-optic effect. As seen in Fig. 2a, we pattern Mo to make three electrical connections to the bottom electrode. The top and bottom connections are connected to ground, whereas the central connection is biased to a constant voltage. The finite resistance of Mo at room temperature leads to Joule heating as current flows through the bottom electrode, modifying the refractive index of the Si_3N_4 waveguide buried directly underneath. The heaters were only used for fast search of devices with adequate mode splitting.

3. Thermal response characterization through cavity-enhanced photothermal spectroscopy

For the purpose of programming the pulse sequence to alleviate thermal effects, we study the thermal response of the present transducer due to optical pump heating. Specifically, thermally induced refractive index change causes a shift in resonance frequency of the Si_3N_4 micro-ring, which is detected using cavity-enhanced pump-probe spectroscopy [13]. The intensity-modulated pump addresses a TE resonance, while the probe measures the resulting side-of-fringe modulation of another TE mode. To ascertain the physical mechanism of each cross phase modulation (XPM) process, we sweep the pump modulation frequency to leverage the separation of time scales. Seen in Supplementary Fig. 3, the response exhibits three plateaus. We associate the process at modulation frequencies $\gtrsim 1 \text{ MHz}$ with Kerr-induced XPM. The two slower processes at $\sim 100 \text{ kHz}$ and $\sim 1 \text{ kHz}$ are characteristic of photothermal XPM in quasi-free-standing microresonators such as spheres [14] and toroids [15]. Here, the former “local” time scale can be attributed to thermalization of the mode volume with the acoustic resonator, while the thermalization of the suspended structure with the rest of the chip constitutes the latter “global” time scale. Therefore, it is expected that for future cryogenic operations, employing a pulse-on time (τ_{on}) shorter than the local time scale to gate the optical pump will significantly reduce thermal occupancy in the cladding acoustic mode. The pulse repetition rate (f_{rep}) and hence pulse-off time can then be chosen based on the available cooling power.



Supplementary Fig. 3. **Thermal response measurement through pump-probe spectroscopy.** **a** Schematic of the setup used for optical pump-probe spectroscopy. A pump laser, addressing a micro-ring resonance and modulated by an intensity modulator (IM), induces photothermal and Kerr cross-phase modulations. As a result, the frequency of a nearby resonance is also modulated, which is detected by a probe laser. The power of the probe laser is adjusted by a variable attenuator (ATT). The pump and probe polarizations are optimized with their respective fiber polarization controllers (FPC). The white boxes represent fiber beam splitters. **b** Measured cross-phase modulation response of a micro-ring as detected at the “Probe” photodetector with the pump filtered. The dotted and dashed lines indicate the 3-dB response bandwidths of the “global” and “local” photothermal processes, respectively.

Appendix D: Characterization of acoustic resonances

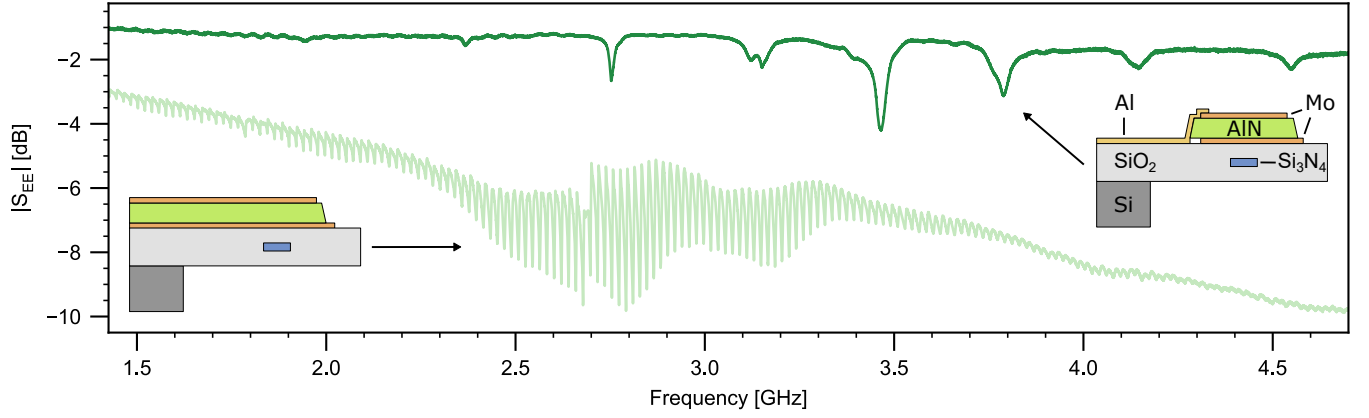
The microwave reflection measurements (also known as S_{11}) shown here were measured at room temperature using custom probes (GGB industries nickel-alloy 40A-GS-135-PC-N Picoprobe) and a vector network analyzer (Rhode&Schwarz ZNB-20). The spectrum shown in Fig. 2d is corrected using a calibration substrate (GGB industries CS-8) to remove the phase delay and attenuation introduced by coaxial cables and probes. However, this calibration procedure was only used to get the S-parameter, and was not applied during the total conversion efficiency measurements shown in Figs. 3 and 4. In a previous version of the transducers and in Ref. [12] where the Band-Aid process (Appendix B) is not employed, the piezoelectric actuator extends beyond the suspended cladding and covers the electrical feedline, causing two detrimental effects for transduction. First, the resulting stray capacitance leads to excess microwave insertion loss. Given the small size of the suspended area ($\sim 1000 \mu\text{m}^2$) compared to that of the feedline ($\gtrsim 10000 \mu\text{m}^2$), this effect is readily seen in Supplementary Fig. 4, where the spectrum without the Band-Aid process exhibits reduced background reflection. Second, the part of the actuator on top of the unreleased cladding leads to resonant excitation of HBAR modes extending over the entire Si substrate. These modes have a smaller free spectral range (FSR) of around 19 MHz and a lower microwave extraction efficiency. The theory presented in Appendix A does not correspond to the case of such a multimode system as it assumes being in the sideband-resolved regime, implying the high density of substrate modes is inappropriate for low-noise frequency conversion. The feedline Band-Aid, discussed in Appendix B and shown in Supplementary Fig. 4, enables the removal of the piezoelectric layer above the substrate, leaving only HBAR modes well confined in the oxide cladding.

We fit the measured complex reflection to obtain the mechanical quality factor $Q_m = 284$ and microwave extraction efficiency $\kappa_{\text{ex},m}/\kappa_m = 0.11$ for the transduction mode [16]. This fit model comprises multiple Fano resonances, given by

$$S_{\text{EE}}[\omega] = A e^{i\alpha - i\omega\tau} \left(1 - \sum_r e^{i\phi_r} \frac{2Q_1^r/Q_c^r}{1 + 2iQ_1^r(\omega - \omega_r)/\omega_r} \right) \quad (\text{D1})$$

and the result is shown in Supplementary Fig. 5a. In addition, we simulate the microwave reflection of the designed stack geometry using finite-element method (FEM). Fitting the simulated spectrum leads to $Q_m = 205$ and $\kappa_{\text{ex},m}/\kappa_m = 0.61$. The discrepancy is attributed to difference in clamping losses and piezoelectric coefficients between the fabricated and simulated devices that lift the degeneracy of electromechanical modes. We extract the effective mass of the transduction HBAR mode from FEM simulations as

$$m_{\text{eff}} = \frac{1}{\max(|\mathbf{x}(\mathbf{r})|^2)} \iiint \rho(\mathbf{r}) \mathbf{x}^*(\mathbf{r}) \mathbf{x}(\mathbf{r}) dV \approx 6 \text{ ng}, \quad (\text{D2})$$



Supplementary Fig. 4. **Elimination of substrate HBAR modes.** Microwave reflection spectra obtained with and without applying the Band-Aid process. The respective stack compositions are indicated as insets.

where $\rho(\mathbf{r})$ is the local mass density and $\mathbf{x}(\mathbf{r})$ is the mechanical displacement, corresponding to $x_{\text{ZPF}} = \sqrt{\hbar/(2m_{\text{eff}}\omega_m)} \approx 2 \times 10^{-8}$ nm.

The admittance is obtained by fitting the microwave reflection measurement to $Y[\omega] = Z_0^{-1}(1 - S_{\text{EE}}[\omega]) / (1 + S_{\text{EE}}[\omega])$, Z_0 being the characteristic impedance of the transmission line. Fitting the admittance to a modified Butterworth–van Dyke model yields a static capacitance of around 200 fF [17]. The small static capacitance is advantageous for reducing microwave insertion loss but comes at the cost of a reduced electromechanical transduction factor. Suspending the cladding decreases the acoustic mode volume, thereby enhancing the electromechanical coupling to the microwave port by improving the confinement in the piezoelectric layer.

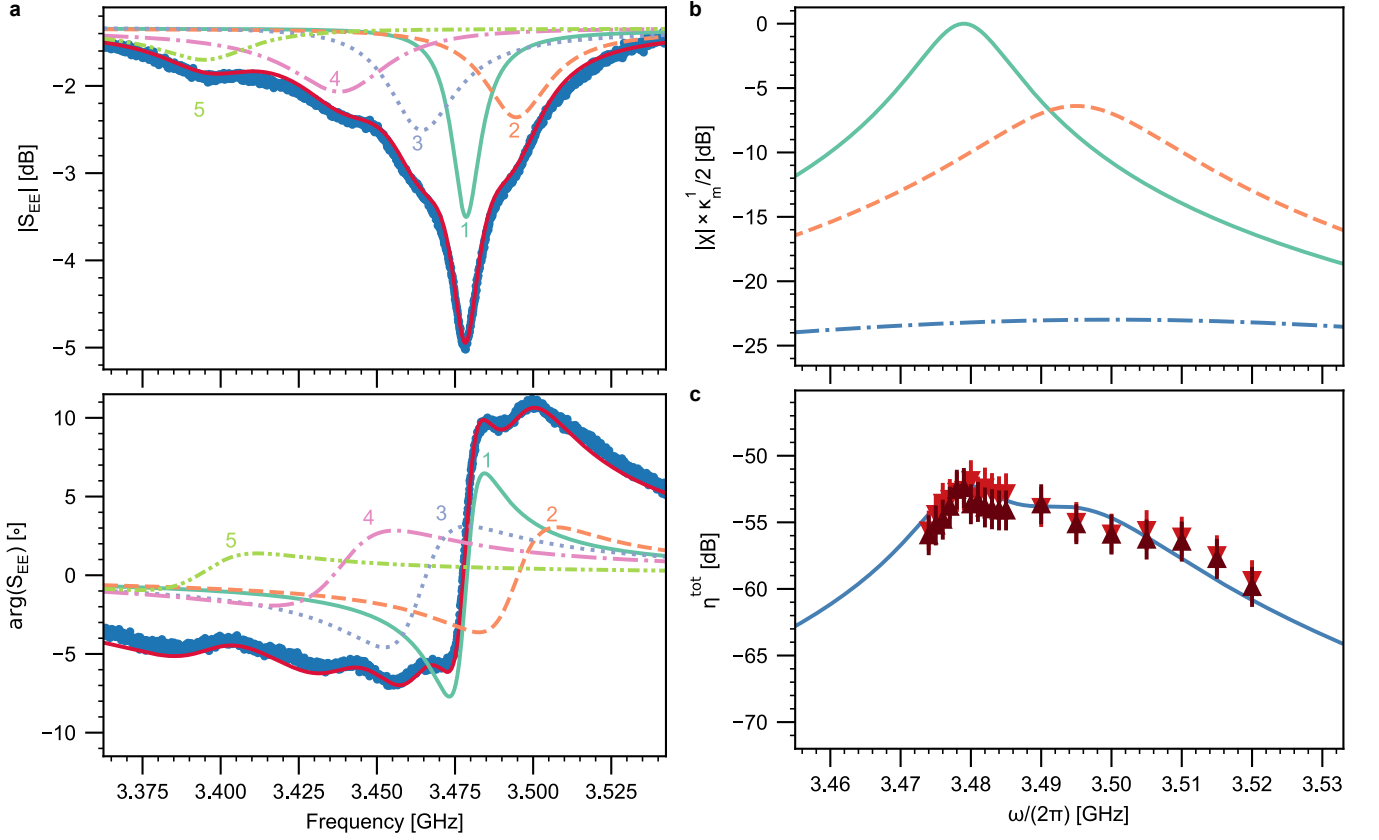
Resonance	$\omega/(2\pi)$ [GHz]	$\kappa/(2\pi)$ [MHz]	$\kappa_{\text{ex}}/\kappa$ [%]	ϕ [°]
1	3.479	13	11	5
2	3.495	27	6	5
3	3.463	28	6	11
4	3.440	40	4	−11
5	3.397	38	2	−11

Supplementary Table I. Parameters for the Fano resonance fit of the microwave reflection corresponding to Eq. D1.

Appendix E: Characterization of bidirectional transduction

1. Non-Lorentzian transduction bandwidth

For an ideal transducer, the transduction lineshape is expected to be Lorentzian in the low-cooperativity limit. Nevertheless, the measured lineshape (Figs. 3b and 3d) is evidently non-Lorentzian. We attribute this deviation to the presence of a second mechanical mode contributing to the transduction process. Supplementary Figure 5b shows the normalized susceptibilities of the principal mechanical mode, the auxiliary mechanical mode as well as the susceptibility of the higher frequency optical mode, considering that the device is operated in the anti-Stokes configuration. Supplementary Figure 5c illustrates the transduction spectrum that would be obtained from the theory exposed in Appendix A considering an additional mode and compares it with the experimental data shown in Fig. 3b. The only free parameters for the fit are the vacuum optomechanical coupling rates ($g_0^1 \approx 2\pi \times 20$ Hz and $g_0^2 \approx 2\pi \times 35$ Hz), the other values being the same as in Appendix E2. The difference with the vacuum optomechanical coupling rate estimated in Appendix E2 comes from the off-resonant contribution of the second mechanical mode.



Supplementary Fig. 5. **Spurious acoustic modes near the main resonance.** **a** Fit of the measured microwave reflection using multiple Fano resonances. The top panel shows the amplitude of the reflection, while the bottom panel depicts its phase after removing the background electrical delay from the cables. The experimental data is shown in blue, the total fit red, and the response of single Fano resonances in the other numbered curves. The parameters used for the fit are indicated in Supplementary Table I. **b** Susceptibilities of the two acoustic resonances relevant for transduction (solid line for the main resonance, dashed line for the secondary mode), as well as the higher frequency optical mode (dashed-dotted line). The absolute value is normalized to the peak of the main acoustic resonance at 3.479 GHz. **c** Simulated total transduction efficiency in the anti-Stokes configuration using the two acoustic resonances from **b**. The triangles correspond to the experimental data from Fig. 3b. The vacuum optomechanical coupling rates are the only free parameters of this fit, which gives $g_0^1 \approx 2\pi \times 20$ Hz and $g_0^2 \approx 2\pi \times 35$ Hz.

2. Estimation of cooperativity and vacuum coupling rate

The up- and down-conversion spectra, described in the main text and illustrated in Supplementary Fig. 6, provide the off-chip efficiency from which the vacuum coupling rate can be estimated. The off-chip efficiency is given by

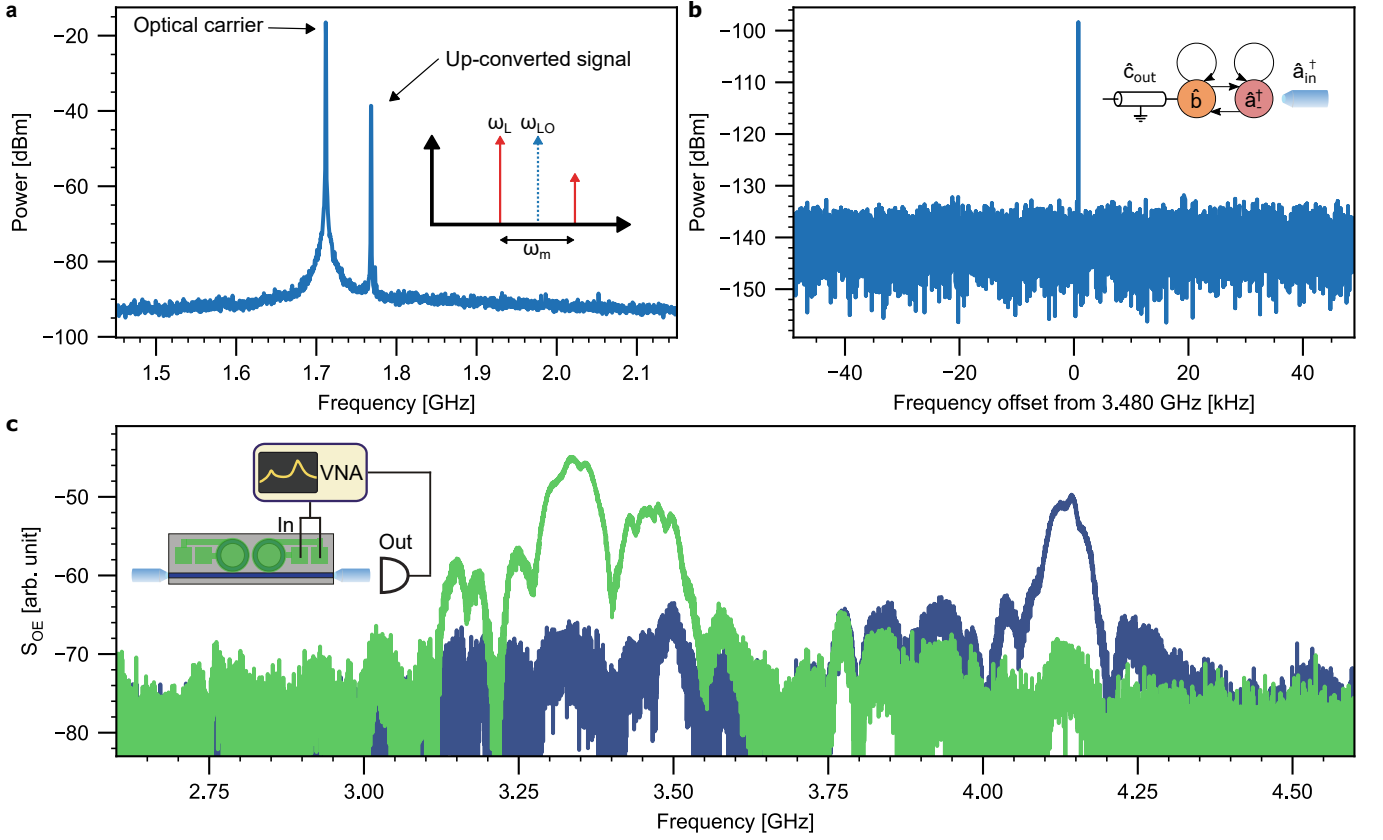
$$\eta^{\text{tot}} = \eta^{\text{probes}} \eta^{\text{fiber-chip}} \eta^{\text{oc}}, \quad (\text{E1})$$

where η^{probes} and $\eta^{\text{fiber-chip}}$ denote the microwave and optical fiber to chip insertion efficiencies, respectively. The on-chip efficiency can be written as

$$\eta^{\text{oc}} = \eta^{\text{ext}} \eta^{\text{int}} = \eta_{\text{o}} \eta_{\text{m}} \eta^{\text{int}}. \quad (\text{E2})$$

The internal efficiency

$$\eta^{\text{int}} = \frac{4C}{(1 \pm C)^2} \approx 4C \quad (\text{E3})$$



Supplementary Fig. 6. **Experimental dataset examples.** **a** Optical heterodyne spectrum for a local oscillator set between the carrier and the up-converted sideband, as indicated in the inset. **b** Microwave probes output measured with an electronic spectrum analyzer. The input optical spectrum contains both the carrier and sidebands generated by an external electro-optic modulator. In this example, the modulation frequency is set to 3.480 GHz. **c** Acousto-optic response for different laser-cavity detunings on the same transducer.

in the low-cooperativity regime, where the plus and minus signs correspond to the anti-Stokes and Stokes configurations, respectively. The optical and microwave extraction efficiencies

$$\eta_o = \frac{\kappa_{\text{ex},o}}{\kappa_o}, \quad (\text{E4})$$

$$\eta_m = \frac{\kappa_{\text{ex},m}}{\kappa_m}. \quad (\text{E5})$$

The single-photon cooperativity C_0 is enhanced by the intracavity photons

$$C = C_0 \bar{n}, \quad (\text{E6})$$

$$\begin{aligned} n_c &= \kappa_{\text{ex}} |\chi_o|^2 \frac{\eta^{\text{fiber-chip}} P_{\text{in}}}{\hbar \omega_L} \\ &\approx \frac{4 \kappa_{\text{ex}}}{\kappa_o^2} \frac{\eta^{\text{fiber-chip}} P_{\text{in}}}{\hbar \omega_L} = \eta_o \frac{4}{\kappa_o} \frac{\eta^{\text{fiber-chip}} P_{\text{in}}}{\hbar \omega_L}, \end{aligned} \quad (\text{E7})$$

where the pump detuning is assumed to be small. The off-chip efficiency with respect to the microwave probe and optical fiber is thus

$$\begin{aligned} \eta^{\text{tot}} &\approx \eta^{\text{probes}} \eta^{\text{fiber-chip}} \times \eta_m \eta_o \times 4 C_0 \times \eta_o \frac{4}{\kappa_o} \frac{1}{\hbar \omega_L} \eta^{\text{fiber-chip}} P_{\text{in}} \\ &= 16 \eta^{\text{probes}} \eta^{\text{fiber-fiber}} \times \eta_m \eta_o^2 \times \frac{C_0 P_{\text{in}}}{\hbar \omega_L \kappa_o}, \end{aligned} \quad (\text{E8})$$

where the fiber-fiber coupling efficiency $\eta^{\text{fiber-fiber}} = \eta^{\text{fiber-chip}} \times \eta^{\text{fiber-chip}}$. Calibration of the RF probes informed about the probes collections efficiency $\eta_{\text{probes}} \approx -3$ dB. A fraction of light is tapped just before and after the coupling lensed fibers for power monitoring. We obtain a typical $\eta^{\text{fiber-fiber}} = -8$ dB and hence $\eta^{\text{fiber-chip}} = -4$ dB. Fitting the microwave (optical) reflection spectrum yields $\eta_{\text{m}} = 11\%$ ($\eta_{\text{o}} = 35\%$). Given that the measured efficiency $\eta^{\text{tot}} = -60$ dB at $P_{\text{in}} = 10$ dBm, it follows that the single-photon cooperativity is

$$C_0 = \frac{\eta^{\text{tot}} \times \hbar\omega_{\text{L}}\kappa_{\text{o}}}{16\eta_{\text{probes}}\eta^{\text{fiber-fiber}} \times \eta_{\text{m}}\eta_{\text{o}}^2 \times P_{\text{in}}} \approx 8 \times 10^{-13}. \quad (\text{E9})$$

The vacuum coupling rate between the acoustic and bare optical modes is then

$$g_0 = 2 \times \sqrt{\frac{\kappa_{\text{o}}\kappa_{\text{m}}C_0}{4}} \approx 2\pi \times 42 \text{ Hz}. \quad (\text{E10})$$

Here $\kappa_{\text{o}} = 2\pi \times 170$ MHz, $\kappa_{\text{m}} = 2\pi \times 13$ MHz and the additional factor of two comes from the hybridization of the optical modes (Eq. A16).

3. Estimation of on-chip and internal transduction efficiencies

We have achieved $\eta^{\text{tot}} = -48$ dB at an input pump power of 21 dBm. With the losses quoted in Appendix E2, the on-chip and internal efficiencies can be estimated. For conversion in either directions, the signal goes through the respective input and output ports exactly once, acquiring an attenuation of $\eta_{\text{probes}}\eta^{\text{fiber-chip}} = -7$ dB. We therefore have $\eta^{\text{oc}} = 7.9 \times 10^{-5}$. Knowing the extraction efficiencies $\eta_{\text{o}} = 35\%$ and $\eta_{\text{m}} = 11\%$, we further obtain $\eta^{\text{int}} = 2 \times 10^{-3}$.

Appendix F: Optimization of optical extraction efficiency

An optimal optical external coupling can be chosen to maximize the total conversion efficiency. Equation E8 can be rewritten in terms of quality factors by using the relation $\kappa = \omega/Q$, which yields

$$\eta^{\text{tot}} \approx 16\eta_{\text{probes}}\eta^{\text{fiber-fiber}} \times \eta_{\text{m}}\eta_{\text{o}}^2 \times g_0^2 \frac{Q_{\text{m}}}{\omega_{\text{m}}} \frac{Q_{\text{o}}}{\omega_{\text{o}}} \times \frac{Q_{\text{o}}}{\omega_{\text{o}}} \frac{P_{\text{in}}}{\hbar\omega_{\text{L}}}. \quad (\text{F1})$$

For the optics, we have $Q_{\text{o}} = Q_{\text{int}} + Q_{\text{ext}}$, where we separate the intrinsic quality factor Q_{int} , oftentimes fabrication-limited and hence not easily adjustable, from the external coupling quality factor Q_{ex} that can be readily engineered through coupler design. The optical extraction efficiency can thus be written in the form

$$\eta_{\text{o}} = \frac{\kappa_{\text{ex}}}{\kappa_{\text{ex}} + \kappa_0} = \frac{Q_{\text{int}}}{Q_{\text{ex}}} \left(1 + \frac{Q_{\text{int}}}{Q_{\text{ex}}}\right)^{-1} = \frac{R}{1 + R}, \quad (\text{F2})$$

where $R = Q_{\text{int}}/Q_{\text{ex}}$ and the total optical quality factor

$$Q_{\text{o}} = \frac{1}{Q_{\text{int}}^{-1} + Q_{\text{ex}}^{-1}} = Q_{\text{int}} \frac{1}{1 + R}. \quad (\text{F3})$$

The total efficiency in the low-cooperativity regime is then given by

$$\eta^{\text{tot}} \approx F \frac{R^2}{(1 + R)^4}, \quad (\text{F4})$$

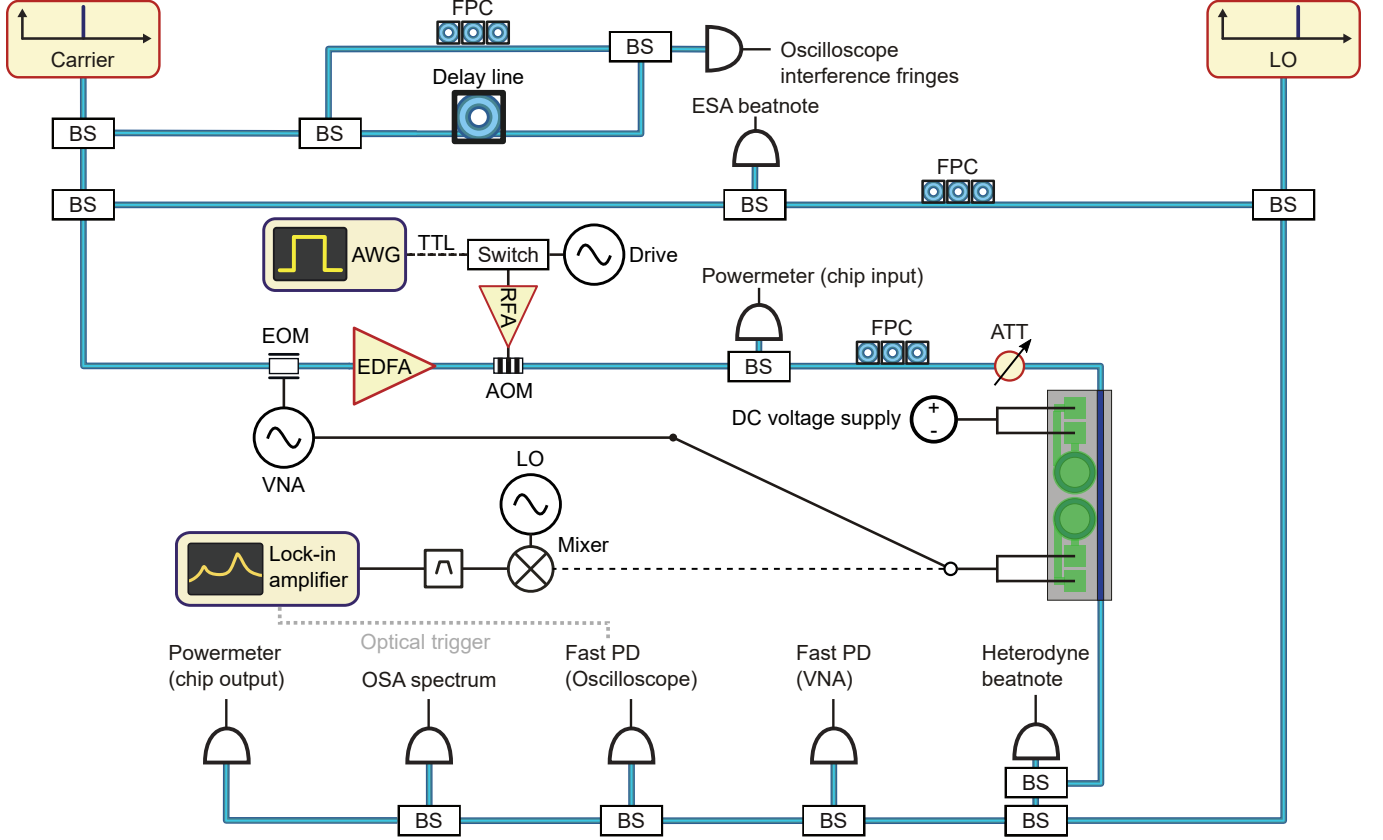
where

$$F = 16\eta_{\text{probes}}\eta^{\text{fiber-fiber}} \times \eta_{\text{m}}g_0^2 \frac{Q_{\text{m}}Q_{\text{int}}^2}{\omega_{\text{m}}\omega_{\text{o}}^2} \times \frac{P_{\text{in}}}{\hbar\omega_{\text{L}}}. \quad (\text{F5})$$

For a given F , the optimal efficiency is achieved at the critical coupling condition $Q_{\text{ex}} = Q_{\text{int}}$, or $R = 1$.

Appendix G: Experimental setup and data acquisition

We employ two Toptica CTL 1550 external cavity diode lasers in our experimental setup, depicted in Supplementary Fig. 7. The first laser (“science laser”), amplified by an erbium-doped fiber amplifier, is used to pump the device and for generating the optical signal for down-conversion, whereas the second laser serves as the local oscillator for optical heterodyne detection. The detuning between the two lasers is monitored from their beat note on a photodetector fed to an electronic spectrum analyzer. Optical transmission spectra are acquired using an oscilloscope while the science laser is scanned slowly. Their frequency axis is calibrated using a Mach-Zehnder interferometer with a long delay line in one arm. Optical powermeters are inserted before and after the chip, while a variable optical attenuator is used to control the power at the chip input. Fiber polarization controllers are used to ensure the science laser address the TE modes of the microresonators.



Supplementary Fig. 7. **Experimental setup.** The setup to measure the conversion efficiency is divided in three parts. First, part of the laser light is picked to calibrate its frequency using a Mach-Zehnder interferometer where a long delay line is introduced in one of the arm. The second laser, which provide the local oscillator, is beaten with the science laser for frequency calibration of the heterodyne signal. Finally, the science laser goes through the chip. Powermeters are placed before and after the chip in order to measure the insertion loss. The vector network analyzer (VNA) is used during the CW measurements, either for generating optical sidebands or for exciting the mechanics. During the pulsed measurements, the VNA is only used during up-conversion, being replaced by the lock-in detection for the down-conversion. Abbreviations used: AOM: acousto-optic modulator; ATT: variable optical attenuator; AWG: arbitrary waveform generator; BS: beam-splitter; EDFA: erbium-doped fiber amplifier; EOM: electro-optic modulator; FPC: fiber polarization controller; LO: local oscillator; OSA: optical spectrum analyzer; PD: photodetector; RFA: radio frequency amplifier; TTL: transistor-transistor logic.

For resonant transduction, the laser is tuned redward into the optical resonance from the blue side [18]. The laser-cavity detuning is chosen to maximize the acousto-optic response, monitored on a vector network analyzer. Regarding the acousto-optic response, one sideband is absorbed by the optical cavity, while the beating between the unsuppressed sideband and the carrier on a photodiode provides the signal. For driving and probing the HBAR, the piezoelectric actuator on the chip are connected to high-frequency RF probes (Picoprobe model 40A from GGB industries). We observed that the microwave reflection spectrum remains stable over the time scale of several months, provided the probe contact condition are maintained. We did not observe any change in resonance or broadening within the

range of microwave probe power up to 13 dBm and optical pump power up to 25 dBm. For the down-conversion demonstration, an external electro-optic modulator creates sidebands on the science laser. Their amplitude is adjusted using the optical heterodyne detection setup. A DC voltage source additionally enables optical mode hybridization using stress-optic effects via the second actuator.

Pulsed pumping is implemented by driving an acoustic-optic modulator that is switched by a transistor-transistor logic signal generated by an arbitrary waveform generator. The pulse sequence is chosen such that the optical resonance does not drift significantly within each pulse, as indicated by a constant steady-state “pulse-on” transmission. Pulsed up-conversion is recorded utilizing the same heterodyne setup as continuous-wave up-conversion. During pulsed down-conversion experiments, the microwave output from the RF probes are routed to a lock-in demodulation setup for detection of weak pulses.

-
- [1] H. Haus and W. Huang, *Proceedings of the IEEE* **79**, 1505 (1991).
 - [2] T. Blésin, H. Tian, S. A. Bhavé, and T. J. Kippenberg, *Phys. Rev. A* **104**, 052601 (2021).
 - [3] C. M. Caves, *Physical Review D* **26**, 1817 (1982).
 - [4] A. Rueda, W. Hease, S. Barzanjeh, and J. M. Fink, *npj Quantum Inf.* **5**, 1 (2019).
 - [5] C. Zhong, Z. Wang, C. Zou, M. Zhang, X. Han, W. Fu, M. Xu, S. Shankar, M. H. Devoret, H. X. Tang, and L. Jiang, *Phys. Rev. Lett.* **124**, 010511 (2020).
 - [6] S. Krastanov, H. Raniwala, J. Holzgrafe, K. Jacobs, M. Lončar, M. J. Reagor, and D. R. Englund, *Phys. Rev. Lett.* **127**, 040503 (2021).
 - [7] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, *Nature* **414**, 413 (2001).
 - [8] M. H. P. Pfeiffer, J. Liu, A. S. Raja, T. Morais, B. Ghadiani, and T. J. Kippenberg, *Optica* **5**, 884 (2018).
 - [9] J. Liu, G. Huang, R. N. Wang, J. He, A. S. Raja, T. Liu, N. J. Engelsen, and T. J. Kippenberg, *Nat. Commun.* **12**, 2236 (2021).
 - [10] H. Tian, J. Liu, B. Dong, J. C. Skehan, M. Zervas, T. J. Kippenberg, and S. A. Bhavé, *Nat. Commun.* **11**, 3073 (2020).
 - [11] A. Siddharth, A. Anastasio, G. Lihachev, J. Zhang, Z. Qiu, S. Kenning, R. N. Wang, S. A. Bhavé, J. Riemensberger, and T. J. Kippenberg, *Hertz-linewidth and frequency-agile photonic integrated extended-DBR lasers* (2023), [arxiv:2306.03184](https://arxiv.org/abs/2306.03184) [physics].
 - [12] H. Tian, J. Liu, A. Siddharth, R. N. Wang, T. Blésin, J. He, T. J. Kippenberg, and S. A. Bhavé, *Nat. Photon.* **15**, 828 (2021).
 - [13] M. Gao, Q.-F. Yang, Q.-X. Ji, H. Wang, L. Wu, B. Shen, J. Liu, G. Huang, L. Chang, W. Xie, S.-P. Yu, S. B. Papp, J. E. Bowers, T. J. Kippenberg, and K. J. Vahala, *Nat. Commun.* **13**, 3323 (2022).
 - [14] V. B. Braginsky, M. L. Gorodetsky, and V. S. Ilchenko, in *Laser Applications*, Vol. 2097 (SPIE, 1994) pp. 283–288.
 - [15] A. Schliesser, P. Del’Haye, N. Nooshi, K. J. Vahala, and T. J. Kippenberg, *Phys. Rev. Lett.* **97**, 243905 (2006).
 - [16] K. Leong and J. Mazierska, *IEEE Trans. Microw. Theory Tech.* **50**, 2115 (2002).
 - [17] J. Larson, P. Bradley, S. Wartenberg, and R. Ruby, in *2000 IEEE Ultrasonics Symposium. Proceedings. An International Symposium (Cat. No.00CH37121)*, Vol. 1 (2000) pp. 863–868 vol.1.
 - [18] T. Carmon, L. Yang, and K. J. Vahala, *Optics Express* **12**, 4742 (2004).