Science Advances

Supplementary Materials for

Superconducting cavity electro-optics: A platform for coherent photon conversion between superconducting and photonic circuits

Linran Fan, Chang-Ling Zou, Risheng Cheng, Xiang Guo, Xu Han, Zheng Gong, Sihao Wang, Hong X. Tang*

*Corresponding author. Email: hong.tang@yale.edu

This PDF file includes:

SECTION S1. THEORY OF CAVITY ELECTRO-OPTICS AND ITS UTILITY FOR MICROWAVE-TO-OPTICAL CONVERSION

The electro-optic effect is a second-order nonlinear process [33], thus the interaction Hamiltonian can be written as

$$H_{\mathrm{I}} = \hbar g_{\mathrm{eo}}(a + a^\dagger)(b + b^\dagger)(c + c^\dagger) \tag{S.1}$$

where a, b, and c are the Bosonic annihilation operators for the two optical modes and microwave mode respectively, and g_{eo} is the vacuum electro-optic coupling strength. As the optical frequency is much higher than the microwave frequency and g_{eo}, the counter rotating terms $a^\dagger b^\dagger + ab$ is neglected. Thus, the interaction Hamiltonian becomes

$$H_{\mathrm{I}} = \hbar g_{\mathrm{eo}}(ab^\dagger + a^\dagger b)(c + c^\dagger) \tag{S.2}$$

And the vacuum coupling strength g_{eo} can be expressed as

$$\hbar g_{\mathrm{eo}} = -\frac{\int(\varepsilon_{a,i}\varepsilon_{b,j}r_{ijk}) \cdot (u_{a,i}u_{b,j}^* u_{c,k})dxdydz}{8\pi\sqrt{\varepsilon_0}\prod_{l=a,b,c}\sqrt{\int \varepsilon_{l,i}u_{l,i}^* u_{l,i}dxdydz/\hbar\omega_l}} \tag{S.3}$$

Here $\varepsilon_{l,i}$ and $u_{l,i}$ ($l \in \{a,b,c\}$, $i \in \{x,y,z\}$) denote relative permittivity and electric field components respectively, ω_l is the angular frequency, r_{ijk} is the electro-optic component, and Einstein summation convention is used.

If we consider a ring structure with radius R, the field distribution of optical whispering gallery modes and microwave modes in the cylindrical coordinator can be expressed as $u_{l,i}(r,z,\theta) = u_{l,\perp,i}(r,z)e^{-im_l\theta}$ with $l \in \{a,b\}$ and $u_{c,i}(r,z,\theta) = \sum_{m_c} x_{c,m_c}u_{c,m_c,\perp,i}(r,z)e^{-im_c\theta}$, respectively, where m_l is the azimuthal number, and $x_{c,m}$ indicates the contribution of different azimuthal numbers for microwave modes. Thus the vacuum coupling rate is

$$\hbar g_{\mathrm{eo}} = \sqrt{\frac{1}{2\pi\varepsilon_0 R}}\frac{\int(\varepsilon_{a,i}\varepsilon_{b,j}r_{ijk}) \cdot (u_{a,\perp,i}u_{b,\perp,j}^* u_{c,m_c,\perp,k})drdz}{8\pi\prod_{l=a,b,c}\sqrt{\int \varepsilon_{l,i}u_{l,\perp,i}^* u_{l,\perp,i}drdz/\hbar\omega_l}} \times x_{c,m_c} \tag{S.4}$$

with $m_c = m_b - m_a$. The expression indicates that the microwave field should have non-zero coefficient for the azimuthal number $m_c = m_b - m_a$.

If mode a is coherently driven with a strong pump, the system Hamiltonian in the resolved sideband regime will become

$$H = \hbar\omega_a a^\dagger a + \hbar\omega_b b^\dagger b + \hbar\omega_c c^\dagger c + H_{\mathrm{I}} + i\sqrt{\frac{\kappa_{a,\mathrm{ex}}P_a}{\hbar\omega_{\mathrm{p}}}}(a^\dagger e^{-i\omega_{\mathrm{p}}t} - ae^{i\omega_{\mathrm{p}}t}) \tag{S.5}$$

where P_a, ω_{p}, and $\kappa_{a,\mathrm{ex}}$ are the pump power, pump frequency, and external coupling rate of mode a, respectively. For a strong coherent pump field, mode a is in steady state and the backaction from the system to the pump mode is negligible, thus we can treat the pump mode as a complex number
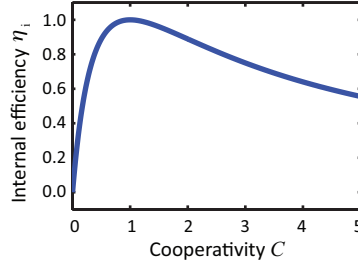
Fig. S1. **Calculated internal conversion efficiency.**

$$\langle a \rangle = \sqrt{\frac{\kappa_{a,\mathrm{ex}}}{\kappa_a^2/4 + \delta_a^2}} \times \sqrt{\frac{P_a}{\hbar \omega_\mathrm{p}}} e^{i\phi} \tag{S.6}$$

with the corresponding pump cavity photon number

$$n_a = \langle a^\dagger a \rangle = |\langle a \rangle|^2 \tag{S.7}$$

Here $\delta_a = \omega_a - \omega_\mathrm{p}$ is the pump detuning, $\kappa_a = \kappa_{a,\mathrm{ex}} + \kappa_{a,\mathrm{i}}$ is the total decay rate, $\kappa_{a,\mathrm{i}}$ is the intrinsic loss rate of mode a, and ϕ is the pump phase. Choosing the pump phase $\phi = 0$ and the pump mode frequency close to $\omega_b - \omega_c$, the system Hamiltonian in the rotation frame of $\hbar \omega_\mathrm{p} b^\dagger b$ under rotating-wave approximation (RWA) can be simplified to

$$H = \hbar \delta_b b^\dagger b + \hbar \omega_c c^\dagger c + \hbar G(b^\dagger c + b c^\dagger) \tag{S.8}$$

with $G = \langle a \rangle g_\mathrm{eo} = \sqrt{n_a} g_\mathrm{eo}$, which is the enhanced coupling rate, and $\delta_b = \omega_b - \omega_\mathrm{p}$. Equation (S.8) has the form of a linear beam splitter Hamiltonian, allows the coherent state transfer between mode b and mode c, which has the potential for quantum state transfer between microwave and optical frequencies. The equation of motion including the input and output of mode b and c can be written as

$$\frac{d}{dt} b = -(i\delta_b + \frac{\kappa_b}{2})b - iGc + \sqrt{\kappa_{b,\mathrm{ex}}} b_\mathrm{in} e^{-i\delta_{b,\mathrm{in}} t} \tag{S.9}$$

$$\frac{d}{dt} c = -(i\omega_c + \frac{\kappa_c}{2})c - iGb + \sqrt{\kappa_{c,\mathrm{ex}}} c_\mathrm{in} e^{-i\omega_{c,\mathrm{in}} t} \tag{S.10}$$

$$b_\mathrm{out} = b_\mathrm{in} - \sqrt{\kappa_{b,\mathrm{ex}}} b \tag{S.11}$$

$$c_\mathrm{out} = c_\mathrm{in} - \sqrt{\kappa_{c,\mathrm{ex}}} c \tag{S.12}$$

where b_in and c_in denote the input signals, $\delta_{b,\mathrm{in}} = \omega_{b,\mathrm{in}} - \omega_\mathrm{p}$ and $\omega_{c,\mathrm{in}}$ are the angular frequency of inputs, κ_b and κ_c are the total decay rates, $\kappa_{b,\mathrm{ex}}$ and $\kappa_{c,\mathrm{ex}}$ represent the external coupling rate of mode b and c, respectively. Solving Eqs. (S.9) - (S.12) in steady state, we can get the relationship between the input and output of mode b and c as

$$\begin{pmatrix} b_\mathrm{out} \\ c_\mathrm{out} \end{pmatrix} = \begin{pmatrix} 1 - \frac{\kappa_{b,\mathrm{ex}}}{i(\delta_b - \delta_{b,\mathrm{in}}) + \frac{\kappa_b}{2} + \frac{G^2}{i(\omega_c - \delta_{b,\mathrm{in}}) + \frac{\kappa_c}{2}}} & \frac{-iG\sqrt{\kappa_{b,\mathrm{ex}} \kappa_{c,\mathrm{ex}}}}{G^2 + [i(\omega_c - \omega_{c,\mathrm{in}}) + \frac{\kappa_c}{2}][i(\delta_b - \omega_{c,\mathrm{in}}) + \frac{\kappa_b}{2}]} \\ \frac{-iG\sqrt{\kappa_{b,\mathrm{ex}} \kappa_{c,\mathrm{ex}}}}{G^2 + [i(\delta_b - \delta_{b,\mathrm{in}}) + \frac{\kappa_b}{2}][i(\omega_c - \delta_{b,\mathrm{in}}) + \frac{\kappa_c}{2}]} & 1 - \frac{\kappa_{c,\mathrm{ex}}}{i(\omega_c - \omega_{c,\mathrm{in}}) + \frac{\kappa_c}{2} + \frac{G^2}{i(\delta_b - \omega_{c,\mathrm{in}}) + \frac{\kappa_b}{2}}} \end{pmatrix} \begin{pmatrix} b_\mathrm{in} \\ c_\mathrm{in} \end{pmatrix} \tag{S.13}$$

Thus, the bidirectional conversion efficiency can be written as

$$\eta = |\frac{b_\mathrm{out}}{c_\mathrm{in}}|^2 = |\frac{c_\mathrm{out}}{b_\mathrm{in}}|^2 = \frac{\kappa_{b,\mathrm{ex}}}{\kappa_b} \frac{\kappa_{c,\mathrm{ex}}}{\kappa_c} \frac{4C}{|C + (1 + \frac{2i(\delta_b - \omega)}{\kappa_b})(1 + \frac{2i(\omega_c - \omega)}{\kappa_c})|^2} \tag{S.14}$$

where $C = \frac{4G^2}{\kappa_b \kappa_c}$ is defined as the system cooperativity, and ω is the input signal angular frequency. The maximum conversion is reached when $\omega = \omega_c$ and the pump frequency matches the frequency difference between mode b and c ($\omega_\mathrm{p} = \omega_b - \omega_c$), and the on-chip conversion efficiency is

$$\eta = \frac{\kappa_{b,\mathrm{ex}}}{\kappa_b} \frac{\kappa_{c,\mathrm{ex}}}{\kappa_c} \times \frac{4C}{(1 + C)^2} \tag{S.15}$$
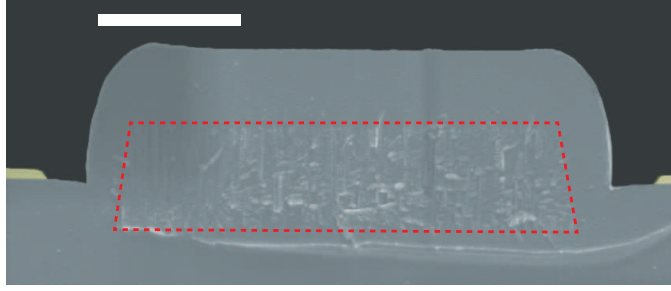
Fig. S2. **Scanning electron microscopy picture of the cross section of a superconducting cavity electro-optic device. AlN (red) is covered with SiO_2 (grey), and NbTiN (yellow) is sputtered on top of SiO_2. The device shown in the SEM picture has a different waveguide width ($3\,\mu$m) with the one used for experiment ($2\,\mu$m). The scale bar is $1\,\mu$m.**

with conversion bandwidth $(1 + C)\kappa_c$. As we can see from Eq. (S.15), the conversion efficiency consists of two parts: the cavity extraction efficiency which is $\frac{\kappa_{b,\text{ex}}}{\kappa_b}\frac{\kappa_{c,\text{ex}}}{\kappa_c}$, and the internal conversion efficiency which is

$$\eta_i = \frac{4C}{(1+C)^2} \tag{S.16}$$

Figure S1 plots the achievable internal conversion efficiency dependence on the cooperativity. The maximum internal conversion efficiency ($\eta_i = 1$) is achieved when the cooperativity $C = 1$. Also the internal conversion efficiency can never be larger than unity, due to the photon number conservation imposed by Eq. (S.8).

SECTION S2. DEVICE FABRICATION PROCEDURE

The 800 nm AlN film is grown on a Si wafer with 2-μm-thick SiO_2 layer by the radio frequency magnetron reactive sputtering, using pure aluminum (99.999%) targets in an argon and nitrogen gas mixture. Optical waveguides are patterned with electron beam lithography (EBL) using hydrogen silsesquioxane (HSQ) resist, subsequently transferred to the AlN layer by chlorine-based reactive ion-etching (RIE). Then 400 nm SiO_2 is deposited with plasma-enhanced chemical vapor deposition (PECVD), followed by device annealing at 940 °C. After annealing, NbTiN superconducting film is sputtered on top of SiO_2 by the magnetron sputtering method. A second EBL is performed to define the microwave resonator pattern with HSQ, which is transferred to the NbTiN layer by chlorine-based RIE. A test device instead of the device used in experiment is cleaved along with the diameter of the AlN optical ring resonator to verify the consistence beween fabricated device and our design. The SEM picture of the device cross section is shown in Fig. S2. As we can see from the SEM picture, the optical waveguide width ($3\,\mu$m) and the gap between optical waveguide and microwave electrode (400 nm) match our design.

SECTION S3. IDENTIFYING PHASE-MATCHING CONDITIONS FOR OPTICAL MODES

We use the r_{13} electro-optic coefficient of AlN, therefore the fundamental TE optical mode, fundamental TM optical mode, and the TE microwave mode are needed in our triple-resonance scheme. As the microwave mode has much smaller frequency and azimuthal number than optical modes, we first identify the condition that TE and TM optical modes have the same frequency and azimuthal number. The waveguide geometry is designed that TE and TM modes have the same phase velocity, thus the same azimuthal number. In a ring structure, the frequency degeneracy between TE and TM optical modes with the same azimuthal number will be broken due to the non-vertical sidewall of AlN waveguides as shown in Fig. S3a. Anti-crossing between TE and TM modes can be observed, with the mode coupling strength (half of the minimum frequency difference) determined by the sidewall angle and the ring radius. At the anti-crossing point, the two optical modes cannot be classified as pure TE or TM mode anymore. Instead, the two optical modes are the mixture of TE and TM modes (Fig. S3b and c).

The mixing between TE and TM modes provides the spectrum signature to identify the phase matching condition. If we tune the optical input in the bus waveguide to be TE mode, we can only observe one group of TE optical modes at the output if there is no mixing between TE and TM modes. If the TE and TM azimuthal numbers approach each other, mixing between TE and TM modes takes place, and we can observe two groups of modes, corresponding to the two mixed modes.
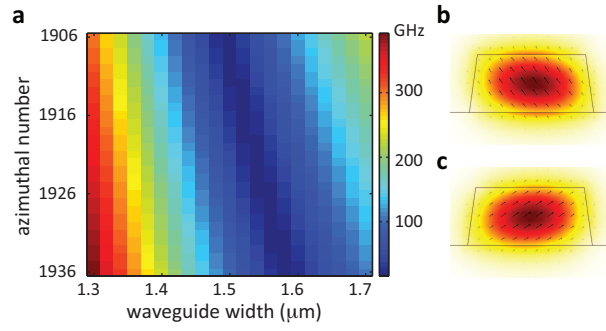
Fig. S3. **Anti-crossing between TE and TM optical modes.** **a** Simulated frequency difference between TE and TM optical modes with the same azimuthal number ($m_a = m_b$) around 194 THz frequency. **b, c** Mode profiles of the hybrid optical modes when the frequency difference is smallest in **a**. Arrows represent the electric field direction, and the color saturation in **b** and **c** shows the mode energy density. Ring radius is 240 μm, and the AlN thickness is 800 nm in simulation.
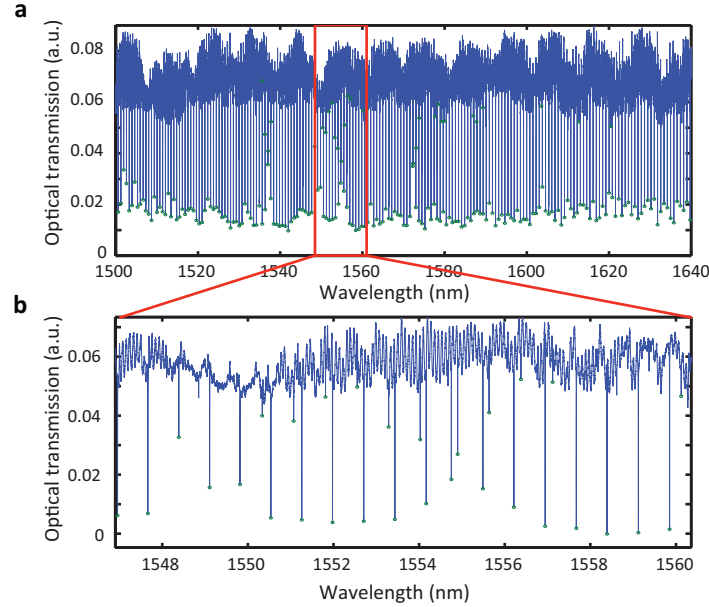

Fig. S4. **Measured spectrum signature of mixing between TE and TM modes.** **a** Optical transmission spectrum with pure TE optical input. **b** Zoom-in of the optical transmission spectrum around 1555 nm. Mode splitting and extinction decrease indicate the mixing between TE and TM modes.

In Fig. S4, we show the measured TE optical spectrum of a AlN ring with radius 240 μm and width 2.1 μm. Beside the fundamental TE optical modes, we observe another group of modes around 1555 nm, showing the mixing between TE and TM modes. In addition, the mode extinction ratio also drops, as the coupling rate between the bus waveguide and ring resonator drops. In Fig. S5, we plot the wavelength difference between two adjacent optical resonances. When there is no mixing between TE and TM modes, the wavelength difference corresponds to the free spectral range of the ring resonator. When the mixing happens, the wavelength difference drops below half of the free spectral range, and the smallest wavelength difference gives us the anti-crossing strength between TE and TM modes. Beside the main anti-crossing around 1555 nm, we can also observe anti-crossing around 1538 nm and 1573 nm, corresponding to the azimuthal number different of 1. As we can see, the anti-crossing strength is much smaller for $|m_a - m_b| = 1$ compared with $m_a = m_b$ (Fig. S6a). Here, the mode mixing between the TE and TM modes for $|m_a - m_b| = 1$ is from the surface roughness and the perturbation induced by the external coupling waveguide.

As the vacuum coupling efficiency is inversely proportional to the square root of the ring radius (Eq. (S.3)), we should minimize the ring radius to maximize the vacuum coupling rate. Therefore, instead of using TE and TM modes with the same azimuthal number ($m_a = m_b$), we use the TE and TM modes with azimuthal number different by 1 ($|m_a - m_b| = 1$), which have a much smaller mode anti-crossing. Accordingly, the microwave azimuthal number should also be designed to be 1, which will be explained in detail in Sect. III.
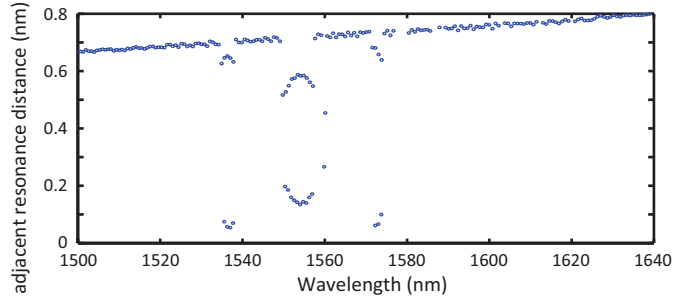
Fig. S5. **Wavelength difference between adjacent resonances.** The resonant wavelengths of all optical modes from Fig. S4 are extracted, and the adjacent resonance distance is computed by subtracting the resonant wavelength of the one optical mode by its previous optical mode.
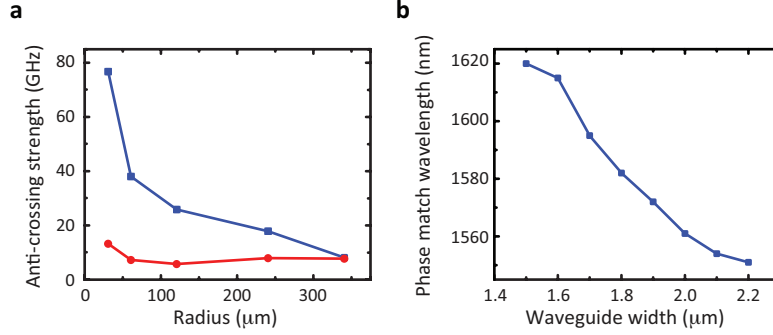
a

b

Fig. S6. **Phase-matching wavelength and anti-crossing strength. a** Anti-crossing strength dependence on ring radius. The cases for $m_a = m_b$ and $|m_a - m_b| = 1$ are shown in blue and red respectively. **b** Phase matching wavelength dependence on the waveguide width, with ring radius 240 μm.

With the same ring radius, the phase matching wavelength can be fine-tuned by the waveguide width, which modifies the phase velocity differently for the TE and TM optical modes. Figure S6b shows that the phase matching wavelength can be tuned by 70 nm with a waveguide width change of 600 nm. Therefore we can precisely control the phase matching wavelength to match different wavelength bands, providing the possibility for wavelength domain multiplexing.

SECTION S4. INFLUENCE OF OPTICAL MODE MIXING ON THE VACUUM COUPLING RATE g_{eo}

As we can see from the last section, the anti-crossing between TE and TM optical modes sets the lowest possible microwave frequency if triple resonance scheme is used. Furthermore, the vacuum coupling rate is decreased due to the mixing between TE and TM optical modes. The system Hamiltonian for the mode mixing is

$$H_{\mathrm{x}} = \hbar g_{\mathrm{x}}(a^\dagger b + ab^\dagger) \tag{S.17}$$

where g_{x} is the mode mixing strength between TE and TM optical modes. Since $g_{\mathrm{x}} \gg g_{\mathrm{eo}}$, we first diagonalize the optical part of the Hamiltonian by introducing the hybrid optical modes as

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{S.18}$$

where A and B are the operators for hybrid optical modes, and $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$, satisfying

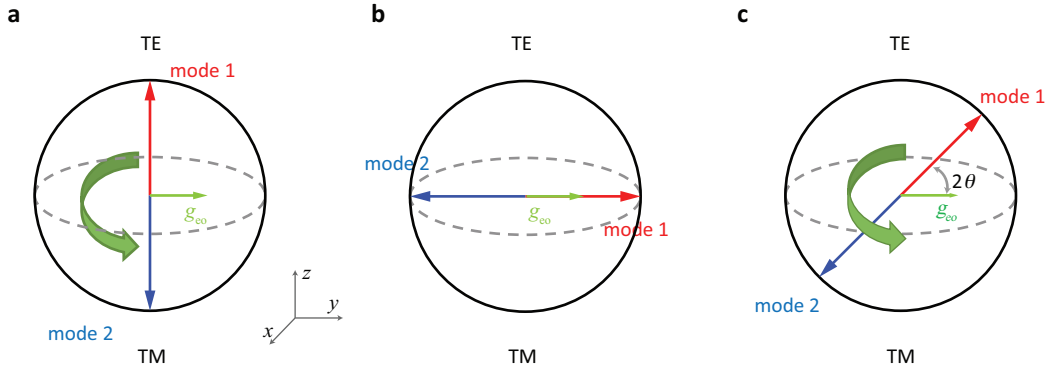$$\tan 2\theta = \frac{2g_{\mathrm{x}}}{\omega_a - \omega_b} \tag{S.19}$$

Fig. S7. **Vacuum coupling rate with hybrid optical modes.** In Bloch sphere representation, the north and south poles are the TE and TM optical modes respectively, and the electro-optic interaction is equivalent to the rotation along y-axis. Three cases are shown: **a,** no mixing; **b,** maximum mixing; **c** partial mixing. The effective coupling rate will be g, 0, and $g \cos 2\theta$ respectively.

The eigenfrequencies of the hybrid optical modes are

$$\omega_A = \frac{\omega_a + \omega_b}{2} + \sqrt{g_{\mathrm{x}}^2 + (\frac{\omega_a - \omega_b}{2})^2} \tag{S.20}$$

$$\omega_B = \frac{\omega_a + \omega_b}{2} - \sqrt{g_{\mathrm{x}}^2 + (\frac{\omega_a - \omega_b}{2})^2} \tag{S.21}$$

The minimum frequency difference is therefore $2g_{\mathrm{x}}$, which corresponding to the frequency gap in the anti-crossing spectrum.

By inverting Eq. (S.18), we can obtain

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \tag{S.22}$$

Plugging Eq. (S.22) into Eq. (S.2), we obtain the modified electro-optic interaction Hamiltonian

$$H_{\mathrm{I,m}} = \hbar g_{\mathrm{eo}} \cos 2\theta (A^\dagger B + AB^\dagger)(c + c^\dagger) \tag{S.23}$$

$$= \hbar g_{\mathrm{eo}}^{\mathrm{eff}} (A^\dagger B + AB^\dagger)(c + c^\dagger) \tag{S.24}$$

Comparing with the original interaction Hamiltonian, the modified Hamiltonian has the same form, but lower effective coupling strength $g_{\mathrm{eo}}^{\mathrm{eff}} = g_{\mathrm{eo}} \cos 2\theta$.

When the detuning between the two original optical modes is zero, thus $\theta = 45°$, the effective electro-optic interaction vanishes. This phenomena can be intuitively shown with the Bloch sphere. Assuming the north and south poles are the TE and TM optical modes respectively, then the electro-optic interaction is equivalent to the rotation along y-axis (Fig. S7a). When there is mode anti-crossing and the detune is zero, the two renormalized optical modes involved are maximally hybridized mode, whose directions are along y-axis, thus the rotation along y-axis will not induce the mode conversion between the two renormalized modes (Fig. S7b). Therefore, the detuning must be non-zero, and the effective vacuum coupling rate is $g_{\mathrm{eo}} \cos 2\theta$ (Fig. S7c). When the detuning is much larger than the anti-crossing strength, the vacuum coupling rate approaches the original value g_{eo}.

SECTION S5. MICROWAVE RESONATOR DESIGN

This planar microwave resonator can be treated as two coupled lumped element microwave resonators, where the central disk and surrounding arms provide the dominant capacitance and inductance respectively. The two lumped microwave resonators are coupled through the connection at the central disk. The resonator has symmetric and anti-symmetric modes due to the coupling between capacitors. The simulated electric field distribution is plotted in Fig. S8a and b. The symmetric mode has uniformly distributed electric field with the azimuthal angle ϕ, thus the azimuthal number $m_c = 0$, satisfying the phase matching condition for $m_a = m_b$. And the anti-symmetric mode has a $\sin\phi$ dependence, thus strong component of $m_c = 1$, satisfying the phase matching condition for $|m_a - m_b| = 1$.
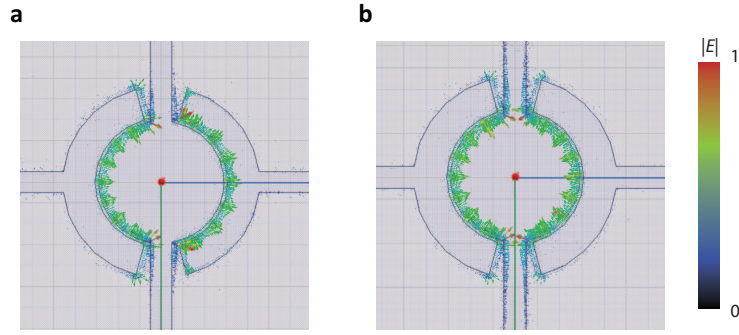
Fig. S8. **Microwave resonator simulation. a** Simulated electric field distribution of the anti-symmetric mode. **b** Simulated electric field distribution of the symmetric mode. The microwave modes shown here are simulated by ANSYS HFSS, and the arrow color shows the relative strength of the elelctric field.
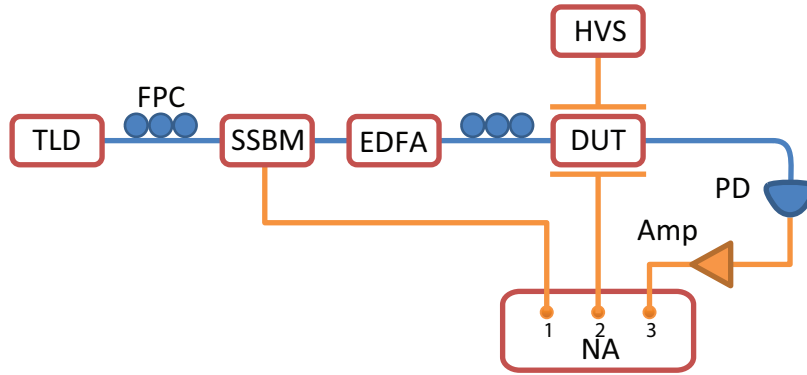

Fig. S9. **Experiment setup for microwave-to-optical photon conversion.** TLD, tunable laser diode; EDFA, erbium doped fiber amplifier; SSBM, single side-band modulator; FPC, fiber polarization controller; DUT, device under test; HVS, high voltage source; PD, photo-detector; Amp, RF amplifier; NA network analyzer.

The lumped element microwave resonator can provide very small electric mode volume, which is critical for enhancing the vacuum coupling rate (Eq. S.3). Also the resonant frequency can be varied by adjusting the arm length of the resonator. Thus the capacitance part of the resonator can be kept fixed, making it easy to match the optical cavity and resonant frequency simultaneously. Another advantage of this design is that the long-arm inductor allows supercurrents to generate magnetic flux far-extended from the chip surface, making it feasible to inductively couple the microwave resonator with an off-chip loop probe for broadband high-efficiency microwave signal input and readout.

SECTION S6. MEASUREMENT SETUP

The experiment setup is shown in Fig. S9. Coherent light from a tunable laser diode (TLD) is used as the control light, which is sent into an single side-band modulator (SSBM) to generate a weak probe light. The modulation RF signal is from a network analyzer (NA). Then a erbium doped fiber amplifier (EDFA) is used to amplify the control light. Light is coupled to and from the on-chip bus waveguide through a pair of grating couplers, which only transmit TE light. The output light is detected by the high frequency photodetector (PD). The output signal of the photodetector is amplified by a RF amplifier (Amp), and de-modulated with network analyzer. The microwave port of the device is directly connect to the network analyzer. A high voltage source (HVS) is used to provide the static electric field across the device for optical resonance tuning.

SECTION S7. DEVICE TEMPERATURE CALIBRATION

The strong control light can heat up the device, which may degrade the microwave resonator performance. In order to calibrated the device temperature, we measure the microwave resonant frequency and linewidth at different
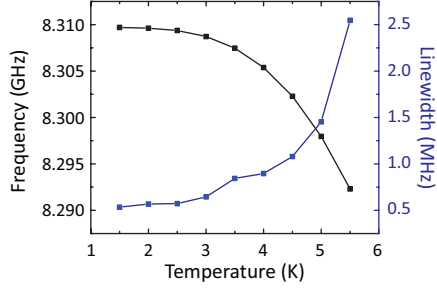
Fig. S10. **Microwave resonator performance under different temperature.**

ambient temperatures without light input. As we can see from Fig. S10, the resonator frequency drops and the linewidth increases with the temperature increase. Therefore, by measuring the microwave resonant frequency and linewidth at base temperature with strong control light, we can infer the effective temperature of the microwave cavity. The linewidths at $2.0\,\mathrm{K}$, $2.5\,\mathrm{K}$, and $3.0\,\mathrm{K}$ are $0.64\,\mathrm{MHz}$, $0.58\,\mathrm{MHz}$, and 0.54 MHz respectively. As shown in Fig. 2E and Fig.4 E in the main text, the microwave resonance linewidth is around $0.55\,\mathrm{MHz}$ under $8\,\mathrm{dBm}$ optical pump, indicating that the effective temperature of the device is around $2.1\,\mathrm{K}$.

SECTION S8. EFFICIENCY CALIBRATION

The cooperativity C of 0.075 ± 0.001 is estimated from the fitting of the blue curve in Fig. 3C in the main text, leading to internal efficiency η_i around $(25.9 \pm 0.3)\%$ based on Eq. (S.16). From the resonance extinction ratio $|R_i|^2$, the photon extraction efficiency can be calculated $\eta_{i,\mathrm{ex}} = \frac{\kappa_{i,\mathrm{ex}}}{\kappa} = \frac{1-|R|}{2}$ assuming under-coupled condition with $i = b$, c for optical signal and microwave modes respectively. From Fig. 2 C & E in the main text, the resonance extinction ratio of optical signal and microwave modes can be extracted, $|R_b|^2 = 0.139 \pm 0.002$ and $|R_c|^2 = 0.229 \pm 0.002$, leading to extraction efficiency $\eta_{b,\mathrm{ex}} = \frac{\kappa_{b,\mathrm{ex}}}{\kappa_b} = 0.313 \pm 0.002$ and $\eta_{c,\mathrm{ex}} = \frac{\kappa_{c,\mathrm{ex}}}{\kappa_c} = 0.261 \pm 0.001$ respectively. Thus the on-chip efficiency can be estimated $\eta = \frac{\kappa_{b,\mathrm{ex}}}{\kappa_b} \frac{\kappa_{c,\mathrm{ex}}}{\kappa_c} \eta_\mathrm{i} = (2.11 \pm 0.03)\%$. Here the error bar of the cooperativity C and extinction ratio $|R_a|$, $|R_b|$ are estimated based on the fitting residual, and passed to the error bar of on-chip efficiency η. We also followed the calibration procedure in Ref. [7]. The complete conversion matrix is measured as shown in Fig. 4 in the main text, including optical reflection S_oo, microwave reflection S_ee, microwave-to-optical conversion S_oe, and optical-to-microwave conversion S_eo. The on-resonance conversion efficiency is normalized by the off-resonance reflection amplitude, thus the gain and loss of the measurement circuit are excluded. Assuming that the conversion efficiency is the same for both directions, we estimate $(2.05 \pm 0.04)\%$ on-chip efficiency, which agrees well with the estimation based on the cooperativity and extraction efficiency. Here the error bar of on-chip efficiency is also estimated based on the fitting residual of Fig. 4.

Based on the cooperativity C and the resonance linewidth, the enhanced coupling strength can be estimated $G_\mathrm{eo} = 2\pi \times (1.76 \pm 0.05)\,\mathrm{MHz}$. The pump photon number inside the pump optical mode is $(3.2 \pm 0.1) \times 10^7$, with the uncertainty determined by the insertion loss. The error of insertion loss is obtained by measuring the transmission variation of multiple devices (supplementary material of Ref. [34]). Therefore the effective vacuum coupling rate estimated from experimental results is

$$g_\mathrm{eo,exp}^\mathrm{eff} = 2\pi \times (310 \pm 10)\,\mathrm{Hz}. \tag{S.25}$$

Using finite-element-method simulation with the actual device geometry estimated from Fig. S2, the field distributions (u_a, u_b, u_c) of the optical pump and signal modes, and the microwave mode can be obtained (Fig. 1D, E, & F). Plugging the field distributions into Eq. (S.4) and assuming $r_{113} = 1\,\mathrm{pm/V}$ [21], the vacuum coupling rate for our device can be calculated as $g_\mathrm{eo} = 2\pi \times (520 \pm 50)\,\mathrm{Hz}$, with the uncertainty determined by the parameter difference between the design and fabricated device estimated from Fig S2. Together with the optical mode mixing strength obtained from Eq. (S.19), the calculated effective vacuum coupling rate is

$$g_\mathrm{eo}^\mathrm{eff} = g_\mathrm{eo} \cos(2\theta) = 2\pi \times (330 \pm 30)\,\mathrm{Hz}, \tag{S.26}$$

which agrees with our experimental estimation.

SECTION S9. ADDED NOISE DURING CONVERSION

In addition to the conversion efficiency η, the added noise during the conversion process is also an important figure of merit. There are mainly two sources of noise, thermal excitation of the microwave cavity and photons generated by the parametric interaction $\hbar G(b^\dagger c^\dagger + bc)$.

By including the counter-rotating term $b^\dagger c^\dagger + bc$, we have the

$$-i\omega b\left(\omega\right) = -\left(i\delta_b + \frac{\kappa_b}{2}\right)b\left(\omega\right) - iGc\left(\omega\right) - iGc^\dagger\left(-\omega\right) + \sqrt{\kappa_{b,\text{ex}}}b_\text{in}\left(\omega\right)\delta\left(\omega - \delta_{b,in}\right) + \sqrt{\kappa_{b,\text{in}}}b_n\left(\omega\right) \tag{S.27}$$

$$-i\omega b^\dagger\left(-\omega\right) = -\left(-i\delta_b + \frac{\kappa_b}{2}\right)b^\dagger\left(-\omega\right) + iGc\left(\omega\right) + iGc^\dagger\left(-\omega\right) + \sqrt{\kappa_{b,\text{ex}}}b_\text{in}^\dagger\left(\omega\right)\delta\left(-\omega - \delta_{b,in}\right) + \sqrt{\kappa_{b,\text{in}}}b_n^\dagger\left(-\omega\right) \tag{S.28}$$

$$-i\omega c\left(\omega\right) = -\left(i\omega_c + \frac{\kappa_c}{2}\right)c\left(\omega\right) - iGb\left(\omega\right) - iGb^\dagger\left(-\omega\right) + \sqrt{\kappa_{c,\text{ex}}}c_\text{in}\left(\omega\right)\delta\left(\omega - \omega_{c,in}\right) + \sqrt{\kappa_{c,\text{in}}}c_n\left(\omega\right) \tag{S.29}$$

$$-i\omega c^\dagger\left(-\omega\right) = -\left(-i\omega_c + \frac{\kappa_c}{2}\right)c^\dagger\left(-\omega\right) + iGb\left(\omega\right) + iGb^\dagger\left(-\omega\right) + \sqrt{\kappa_{c,\text{ex}}}c_\text{in}^\dagger\left(-\omega\right)\delta\left(-\omega - \omega_{c,in}\right) + \sqrt{\kappa_{c,\text{in}}}c_n^\dagger\left(-\omega\right) \tag{S.30}$$

Here, $b_n\left(\omega\right)$ and $c_n\left(\omega\right)$ are the noise from the intrinsic loss channels of the optical and microwave cavities respectively. The noise correlators can be expressed as following

$$\left\langle b_n^\dagger\left(\omega\right)b_n\left(\Omega\right)\right\rangle = 0 \tag{S.31}$$

$$\left\langle b_n\left(\Omega\right)b_n^\dagger\left(\omega\right)\right\rangle = \delta\left(\omega - \Omega\right) \tag{S.32}$$

$$\left\langle c_n^\dagger\left(\omega\right)c_n\left(\Omega\right)\right\rangle = n_c\delta\left(\omega - \Omega\right) \tag{S.33}$$

$$\left\langle c_n\left(\Omega\right)c_n^\dagger\left(\omega\right)\right\rangle = \left(n_c + 1\right)\delta\left(\omega - \Omega\right) \tag{S.34}$$

By solving Eq. (S.27) - (S.30) and the input-output relations Eq. (S.11) & Eq. (S.12), we can obtain the output signal at phase matching condtion $\delta_b = \omega_c$

$$\left\langle b_\text{out}^\dagger\left(\omega\right)b_\text{out}\left(\omega\right)\right\rangle = \eta_{c\to b}\left\langle c_\text{in}^\dagger\left(\omega\right)c_\text{in}\left(\omega\right)\right\rangle + N_{b,add} + N_{b,para} \tag{S.35}$$

$$\left\langle c_\text{out}^\dagger\left(\omega\right)c_\text{out}\left(\omega\right)\right\rangle = \eta_{b\to c}\left\langle b_\text{in}^\dagger\left(\omega\right)b_\text{in}\left(\omega\right)\right\rangle + N_{c,add} + N_{c,para} \tag{S.36}$$

with the conversion efficiency

$$\begin{aligned}
\eta_{c\to b} &= \frac{\kappa_{b,\text{ex}}\kappa_{c,\text{ex}}}{\kappa_b\kappa_c}\frac{4C}{\left(1+C\right)^2} \\
&+ \frac{\kappa_{b,\text{ex}}\kappa_{c,\text{ex}}}{\kappa_b\kappa_c}\frac{C}{\left(C+1\right)^4}\left[\left(C^2 + 2C\right)\frac{\kappa_b^2}{4\omega_c^2} + \left(1 + 4C + 2C^2\right)\frac{\kappa_c^2}{4\omega_c^2} + 2C\frac{\kappa_c\kappa_b}{4\omega_c^2}\right] + O(1/\omega_c^3)
\end{aligned} \tag{S.37}$$

$$\begin{aligned}
\eta_{b\to c} &= \frac{\kappa_{b,\text{ex}}\kappa_{c,\text{ex}}}{\kappa_b\kappa_c}\frac{4C}{\left(1+C\right)^2} + \\
&+ \frac{\kappa_{b,\text{ex}}\kappa_{c,\text{ex}}}{\kappa_b\kappa_c}\frac{C}{\left(C+1\right)^4}\left[\left(C^2 + 2C\right)\frac{\kappa_c^2}{4\omega_c^2} + \left(1 + 4C + 2C^2\right)\frac{\kappa_b^2}{4\omega_c^2} + 2C\frac{\kappa_c\kappa_b}{4\omega_c^2}\right] + O(1/\omega_c^3)
\end{aligned} \tag{S.38}$$

The added noise due to the thermal environment is

$$N_{b,add} = \frac{4C}{\left(1+C\right)^2}\frac{\kappa_{b,\text{ex}}\kappa_{c,\text{in}}}{\kappa_b\kappa_c}n_c \tag{S.39}$$

$$N_{c,add} = \frac{4}{\left(1+C\right)^2}\frac{\kappa_{c,\text{in}}\kappa_{c,\text{ex}}}{\kappa_c^2}n_c \tag{S.40}$$

And the noise due the counter-rotating term is

$$N_{b,para} = \frac{C}{\left(C+1\right)^2}\frac{\kappa_{b,\text{ex}}}{\kappa_b}\frac{C\kappa_b^2 + \kappa_c^2}{4\omega_c^2} \tag{S.41}$$

$$N_{c,para} = \frac{C}{\left(C+1\right)^2}\frac{\kappa_{c,\text{ex}}}{\kappa_c}\frac{\kappa_b^2 + C\kappa_c^2}{4\omega_c^2} \tag{S.42}$$

Compare with Eq. (S.15), the additional term in the conversion efficiency is the counter-rotating term induced amplification. For our experiment parameters with $\kappa_b/2\pi = 294\,\mathrm{MHz}$, $\kappa_c/2\pi = 0.55\,\mathrm{MHz}$, and $\omega_c/2\pi = 8.31\,\mathrm{GHz}$, we estimated the modification of the conversion efficiency $\eta \rightarrow \eta\,(1 + \mathcal{A})$ due to the parametric amplification factor $\mathcal{A} \sim \frac{(C^2 + 2C)}{4(1+C)^2}\frac{\kappa_b^2}{4\omega_c^2} = 1.0 \times 10^{-5}$ for the achieved $C \sim 0.075$. This indicates that the RWA is valid for our experiments.

For the added noise, we have the thermal excitation $n_c \sim 4.5$ for $T \sim 2\,\mathrm{K}$, thus $N_{b,add} \sim 0.3$ and $N_{c,add} \sim 3$, which can be further reduced by working at lower temperature. In addition, the noise due to parametric amplification can be estimated, $N_{b,para} \sim 5 \times 10^{-7}$ and $N_{c,para} \sim 5 \times 10^{-6}$, which are negligible in our current experiments.