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 and Hong X. Tang* 
Department of Electrical Engineering, Yale University, New Haven, Connecticut 06511, USA

Leveraging the quantum information processing ability of superconducting circuits and long-distance distribution ability of optical photons promises the realization of complex and large-scale quantum networks. In such a scheme, a coherent and efficient quantum transducer between superconducting and photonic circuits is critical. However, such quantum transducer is still challenging since the use of intermediate excitations in current schemes introduces extra noise and limits bandwidth. Here we realize direct and coherent transduction between superconducting and photonic circuits based on triple-resonance electro-optics principle, with integrated devices incorporating both superconducting and optical cavities on the same chip. Electromagnetically induced transparency is observed, indicating the coherent interaction between microwave and optical photons. Internal conversion efficiency of $25.9 \pm 0.3\%$ has been achieved, with $2.05 \pm 0.04\%$ total efficiency. Superconducting cavity electro-optics offers broad transduction bandwidth and high scalability, and represents a significant step towards the integrated hybrid quantum circuits and distributed quantum computation.

Introduction

The hybrid approach of combining superconducting and photonic quantum technologies promises to realize large scale quantum networks [1–4]. In superconducting quantum circuits, the low-loss single quanta nonlinearity at microwave frequencies inherent to Josephson effect allows efficient and fast quantum operations [5]. However, it is challenging to directly transmit quantum states at microwave frequencies over long distance due to the high attenuation and thermal noise at room temperature. On the other hand, optical photons show complementary features. The weak single photon nonlinearity prevents the development of high fidelity quantum gates at optical frequency [6]. However, low decoherence and dissipation rates make optical photons the ideal information carrier for quantum communication [2, 3]. As a result, it is beneficial to develop hybrid quantum platform where quantum information is processed by superconducting circuits, and transmitted with optical photons. Thus, the quantum transducer which can coherently interface superconducting and photonic circuits with high conversion efficiency is highly demanded [7–20].

Coherent conversion of photons between microwave and optical frequencies has been proposed utilizing various intermediate excitations, including collective spin in atom ensembles [18], phonon in electro-optomechanics [7–10], and magnon in magneto-optics [19, 20]. Currently, the highest conversion efficiency is demonstrated based on electro-optomechanical systems where a compliant mechanical resonator couples to microwave and optical cavities simultaneously [7]. The use of intermediate low-frequency (MHz) excitations inevitably introduces extra noise channels, limits conversion bandwidth, and complicates operation with impedance matching between optical and microwave ports. The electro-optic approach can overcome these obstacles by excluding intermediate excitations in the conversion process, as proposed by Tsang recently [12, 13]. Great improvement of electro-optic coupling strength is proposed by Javerzac-Galy et al. utilizing coplanar microwave structure and integrated optical resonators, showing the feasibility of near-unity conversion efficiency with practical device parameters [15]. Even though the conversion efficiency of electro-optical systems has been improved dramatically, it is still limited to ~0.1% due to the large ohmic loss of non-superconducting material, and small coupling rates resulting from large mode volumes [16]. Moreover, the coherence and bidirectionality of the conversion process remain to be proved.

In this paper, we report the experimental demonstration of the coherent conversion between microwave and optical photons based on the electro-optic effect within a hybrid superconducting-photonic device, where planar superconducting resonators are integrated with aluminum nitride (AlN) optical cavities on the same chip. We observe the electromagnetically induced transparency effect in electro-optic systems, as a signature of coherent conversion between microwave and optical photons. Internal conversion efficiency of $25.9 \pm 0.3\%$ and on-chip efficiency of $2.05 \pm 0.04\%$ are realized. A major challenge we have addressed is to realize the energy and phase conservation of the triple-resonance condition with ultra-small mode volumes, boosting the pump photon number and vacuum coupling rate simultaneously to enhance the coherent conversion. Moreover, our device is ready to incorporate other superconducting and photonic quantum devices on the same chip, providing the scalable platform for hybrid quantum network.

Results
ports low loss optical modes and provides high electro-optic coefficients simultaneously [22] (Fig. 1B). Superconducting microwave resonators are placed on top of a thin buffer layer, and the capacitor shape is designed to match the optical cavity to maximize the field overlap between microwave and optical modes [15, 17].

We employ the triple-resonance scheme to enhance the coupling as shown Fig. 1C, with the interaction Hamiltonian written as

$$H_1 = \hbar g_{eo}(ab^\dagger c + a^\dagger bc^\dagger)$$

where $a$, $b$, $c$ are the annihilation operators for the optical pump and signal modes, and microwave mode respectively, and $g_{eo}$ is the vacuum electro-optic coupling rate. Cavity electro-optic systems with triple resonances have been proposed and demonstrated recently [15, 16]. However, the device geometry is limited to above several millimeters, as the pump and signal optical modes are from the same mode group, and the free spectral range (FSR) needs to match the microwave frequency. Large mode volume inevitably leads to small vacuum coupling rate, thus low conversion efficiency (Supplementary Section I). In contrast, our integrated approach uses the transverse-electric (TE) and transverse-magnetic (TM) optical modes as pump and signal modes respectively (Fig.1E & F), whose frequency difference equals the microwave frequency [23–26]. Thus the device size and mode volume can be further reduced without the limitation imposed by FSR. In this case, $r_{13}$ electro-optic coefficient is used, which also enables the use of TE-polarized microwave mode, thus the heterogeneous integration of planar microwave resonators with optical cavities (Fig. 1D). During experiments, a strong coherent field is applied to the pump mode ($a$) to stimulate the coherent coupling between signal mode ($b$) and microwave mode ($c$), and photons can be bidirectionally converted between optical and microwave frequencies with on-chip efficiency

$$\eta = \frac{\kappa_{b,ex} \kappa_{c,ex}}{\kappa_b \kappa_c} \frac{4C}{(1 + C)^2},$$

where $\kappa_{b,ex}$, $\kappa_b$, $\kappa_{c,ex}$, $\kappa_c$ are the external coupling and total loss rates for signal and microwave modes respectively, and $C = \frac{4n_a\varepsilon_0}{\kappa_a \kappa_c}$ is the cooperativity with $n_a$ the photon number in the pump mode (Supplementary Section I).

In experiments, the optical cavity is fabricated from an AlN layer on a silicon dioxide cladding on silicon wafer (Fig. 2A). The optical ring cavity has a radius of 120 $\mu$m, and cross-section of 2.0 $\mu$m x 0.8 $\mu$m (See Supplementary Section II for fabrication procedure and device cross-section). An azimuthal number difference of 1 between the pump and signal optical modes is chosen to mitigate the optical mode mixing induced by the non-vertical waveguide sidewalls (See Supplementary Section III for

Figure 1. **Coherent conversion with cavity electro-optics.** (A) Schematic of cavity electro-optic systems. The optical cavity is made of materials with Pockel nonlinearity ($\chi^{(2)}$), and placed in the capacitor of the LC circuit. At the same time, optical and microwave cavities are coupled to optical and microwave bus waveguides respectively. (B) Integrated superconducting cavity electro-optic device. The red part is the optical cavity and coupling waveguide, and the yellow part is the superconducting microwave cavity. A buffer layer (semi-transparent) is placed between optical devices and the superconducting cavity to prevent metallic absorption of optical photons. (C) Diagram of frequencies in the conversion process. Strong control light is applied to the TE optical mode (pump mode), and photons can be converted between the microwave mode and TM optical mode (signal mode). Microwave photons are converted to optical photons through sum frequency generation, and optical photons are converted to microwave photons through difference frequency generation as shown in insets. The mode distribution in the cross-section is shown for the microwave mode (D), TE optical mode (E), and TM optical mode (F). Arrow direction and length represent the electric field direction and strength in log scale respectively. Colors in (D) represent the voltage distribution, and colors in (E) and (F) represent the energy density. In simulation, the optical waveguide is 2 $\mu$m wide and 800 nm thick, and the sidewall angle is 8 deg. The distance between microwave electrodes is 2.8 $\mu$m. The material boundary is plotted in grey.

The principle of quantum transducers based on superconducting cavity electro-optics is illustrated in Fig.1A. The optical cavity, consisting of materials with Pockel nonlinearity, is placed inside the capacitor of the LC microwave resonator. The electric field across the capacitor changes the refractive index of the optical cavity, thus modulates the optical resonant frequency. Reversely, modulated optical fields can generate microwave field due to the optical mixing (rectification) in Pockels materials [21]. To implement the quantum transducer, we use integrated optical microring cavities made of AlN, which supports low loss optical modes and provides high electro-optic coefficients simultaneously [22] (Fig. 1B). Superconducting microwave resonators are placed on top of a thin buffer layer, and the capacitor shape is designed to match the optical cavity to maximize the field overlap between microwave and optical modes [15, 17].

We employ the triple-resonance scheme to enhance the coupling as shown Fig. 1C, with the interaction Hamiltonian written as

$$H_1 = \hbar g_{eo}(ab^\dagger c + a^\dagger bc^\dagger)$$

where $a$, $b$, $c$ are the annihilation operators for the optical pump and signal modes, and microwave mode respectively, and $g_{eo}$ is the vacuum electro-optic coupling rate. Cavity electro-optic systems with triple resonances have been proposed and demonstrated recently [15, 16]. However, the device geometry is limited to above several millimeters, as the pump and signal optical modes are from the same mode group, and the free spectral range (FSR) needs to match the microwave frequency. Large mode volume inevitably leads to small vacuum coupling rate, thus low conversion efficiency (Supplementary Section I). In contrast, our integrated approach uses the transverse-electric (TE) and transverse-magnetic (TM) optical modes as pump and signal modes respectively (Fig.1E & F), whose frequency difference equals the microwave frequency [23–26]. Thus the device size and mode volume can be further reduced without the limitation imposed by FSR. In this case, $r_{13}$ electro-optic coefficient is used, which also enables the use of TE-polarized microwave mode, thus the heterogeneous integration of planar microwave resonators with optical cavities (Fig. 1D). During experiments, a strong coherent field is applied to the pump mode ($a$) to stimulate the coherent coupling between signal mode ($b$) and microwave mode ($c$), and photons can be bidirectionally converted between optical and microwave frequencies with on-chip efficiency

$$\eta = \frac{\kappa_{b,ex} \kappa_{c,ex}}{\kappa_b \kappa_c} \frac{4C}{(1 + C)^2},$$

where $\kappa_{b,ex}$, $\kappa_b$, $\kappa_{c,ex}$, $\kappa_c$ are the external coupling and total loss rates for signal and microwave modes respectively, and $C = \frac{4n_a\varepsilon_0}{\kappa_a \kappa_c}$ is the cooperativity with $n_a$ the photon number in the pump mode (Supplementary Section I).

In experiments, the optical cavity is fabricated from an AlN layer on a silicon dioxide cladding on silicon wafer (Fig. 2A). The optical ring cavity has a radius of 120 $\mu$m, and cross-section of 2.0 $\mu$m x 0.8 $\mu$m (See Supplementary Section II for fabrication procedure and device cross-section). An azimuthal number difference of 1 between the pump and signal optical modes is chosen to mitigate the optical mode mixing induced by the non-vertical waveguide sidewalls (See Supplementary Section III for
Figure 2. Integrated superconducting cavity electro-optic device. (A) Optical image of the superconducting cavity electro-optic device. The scale bar is 100 μm. (B) TE optical spectrum with different DC voltages. The azimuthal number difference between the TE and TM optical modes is 1. The mode anti-crossing gap is $2g_x \sim 6.1$ GHz, and the original dissipation rates for TM and TE optical modes without optical mode mixing are 190 MHz and 480 MHz, respectively. (C) TE optical spectrum with DC voltages of -400 V (green), 300 V (red), 900 V (orange), corresponding to the green, red, and orange dash lines in (B) respectively. (D) Schematic of the microwave resonator and electric field distribution of the microwave mode, as well as the equivalent circuit. (E) Measured reflection spectrum of the microwave cavity.

The microwave resonator is made of NbTiN superconducting film with critical temperature around 14 K. The device is placed in a cryostat and cooled down to 2 K, and the device surface is covered by superfluid helium to introduce fast heat dissipation [27]. To allow electro-optic phase matching, it is important to shape the microwave resonator to have azimuthal number of 1, to match the azimuthal number difference between the pump and signal modes (Fig. 2D & Supplementary Section V). The capacitance part of the microwave resonator has a radius of 120 μm to match the optical cavity, and the distance between electrodes is 2.8 μm. Each inductance arm has a length of 1.5 mm, allowing far-field magnetic coupling to an off-chip loop probe for broadband microwave signal input and readout. The microwave mode has resonance around $\omega_c/2\pi = 8.31$ GHz with the decay rate of $\kappa_c/2\pi = 0.55$ MHz at 2 K (Fig. 2E). When the DC voltage is tuned to 297 V, the frequency difference between pump and signal modes is also around 8.31 GHz (Fig. 2C). Therefore, the phase matching and energy conservation are fulfilled simultaneously.

The coherent conversion of our device is first characterized with optical reflection spectrum. Strong control light is applied to the pump mode, and a weak probe light, derived from the control light by single side-band modulation, is sent to the signal mode (Supplementary Section VI). No obvious temperature change of the superconducting microwave resonator is observed (Supplementary Section VII). Figure 3A presents the probe light transmission spectrum sweeping across the signal mode, with a fixed control light on resonance with the pump mode. By tuning the DC voltage, the broad Lorentzian
Figure 3. Electromagnetically induced transparency with cavity electro-optics. (A) Measured optical reflection spectrum as a function of the modulation frequency. Each spectrum in (A) and (B) corresponds to a different DC voltage, thus different frequency detuning between pump and signal modes. Spectrums are offset for clarity. (C) Transparency window with the control light power of 8 dBm (blue), 5 dBm (orange), and 0 dBm (green). Circles are measured data, and solid lines are fitted spectra with Eq. S13 in Supplementary Section I. (D) Cooperativity and internal conversion efficiency versus control light power. The blue, orange, and green points correspond to the blue, orange, and green curves in (C) respectively. Grey lines are the fitted result based on measured data.

dip corresponding to the signal mode is shifted, with a sharp modification of the spectrum at a fixed frequency $\omega = \omega_c$. This modification originates from the destructive interference between two pathways for the probe light: directly passing through the optical cavity, and converting to microwave photons and then back to optical photons [28, 29]. As shown in the enlarged spectrum (Fig. 3B), when the signal mode frequency matches $\omega_c$, there is a sharp transparency window, with the bandwidth equal to $(1 + C) \kappa_c = (0.59 \pm 0.01)$ MHz. If the signal mode is detuned, the interference gives an asymmetric Fano-shape spectrum. By fitting the transparency window, the conversion cooperativity can be extracted, and internal conversion efficiency can be inferred (Fig. 3C & D) [30]. An internal efficiency as high as $(25.9 \pm 0.3)\%$ is achieved with our device under 8 dBm control light power (14 dBm total off-chip optical power with 6 dB insertion loss).

The bidirectional conversion is characterized by measuring the complete conversion matrix (Fig. 4A) with the DC voltage fixed at 297 V. By injecting the optical probe and monitoring the microwave output, we measure the microwave-to-optical coefficient $S_{oe}$, which is a Lorentzian lineshape centered at the microwave resonant frequency (Fig. 4C). The optical-to-microwave coefficient $S_{eo}$ is measured by reversing the input and output signal, which has the same spectrum shape with microwave-to-optical coefficient $S_{oe}$, indicating that the conversion is bidirectional (Fig. 4D). The optical and microwave reflection spectra are also measured for calibration (Fig. 4B & E), and the on-chip efficiency is estimated to be $(2.05 \pm 0.04)\%$ (Supplementary Section VIII). The main noise source during the conversion process is the thermal excitation of the microwave cavity, which can be reduced by working at lower temperature. And the noise generated by the parametric amplification process is negligible because of the deep resolved sideband condition (Supplementary Section IX).

Discussion

An ideal quantum transducer demands the coherent conversion efficiency approaching 100%, when the cooperativity equals unity and both microwave and optical modes are deeply over-coupled. By optimizing the fabrication process and material properties, the efficiency of our device can be further increased. For instance, optical quality factors above 2 million have been demonstrated with single crystalline AlN [31], therefore intrinsic loss for optical modes can be reduced to $\kappa_{b,i} = 2\pi \times 100$ MHz. Then, pump photon number can be increased by $\sim 100$ times with the same pump power, and the enhanced coupling rate $g_{eo}$ can reach 16 MHz. With an improved microwave intrinsic loss rate of $\kappa_{c,i} = 2\pi \times 10$ kHz [32], for example by using sapphire substrate, the optimal on-chip efficiency exceeding $\eta = 92\%$ can be achieved by choosing the external optical coupling rate $\kappa_{b,ex} = 2\pi \times 2.9$ GHz.
Figure 4. Bidirectional Frequency Conversion. (A) Schematic showing the full conversion process. The optical reflection $S_{oo}$ (B), microwave-to-optical conversion $S_{oe}$ (C), optical-to-microwave conversion $S_{eo}$ (D), and microwave reflection $S_{ee}$ (E) are measured in order to calibrate the on-chip conversion efficiency. The control light power is 8dBm, and the DC voltage is 297V. All conversion matrix coefficients are normalized to the RF output power of the network analyzer (Supplementary Section VI).

and microwave coupling rate $\kappa_{c,ex} = 2\pi \times 0.29 \text{ MHz}$. Therefore, with future development of single crystalline AlN on sapphire system and its adaption for superconducting resonators, the approach presented here is promising to realize high fidelity quantum state transduction between superconducting and photonic circuits.

**Conclusion**

We have demonstrated the coherent photon transduction between integrated superconducting and photonic circuits. High transduction efficiency is realized based on triple-resonance electro-optics principle. Besides high efficiency, low noise, and large bandwidth, the large frequency tunability makes it easy to interface with different quantum systems, and the planar structure allows the circuit-level integration of different quantum devices. All these features not only make our device an ideal quantum transducer, but also provide a scalable platform to synthesize different quantum systems, paving the route towards large scale hybrid quantum networks.

**Materials and Methods**

**Polycrystal AlN film preparation** 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. The sputtered AlN film is polycrystalline with $c$-axis highly oriented perpendicular to the substrate. As AlN has wurtzite crystal structure, the electro-optic coefficient has no direction dependence in the plane perpendicular to $c$-axis. Thus the $r_{13}$ coefficient in the sputtered AlN film can be used without considering the in-plane crystal direction change [22].

**NbTiN film preparation** NbTiN films are sputtered by the RF magnetron method at room temperature using 70% Nb and 30% Ti alloy target in an argon and nitrogen gas mixture. Anomalous NbTiN film is formed uniformly on the surface, and the superconducting critical temperature $T_c$ around 13.8 K is achieved.

**Supplementary Materials**

Theory of cavity electro-optics and its utility for microwave-to-optical conversion

Device fabrication procedure

Identifying phase matching conditions for optical modes

Influence of optical mode mixing on the vacuum coupling rate $g_{eo}$

Microwave resonator design

Measurement setup

Device temperature calibration

Efficiency calibration

Added noise during conversion

Fig.S1. Calculated internal conversion efficiency

Fig.S1. SEM picture of the cross section of a superconducting cavity electro-optic device

Fig.S3. Anti-crossing between TE and TM optical modes.

Fig.S4. Measured spectrum signature of mixing between TE and TM modes

Fig.S5. Wavelength difference between adjacent resonances

Fig.S6. Phase matching wavelength and anti-crossing strength

Fig.S7. Vacuum coupling rate with hybrid optical modes

Fig.S8. Microwave resonator simulation

Fig.S9. Experiment setup for microwave-to-optical photon conversion

Fig.S10. Microwave resonator performance under different temperature

**References and Notes**

* hong.tang@yale.edu

[1] R. Schoelkopf and S. Girvin, “Wiring up quantum systems,” Nature 451, 664 (2008).

[2] J. L. O’Brien, A. Furusawa, and J. Vučković, “Photonic quantum technologies,” Nature Photonics 3, 687 (2009).
[3] A. I. Lvovsky, B. C. Sanders, and W. Tittel, “Optical quantum memory,” Nature photonics 3, 706 (2009).
[4] Z.-L. Xiang, S. Ashhab, J. You, and F. Nori, “Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems,” Reviews of Modern Physics 85, 623 (2013).
[5] J. Clarke and F. K. Wilhelm, “Superconducting quantum bits,” Nature 453, 1031 (2008).
[6] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, “Linear optical quantum computing with photonic qubits,” Reviews of Modern Physics 79, 135 (2007).
[7] R. W. Andrews, R. W. Peterson, T. P. Purdy, K. Ci-cak, R. W. Simmonds, C. A. Regal, and K. W. Lehnert, “Bidirectional and efficient conversion between microwave and optical light,” Nature Physics 10, 321 (2014).
[8] J. Bohmann, A. Vainsencher, D. D.Awschalom, and A. N. Cleland, “Nanomechanical coupling between microwave and optical photons,” Nature Physics 9, 712 (2013).
[9] K. C. Balram, M. I. Davanço, J. D. Song, and K. Srinivasan, “Coherent coupling between radiofrequency, optical and acoustic waves in piezo-optomechanical circuits,” Nature photonics 10, 346 (2016).
[10] S. Barzanjeh, M. Abdí, G. J. Milburn, P. Tombesi, and D. Vitali, “Reversible optical-to-microwave quantum interface,” Physical Review Letters 109, 130503 (2012).
[11] F. Lecocq, J. B. Clark, R. W. Simmonds, J. Aumentado, and J. D. Teufel, “Mechanically mediated microwave frequency conversion in the quantum regime,” Physical review letters 116, 043601 (2016).
[12] M. Tsang, “Cavity quantum electro-optics,” Physical Review A 81, 063837 (2010).
[13] M. Tsang, “Cavity quantum electro-optics. ii. input-output relations between traveling optical and microwave fields,” Physical Review A 84, 043845 (2011).
[14] S. Huang, “Quantum state transfer in cavity electro-optic modulators,” Phys. Rev. A 92, 043845 (2015).
[15] C. Javerzac-Galy, K. Plekhanov, N. Bernier, L. Toth, A. Feroanov, and T. Kippenberg, “On-chip microwave-to-optical quantum coherent converter based on a superconducting resonator coupled to an electro-optic microresonator,” Physical Review A 94, 053815 (2016).
[16] A. Rueda, F. Sedlmeir, M. C. Collodo, U. Vogl, B. Stiller, G. Schunk, D. V. Strekalov, C. Marquardt, J. M. Fink, O. Painter, G. Leuchs, and H. G. L. Schwefel, “Efficient microwave to optical photon conversion: an electro-optical realization,” Optica 3, 597 (2016).
[17] M. Soltani, M. Zhang, C. Ryan, G. J. Ribeill, C. Wang, and M. Loncar, “Efficient quantum microwave-to-optical conversion using electro-optic nanophotonic coupled resonators,” Physical Review A 96, 043808 (2017).
[18] M. Hafezi, Z. Kim, S. Rolston, L. Orozco, B. Lev, and J. Taylor, “Atomic interface between microwave and optical photons,” Physical Review A 85, 020302 (2012).
[19] R. Hisatomi, A. Osada, Y. Tabuchi, T. Ishikawa, A. Noguchi, R. Yamazaki, K. Usami, and Y. Nakamura, “Bidirectional conversion between microwave and light via ferromagnetic magnons,” Phys. Rev. B 93, 174427 (2016).
[20] L. A. Williamson, Y.-H. Chen, and J. J. Longdell, “Magneto-optic modulator with unit quantum efficiency,” Physical review letters 113, 203601 (2014).

Acknowledgments We acknowledge funding support from an LPS/ARO grant (W911NF-14-1-0563), AFOSR MURI grant (FA9550-15-1-0029), NSF EFRI grant (EFMA-1640959) and DARPA ORCHID program through a grant from the Air Force Office of Scientific Research (AFOSR FA9550-10-1-0297), and the Packard Foundation. Facilities used were supported by Yale Institute for Nanoscience and Quantum Engineering and
NSF MRSEC DMR 1119826. The authors thank L. Jiang for discussion, and M. Power, M. Rooks and L. Frunzio for assistance in device fabrication. **Author contributions** H.X.T., L.F., C.-L.Z. & X.H. conceived the experiment; L.F. fabricated the device; L.F., R.C., X.G., Z.G., & S.W. performed the experiment; L.F. & C.-L.Z. analyzed the data. L.F. & C.-L.Z. wrote the manuscript, and all authors contribute to the manuscript. H.X.T. supervised the work. **Competing interests** The authors declare that they have no competing interests. **Data and materials availability** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.
Supplementary Information

CONTENTS

References 5

I. Theory of cavity electro-optics and its utility for microwave-to-optical conversion S1

II. Device fabrication procedure S3

III. Identifying phase matching conditions for optical modes S3

IV. Influence of optical mode mixing on the vacuum coupling rate $g_{eo}$ S5

V. Microwave resonator design S6

VI. Measurement setup S7

VII. Device temperature calibration S7

VIII. Efficiency calibration S8

IX. Added noise during conversion S9

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

The electro-optic effect is a second-order nonlinear process [30], thus the interaction Hamiltonian can be written as

$$H_I = \hbar g_{eo} (a + a^\dagger)(b + b^\dagger)(c + c^\dagger) \quad (S.1)$$

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 $a^\dagger b + ab$ is neglected. Thus, the interaction Hamiltonian becomes

$$H_I = \hbar g_{eo} (ab^\dagger + a^\dagger b)(c + c^\dagger) \quad (S.2)$$

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

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

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

$$H = \hbar \omega_a a^\dagger a + \hbar \omega_b b^\dagger b + \hbar \omega_c c^\dagger c + H_1 + i \sqrt{\kappa_{a,\text{ex}} P_a / \hbar \omega_p} (a^\dagger e^{-i\omega_p t} - ae^{i\omega_p t})$$

(S.5)

where $P_a$, $\omega_p$, and $\kappa_{a,\text{ex}}$ 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

$$\langle a \rangle = \sqrt{\kappa_{a,\text{ex}} / \kappa_a^2 / 4 + \delta_a^2} \times \sqrt{P_a / \hbar \omega_p} e^{i\phi},$$

(S.6)

with the corresponding pump cavity photon number

$$n_a = \langle a^\dagger a \rangle = |\langle a \rangle|^2$$

(S.7)

Here $\delta_a = \omega_a - \omega_p$ is the pump detuning, $\kappa_a = \kappa_{a,\text{ex}} + \kappa_{a,\text{i}}$ is the total decay rate, $\kappa_{a,\text{i}}$ is the intrinsic loss rate of mode $a$, and $\phi$ 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 $\hbar \omega_p b^\dagger b$ under rotating-wave approximation (RWA) can be simplified to

$$H = \hbar \delta_b b^\dagger b + \hbar \omega_c c^\dagger c + \hbar G(b^\dagger c + bc^\dagger)$$

(S.8)

with $G = \langle a \rangle g_{eo} = \sqrt{n_a g_{eo}}$, which is the enhanced coupling rate, and $\delta_b = \omega_b - \omega_p$. 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

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

(S.9)

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

(S.10)

$$b_{out} = b_{in} - \sqrt{\kappa_{b,\text{ex}}} b$$

(S.11)

$$c_{out} = c_{in} - \sqrt{\kappa_{c,\text{ex}}} c$$

(S.12)

where $b_{in}$ and $c_{in}$ denote the input signals, $\delta_{b,\text{in}} = \omega_b,\text{in} - \omega_p$ and $\omega_{c,\text{in}}$ are the angular frequency of inputs, $\kappa_b$ and $\kappa_c$ are the total decay rates, $\kappa_{b,\text{ex}}$ and $\kappa_{c,\text{ex}}$ 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

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

(S.13)

Thus, the bidirectional conversion efficiency can be written as

$$\eta = \left| \frac{b_{out}}{b_{in}} \right|^2 = \left| \frac{c_{out}}{c_{in}} \right|^2 = \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{ex}}}{\kappa_b \kappa_c} \left( C + \frac{4C}{\kappa_b} \right) \left( 1 + \frac{2i(\delta_b - \omega_p)}{\kappa_b} \right) \left( 1 + \frac{2i(\omega_c - \omega_p)}{\kappa_c} \right)^2$$

(S.14)

Figure S1. Calculated internal conversion efficiency.
Figure S2. SEM 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 \(\mu m\).

where \(C = \frac{4G^2}{\kappa_b \kappa_c}\) is defined as the system cooperativity, and \(\omega\) 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 \(c\) \((\omega_p = \omega_b - \omega_c)\), and the on-chip conversion efficiency is

\[
\eta = \frac{\kappa_{b,ex} \kappa_{c,ex}}{\kappa_{b} \kappa_{c}} \times \frac{4C}{(1 + C)^2},
\]

(S.15)

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 \(\frac{\kappa_{b,ex} \kappa_{c,ex}}{\kappa_{b} \kappa_{c}}\), and the internal conversion efficiency which is

\[
\eta_i = \frac{4C}{(1 + C)^2}.
\]

(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).

II. DEVICE FABRICATION PROCEDURE

The 800\(\text{nm}\) AlN film is grown on a Si wafer with 2-\(\mu m\)-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\(\text{nm}\) SiO\(_2\) is deposited with plasma-enhanced chemical vapor deposition (PECVD), followed by device annealing at 940 \(^\circ\text{C}\). 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 between 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 \(\text{nm}\)) match our design.

III. 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
Figure 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.

Figure 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.

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.

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 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.

Figure 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 \( \mu m \).

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.

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.

**IV. 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

\[
H_x = \hbar g_x (a^\dagger b + ab^\dagger).
\]

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

\[
\begin{pmatrix}
A \\
B
\end{pmatrix} = \begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix} \begin{pmatrix}
a \\
b
\end{pmatrix}
\]

(S.18)
Figure 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.

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

$$\tan 2\theta = \frac{2g_x}{\omega_a - \omega_b}.$$ (S.19)

The eigenfrequencies of the hybrid optical modes are

$$\omega_A = \frac{\omega_a + \omega_b}{2} + \sqrt{g_x^2 + \left(\frac{\omega_a - \omega_b}{2}\right)^2},$$ (S.20)

$$\omega_B = \frac{\omega_a + \omega_b}{2} - \sqrt{g_x^2 + \left(\frac{\omega_a - \omega_b}{2}\right)^2}.$$ (S.21)

The minimum frequency difference is therefore $2g_x$, which corresponding to the frequency gap in the anti-crossing spectrum.

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

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

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

$$H_{1,m} = \hbar g_{eo} \cos 2\theta (A^\dagger B + AB^\dagger)(c + c^\dagger).$$ (S.23)

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

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

When the detuning between the two original optical modes is zero, thus $\theta = 45^\circ$, 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 $g_{eo} \cos 2\theta$ (Fig. S7c). When the detuning is much larger than the anti-crossing strength, the vacuum coupling rate approaches the original value $g_{eo}$.

V. **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 $\phi$, 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 $|m_a - m_b| = 1$.

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.

VI. 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.

VII. 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 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 2.0K, 2.5K, and 3.0K are 0.64MHz, 0.58MHz, and 0.54 MHz respectively. As shown in Fig.2E and Fig.4E in the main text, the microwave resonance linewidth is around 0.55MHz under 8dBm optical pump, indicating that the effective temperature of the device is around 2.1K.
The cooperativity $C$ of $0.075 \pm 0.001$ is estimated from the fitting of the blue curve in Fig. 3C in the main text, leading to internal efficiency $\eta_i$ 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 $\eta_{i,ex} = \frac{\kappa_{i,ex}}{\kappa_i} = \frac{1-|R_i|^2}{2}$ assuming under-coupled condition with $i = b, c$ for optical signal and microwave modes respectively. From Fig. 2C & 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 $\eta_{b,ex} = \frac{\kappa_{b,ex}}{\kappa_b} = 0.313 \pm 0.002$ and $\eta_{c,ex} = \frac{\kappa_{c,ex}}{\kappa_c} = 0.261 \pm 0.001$ respectively. Thus the on-chip efficiency can be estimated $\eta = \eta_{i,ex} = (2.11 \pm 0.03)\%$. 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.

Based on the cooperativity $C$ and the resonance linewidth, the enhanced coupling strength can be estimated $G_{eo} = 2\pi \times (1.76 \pm 0.05)\text{MHz}$. 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 (supplementary material of Ref. [34]). Therefore the effective vacuum coupling rate estimated from experimental results is

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

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 $r_{113} = 1 \text{pm/V}$ [21], the vacuum coupling rate for our device can be calculated as $g_{eo} = 2\pi \times (520 \pm 50)\text{Hz}$, 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
\[ g_{\text{eff}} = g_{\text{eo}} \cos(2\theta) = 2\pi \times (330 \pm 30) \text{ Hz}, \]  
(S.26)
which agrees with our experimental estimation.

IX. ADDED NOISE DURING CONVERSION

In addition to the conversion efficiency \( \eta \), 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 \( hG(b^\dagger c + bc) \).

By including the counter-rotating term \( b^\dagger c^\dagger + bc \), we have the
\[
- i \omega_b (\omega) = - \left( i \delta_b + \frac{\kappa_b}{2} \right) b(\omega) - i Gc(\omega) - i Gc^\dagger (-\omega) + \sqrt{\kappa_{b,\text{ex}}} b_{\text{in}}(\omega) \delta (\omega - \delta_{b,\text{in}}) + \sqrt{\kappa_{b,\text{in}}} b_n(\omega),
\]
\[ \text{(S.27)} \]
\[
- i \omega_{b^\dagger} (-\omega) = - \left( -i \delta_b + \frac{\kappa_b}{2} \right) b^\dagger (-\omega) + i Gc(\omega) + i Gc^\dagger (-\omega) + \sqrt{\kappa_{b,\text{ex}}} b_{\text{in}}^\dagger (\omega) \delta (\omega - \delta_{b,\text{in}}) + \sqrt{\kappa_{b,\text{in}}} b_n^\dagger (-\omega),
\]
\[ \text{(S.28)} \]
\[
- i \omega_c (\omega) = - \left( i \omega_c + \frac{\kappa_c}{2} \right) c(\omega) - i Gb(\omega) - i Gb^\dagger (-\omega) + \sqrt{\kappa_{c,\text{ex}}} c_{\text{in}}(\omega) \delta (\omega - \omega_{c,\text{in}}) + \sqrt{\kappa_{c,\text{in}}} c_n(\omega),
\]
\[ \text{(S.29)} \]
\[
- i \omega_{c^\dagger} (-\omega) = - \left( -i \omega_c + \frac{\kappa_c}{2} \right) c^\dagger (-\omega) + i Gb(\omega) + i Gb^\dagger (-\omega) + \sqrt{\kappa_{c,\text{ex}}} c_{\text{in}}^\dagger (-\omega) \delta (\omega - \omega_{c,\text{in}}) + \sqrt{\kappa_{c,\text{in}}} c_n^\dagger (-\omega),
\]
\[ \text{(S.30)} \]

Here, \( b_n(\omega) \) and \( c_n(\omega) \) are the noise from the intrinsic loss channels of the optical and microwave cavities respectively. The noise correlators can be expressed as following
\[
\langle b^\dagger_n(\omega) b_n(\Omega) \rangle = 0 \tag{S.31}
\]
\[
\langle b_n(\Omega) b^\dagger_n(\omega) \rangle = \delta (\omega - \Omega) \tag{S.32}
\]
\[
\langle c^\dagger_n(\omega) c_n(\Omega) \rangle = n_c \delta (\omega - \Omega) \tag{S.33}
\]
\[
\langle c_n(\Omega) c^\dagger_n(\omega) \rangle = (n_c + 1) \delta (\omega - \Omega) \tag{S.34}
\]

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 condition \( \delta_b = \omega_c \)
\[
\langle b^\dagger_{\text{out}}(\omega) b_{\text{out}}(\omega) \rangle = \eta_{c\rightarrow b} \langle c^\dagger_{\text{in}}(\omega) c_{\text{in}}(\omega) \rangle + N_{b,\text{add}} + N_{b,\text{para}}, \tag{S.35}
\]
\[
\langle c^\dagger_{\text{out}}(\omega) c_{\text{out}}(\omega) \rangle = \eta_{b\rightarrow c} \langle b^\dagger_{\text{in}}(\omega) b_{\text{in}}(\omega) \rangle + N_{c,\text{add}} + N_{c,\text{para}}, \tag{S.36}
\]

with the conversion efficiency
\[
\eta_{c\rightarrow b} = \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{ex}}}{\kappa_b \kappa_c} \frac{4C}{(1 + C)^2} + \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{ex}}}{\kappa_b \kappa_c} \frac{C}{(C + 1)^4} \left[ (C^2 + 2C) \frac{\kappa_b^2}{4\omega_b^2} + (1 + 4C + 2C^2) \frac{\kappa_c^2}{4\omega_c^2} + 2C \frac{\kappa_{c,b}}{4\omega_c^2} \right] + O(1/\omega_c^3), \tag{S.37}
\]
\[
\eta_{b\rightarrow c} = \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{ex}}}{\kappa_b \kappa_c} \frac{4C}{(1 + C)^2} + \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{ex}}}{\kappa_b \kappa_c} \frac{C}{(C + 1)^4} \left[ (C^2 + 2C) \frac{\kappa_b^2}{4\omega_b^2} + (1 + 4C + 2C^2) \frac{\kappa_c^2}{4\omega_c^2} + 2C \frac{\kappa_{c,b}}{4\omega_c^2} \right] + O(1/\omega_c^3), \tag{S.38}
\]

The added noise due to the thermal environment is
\[
N_{b,\text{add}} = \frac{4C}{(1 + C)^2} \frac{\kappa_{b,\text{ex}} \kappa_{c,\text{in}}}{\kappa_b \kappa_c} n_c, \tag{S.39}
\]
\[
N_{c,\text{add}} = \frac{4}{(1 + C)^2} \frac{\kappa_{c,\text{in}} \kappa_{c,\text{ex}}}{\kappa_c^2} n_c. \tag{S.40}
\]
And the noise due the counter-rotating term is

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

\[
N_{c,\text{para}} = \frac{C}{(C + 1)^2} \frac{\kappa_{c,\text{ex}}}{\kappa_c} \frac{C\kappa_c^2 + C\kappa_b^2}{4\omega_c^2},
\]  
(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 \(\kappa_b/2\pi = 294\,\text{MHz}\), \(\kappa_c/2\pi = 0.55\,\text{MHz}\), and \(\omega_c/2\pi = 8.31\,\text{GHz}\), we estimated the modification of the conversion efficiency \(\eta \to \eta (1 + A)\) due to the parametric amplification factor \(A \sim \frac{(C^2 + 2C)\kappa_b^2}{4(1+C)^2 \omega_c^2} = 1.0 \times 10^{-5}\) 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 \(T \sim 2\,\text{K}\), thus \(N_{b,\text{add}} \sim 0.3\) and \(N_{c,\text{add}} \sim 3\), which can be further reduced by working at lower temperature. In addition, the noise due to parametric amplification can be estimated, \(N_{b,\text{para}} \sim 5 \times 10^{-7}\) and \(N_{c,\text{para}} \sim 5 \times 10^{-6}\), which are negligible in our current experiments.