Silicon-chip source of bright photon pairs

Wei C. Jiang,1,5 Xiyuan Lu,2,5 Jidong Zhang,3 Oskar Painter,4 and Qiang Lin1,3,∗

1Institute of Optics, University of Rochester, Rochester, New York, USA
2Department of Physics and Astronomy, University of Rochester, Rochester, New York, USA
3Department of Electrical and Computer Engineering, University of Rochester, New York, USA
4Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California, USA
5These authors contributed equally to this work
∗qiang.lin@rochester.edu

Abstract: Integrated quantum photonics relies critically on the purity, scalability, integrability, and flexibility of a photon source to support diverse quantum functionalities on a single chip. Here we report a chip-scale photon-pair source on the silicon-on-insulator platform that utilizes dramatic cavity-enhanced four-wave mixing in a high-Q silicon microdisk resonator. The device is able to produce high-quality photon pairs at different wavelengths with a high spectral brightness of $6.24 \times 10^7$ pairs/s/mW$^2$/GHz and photon-pair correlation with a coincidence-to-accidental ratio of $1386 \pm 278$ while pumped with a continuous-wave laser. The superior performance, together with the structural compactness and CMOS compatibility, opens up a great avenue towards quantum silicon photonics with capability of multi-channel parallel information processing for both integrated quantum computing and long-haul quantum communication.

© 2015 Optical Society of America

OCIS codes: (130.3120) Integrated optics devices; (140.3948) Microcavity devices; (190.4380) Nonlinear optics, four-wave mixing; (270.0270) Quantum optics.

References and links


Quantum photonic technology has found broad applications ranging from quantum communication [1], quantum computing [2], to quantum metrology [3]. Recent advances in integrated quantum photonics [4–6] show great promise for chip-scale quantum information processing with tremendous complexity. A bright, single-mode, high-purity, and integrated source of single photons and/or entangled photon pairs is essential for all these applications [7–10], particularly for integrated quantum photonic circuits which rely critically on the purity, scalability, integrability, and flexibility of the photon source to support diverse quantum functionalities on a single chip.

To date, nearly all photon-pair sources are based upon spontaneous parametric down conversion (SPDC) or four-wave mixing (FWM) in nonlinear optical crystals/waveguides [8–10]. Bulk crystals emit photon pairs into a multimode spatial profile, resulting in a fairly low photon generation/collection efficiency [10–12]. Significant efforts have been devoted in recent years to developing waveguide sources for single-mode emission [13–17]. However, the pro-
duced photon pairs generally exhibit a non-factorable spectrum which degrades considerably the quantum-state purity [8]. FWM in silica optical fibers/waveguides recently appears as a promising approach with a great flexibility of engineering photon spectrum [18–25], which, unfortunately, suffers seriously from broadband Raman scattering of silica [21, 22, 26].

Silicon exhibits a strong Kerr nonlinearity for nonlinear optical interaction and a large refractive index enabling tight mode confinement, which has been explored intensively recently for a variety of applications [27–30]. In particular, single-crystalline silicon has a clean phonon spectrum with Brillouin-zone-center phonons of a well-defined frequency of 15.6 THz and a narrow linewidth of \(\sim 105\) GHz, which eliminates the deleterious broadband Raman noises [27, 31]. These superior features together with mature nanofabrication technology make silicon an ideal platform for integrated quantum photonic application [32–47].

Here we propose and demonstrate a bright high-purity chip-scale photon-pair source on the silicon-on-insulator (SOI) platform. By taking advantage of the dramatic cavity enhanced FWM in a high-quality silicon microdisk resonator, we are able to achieve a spectral brightness of \(6.24 \times 10^7\) pair/s/mW/GHz which is orders of magnitude larger than other FWM-based photon-pair sources [18–25,32–47], and a strong photon-pair correlation with a coincidence-to-accidental ratio of \(1386 \pm 278\) while pumped with a continuous-wave (CW) laser. The unique device characteristics enable simultaneous generation of multiple photon pairs at different wavelengths, thus significantly extending the wavelength management capability for integrated quantum photonics. The superior device performance together with its CMOS compatibility opens the door towards quantum silicon photonics with great potential for multi-channel parallel operation of novel quantum photonic functionalities on chip.

The device structure we employ is a compact silicon microdisk resonator sitting on a silica pedestal (Fig. 1(a)). An excellent feature of a microdisk is that its group-velocity dispersion is dominantly determined by the disk thickness, allowing flexible engineering of the zero-dispersion wavelength (ZDWL) to the useful telecom band (Fig. 1(b)). Detailed analysis shows that the dispersion characteristics support phase matching for the FWM process over a broad spectrum \(\sim 100\) nm, thus enabling simultaneous generation of multiple photon pairs over all phase-matched cavity modes. In particular, the microdisk resonator uniquely supports multiple mode families with different radial orders (Fig. 1(b), inset). As they experience different mode confinement introduced by the whispering-gallery geometry, these mode families exhibit ZDWLs shifted by about 30–40 nm from each other (Fig. 1(b)). Consequently, a variety of desired photon-pair modes can be produced inside a single device simply by pumping at different mode families. Moreover, all photon modes produced transmit out of the cavity into the same single mode of the coupling waveguide (Fig. 1(a)), regardless of the mode family employed inside the microdisk.

Most importantly, the high-Q cavity drastically modifies the density of states of vacuum, resulting in significant Purcell enhancement on cavity modes which not only dramatically increases the generation efficiency of photon pairs but also ensures them to be created at discrete single frequencies with a high state purity [48]. With \(N_p\) pump photons inside the cavity, the probability of emitting a pair of signal and idler photons at time \(t_s\) and \(t_i\) into the coupling waveguide (Fig. 1(a)) is given by (see Appendix A)

\[
p_{s,i}(t_s,t_i) = \frac{\Gamma_{es}\Gamma_{ei}}{2} (gN_p)^2 e^{-\frac{1}{2} \Gamma_{ei}|t_s-t_i|},
\]

with \(\Gamma_{ij} = \Gamma_{is} \text{ (or } \Gamma_{ii})\) on the exponent when \(t_i \geq t_s\) (or \(t_s < t_i\), where \(\Gamma_{ej}\) and \(\Gamma_{ij}\) (j = s, i) are the external coupling rate and photon decay rate of the loaded cavity, for the signal and idler photons, respectively, and \(\bar{\Gamma} = (\Gamma_{is} + \Gamma_{ii})/2\) represents the average. \(g\) is the vacuum coupling rate of the FWM process, which is proportional to \(\frac{1}{V}\) where \(n_2\) and \(V\) are the Kerr nonlinear coefficient and the effective mode volume, respectively.
Fig. 1. a, Schematic of generating photon pairs from a silicon microdisk resonator sitting on a silica pedestal. All photon modes couple to the same propagation mode inside the delivery coupling waveguide, which is a single-mode tapered silica fiber. b, Group-velocity dispersion of vertically polarized transverse-magnetic-like (TM-like) cavity modes with four different radial orders, termed as $P_0$, $P_1$, $P_2$, and $P_3$, respectively, for a silicon microdisk with a thickness of 260 nm and a radius of 4.9 µm, simulated by the finite element method. The insets show the simulated optical field profiles.

As indicated by Eq. (1), the signal and idler photons are created in a time scale of the cavity photon lifetime within which the photon pair remains highly correlated. A long photon lifetime leads to a large pair emission probability, with a photon-pair emission rate given by

$$R_c = \frac{\Gamma_{es} \Gamma_{ei}}{\Gamma_{ts} \Gamma_{ti}} \frac{2 (gN_p)^2}{\bar{\Gamma}}. \quad (2)$$

FWM creates photon pairs inside the cavity at a rate of $2 (gN_p)^2 / \bar{\Gamma}$, which are delivered to the coupling waveguide with a pair extraction efficiency of $\Gamma_{es} \Gamma_{ei} / (\Gamma_{ts} \Gamma_{ti})$. A high cavity Q, a large Kerr nonlinearity, and a small effective mode volume would result in strong photon-pair emission.

The silicon microdisk resonator is fabricated from a standard SOI wafer by use of e-beam lithography to define the device pattern, fluorine-based plasma to etch the silicon layer, and hydrofluoric acid to undercut the silica pedestal (see the inset of Fig. 2). The device is tested using the experimental setup shown in Fig. 2. The pump laser is coupled into the device via a single-mode tapered silica fiber. Before coupling into the device, the pump laser passes through a bandpass filter and a coarse wavelength-division multiplexer (CWDM MUX) to cut the laser noises. The CWDM MUX has a 3-dB bandwidth of 17 nm for each of its transmission bands whose center wavelengths are separated by 20 nm apart with a band isolation of $> 120$ dB. The bandpass filter is identical to one band of the CWDM MUX. For easy separation of individual photon modes within the generated photon pairs, the device is designed with a free spectral range (FSR) close to 20 nm such that the cavity resonances fit directly into the transmission bands of the CWDM. The produced multiple photon pairs are thus separated into individual photon modes by the CWDM DEMUX which is identical to the CWDM MUX used at the input end. The photoluminescence spectra of the photon pairs are recorded at each transmission port of the CWDM DEMUX for easy suppression of the pump wave. For coincidence counting, the photon pairs are recorded by gated InGaAs single photon detectors (SPDs). Two tunable bandpass filters with a 3-dB bandwidth of 1.2 nm are used to cut the Raman noises produced by the delivery silica fiber (see Appendix C). The pump power is controlled by variable optical...
Fig. 2. Schematic of the experimental setup for optical spectroscopy and coincidence photon counting. The inset shows the scanning electron microscopic image of the fabricated silicon microdisk with a thickness of 260 nm and a radius of 4.9 µm.

The constant FSR together with a high optical Q ensures efficient generation of photon pairs. As shown by the photoluminescence (PL) spectra in Fig. 4a, pumping at the P₂ mode located at 1515.6 nm is able to produce a set of four clean photon modes within the P₂ mode family (Fig. 4(a)). Moreover, switching to pump at the P₁ mode at 1546.3 nm is able to produce another set of clean photon modes within the P₁ mode family (Fig. 4(b)). For the simplicity of notation, we term the photon mode pairs in the two combs as P₁, P₂, W, and P₀, respectively (Fig. 4(a) and 4(b)). The spectrum of each photon mode is so sharp that it is beyond the resolution of our spectrometer (∼ 0.135 nm), implying the high coherence of generated photons. The amplitude difference within a photon comb is primarily due to different external couplings of cavity modes to the tapered fiber. When the pump mode is critically coupled to the cavity, the signals at shorter wavelengths are under coupled while the idlers at longer wavelengths are over coupled, resulting in a higher photon extraction efficiency for the idler within each mode pair (Fig. 4(a) and 4(b)). Note that the PL spectra also show some other modes with small amplitudes whose physical origin is not clear at this moment. Potential origins are multiphonon scattering in silicon [49] or light scattering from certain defects on the device surface. As these tiny modes are well separated from the FWM-created photon modes, they have negligible effect on the photon-pair performance.
Fig. 3. The transmission spectrum of the passive cavity scanned by two tunable lasers operating at different spectral ranges (indicated as blue and green). The two sets of arrows indicate cavity modes with constant mode spacings within the $P_1$ and $P_2$ mode families, respectively. The insets show detailed transmission spectra of a $P_2$ mode at 1497.2 nm and a $P_1$ mode at 1564.9 nm, with theoretical fitting shown in red. See Appendix A for the discussion of cavity mode splitting.

Fig. 4. a and b, The photoluminescence spectra of the $P_2$ and $P_1$ photon-pair comb, respectively, with the pump nearly critically coupled to the cavity. The colors indicate the individual spectra recorded at different transmission ports of the CWDM DEMUX. The four mode pairs in the two combs are termed as $P_2^I$, $P_2^{II}$, $P_1^I$, and $P_1^{II}$, respectively, as indicated on the figures.
To characterize the performance of photon pairs, we carried out coincidence photon counting for each mode pair when the device is pumped with a CW laser (with a linewidth of ∼200kHz). As shown by Fig. 5(e)-5(h), the coincidence counts for all mode pairs depend quadratically on the pump power, indicating they are all produced through the FWM process. The experimentally recorded pair fluxes (blue dots) are close to the theoretical predictions (red curves) (see Appendix A and B). No significant saturation is observed, indicating the negligible impact of free-carrier absorption and two-photon absorption. In particular, a small input pump power of 79 µW is able to generate a photon-pair flux of $8.55 \times 10^{4}$ pair/s at $P_{I}^{2}$ (Fig. 5(e)), clearly showing the high efficiency of the device. A metric characterizing the photon-pair generation efficiency is the spectral brightness, defined as the pair flux per unit spectral width per unit pump power square (since FWM depends quadratically on the pump power). The spectral width of each photon mode is determined by the linewidth of the loaded cavity, which is 1.59 and 2.80 GHz for the $P_{I}^{1}$ and $P_{II}^{1}$ mode pairs at 1497.2 and 1534.3 nm, respectively. It infers that the spectral brightness of the generated photon pairs is $6.24 \times 10^{7}$ pair/s/mW$^2$/GHz, which is orders of magnitude larger than any other FWM-based photon-pair sources reported up to date [18–25, 32–47]. $P_{I}^{1}$, $P_{I}^{2}$ and $P_{II}^{1}$ exhibit similar spectral brightness, with magnitudes of $1.78 \times 10^{7}$, $2.32 \times 10^{7}$, and $2.57 \times 10^{7}$ pair/s/mW$^2$/GHz, respectively. Only the state-of-the-art cavity-enhanced SPDC sources [50–53] are able to exhibit a higher spectral brightness, which, however, are based upon either macroscopic optical cavities or millimeter-size resonator that are challenging for chip-scale integration.

An important figure of merit characterizing the quantum correlation between the photon pairs is the coincidence-to-accidental ratio (CAR) [15, 20, 22, 35], which is directly related to the quantum-interference visibility of the entangled states constructed from the photon pairs [26]. As shown in Fig. 5(i)-5(l), all photon mode pairs exhibit very high CARs (blue dots) at all power levels. For the $P_{I}^{2}$ mode pair, the device is able to achieve CAR with a value of $274 \pm 10$ at a pair flux of $8.55 \times 10^{4}$ pair/s and a peak value of $1386 \pm 278$ at a pair flux of $2.35 \times 10^{4}$ pair/s. Similar CAR is achieved for $P_{II}^{1}$, with values of $1097 \pm 165$ and $244 \pm 6$ at pair fluxes of $2.43 \times 10^{4}$ and $6.31 \times 10^{4}$ pair/s, respectively. At the same time, $P_{II}^{1}$ and $P_{II}^{2}$ exhibit as well relatively high CAR, with peak values of $733 \pm 106$ and $679 \pm 74$, respectively. For all photon mode pairs, the power-dependent trend of CAR follows closely the theory (red curves). The saturation of CAR at low photon-pair flux is due to detector dark counts as well as the Raman noises produced inside the delivery fibers (see Appendix C). The achieved CARs are significantly higher than those of other FWM-based photon-pair sources [18–25, 32–47], even without using pulsed pump to suppress noise photons or delicate superconducting SPDs to reduce detector dark counts, thus showing the superior quality of our generated photon pairs.

The high pair correlation and high coherence of the photon pairs allow resolving the temporal correlation between the signal and idler photons. This is particularly enabled by the tapered-fiber coupling system which provides flexible control of the external light coupling and thus the cavity photon lifetime. Figure 6 shows the normalized coincidence spectra of $P_{I}^{2}$ with two different external coupling conditions. When the tapered fiber is close to the device, the external coupling is strong which results in a short photon lifetime of 65 ps (averaged between signal and idler). As a result, the coincidence spectrum (blue curve) shows a full width at half maximum (FWHM) of 370 ps, which is 19% larger than the time resolution of the instrument response function (IRF, gray curve) of our coincidence counting system. However, moving the tapered fiber away from the device increases the photon lifetime considerably to 171 ps, leading to a coincidence spectrum (green curve) with a FWHM of 453 ps which is 45% larger than the time resolution of IRF. The experimentally observed time correlation is explained well by the theory (Appendix B). To the best of our knowledge, this is the first time to resolve the temporal photon correlation for a chip-scale photon-pair source, although it was observed before in
Fig. 5. Recorded photon emission fluxes, pair emission fluxes, and CARs for different photon-pair modes (P_I^I, P_{II}^I, P_I^I, and P_{II}^I). a-d, Recorded emission fluxes of individual photon modes within each mode pair. The signal (idler) at the shorter (longer) wavelength is shown as blue (red). Solid curves show the theoretical prediction of pure FWM-created photon fluxes. e-h, Recorded emission fluxes of correlated photon pairs. In each figure, the left and right vertical axes show the emission flux and the corresponding true coincidence count per gate, respectively, indicated by brown and green colors. The solid curves show the theoretical prediction. i-l, Recorded CAR for each mode pair. No accidental coincidence is subtracted. The red solid curves show the theoretical prediction. The black dashed curves are calculated from independently recorded true coincidence counts and photon fluxes of individual modes (Appendix B). In all photon counting measurements, the detector gating width is set to 10 ns and the deadtime is equal to the clock period. The InGaAs SPDs have dark counts of $5.09 \times 10^{-5}$ and $3.18 \times 10^{-5}$ per gate, respectively, at a quantum efficiency of 15% and a clock frequency of 250 kHz. The after-pulsing probability is less than 8%.
Fig. 6. Normalized coincidence spectra for P\textsubscript{I}\textsuperscript{2}. The blue and green curves show the cases with different external coupling conditions and thus different averaged photon lifetimes. The gray curve shows the instrument response function (IRF) of our coincidence counting system, with a FWHM of 312 ps which primarily come from the timing jitters of the two InGaAs SPDs. The red and purple curves show the theoretical prediction. The coincidence spectra were recorded with a detector gating width of 10 ns, a quantum efficiency of 25%, and a clock frequency of 125 kHz. The coincidence counter has a time resolution of 4 ps.

Cavity-enhanced SPDC sources [50–53]. Quantum coherence between correlated photon pairs is essential for constructing a variety of quantum functionalities [1, 2, 9]. The photon lifetime of 171 ps in our device corresponds to a propagation length of \(\sim 1.5\) cm, which is about the footprint of a photonic chip. Therefore, the produced photon pairs are able to share temporal coherence over the entire chip, which would significantly improve the quality of quantum functionalities built upon.

In summary, we have demonstrated a bright high-purity on-chip photon pair source on the SOI platform. The superior performance of the demonstrated device offers a great opportunity for realizing a variety of high-quality quantum functionalities on chip. Straightforward applications include producing quantum entanglement in various degrees of freedom, which are essential for quantum communication and quantum computing [1, 2]. Compared with conventional approaches, the high photon coherence in the device would significantly improve the purity of the constructed quantum states. As the cavity enhancement only occurs within the cavity photon lifetime, the demonstrated approach is able to operate with a repetition rate above gigahertz by use of pulsed pumping. Further increase of photon-pair coherence and spectral brightness can be obtained by improving optical Q. Such a high coherence would enable teleporting quantum entanglement intra-chip, inter-chip, or off-chip over far distance free from the stringent synchronization requirement in conventional schemes [1, 9]. The direct photon coupling into optical fibers with negligible losses (Fig. 2) is particularly suitable for this application. Moreover, the simultaneous generation of multiple photon pairs makes readily available wavelength-division multiplexing in quantum regime, and thus significantly extends the spectral efficiency and wavelength management capability of integrated quantum photonics. In particular, it exhibits great potential for realizing multipartite entanglement [9] and cluster state [55] inside a single device, which would dramatically enhance the capability of integrated quantum comput-
ing. On the other hand, the high coherence and spectral flexibility of the demonstrated approach enable excellent interfacing with other chip-scale information processing/storage elements [7]. These features together with the developed silicon photonic technology [29, 30] offer a great opportunity to ultimately form a complete CMOS compatible chip-scale platform for quantum information processing.

Appendix

A. Theory of photon-pair generation in a microdisk resonator

In this section, we provide the theory describing photon pair generation in a silicon microdisk resonator. We consider only the Kerr nonlinear optical interaction and neglect two-photon absorption and free-carrier effect because of their negligible effect in our device. We consider first the case shown in Fig. 7 where a pump wave launched into the resonator produces a pair of signal and idler photon which are transmitted out into the coupling waveguide. The Kerr nonlinear interaction among the three cavity modes at frequencies $\omega_0$ ($j = p, s, i$) can be described by a Hamiltonian $H = H_0 + H_I$, where $H_0$ describes the passive cavity modes coupled to the bus waveguide and $H_I$ describes the Kerr nonlinear interaction [56] inside the cavity, with the following forms

$$H_0 = \sum_{j=p,s,i} \left\{ \hbar \omega_0 a_j^\dagger a_j - \hbar \sqrt{\Gamma_{ej}} \left[ a_j^\dagger b_j e^{-i\omega_0 t} + b_j^\dagger a_j e^{i\omega_0 t} \right] \right\}, \quad (3)$$

$$H_I = -\frac{\hbar}{2} g_p |a_p|^2 a_p^2 - 2\hbar a_p^\dagger a_p \left( g_{ps} a_s^\dagger a_s + g_{pi} a_i^\dagger a_i \right) - \hbar g_{psi} a_p^\dagger a_s^\dagger a_i^\dagger a_p^2 - \hbar g_{psi}^* (a_p^\dagger)^2 a_s a_i, \quad (4)$$

where the intracavity field operator $a_j$ ($j = p, s, i$) is normalized such that $a_j^\dagger a_j$ represents the photon number operator and $b_j$ is the field operator of the incoming wave at carrier frequency $\omega_0$ inside the coupling waveguide normalized such that $b_j^\dagger b_j$ represents the operator of the input photon flux. $\langle b_j^\dagger b_j \rangle = \langle b_j^\dagger b_j \rangle = 0$ since only the pump is launched into the cavity. $b_j$ satisfies the commutation relation of $[b_j(t), b_j^\dagger(t')] = \delta(t-t')$. $\Gamma_{ej}$ is the external coupling rate of the cavity mode at $\omega_0$.

In Eq. (4), the first and second terms describe the self-phase and cross-phase modulation (SPM and XPM) from the pump mode, respectively, and the third and fourth terms govern the four-wave mixing (FWM) process. We have assumed the signal and idler are much weaker than the pump so that we can neglect the SPM and XPM introduced by them. $g_p, g_{pj}$ ($j = s, i$), and $g_{psi}$ are the vacuum coupling rate for the SPM, XPM, and FWM, respectively. They are given

Fig. 7. Schematic of photon pair generation in a silicon microdisk resonator.
by the following general expression

\[
g_{ijkl} = \frac{3\hbar}{4\epsilon_0 n_i n_j n_k n_l V_{ijkl}} \chi^{(3)} (-\omega_{ki}; \omega_{ij}, -\omega_{jk}, \omega_{kl}), \tag{5}
\]

where \(n_v (v = i, j, k, l)\) is the refractive index of silicon material at \(\omega_{kv}\) and \(\chi^{(3)}\) is the third-order nonlinear susceptibility. \(V_{ijkl} = (V/V_l V_i V_j) \) is the average effective mode volume and \(V_r (v = i, j, k, l)\) is that at individual frequency \(\omega_{kv}\) given as [57]

\[
V_r = \left\{ \frac{\int d\boldsymbol{r} \varepsilon_r (\boldsymbol{r}, \omega_{kv}) |\vec{E}_v(\boldsymbol{r}, \omega_{kv})|^2}{\int d\boldsymbol{r} \varepsilon_r^2 (\boldsymbol{r}, \omega_{kv}) |\vec{E}_v(\boldsymbol{r}, \omega_{kv})|^4} \right\}^{1/4},
\tag{6}
\]

where \(\varepsilon_r (\boldsymbol{r}, \omega_{kv})\) is the dielectric constant of material and the subscript \(si\) in the integral denotes the integral over the silicon core. \(\eta_{ijkl}\) describes the spatial mode overlap given by [57]

\[
\eta_{ijkl} \equiv \left\{ \frac{\int d\boldsymbol{r} (\varepsilon_{ri} \varepsilon_{jk} \varepsilon_{ki})^{1/2} \bar{E}_i \bar{E}_j \bar{E}_k \bar{E}_l}{(\Pi_{s=i,j,k,l} \int d\boldsymbol{r} \varepsilon_r^2 |\vec{E}_v|^4)^{1/4}} \right\},
\tag{7}
\]

where \(\varepsilon_{ri} = \varepsilon_r (\boldsymbol{r}, \omega_{kv})\) and \(\bar{E}_v = \bar{E}_v (\boldsymbol{r}, \omega_{kv})\) is the electrical field profile for the cavity resonant mode at \(\omega_{kv}\).

The vacuum coupling rates for various nonlinear effects are thus obtained from Eq. (5) as \(g_p = g_{pppp}\), \(g_{pj} = g_{pjp}\) \((j = s, t)\), and \(g_{psi} = g_{spip}\). In our device, the three cavity modes are of close frequencies and similar mode profiles, resulting in

\[
g_p \approx g_{pj} \approx g_{psi} = g = \frac{c \eta n_2^2 \hbar \omega_p \sqrt{\omega_{ps} \omega_{sp}}}{n_j n_i V},
\tag{8}
\]

where \(\eta\) and \(\bar{V}\) denote \(\eta_{pip}\) and \(V_{pip}\), respectively, and we have used the definition of Kerr nonlinear coefficient [27]: \(n_2 = \frac{3\chi^{(3)}}{4\epsilon_0 c^2 \eta^2}\) at the pump frequency. Using Eqs. (3) and (4) and counting in the intrinsic cavity loss [58], we obtain the following equations of motion in the Heisenberg picture governing the wave dynamics inside the cavity:

\[
d\alpha_p/dt = (-i\omega_{kp} - \Gamma_{kp}/2)\alpha_p + ig\alpha_j^+ \alpha_p^2 + i\sqrt{\Gamma_{kp}} b_p e^{-i\omega_pt} + i\sqrt{\Gamma_{0p}} u_p,
\tag{9}
\]

\[
d\alpha_s/dt = (-i\omega_{ks} - \Gamma_{ks}/2)\alpha_s + 2ig\alpha_j^+ \alpha_s^2 + ig\alpha_j^+ \alpha_p^2 + i\sqrt{\Gamma_{ks}} b_s e^{-i\omega_st} + i\sqrt{\Gamma_{0s}} u_s,
\tag{10}
\]

\[
d\alpha_i/dt = (-i\omega_{ki} - \Gamma_{ki}/2)\alpha_i + 2ig\alpha_j^+ \alpha_i^2 + i\sqrt{\Gamma_{ki}} b_i e^{-i\omega_it} + i\sqrt{\Gamma_{0i}} u_i,
\tag{11}
\]

where \(\Gamma_{ip} = \Gamma_{ip} + \Gamma_{ej}(j = p, s, i)\) are the photon decay rates of the intrinsic and loaded cavity, respectively. \(u_j\) is the noise operator associated with intrinsic cavity loss, which satisfies the commutation relation of \([u_j(t), u_j^\dagger(t')] = \delta(t - t')\). In Eqs. (9)-(11), we have neglected the pump depletion and the signal/idler-induced SPM and XPM. The transmitted field is given by [58]

\[
f_j = b_j + i \sqrt{\Gamma_{ej}} \alpha_j.
\tag{12}
\]

In general, the pump mode can be treated as a classical field and Eqs. (9)-(11) can be easily solved to find the solutions for \(\alpha_p\) and \(\alpha_i\), from which, together with Eq. (12), we find the
emitted photon fluxes of signal and idler as
\[
\langle f_s^* (t) f_i(t) \rangle = \frac{\Gamma_{es} \Gamma_{it}}{2\pi} \int_{-\infty}^{+\infty} |B(\omega)|^2 d\omega,
\]
and the pair correlation as
\[
p_c(t_s, t_i) = \langle f_s^* (t_s) f_i^* (t_i) f_s(t_s) f_i(t_i) \rangle - \langle f_s^* (t_s) f_i(t_i) \rangle \langle f_s^* (t_i) f_i(t_s) \rangle
= \Gamma_{es} \Gamma_{it} \left| \frac{1}{2\pi} \int_{-\infty}^{+\infty} B(-\omega) \left| \Gamma_{it} A(\omega) - 1 \right| e^{-i\omega(t_i-t_s)} d\omega \right|^2,
\]
where \( A(\omega) \) and \( B(\omega) \) have the following expressions
\[
A(\omega) = \frac{\Gamma_{ui}/2-i\omega}{(\Gamma_{ts}/2-i\omega)(\Gamma_{ui}/2-i\omega)-(gN_p)^2},
\]
\[
B(\omega) = \frac{-iga_p^2}{(\Gamma_{ts}/2-i\omega)(\Gamma_{ui}/2-i\omega)-(gN_p)^2},
\]
where \( N_p = \langle a_p^* a_p \rangle \) is the average photon number of the pump wave inside the cavity.

Equations (13)-(15) are quite general since they include the multiphoton generation induced by the stimulated FWM. For single photon-pair generation, Eqs. (13) and (14) reduce to
\[
\langle f_s^* (t) f_i(t) \rangle \approx \frac{2\Gamma_{es}(gN_p)^2}{\Gamma_{ts}\Gamma},
\]
\[
\langle f_s^* (t) f_i(t) \rangle \approx \frac{2\Gamma_{es}(gN_p)^2}{\Gamma_{ui}\Gamma},
\]
where \( \Gamma = (\Gamma_{ts} + \Gamma_{ui})/2 \) represents the average photon decay rate of the loaded cavity. The photon-pair correlation reduces to
\[
p_c(t_s, t_i) \approx \begin{cases} 
\frac{\Gamma_{es} \Gamma_{it} (gN_p)^2 e^{-\Gamma_{ui}(t_i-t_s)}}{\Gamma_{ts} \Gamma} & \text{if } t_s \geq t_i \\
\frac{\Gamma_{es} \Gamma_{it} (gN_p)^2 e^{-\Gamma_{ui}(t_i-t_s)}}{\Gamma_{ui} \Gamma} & \text{if } t_s < t_i 
\end{cases}
\]
Equation (20) can be regarded as the joint probability density of emitting a photon pair into the coupling waveguide. It is given as Eq. (1) in the main text. Integrating Eq. (20) over \( \tau = t_i - t_s \), we obtain the average photon-pair emission flux as
\[
R_c = \frac{\Gamma_{es} \Gamma_{it} 2(gN_p)^2}{\Gamma_{ts} \Gamma},
\]
which is Eq. (2) in the main text. Comparing Eq. (21) with Eqs. (18) and (19), we can find that the photon-pair generation rate inside the cavity is \( 2(gN_p)^2/\Gamma \) and the photon extraction efficiency is \( \Gamma_{ej}/\Gamma_{ij} \) \( (j = s, i) \) for the individual signal and idler mode while it is \( \Gamma_{es} \Gamma_{ei}/(\Gamma_{ts} \Gamma_{ui}) \) for the correlated photon pair. As \( \Gamma_{ej}/\Gamma_{ij} \) is less than 1, the photon-pair flux is smaller than the individual photon fluxes. This is experimentally observed in Fig. 5 in the main text.

In practice, the situation is much more complicated than the ideal situation described above, since, due to the light scattering, the optical wave inside the cavity will couple to the degenerate mode propagating backward in the opposite direction (Fig. 8). For a passive cavity, this
phenomenon manifests as a doublet in the transmission spectrum (Insets of Fig. 3 in the main text). In this case, The Hamiltonians $H_0$ and $H_I$ become

$$
\begin{align*}
H_0 &= \sum_{j=p,b,f} \left\{ \hbar \omega_{j} \left( a_{j}^\dagger a_{j} + a_{j}^\dagger a_{j} \right) - \left( \hbar \beta_{j} a_{j}^\dagger a_{j} + \hbar \beta_{j}^* a_{j}^\dagger a_{j} \right) \right. \\
&\quad - \hbar \sqrt{\gamma_{ij}} \left[ \left( a_{j}^\dagger b_{j} + a_{j}^\dagger b_{j} \right) e^{-i\omega_{j}t} + \left( b_{j}^\dagger a_{j} + b_{j}^\dagger a_{j} \right) e^{i\omega_{j}t} \right] \right\}, \\
H_I &= -\frac{\hbar g_p}{2} \left[ (a_{pf}^\dagger)^2 a_{pf}^2 + (a_{pb}^\dagger)^2 a_{pb}^2 + 4a_{pf}^\dagger a_{pf} a_{pb} a_{pb} \right] \\
&\quad - 2\hbar g_{ps} \left[ (a_{pf}^\dagger a_{pf} + a_{pb}^\dagger a_{pb} \right) \left[ g_{ps} (a_{sf} a_{sf} + a_{sb}^\dagger a_{sb} + g_{ps} (a_{sf} a_{sf} + a_{sb}^\dagger a_{sb}) \right] \\
&\quad - \hbar g_{ps} \left[ (a_{pf}^\dagger)^2 a_{sf} a_{sf} + (a_{pb}^\dagger)^2 a_{sb} a_{sb} \right],
\end{align*}
$$

where $a_{j}$ and $a_{j}^\dagger$ ($j = p, s, f$) represent the optical fields propagating clockwise and counter-clockwise, respectively, inside the cavity (Fig. 8), and $\beta_{j}$ is the coupling coefficient between them. $b_{j}$ and $b_{j}^\dagger$ represent the input fields propagating forward (input from the left end) and backward (input from the right end), respectively, inside the coupling waveguide. The transmitted fields now become $f_{j} = b_{j} + i\sqrt{\gamma_{ij}} a_{j}$. Using Eqs. (22) and (23) and following the same procedure, we can find the equations of motion as

$$
\begin{align*}
\frac{da_{pf}}{dt} &= (-i\omega_{pf} - \Gamma_{pf}/2) a_{pf} + i\beta_{pf} a_{ph} + ig (a_{pf}^\dagger a_{pf} + 2a_{pf}^\dagger a_{pb}) a_{pf} + i\zeta_{pf}, \\
\frac{da_{ph}}{dt} &= (-i\omega_{ph} - \Gamma_{ph}/2) a_{ph} + i\beta_{ph} a_{pf} + ig (a_{ph}^\dagger a_{ph} a_{pb} + 2a_{ph}^\dagger a_{ph} a_{ph}) a_{pf} + i\zeta_{ph}, \\
\frac{da_{sf}}{dt} &= (-i\omega_{sf} - \Gamma_{sf}/2) a_{sf} + i\beta_{sf} a_{sf} + 2ig (a_{sf}^\dagger a_{pf} + a_{sf}^\dagger a_{sf} a_{pf} a_{ph}) a_{sf} + ig a_{sf}^\dagger a_{sf} + i\zeta_{sf}, \\
\frac{da_{sb}}{dt} &= (-i\omega_{sb} - \Gamma_{sb}/2) a_{sb} + i\beta_{sb} a_{sf} + 2ig (a_{sb}^\dagger a_{pf} + a_{sb}^\dagger a_{pf} a_{pb}) a_{sb} + ig a_{sb}^\dagger a_{sb} + i\zeta_{sb}, \\
\frac{da_{pf}}{dt} &= (-i\omega_{pf} - \Gamma_{pf}/2) a_{pf} + i\beta_{pf} a_{ph} + ig (a_{pf}^\dagger a_{pf} + 2a_{pf}^\dagger a_{pb}) a_{pf} + i\zeta_{pf}, \\
\frac{da_{ph}}{dt} &= (-i\omega_{ph} - \Gamma_{ph}/2) a_{ph} + i\beta_{ph} a_{pf} + ig (a_{ph}^\dagger a_{ph} a_{pb} + 2a_{ph}^\dagger a_{ph} a_{ph}) a_{pf} + i\zeta_{ph}, \\
\frac{da_{sf}}{dt} &= (-i\omega_{sf} - \Gamma_{sf}/2) a_{sf} + i\beta_{sf} a_{sf} + 2ig (a_{sf}^\dagger a_{pf} + a_{sf}^\dagger a_{sf} a_{pf} a_{ph}) a_{sf} + ig a_{sf}^\dagger a_{sf} + i\zeta_{sf}, \\
\frac{da_{sb}}{dt} &= (-i\omega_{sb} - \Gamma_{sb}/2) a_{sb} + i\beta_{sb} a_{sf} + 2ig (a_{sb}^\dagger a_{pf} + a_{sb}^\dagger a_{pf} a_{pb}) a_{sb} + ig a_{sb}^\dagger a_{sb} + i\zeta_{sb},
\end{align*}
$$

where $\zeta_{j} \equiv \sqrt{\gamma_{ij}} b_{j} e^{-i\omega_{j}t} + \sqrt{\gamma_{ij}} a_{j} (j = p, s, i$ and $v = f, b)$.

Similar to the previous case, Eqs. (24)-(29) can be used to find the solution of signal and idler fields inside the cavity $a_{j}$, ($j = s, i$ and $v = f, b$), through which we obtain the signal and
idler fluxes as
\[
\langle f_s^i(t) f_i(t) \rangle = \frac{\Gamma_{es} \Gamma_{ei}}{2\pi} \int_{-\infty}^{+\infty} \left( |M_{13}(\omega)|^2 + |M_{14}(\omega)|^2 \right) d\omega,
\]
and the pair correlation as
\[
\rho_c(t_1, t_i) = \langle f_s^i(t_i) f_s^i(t_i) f_i(t_1) f_i(t_i) \rangle - \langle f_s^i(t_i) f_i(t_1) \rangle \langle f_i(t_1) f_i(t_i) \rangle
\]
\[
= \Gamma_{es} \Gamma_{ei} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \{ \Gamma_{ts} [M_{11}(\omega) M_{31}(\omega) + M_{12}(\omega) M_{32}(\omega)] - M_{31}^+(\omega) \} e^{-i\omega(t_i - t_1)} d\omega \right|^2,
\]
where $M_{jk}$ is the elements of the matrix $M(\omega)$ which is given by
\[
M^{-1}(\omega) = -\begin{pmatrix}
\frac{i\omega - \Gamma_{ts}/2}{i\beta_i} & \frac{-iga_{pl}^2}{-i\beta_i} & 0 \\
0 & \frac{i\omega - \Gamma_{ts}/2}{i\beta_i} & \frac{-iga_{pl}^2}{-i\beta_i} \\
0 & \frac{iga^*(a_{pl})^2}{-i\beta_i} & \frac{i\omega - \Gamma_{ts}/2}{i\beta_i}
\end{pmatrix}.
\]
Equations (30) and (31) provide the theoretical curves of individual signal and idler fluxes given in Fig. 5(a)-5(d). Equation (32) shows that the photon-pair correlation is only a function of time difference $\tau = t_i - t_1$, $\rho_c(t_i, t_1)$, which is given as the theoretical curves in Fig. 5(e)-5(h). The complete expressions of Eq. (5)-(7) are used in the theoretical calculation, where the effective mode volumes and the mode overlap factors are obtained from the finite element simulation of the device. A Kerr nonlinear coefficient, $n_2 = 5.0 \times 10^{-5}$ cm$^2$/GW, is used in the theoretical calculation. The cavity parameters, such as $\Gamma_{0j}$, $\Gamma_{es}$, and $\beta_j$ $(j = p, s, i)$, are measured from the laser-scanned transmission spectrum of the passive cavity. The P$_1$ mode at 1584.1 nm is outside the scanning range of our lasers and its cavity parameters cannot be obtained through independent measurement. We approximate its intrinsic Q and mode splitting with that of the P$_2$ mode at 1553.6 nm because of their similar external couplings as observed from the PL spectra shown in Fig. 4(a) and 4(b). Figure Fig. 5(d), 5(h), and 5(i) show that these parameters provide a good approximation for the cavity properties of the P$_1$ mode at 1584.1 nm.

B. Statistics of coincidence photon counting

In this section, we provide the theoretical description of the statistics of coincidence photon counting with gated single photon detectors (SPDs). Assume a signal and idler photon strike the detectors at time $t_1$ and $t_i$ with a probability density of $p_{ph}(t_1, t_i)$. The SPDs then produce a pair of photoelectric events which are recorded by the coincidence counter at a later time $t_2 = t_i + t_2$ (Fig. 9), where $t_1$ and $t_2$ include the electron transit time inside the SPDs, the propagation time from the SPDs to the coincidence counter, and time delays introduced by the coincidence counter, with a joint probability density of $p_{ij}(t_1, t_2)$. The joint probability density of the photoelectric events is thus given by
\[
\rho_{SI}(t_1, t_2) = \xi_1 \xi_2 \int_{-\infty}^{+\infty} p_{ph}(t_1, t_i) p_{ij}(t_1, t_2) d\tau_1 d\tau_2
\]
\[
= \xi_1 \xi_2 \int_{-\infty}^{+\infty} p_{ph}(t_2 - t_1 - \tau_D) p_{IF}(\tau_D) d\tau_D.
\]
where $\xi_j$ ($j = 1, 2$) are the detection efficiencies of the two SPDs, and we have used the fact that $p_{ph}(t_s, t_i) = p_{ph}(t_i - t_s)$ which is only a function of the time difference between the signal and idler arrival. For a correlated photon pair, Eq. (32) indicates that $p_{ph}(t_s, t_i) = \alpha_s \alpha_i \langle f_1^\dagger(t_i) f_2^\dagger(t_s) f_2(t_i) f_1(t_s) \rangle$ where $\alpha_s$ and $\alpha_i$ are the propagation losses of signal and idler photons from the coupling waveguide (a tapered fiber in our case) to the SPDs.

![Fig. 9](image_url)

**Fig. 9.** a. Schematic of coincidence photon counting. A signal and idler photon arrive at the detectors at time $t_s$ and $t_i$, respectively, with a probability density of $p_{ph}(t_s, t_i)$. The two detectors produce a pair of events which are recorded by the coincidence counter at time $t_1$ and $t_2$, with a probability density of $p(t_1, t_2)$. b. The time relationship of the events arriving at the coincidence counter. Each event arrives within a gating window centered at $t_0$ ($j = 1, 2$) and with a width of $T_G$. The two input channels of the coincidence counter has a time delay of $T_d$.

In Eq. (34), $p_{IRF}(\tau)$ is the instrument response function (IRF) of the coincidence counting system, given as

$$p_{IRF}(\tau) = \int_0^{\infty} p_I(\tau_1, \tau_1 + \tau) d\tau_1.$$  

(35)

In out coincidence counting system, the IRF is dominantly induced by the timing jitters of the two SPDs. As the detector timing jitter varies with its quantum efficiency (QE), the IRF is recorded at different QEs for specific experiments. The IRF is independently measured by launching identical ultrashort pulses onto the SPDs and recording the coincidence spectrum. The ultrashort pulses come from a mode-locked fiber laser with a pulse width (FWHM) of 240 fs. They are broadened to about 5 ps by the delivery fibers before they hit on the SPDs. The pulse width is much shorter than the timing jitters of the SPDs, thus providing a fairly accurate approach to characterize the IRF.

As a detection event can be either a photoelectric event fired by a photon or a detector dark count, the joint probability density, $p(t_1, t_2)$, that the two detectors produce a pair of events at time $t_1$ and $t_2$ (Fig. 9) is thus given by

$$p(t_1, t_2) = p_{SI}(t_1, t_2) + p_{SD}(t_1, t_2) + p_{DI}(t_1, t_2) + p_{DD}(t_1, t_2),$$  

(36)

where $p_{SI}$, $p_{DI}$, and $p_{DD}$ are the joint probability densities of an event pair of signal–dark count, dark count–idler, and dark count–dark count, respectively, produced in the two separate SPDs. The generation of a detector dark count is independent of an event generated in another detector (either a photoelectric event produced by an incoming photon or a detector dark count). As a result, $p_{SI}(t_1, t_2) = p_{SI}(t_1)p_{DI}(t_2)$, $p_{DI}(t_1, t_2) = p_{DI}(t_1)p_{DI}(t_2)$, and $p_{DD}(t_1, t_2) = p_{DI}(t_1)p_{DI}(t_2)$, where $p_{SI}(t)$, $p_{DI}(t)$, and $p_{DD}(t)$ ($j = 1, 2$) are the probability densities of detection events from signal, idler, and dark counts, respectively, produced in the two detectors. In experiment, $p_{SI}(t)$, $p_{DI}(t)$, and $p_{DD}(t)$ are equal to the average detection rates of the related events, respectively, which can be measured independently. Consequently, Eq. (36) becomes

$$p(t_1, t_2) = p_{SI}(t_2 - t_1) + p_{SI}(t_1)p_{DI}(t_2) + p_{DI}(t_1)p_{DI}(t_2) + p_{DI}(t_1)p_{DD}(t_2).$$  

(37)
Therefore, the coincidence at \( \tau = t_2 - t_1 \) within a small delay-time interval of \( d\tau \) is given by

\[
\rho(\tau)d\tau = \int_{t_{10}+T_G/2}^{t_{10}-T_G/2} dt_1 \int_{t_{20}+T_G/2}^{t_{20}-T_G/2} dt_2 \rho(t_1, t_2) = d\tau \langle T_G - |\tau - T_d| \rangle [p_{SI}(\tau - T_d) + p_{SPD2} + p_{DI} p_I + p_{DI} p_{D2}],
\]

(38)

for \( |\tau - T_d| \leq T_G \) and \( \rho(\tau) = 0 \) otherwise. In experiment, \( d\tau \) corresponds to the time resolution of the coincidence counter, which is 4 ps in our system. \( t_{10} \) and \( t_{20} \) are the centers of the gating windows in the two channels (Fig. 9). As a result, the delay-time-dependent coincidence counting histogram (we term as the coincidence spectrum here for simplicity) is obtained accordingly as

\[
C(\tau)d\tau = R_G T \rho(\tau)d\tau,
\]

(39)

where \( R_G \) is the clock frequency of the SPDs and \( T \) is the data collection time. Equations (34), (38), and (39) provide a complete description of coincidence photon counting. In practice, the gating window \( T_G \) of SPDs is always set to be much larger than the temporal width of the coincidence spectrum and Eq. (38) becomes

\[
\rho(\tau)d\tau = d\tau T_G [p_{SI}(\tau - T_d) + p_{SPD2} + p_{DI} p_I + p_{DI} p_{D2}],
\]

(40)

for \( |\tau - T_d| \ll T_G \).

From Eqs. (34)-(39), we obtain the true coincidence spectrum \( C_T(\tau) \) and the accident coincidence spectrum \( C_A(\tau) \) as

\[
C_T(\tau)d\tau = R_G T d\tau (T_G - |\tau - T_d|) [p_{SI}(\tau - T_d) - p_{SI} p_I]
\]

\[
= R_G T d\tau (T_G - |\tau - T_d|) J_1 \xi_1 \xi_2 \alpha_c\alpha_t \int_{-\infty}^{+\infty} p_c(\tau - T_d - \tau_d) p_{IRF}(\tau_d)d\tau_d,
\]

(41)

\[
C_A(\tau)d\tau = R_G T d\tau (T_G - |\tau - T_d|) (p_S + p_{DI})(p_I + p_{D2}),
\]

(42)

where \( p_c(\tau) \) is given in Eq. (32). As a result, the theoretical coincidence-to-accidental ratio (CAR) shown as the red curves in Fig. 5(e)-5(h) is obtained as

\[
\text{CAR} = \frac{\int_{-T_{FWHM}/2}^{+T_{FWHM}/2} C_T(\tau + T_d)d\tau}{\int_{-T_{FWHM}/2}^{+T_{FWHM}/2} C_A(\tau + T_d)d\tau},
\]

(43)

where \( T_{FWHM} \) is the full width at half maximum of the coincidence spectrum. In Eqs. (41) and (42), \( p_S = \xi_1 \alpha_c \langle f_S^f(t) f_S(t) \rangle + p_{IR} \) and \( p_I = \xi_2 \alpha_t \langle f_I^f(t) f_I(t) \rangle + p_{IR} \), where \( \langle f_S^f(t) f_S(t) \rangle \) and \( \langle f_I^f(t) f_I(t) \rangle \) are given by Eqs. (30) and (31), and \( p_{IR} \) and \( p_{IR} \) are the probability densities of detection events from Raman noise photons at the signal and idler channels, respectively (See the next section for the details of characterizing Raman noises). On the other hand, Eq. (42) shows clearly that the accidental coincidence can be accurately calculated from independently measured \( p_S, p_I \), and \( p_{Dj} \) (\( j = 1, 2 \)). Therefore, using these measured parameters together with independently recorded true coincidence we can reproduce the experimentally recorded CAR.

This theoretical approach explains well the saturation of CAR at low pair fluxes and is commonly used in literature [35]. It is shown as the black dashed curves in Fig. 5(i)-5(l).

The experimental value of CAR is obtained via the following procedure: First, at each pump power, a coincidence spectrum is obtained with zero time delay between the two SPDs. The coincidence spectrum is integrated over its full width at half maximum \( T_{FWHM} \) to obtain the total coincidence \( C \). The statistical error of \( C \) is given by \( E_C = \sqrt{C} \) since the photon counting follows a Poisson distribution. To find the accidental coincidence, a series of coincidence spectra are
obtained with a time delay of $jT_0$ between the two SPDs, where $T_0$ is the detector clock period and $j = 1, 2, ..., 200$. Integrate each of them within the same $T_{FWHM}$ to obtain an accidental coincidence $A_j$. The average and the standard deviation of $A_j$ give the accidental coincidence $A$ and its statistical error $E_A$. Note that no accidental coincidence is subtracted. Finally, $CAR = (C - A)/A$ and its statistical error $E_{CAR}$ is given by $E_{CAR} = \sqrt{\left(\frac{E_C}{C}\right)^2 + \left(\frac{E_A}{A}\right)^2}$. To reduce the statistical error, each data point in Fig. 5 was recorded with a time period of $5400 - 18900$ seconds depending on the magnitude of photon-pair flux. With such amount of photon counting data, $E_{CAR}$ is very small (< 2%) and $E_{CAR}$ is dominantly determined by $\frac{E_A}{A}$ since the accidental coincidence is very small in our device.

To find the photon-pair emission flux shown in Fig. 5(e)-5(h), a same procedure is carried out but now the integration is performed over the entire time window of coincidence spectra (20 ns, with a detector gating width of 10 ns) to obtain the total coincidence $C_t$, the accidental coincidence $A_t$, and their statistical errors $E_{C_t}$ and $E_{A_t}$. The true coincidence is thus given by $C_T = C_t - A_t$. Its statistical error $E_{C_T}$ is given by $E_{C_T} = \sqrt{(E_{C_t})^2 + (E_{A_t})^2}$. Dividing $C_T$ by the total number of detector gates within the data collection time we find the true coincidence counts per gate (which is shown on the right axes of Fig. 5(e)-5(h)). The photon-pair emission flux is obtained by calibrating the true coincidence counts per gate with the clock frequency, duty cycle, and quantum efficiency of the detectors as well as the propagation losses from the device to the detectors for the photon pairs. The statistical error of pair emission flux is obtained accordingly from $E_{C_T}$.

The theoretical curves shown in Fig. 6 are described by Eq. (39) (normalized by its peak value), in which $p_{pa}(t_i, t_s) = \alpha_0 \alpha_0 (f^s_i(t_i) f^s_i(t_s))$ obtained from Eq. (32) and $p_{IRF}(\tau)$ is given by a Gaussian fitting to the experimentally recorded IRF. The cavity parameters used in $(f^s_i(t_i))$, $A$ $s$ $i$, and $\beta_j$ ($j = p, s, i$) were independently measured from the laser-scanned transmission spectrum of the passive cavity.

**C. Characterization of Raman noise produced in delivery fibers**

![Fig. 10. Schematic of the experimental setup (details can be found in Fig. 2).](image)

Figure 5(a)-5(d) show that the fluxes of individual photon modes deviate slightly away from the quadratic power dependence, which is due to the broadband Raman noises generated inside the delivery fibers. To characterize the noise for each photon mode, we recorded the photon flux when the center wavelength of the tunable bandpass filter was tuned away from the photon mode so that the latter was completely blocked (see Fig. 10). The pumping condition remained the same as what we used to characterize the photon pairs, with the tapered fiber touched down to the nanoforks and the pump wavelength on the cavity resonance. Such an experimental condition ensures that the pump power distribution inside the delivery fibers remains unchanged. The photon flux recorded under this condition is not produced from the device since the photon wavelength does not coincide with any cavity resonance.

Figure 11 shows an example of recorded photon fluxes for $P_i$, where the solid and open
circles show the cases when the filter centers were tuned onto and away from the photon mode pair, respectively. The noise photon fluxes exhibit a linear dependence on the pump power. Figure 11 shows clearly that the observed noise is the dominant noise source accompanied with the photon pairs. To further verify that the observed noise does not come from the device, we directly lifted the taper up so that the tapered fiber did not couple to the device and the pump wave propagated only inside the delivery fibers (as well as the optical components in the experimental system. See Fig. 10). Kept other conditions unchanged and recorded again the photon fluxes, we obtained the open triangles shown in Fig. 11, which are very similar to the open circles. This clearly verifies that the noise photons are not generated from the device but produced inside the experimental testing system.

![Figure 11](image_url)

Fig. 11. Photon fluxes recorded for P\textsubscript{I\textsubscript{2}}, with detector dark counts subtracted. The solid and open circles show the photon fluxes when the center wavelengths of the bandpass filters are tuned onto and away from the P\textsubscript{I\textsubscript{2}} mode pair, respectively, with the tapered fiber touched down to the nanoforks. The open triangles are recorded when the tapered fiber is lifted up and thus does not couple to the device. The blue and red show the signal at shorter wavelength and idler at longer wavelength, respectively. The dashed lines are only for eye guidance.

In our testing system (Fig. 10), the CWDM MUX has a band isolation greater than 120 dB. As a result, any broadband noise produced before the MUX (e.g., in fiber 0) will be completely blocked. On the other hand, the CWDM DEMUX separates the pump wave from the photon pairs and no noise will be produced in fiber 3 after the DEMUX. Moreover, when the pump wave is nearly critically coupled into the device, the pump power in fiber 2 after the device is small. Therefore, the noise should be dominantly produced inside fiber 1 before the device, which has a rough length of \(\sim 5\) m. The open triangles in Fig. 11 have higher magnitudes than the open circles simply because, when the taper is lifted up, the pump power will increase in fiber 2 and will start to produce noise.

Figure 11 shows that the idler channel at the Stokes side of the pump has a larger noise photon flux than the signal channel at the anti-Stokes side. This feature together with their linear power dependence implies that the noise is likely to be the Raman noise produced inside fiber 1. To verify this, we incremented the length of fiber 1 (with the tapered fiber not coupled to the device) and monitored the noise spectrum. Figure 12a shows an example of the noise spectra recorded with three different fiber lengths of \(\sim 5\) m, \(\sim 29\) m, and \(\sim 1\) km. The similarity of the spectra among these three cases indicates clearly that they share an exactly same physical...
origin. The noise spectra show the Raman noise characteristics of silica [59], with the Boson peak clearly visible at a frequency shift of \( \sim 1.5 \) THz away from the pump.

Fig. 12. **a.** Raman noise spectra recorded with different lengths of fiber 1, with an input pump power of about 270 \( \mu \)W. The spectra were recorded at different transmission ports of the CWDM DEMUX (indicated by color), for easy suppression of the pump. The solid curves and dotted curves show the cases with a fiber length of \( \sim 29 \) m and \( \sim 1 \) km, respectively. The amplitudes of the dotted curves were reduced by a factor of 10, for easy comparison with the thin solid curves. The inset shows the case with a fiber length of \( \sim 5 \) m, same as what we used to characterize correlated photon pairs. The pump power and wavelength are the same for these three cases. The spectral cut-off beyond \( \sim 1590 \) nm is due to the quantum-efficiency cut-off of the detector used in our spectrometer. **b.** Raman noise spectra with the pump located at 1470 nm, showing the Stokes side. Fiber 1 has a length of \( \sim 1 \) km. **c.** Raman noise spectra with the pump located at 1570 nm, showing mostly the anti-Stokes side. Fiber 1 has a length of \( \sim 1 \) km.

To reveal more characteristics of the noise spectra, we tuned the pump wavelength to the left end of the CWDM and recorded again the noise spectra (Fig. 12(b)). In this case, the noise spectra primarily show the Stokes side. Both the Boson peak and Raman gain peak are clearly visible, with a frequency shift of about 1.5 and 13 THz, respectively, away from the pump frequency. Tuning the pump wavelength to 1570 nm we recorded the noise spectra now mostly on the anti-Stokes side (Fig. 12(c)). Again, we observed both peaks but with smaller amplitudes. These observations clearly verify that the observed noise indeed comes from the spontaneous Raman scattering inside the delivery fiber.

As the Raman noise is very broadband, in experiment, we used a pair of tunable bandpass filters with a 3-dB bandwidth of 1.2 nm in front of the SPDs (Fig. 2 and Fig. 10) to cut the Raman noise. The Raman noise fluxes associated with each mode pair were measured with the filter centers tuned away from the mode pair far enough that the latter were completely blocked. The recorded Raman noise fluxes were used in Eq. (42) for finding the theoretical CAR, as
discussed in the previous section. Due to the strong frequency dependence of Raman noise spectrum, a certain uncertainty may be accompanied with the recorded Raman noise fluxes. Note that the Raman noise can be further suppressed in the future by using narrower filters since the correlated photon pairs have much narrower spectra. The noise can also be reduced by shortening the length of fiber 1. Such Raman noise will be completely absent if an on-chip laser [29] is used as the pump.

Acknowledgments

The authors would like to thank Chris Michael for his help during the initial stage of this project. This work was supported in part by the Air Force Office of Scientific Research under grant No. FA9550-12-1-0419 and by the National Science Foundation under grant ECCS-1351697. It was performed in part at the Cornell NanoScale Facility, a member of the National Nanotechnology Infrastructure Network.