Generation rate scaling: the quality factor optimization of microring resonators for photon-pair sources

Guo, Kai; Shi, Xiaodong; Wang, Xiaolin; Yang, Junbo; Ding, Yunhong; Ou, Haiyan; Zhao, Yijun

Published in:
Photonics Research

Link to article, DOI:
10.1364/PRJ.6.000587

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
To achieve photon-pair generation scaling, we optimize the quality factor of microring resonators for efficient continuous-wave-pumped spontaneous four-wave mixing. Numerical studies indicate that a high intrinsic quality factor makes high pair rate and pair brightness possible, in which the maximums take place under overcoupling and critical-coupling conditions, respectively. We fabricate six all-pass-type microring resonator samples on a silicon-on-insulator chip involving gap width as the only degree of freedom. The signal count rate, pair brightness, and coincidence rate of all the samples are characterized, which are then compared with the modified simulations by taking the detector saturation and nonlinear loss into account. Being experimentally validated for the first time to the best of our knowledge, this work explicitly demonstrates that reducing the round-trip loss in a ring cavity and designing the corresponding optimized gap width are more effective to generate high-rate or high-brightness photon pairs than the conventional strategy of simply increasing the quality factor. © 2018 Chinese Laser Press

1. INTRODUCTION

A quantum-correlated photon-pair source is a key resource in research of quantum optics such as quantum information processing [1] and quantum communication [2]. Thereinto, the most mature technology is quantum key distribution (QKD), where the nature of correlated photon pairs is applied to suffice high-security distribution [3–7]. Sources capable of QKD are required to emit single photons in a probabilistic manner with low noise, preferably in the telecom wavelength range, to benefit from the compatibility of optical fiber networks [8]. Moreover, the single photon generated from spontaneous nonlinear processes has a naturally correlated twin photon, which makes it possible to apply the detection of one photon (signal) to herald the existence of the other (idler). While the initially heralded photon-pair sources have been demonstrated via spontaneous downconversion in optical crystals [9] or quasi-matched waveguides [10], and via spontaneous four-wave mixing (SpFWM) in optical fibers [8,11], a number of experiments were carried out in the past decade via SpFWM in integrated waveguide platforms, of which the material can be crystalline silicon [12–14], amorphous silicon [15], silica [16], silicon nitride [17], and AlGaAs [18]. Integrated waveguides often have large refractive index contrast leading to strong light confinement and high nonlinear interaction, which enables efficient photon-pair generation within a few millimeters. Moreover, by employing either butt-coupled [19] or vertical-coupled approaches [20], integrated waveguides are compatible with fiber-based systems; thus, it can be applied as the nonlinear medium of photon-pair sources instead of optical fibers, to avoid broadband spontaneous Raman scattering noise [11]. In addition, thanks to the mature fabrication methods of semiconductors and integrated circuits, which enable a variety of functionalities [21], efficient quantum communication systems are in the progress of on-chip integration [22,23].

Although photon-pair sources using SpFWM in integrated waveguides are often driven by a pulsed pump, the continuous-wave (CW) pump, with advantages of cheaper, more stable, and especially easier on-chip integration, is also widely employed. The narrow linewidth of CW pumps gives rise to strong spectral anticorrelation and projects the generated photon pairs into a classical spectral mixture [24,25], which is desired for special applications such as time–energy entanglement [26–28], wavelength-multiplexed quantum communication [29,30], and covert quantum communication [31,32]. Moreover, photon-pair sources capable of long-distance quantum communication are required to have a high pair rate, which corresponds...
to vast photon pairs for cryptography coding, huge loss tolerance for long-distance transmission, and enough photon pairs for cryptography decoding. High pair brightness, corresponding to high pair rate spectral density driven by specific pump power, is also desired because a narrow photon-pair bandwidth does not only make easier entanglement-based QKD but also makes the promotion of quantum key rate possible by applying dense wavelength division multiplexing. Noteworthy, the anticorrelation for CW-pumped sources can be avoided by using narrow bandwidth filtering. However, it comes at the cost of reducing photon pairs, where the pair brightness may become lower. As shown in Ref. [14], by applying narrow-bandwidth (0.4 nm) filtering to a CW-pumped source using a silicon strip waveguide, the highest pair rate can reach $1.6 \times 10^6$ Hz; however, the corresponding pair brightness of $4.0 \times 10^5$ (s $\cdot$ mW $\cdot$ nm)$^{-1}$ remains the same order of magnitude as that in other studies [13,33]. A valid approach facilitating high pair brightness is to use microring resonators (MRRs) instead of straight waveguides, which provides not only narrow-bandwidth filtering but also strong cavity enhancement. Therefore, a number of experiments were carried out using different MRR designs, which achieved 1 to 2 orders of magnitude higher pair brightnessness but with an ultrasmall footprint [34–41].

Although MRR has shown the capability of photon-pair generation, it lacks a normative evaluation of the waveguide design, for photon-pair generation in MRR of high pair rate and high pair brightness. Moreover, a solid understanding of the key parameters of different MRR structures, especially the quality factor, is significant, based on which the optimization may put forward an approach of generation rate scaling. The quality factor is given by [42]

$$Q = \frac{\lambda_{\text{res}}}{\Delta \lambda}, \quad (1)$$

where $\lambda_{\text{res}}$ and $\Delta \lambda$ denote the resonance wavelength and its full width at half-maximum (FWHM), respectively. More specifically, for an all-pass-type MRR consisting of a bus waveguide and a ring cavity, the total quality factor obtained from the transmittance of the bus waveguide is jointly determined by the round-trip loss in the ring cavity, which is quantified by the intrinsic quality factor

$$Q_i = \frac{\omega}{\alpha v_g} \quad (2)$$

and the coupling efficiency between two components, which is quantified by the external quality factor

$$Q_e = \frac{2 \alpha \omega R}{|k|^2 v_g} \quad (3)$$

where $\alpha$ denotes the round-trip loss coefficient, $v_g$ denotes the light group velocity in the ring cavity, $R$ denotes the radius of the ring cavity, and $k$ denotes the coupling coefficient [43]. The quality factor given by Eq. (1) follows

$$\frac{1}{Q} = \frac{1}{Q_i} + \frac{1}{Q_e}. \quad (4)$$

From previous studies [41,44,45], the pair rate $N_c$ has a third-order polynomial dependence on $Q$, which indicates the larger quality factor, the better performance. However, these studies omit the impact of round-trip loss that results in $Q = Q_e$ and present an approach of pair rate scaling by simply increasing the gap width $g$. Although Ref. [40] shows that $N_c$ has a seventh-order polynomial dependence on $Q$ by taking all types of loss into account and demonstrates good agreement between simulations and measurements, their discussion based on only one MRR does not present the potential quality factor optimization that facilitates a higher pair rate. In addition, by trading off pair rate and photon-pair bandwidth, it is valid to achieve higher pair brightness $B$.

In this paper, we demonstrate the scaling approaches of pair rate and pair brightness, respectively, by involving $Q_e$ and $Q_i$ as degrees of freedom [46]. We fabricate six all-pass-type MRRs on a silicon-on-insulator (SOI) chip with different $g$ and characterize the photon-pair sources using all samples to verify the numerical predictions, taking the impact of non-linear loss in SOI platforms and detector saturation into consideration. In the end, the future direction for efficient photon-pair generation in all types of microcavity platforms is presented.

2. THEORETICAL APPROACHES

Based on SpFWM in a microcavity, the pair rate $N_c$ is quadratic in the circling pump power $P_c$, [47], which is defined as

$$P_c = P_p |F(\omega_p)|^2, \quad (5)$$

where $P_p$ denotes the incident pump power in the bus waveguide. The enhancement factor $F(\omega_p)$ follows

$$|F(\omega_p)|^2 = \frac{2v_g^2 Q_e^2}{\pi R \omega_p Q_i [1 + 4Q_e^2(\omega_p - \omega_{\text{res}})^2/\omega_{\text{res}}^2]}, \quad (6)$$

which reaches the maximum

$$|F(\omega_p)|^2_{\max} = \frac{2v_g^2 Q_e^2}{\pi R \omega_p Q_i^2}, \quad (7)$$

when $\omega_p = \omega_{\text{res}}$, that is, the pump is on-resonance. Assume that only the signal/idler photons generated from one resonance are counted, meanwhile the pump, signal, and idler approximately have the same frequency of $\omega_{\text{res}}$; then, the pair rate in signal/idler arms is calculated through

$$N_{c,\text{id}} = (\gamma P_c 2\pi R)^2 \int_{\omega_{\text{id}}} |F(\omega_{\text{res}} - \Delta \omega)|^2 |F(\omega_{\text{res}} + \Delta \omega)|^2 d\Delta \omega, \quad (8)$$

where $\gamma$ denotes the nonlinear coefficient. By substituting Eq. (7) into Eq. (8), the pair rate in the on-resonance regime is given by

$$N_{c,\text{id}} = 8\pi^6 \gamma^2 P_p^2 Q_s Q_i^3 Q_j^{\text{id}} / (\omega_{\text{res}} \pi R^3 \omega_p Q_s Q_j^{\text{id}}), \quad (9)$$

where $Q_s$ and $Q_j^{\text{id}}$ denote the total quality factor, $Q_{c,p}$ and $Q_{c,\text{id}}$ denote the external quality factor, corresponding to the pump and signal/idler, respectively. Furthermore, by using the definition of $B = N_c / (P_p \Delta \lambda)$ with a unit of (s $\cdot$ mW $\cdot$ nm)$^{-1}$ [27], the pair brightness becomes.
and we assume that the source operates in the on-resonance
However, for a given 
estimated via finite-difference mode solver [48], the group veloc-
both wavelength independent; hence, Eqs. (9) and (10) can be

\[
Q_e = \frac{8v_0^2 R^2 Q^7}{\alpha_0^2 \pi^2 R^2 Q^7},
\]

(11)

\[
B = \frac{4v_0^2 P_p Q^8}{\alpha_0^2 \pi^2 R^2 c Q^8},
\]

(12)

which take similar forms known from Ref. [40]. Therefore, we acquire the relations, \( N_e \propto Q^7/Q_i^4 \) and \( B \propto Q^8/Q_i^4 \). As Eq. (6) takes the round-trip loss of both the pump and the generated photon pairs into account, when the round-trip loss is negligible in the enhancement factor definition [44,45], the foregoing relations become \( N_e \propto Q^7 \) and \( B \propto Q^4 \).

By utilizing Eqs. (11) and (12), we simulate \( N_e \) and \( B \) against both \( Q_i \) and \( Q_e \) for a specific source using silica-cladded, all-pass-type MRRs, of which the ring-cavity radius \( R \) is 110 \( \mu \)m with a cross-sectional dimension of 250 nm \( \times \) 450 nm. Being estimated via finite-difference mode solver [48], the group velocity \( v_g \) at a pump wavelength of 1550 nm is \( 7 \times 10^7 \) m/s, and the nonlinear coefficient \( \gamma \) is 300 mW\(^{-1} \) [49], for the fundamental transverse-electric mode. The incident pump power \( P_p \) is 1 mW, and we assume that the source operates on the resonance scheme. As shown in Fig. 1(a), for a given \( Q_i \), an increase of \( Q_e \) gives rise to an increase of \( Q_i \) and finally makes \( N_e \) higher. However, for a given \( Q_i \), the initial increase of \( Q_e \) makes \( N_e \) higher until the increase of \( Q^2 \) in Eq. (11) has more of an impact than does the increase of \( Q^7 \). The black solid denotes the optimized \( Q_e \) for a given \( Q_i \), which facilitates the highest \( N_e \) with a slope of 3/4, which can be as well obtained from Eq. (11), involving \( Q_e/Q_i \) as a variable. Moreover, as shown in Fig. 1(b), \( B \) increases with \( Q_i \) increasing more rapidly for a given \( Q_e \) because of the eighth-order polynomial dependence. For a given \( Q_i \), there also exists an optimized \( Q_e \), which facilitates the highest \( B \) but has a slope of 1 instead.

From the simulations, the primary key point of achieving high pair rate and high pair brightness is to propose high \( Q_e \) corresponding to low \( \alpha \), which is experimentally achieved by reducing the scattering loss induced by the side-wall roughness, bending loss determined by the ring-cavity radius, coupling loss determined by the gap width [50], and nonlinear loss, including two-photon absorption (TPA) and free-carrier absorption (FCA), especially in SOI platforms [51,52].

Fig. 1. (a) \( N_e \) versus \( Q_i \) and \( Q_e \). Black solid denotes the optimized \( Q_e \) for a given \( Q_i \) that facilitates the highest \( N_e \) with a unit of Hz. (b) \( B \) versus \( Q_i \) and \( Q_e \). Black solid denotes the optimized \( Q_e \) for a given \( Q_i \) that facilitates the highest \( B \) with a unit of \((s \cdot mW \cdot nm)^{-1}\).
and $Q_i$ are much higher than 1. We used tunable-laser sweeping to measure the transmittance at low power, from which the quality factors of pump and signal/idler resonances are estimated. Because the initial resonance of each sample is detuned from the exact wavelength of pump, signal, and idler, a temperature controller was employed to achieve the on-resonance corresponding to the minimal output power. As shown in Table 1, $Q_e$ becomes higher with $g$ increasing, which results in MRR1–MRR4 in the overcoupling and MRR3–MRR6 in the undercoupling. Moreover, only when $g$ is larger than 270 nm does $Q_i$ receive an approximate value of $1 \times 10^5$ because a smaller gap width brings significant coupling loss; that is, the smaller $g$, the lower $Q_i$.

By using the parameters in Table 1, we calculate $N_c$ and $B$ for all samples by numerical simulations, according to Eqs. (9) and (10). We also simulate $N_c$ and $B$ versus $Q_i$ using $Q_i$ of 4.76 $\times$ 10^4, 5.59 $\times$ 10^4, 7.54 $\times$ 10^4, and 1.04 $\times$ 10^5. As shown in Fig. 2(a), $N_c$ becomes higher for MRR1–MRR4 because the increase of the $Q_i$ in Eq. (11) contributes more than that of the $Q_e$, which is more essentially attributed to the higher $Q_i$ corresponding to lower round-trip loss of photon pairs. For MRR4–MRR6 with approximate $Q_i$, $N_c$ becomes lower with $Q_i$ increasing because the increase of the $Q_i$ contributes more than that of the $Q_e$, which reveals the fact that both pump power coupling into the ring cavity and photon pairs coupling into the bus waveguide suffer from the low coupling efficiency. By ranking all samples in $N_c$ order from high to low, we predict MRR4 > MRR3 > MRR5 > MRR2 > MRR1 > MRR6. Note that MRR6 with a higher $Q_i$ of 9.15 $\times$ 10^4 even generates fewer pairs than MRR1 with a lower $Q_i$ of 1.83 $\times$ 10^4, which contradicts the conventional understanding of using MRR with higher $Q$ for more pairs. As shown in Fig. 2(b), $B$ becomes higher for MRR1–MRR4, which is jointly contributed by the increase of $N_c$ and the decrease of $\Delta \lambda$, while for MRR4–MRR6, $B$ becomes lower with $Q_i$ increasing. By ranking all samples in $B$ order from high to low, we obtain MRR4 > MRR5 > MRR3 > MRR6 > MRR2 > MRR1, where MRR6 with lower $N_c$ enables higher $B$ than MRR1. In addition, MRR4 with almost the highest $Q_i$ and the closest $Q_e$ to the optimized value facilitates the highest $N_c$ and $B$ simultaneously among all the samples, which can be potentially higher when $Q_e$ reduces to 7.8 $\times$ 10^4 and 1.04 $\times$ 10^3, respectively.

### 3. Experimental Characterization

We demonstrate the experimental characterization using the setup in Fig. 3. A CW laser at 1554.8 nm was utilized as the pump, which was power-amplified by an erbium-doped fiber amplifier (EDFA). Tunable bandpass filters (TBPFs) were applied to suppress both sideband noise and amplified spontaneous emission. A tunable attenuator (ATT) was utilized to control the incident pump power; meanwhile, the signal-to-noise ratio at a wavelength detuning of ±5.6 nm remained 140 dB. Photonic crystal-based grating couplers (PCGCs) [20] were used for beam-coupling between the bus waveguide and fibers, with a total insertion loss of around 12 dB for all samples. Input and output powers were monitored using two 99%–1% couplers and two power-meters (PMs) for precise $P_e$ estimation. Cascaded arrayed waveguide gratings (AWGs) were utilized to suppress the leaked pump field down to −100 dBm; meanwhile, we separated the signal and idler

### Table 1. Key Parameters of MRR Samples

<table>
<thead>
<tr>
<th>Marker</th>
<th>$g$ (nm)</th>
<th>$Q_i$ (P)</th>
<th>$Q_e$ (P)</th>
<th>$Q_e$ (S)</th>
<th>$Q_i$ (S)</th>
<th>$Q_i$ (I)</th>
<th>$Q_e$ (I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRR1</td>
<td>180</td>
<td>1.83 $\times$ 10^4</td>
<td>2.97 $\times$ 10^4</td>
<td>4.76 $\times$ 10^4</td>
<td>2.92 $\times$ 10^4</td>
<td>4.79 $\times$ 10^4</td>
<td>2.82 $\times$ 10^4</td>
</tr>
<tr>
<td>MRR2</td>
<td>210</td>
<td>2.59 $\times$ 10^4</td>
<td>4.83 $\times$ 10^4</td>
<td>5.59 $\times$ 10^4</td>
<td>5.23 $\times$ 10^4</td>
<td>5.35 $\times$ 10^4</td>
<td>4.94 $\times$ 10^4</td>
</tr>
<tr>
<td>MRR3</td>
<td>240</td>
<td>4.52 $\times$ 10^4</td>
<td>1.01 $\times$ 10^5</td>
<td>7.54 $\times$ 10^4</td>
<td>1.24 $\times$ 10^4</td>
<td>8.84 $\times$ 10^4</td>
<td>1.01 $\times$ 10^5</td>
</tr>
<tr>
<td>MRR4</td>
<td>270</td>
<td>6.21 $\times$ 10^4</td>
<td>1.55 $\times$ 10^5</td>
<td>1.04 $\times$ 10^4</td>
<td>1.52 $\times$ 10^4</td>
<td>1.12 $\times$ 10^4</td>
<td>1.37 $\times$ 10^5</td>
</tr>
<tr>
<td>MRR5</td>
<td>300</td>
<td>7.40 $\times$ 10^4</td>
<td>2.82 $\times$ 10^5</td>
<td>1.01 $\times$ 10^4</td>
<td>2.95 $\times$ 10^4</td>
<td>9.88 $\times$ 10^4</td>
<td>2.67 $\times$ 10^5</td>
</tr>
<tr>
<td>MRR6</td>
<td>330</td>
<td>9.15 $\times$ 10^4</td>
<td>5.70 $\times$ 10^5</td>
<td>1.09 $\times$ 10^4</td>
<td>5.70 $\times$ 10^4</td>
<td>1.09 $\times$ 10^4</td>
<td>4.19 $\times$ 10^5</td>
</tr>
</tbody>
</table>

![Fig. 2](image_url)

**Fig. 2.** (a) $N_c$ versus $Q_e$. (b) $B$ versus $Q_i$, when $Q_i$ is 4.76 $\times$ 10^4 (black solid), 5.59 $\times$ 10^4 (red solid), 7.54 $\times$ 10^4 (green solid), and 1.04 $\times$ 10^5 (blue solid). MRR1, black circle; MMR2, red circle; MRR3, green circle; MRR4, blue circle; MRR5, cyan diamond; MRR6, magenta triangle.
measured signal count rate of lower than 30 kHz (see black dash) is around 1.66 for each sample. Because $N_m$ is theoretically quadratic in $P_p$, which accounts for a slope of 2, the measured photon pairs generated from SpFWM are dominant, especially at low power. Moreover, by taking detector saturation into account, $N_m'$ can be modified through [40]

$$N_m' = \frac{N_m - D}{1 - \tau_0 N_m'^2}$$  \hspace{1cm} (14)

where $D$ and $\tau_0$ denote dark count rate and dead-time, respectively. Pumped at 3.98 mW, $N_m'$ for MRR1–MRR6 are measured at 35, 38, 56, 64, 44, and 19 kHz and modified at 54, 61, 127, 178, 79, and 23 kHz, respectively, where $N_m'^{\Delta}\lambda$ for MRR3 and MRR4 even exceed the maximal detection rate of our detector, which is 100 kHz. By applying Eq. (14) to all data, the fitted slopes increase to around 1.82, indicating that the detector saturation results in an underestimation of the measured signal count rate generated from SpFWM. Furthermore, the intrinsic nonlinear loss in SOI platforms also contributes to the signal count rate saturation because the enlarging $\alpha$ reduces $Q_s$ significantly at high power. The ratio of the signal count rate with nonlinear loss to that without nonlinear loss equals the ratio of $Q_s^{\Delta}\lambda$ with nonlinear loss to that without nonlinear loss [40].

Hence, through a simple calculation, the maximal nonlinear-loss-free $N_m^{\Delta}\lambda$ for each sample reaches 63, 83, 251, 608, 223, and 53 kHz, respectively. In addition, by ranking all samples in $N_m'$ (or fitted constant term) order from high to low, the result, MRR4 > MRR3 > MRR5 > MRR2 > MRR1 > MRR6, agrees well with our prediction.

As shown in Table 2, we calculate the measured pair brightness $B'$ at $P_p = 1$ mW using $N_m'$ and $\Delta\lambda$ for each sample, according to $B' = N_m'/ (P_p \Delta\lambda)$. $B'$ for MRR4 is 2 orders of magnitude higher than that for MRR1 because of smaller $\Delta\lambda$ and higher $N_m'$. Although $N_m'$ for MRR6 is lower than that for MRR1 and MRR2, the much smaller $\Delta\lambda$ enables $B'$ for MRR6 to be higher. By ranking all samples in $B'$, order from high to low, the result, MRR4 > MRR5 > MRR3 > MRR6 > MRR2 > MRR1, also agrees well with our prediction. Additionally, the saturation can also affect pair brightness, where $B'$ for MRR4 of $6.4 \times 10^5$ (s $\cdot$ mW $\cdot$ nm)$^{-1}$ at $P_p = 3.98$ mW becomes even lower than that at $P_p = 1$ mW. By omitting detector saturation and nonlinear loss, the modified pair brightness $B_m' = N_m^{\Delta}\lambda / (P_p \Delta\lambda)$ can reach $5.0 \times 10^6$ (s $\cdot$ mW $\cdot$ nm)$^{-1}$, which demonstrates $B \propto P_p^{\Delta}\lambda$ in Eq. (12) and appears an order of magnitude higher than that shown in Ref. [14].

The coincidence rate $N_m'$ is of great interest, especially in the research that requires photon heralding, e.g., the time-energy entanglement [26–28]. In a temporal histogram for detection

<table>
<thead>
<tr>
<th>Number</th>
<th>$N_m'$ [kHz]</th>
<th>$\Delta\lambda$ [nm]</th>
<th>$B'$ [(s $\cdot$ mW $\cdot$ nm)$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRR1</td>
<td>4.6</td>
<td>0.085</td>
<td>$5.4 \times 10^5$</td>
</tr>
<tr>
<td>MRR2</td>
<td>6.9</td>
<td>0.060</td>
<td>$1.1 \times 10^5$</td>
</tr>
<tr>
<td>MRR3</td>
<td>15.2</td>
<td>0.036</td>
<td>$4.2 \times 10^5$</td>
</tr>
<tr>
<td>MRR4</td>
<td>30.0</td>
<td>0.025</td>
<td>$1.2 \times 10^6$</td>
</tr>
<tr>
<td>MRR5</td>
<td>9.6</td>
<td>0.021</td>
<td>$4.8 \times 10^5$</td>
</tr>
<tr>
<td>MRR6</td>
<td>2.3</td>
<td>0.017</td>
<td>$1.4 \times 10^5$</td>
</tr>
</tbody>
</table>
events in two arms of CW-pumped sources, the peak, corresponding to the exact time delay between the signal and idler photons, reveals \( N_{cc} \), which is resolved in the SPDs’ jitter time (300 ps). The average of other coincidence counts represents the accidental coincidence rate \( N_{ac} \). As shown in Figs. 5(a)–5(f), the measured \( N_{cc} \) (red triangle) for all samples becomes higher with \( P_p \) increasing because \( N_{cc} \propto N_c \) and saturates obviously at high power. The measured \( N_{cc} \) for MRR1–MRR6 at \( P_p = 3.98 \) mW reach the maximum of 14, 16, 23, 27, 16, and 7 Hz, respectively, which are several orders of magnitude lower than the corresponding \( N_c \) due to \( N_{cc} = N_c \eta_i \eta_i \). As a comparison, we present the simulation of coincidence rate (red solid), where \( N_c \) is calculated by Eq. (9), and \( \eta_i = \eta_i = 0.006 \) is estimated from the experiments. The simulated \( N_{cc} \) is quadratic in \( P_p \), which agrees well with the measured \( N_{cc} \) at low power. However, the simulated \( N_{cc} \) becomes significantly higher than the measured \( N_{cc} \), where the potential maximum could reach 41, 60, 138, 330, 80, and 23 Hz for MRR1–MRR6, respectively. The disagreement between simulation and measurement at high power is primarily because the general model does not take detector saturation and nonlinear loss into account. By introducing detector saturation using Eq. (14) into \( N_c \) estimation, the modified \( N_{cc} \) (red dotted) overlaps with the unmodified simulation at low power and meanwhile partly saturates at high power. At \( P_p = 3.98 \) mW, the simulated \( N_{cc} \) with detector saturation reaches 24, 32, 55, 101, 52, and 19 Hz for MRR1–MRR6, respectively, demonstrating that large \( N_{cc} \) at high power comes at the cost of huge accidental coincidence, which may miss the true coincidence detection.

Moreover, we assume that the external quality factor does not change with pump power, thus nonlinear loss only has an impact on the intrinsic quality factor. Note that TPA and FCA depend on the circling power \( P_c \), which describes round-trip enhancement for both nonlinear coefficient and nonlinear loss, instead of incident power \( P_p \). Thus, Eq. (2) in a silicon ring cavity takes the form of

\[
\alpha_c = \alpha_i + \beta_T P_c + 6.04 \times 10^{-10} \frac{3^2 \beta_T P_c^2 \tau}{2 \hbar \omega_p A_{eff}^2},
\]

where \( \alpha_i \) denotes the linear loss corresponding to the intrinsic quality factor at lower power, \( \beta_T \) denotes the TPA coefficient valued at \( 5.6 \times 10^{-12} \) m/W [52], \( A_{eff} \) denotes the effective mode area estimated at \( 6 \times 10^{-14} \) m² for our sample [49], and \( \tau \) denotes the free carrier lifetime valued at 10 ns [51]. By substituting Eqs. (2), (4), (5), and (15) into Eq. (6), we get

\[
P_p = \frac{\pi R A Q_e}{2 \nu_g} \left[ \left( \frac{\omega_p - \omega_{res}}{\omega_{res}} \right)^2 + \left( \frac{1}{Q_e} + \frac{\alpha_{pc} \nu_g}{\omega_{res}} \right)^2 \right] P_c.
\]

By solving this quintic equation for a given \( P_p \), the corresponding \( P_c \) that takes nonlinear loss into account can be acquired. By applying \( P_c \) to Eq. (15), we get ring-cavity loss and then estimate the intrinsic quality factor by Eqs. (5) and (7). Hence, the further modified \( N_{cc} \) (red dashed) gets closer to the measured \( N_{cc} \), which have the maximums of 21, 23, 29, 18, and 8 Hz for MRR1–MRR6, respectively. Note that \( N_{cc} \) saturation induced by nonlinear loss behaves more dominantly for MRR4–MRR6 because the round-trip loss is much smaller. The resulting sequence of all samples in \( N_{cc} \) order from high to low agrees well with the general model at low power but does not fit at high power because of the nonlinear loss. In addition, the slight disagreement between the modified simulation and measurement is attributed to the uncertainty induced by the pump off-resonance, the different \( \eta_i \) for each sample, the
alignment drift during each measurement, and the slightly different parameters of each sample in calculating $N_c$.

Although the filtering brings insertion loss that reduces the coincidence rate, the suppression of the accidental coincidence counts is efficient. To quantify the signal-to-noise ratio of the coincidence rate, the suppression of the accidental coincidence rate is measured. As shown in Figs. 5(a)–5(f) (left, black), CAR decreases with $P_p$ and $N_{cc}$ increasing. Because $N_{acc} \propto (N_c + N_{pn} + D/\eta_i)(N_c + N_{pn} + D/\eta_i)$, where $N_c$ is quadratic in $P_p$, $N_{pn}$ represents photon counts from sideband noise and the leaked pump field is linear in $P_p$, $D$ is a constant, and $N_{acc}$ has a fourth-order polynomial dependence on $P_p$, at low power, $N_{acc}$ is often 1 to 2 orders of magnitude lower than $N_{cc}$, but at high power $N_{acc}$ almost approaches to $N_{cc}$, which greatly reduces CAR and makes it difficult to identify the true coincidence counts from the histogram. The maximal CARs for all samples are over 400, demonstrating that all sources are operating in the low-noise regime, particularly the highest CAR of 892, which is achieved in MRR3 with $N_{cc}$ of 0.6 Hz.

During all experiments, a temperature controller was applied to tune the resonance wavelength such that the measurement for each sample could be carried out at on-resonance condition. Being quadratic in the circling power $P_p$ and further biquadratic in the enhancement factor $F$, the highest signal count rate (pair brightness, coincidence rate) takes place when the pump-resonance detuning $\lambda_{pr}$ is zero; that is, the minimal transmittance corresponds to the maximal photon pairs. Note that the FWHM of the filtering at signal/idler wavelength is larger than that of the corresponding resonance, $\lambda_{pr}$ only determines the circling pump power; then, we can use the biquadrate of the normalized enhancement factor $|F_s(\lambda_{pr})|^4$ to describe the gain of spontaneous four-wave mixing. As shown in Figs. 6(a)–6(f), where the normalized transmittance of the pump resonance measured at quite low power is demonstrated (black solid), $N_{cc}$ at $P_p = 3.98$ mW (blue circle) becomes significantly higher when $|\lambda_{pr}|$ gets closer to zero, due to the dramatic increase of $|F_s|^4$ (red solid). Hence, the highest $N_{cc}$ for each sample represents the on-resonance condition and so does the signal count rate. However, the detector saturation prevents $N_{cc}$ from reaching the expected value, even when the nonlinear loss term is taken into $|F_s|^4$, which makes the on-resonance condition less important. The highest $N_{cc}$ for MRR1–MRR4 take approximate values of around 14, 16, 23, and 26 Hz, respectively, when $\lambda_{pr}$ is in between ±10 pm. Moreover, the thermal-based bistability is more obvious for the samples with a high quality factor; for example, the highest $N_{cc}$ reach 16 Hz and 7 Hz for MRR5 and MRR6, corresponding to $\lambda_{pr} = -6$ pm and $\lambda_{pr} = -4$ pm, respectively. Additionally, because we have proved that MRR5–MRR6 with the higher quality factor generate less photon pairs than MRR4, it remains a challenge to achieve stable photon emission due to relatively small FWHM and huge thermal-based bistability.

To ensure that the photon pairs are generated in the single-photon regime, we measure the heralded second-order correlation function for the source using MRR4. The photons in the idler arm pass through a 50% coupler to be separated into two arms, A and B, while the photons in the signal arm H, are used for heralding. By carefully adjusting the fiber length in each arm to ensure the same arrival time of the photons at the detectors, the zero-delayed heralded second-order correlation function [33]
is calculated by the photon counts in $H(N_H)$, the coincidence counts between $H$ and $A/B$ ($N_{HA}/N_{HB}$), and the triple coincidence counts among all arms ($N_{HAB}$), in a time window of 2.4 ns. As shown in Fig. 7, $g_{H}^{(2)}(0)$ becomes higher with $P_p$ increasing, that is, the proportion of noise photons, unheralded photon pairs and heralded photon pairs operating in the multiphoton regime, becomes larger. All $g_{H}^{(2)}(0)$ of lower than 0.5 indicates that the heralded photon pairs operating in the single-photon regime are dominant, especially at $P_p = 0.72$ mW, where the minimal $g_{H}^{(2)}(0)$ reaches 0.13. A lower $g_{H}^{(2)}(0)$ is expected by turning down $P_p$, but fewer triple coincidence counts can be detected during a long measurement. Additionally, as $N_{HA} \propto \eta_i \eta_i$, $N_{HB} \propto \eta_i \eta_i$, and $N_{H} \propto \eta_i$, it is valid to achieve lower $g_{H}^{(2)}(0)$ by increasing $\eta_i$, which can be achieved by using the same approach of coincidence rate scaling as previously discussed.

4. DISCUSSION

A general model is given to simulate the pair rate and pair brightness for photon-pair sources using spontaneous four-wave mixing in microring resonators and can be applied to the sources using other types of resonators [28, 53, 54]. The key strategy of the quality factor optimization for generation scaling is to separate the intrinsic quality factor corresponding to round-trip loss and the external quality factor corresponding to the coupling ratio. We conclude that a high intrinsic quality factor is always useful, while the external quality factor needs to be particularly designed; that is, for a given intrinsic quality factor, the highest pair rate and the highest pair brightness take place in the overcoupling regime and the critical-coupling regime, respectively. Hence, the conventional understanding of using a higher-quality factor may still be valid, but, to achieve it, special attention should be given to reduce the round-trip loss in a ring cavity instead of enlarging the gap width. From the general model, potential generation scaling can be also achieved by enlarging the nonlinear coefficient, the group velocity, or reducing the ring-cavity radius. Note that the final approach is arguable because a small radius microcavity brings large bending loss, which contradicts the intrinsic quality factor; meanwhile, its large free spectral range also limits applications, especially in quantum dense wavelength division multiplexing. We fabricate six all-pass-type microring resonators in a silicon-on-insulator chip, whose structures are the same except for the gap width, in order to characterize the photon-pair generation. The measurement demonstrates the rankings of all samples in both signal count rate and pair brightness orders from high to low that agree well with the predictions. The most explicit proof of the quality factor optimization is that the measured signal count rate ratio, by using the sample with quality factor of $9.15 \times 10^4$ to $6.21 \times 10^4$, should have been 3.18 from conventional strategies [41, 44, 45] but becomes 0.077 in our demonstration. Thus, being experimentally validated, this work shows that reducing the round-trip loss in the ring cavity and designing a suitable gap are more effective in generation rate scaling than simply increasing the quality factor.

It is worth noting that the measured signal count rate suffers from huge saturation at high power, which is primarily attributed to the detector saturation. The saturation affects the coincidence detection strongly, which makes the on-resonance of the pump less important at high power. This issue can be improved by using the detectors with higher detection rate, shorter dead-time, and near-unity efficiency in future works [55, 56]. As the modified simulation also takes nonlinear loss of silicon waveguides into account, which well describes the limitation of the coincidence rate scaling, future direction can focus on materials such as AlGaAs and silicon nitride, which have potential to achieve a high intrinsic quality factor even at high power. Moreover, the characterization can be more sufficient when the experiments are carried out in the whispering gallery mode microcavities with an ultrahigh quality factor [57–59]. The external quality factor can be continuously adjusted by moving fiber tapers [60], and the intrinsic quality factor can be flexibly controlled by transferring individual polystyrene nanoparticles [61]. More significantly for future applications, the breathtaking merits of momentum transformation ensure ultra-broadband coupling [62], and the on-chip lasing makes loss-free pump incidence possible [63]. In addition, higher total efficiency also enables a higher coincidence rate and results in a smaller non-zero-delayed heralded second-order correlation function, representing that more photon pairs operate in the low-noise single-photon scheme. To sum up, this work demonstrates the approaches of photon-pair generation scaling, especially with a suitable microring resonator design, which further benefits applications of on-chip quantum optics.

Funding. National Natural Science Foundation of China (NSFC) (60907003); Natural Science Foundation of Hunan Province, China (13J13001); Program for New Century Excellent Talents in University (NCET), China (NCET-12-0142); Danmarks Grundforskningsfond (DNRF) (DNRF123); China Scholarship Council (CSC).

Acknowledgment. The authors would like to thank Prof. Karsten Rottwitt, Dr. Davide Bacco, Dr. Mikkel Heuck, Mr. Jesper B. Christensen, Mr. Erik N. Christensen, Mr. Minghong Gao, and Mr. Dingbo Chen for their useful suggestions.


