Frequency-domain Hong-Ou-Mandel interference

Toshiki Kobayashi[1], Rikizo Ikuta[1], Shuto Yasui[1], Shigehito Miki[2], Taro Yamashita[2], Hirotaka Terai[2], Takashi Yamamoto[1], Masato Koashi[3], and Nobuyuki Imoto[1]

[1] *Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[2] *Advanced ICT Research Institute, National Institute of Information and Communications Technology (NICT), Kobe 651-2492, Japan*
[3] *Photon Science Center, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*

Hong-Ou-Mandel (HOM) interference[1] unveils a distinct behavior of identical particles which cannot be distinguished from each other. Especially for bosons, two separated identical particles passing through a beamsplitter always go together into one of the output ports, but that is not the case with other particles including fermions or classical ones. So far many elemental properties of quantum physics and information[2] have been discovered through the concatenated HOM effects, which has been demonstrated in photons [1, 3, 4, 5, 6, 7, 8] and recently in plasmons[9, 10], atoms[11] and phonons[12]. However, all demonstrations in optical region employed two particles in different spatial modes. Here we first report the HOM interference between two photons in a single spatial mode with different frequencies (energies) by using a partial frequency conversion. The demonstrated frequency-domain interferometer allows us to replace spatial optical paths by optical frequency multiplexing, which opens up a distinct architecture of the quantum interferometry.

In the past three decades since the HOM interference has been proposed and demonstrated with two photons from spontaneous parametric down-conversion (SPDC) process [1], a huge varieties of experiments based on the HOM interference revealed fundamental properties in quantum physics, especially in quantum optics[2], and its applications are widely spreading over quantum information processing, such as quantum computation[13, 14, 15, 16], quantum key distribution[17, 18], quantum repeater[19, 20, 21] and quantum-optical coherence tomography[22]. HOM interference has been observed with photons generated not only from nonlinear optical phenomenon but also from quantum dots[3, 4], trapped neutral atoms[5], trapped ions[6], NV centers[7] and SiV centers[8] in diamond. Furthermore not only photons but also other bosonic particles, e.g., surface plasmons[9, 10], Helium 4 atoms[11] and phonons[12] show the HOM interference. In spite of such demonstrations using various kinds of physical systems, to the best of our knowledge, all of them essentially used the spatial or polarization degree of freedom for the HOM interference, including the use of polarization modes of photons that are easily converted to and from spatial modes. The demonstrations use the beamsplitter (BS) which mixes the two particles in different spatial/polarization modes.

In this letter, we report the first observation of the HOM interference between two photons with different frequencies in optical region. In contrast to the spatial interferometer, the frequency-domain HOM interferometer is implemented in a single spatial mode with a nonlinear optical frequency conversion[23, 24, 25]. In the experiment, we input a 780 nm photon and a 1522 nm photon to the frequency converter that partially converts the wavelengths of the photons between 780 nm and 1522 nm as shown in Fig. 1a. We measured coincidence counts between the output photons at 780 nm and those at 1522 nm from the frequency converter. The observed HOM interference between the two photons in a single spatial mode at different frequencies clearly indicates the nonclassical property.

The frequency-domain HOM interference between the photons with different frequencies in this paper is performed by using a partial frequency converter based on the second-order nonlinear optical effect [24, 26]. Suppose that an upper angular frequency ω_s, a lower angular frequency ω_i and an angular frequency ω_p of the pump light satisfy $\omega_i = \omega_s - \omega_p$. When the pump light is sufficiently strong, the effective Hamiltonian of the nonlinear optical process is described by

$$\hat{H} = i\hbar \left(g^* \hat{a}_i^\dagger \hat{a}_s - g \hat{a}_i \hat{a}_s^\dagger \right), \tag{1}$$

where $\hat{a}_s$ and $\hat{a}_i$ are annihilation operators of the upper and the lower frequency modes, respectively. $g = |g|e^{i\phi}$ is proportional to the complex amplitude of the pump light, where ϕ represents the phase of the pump
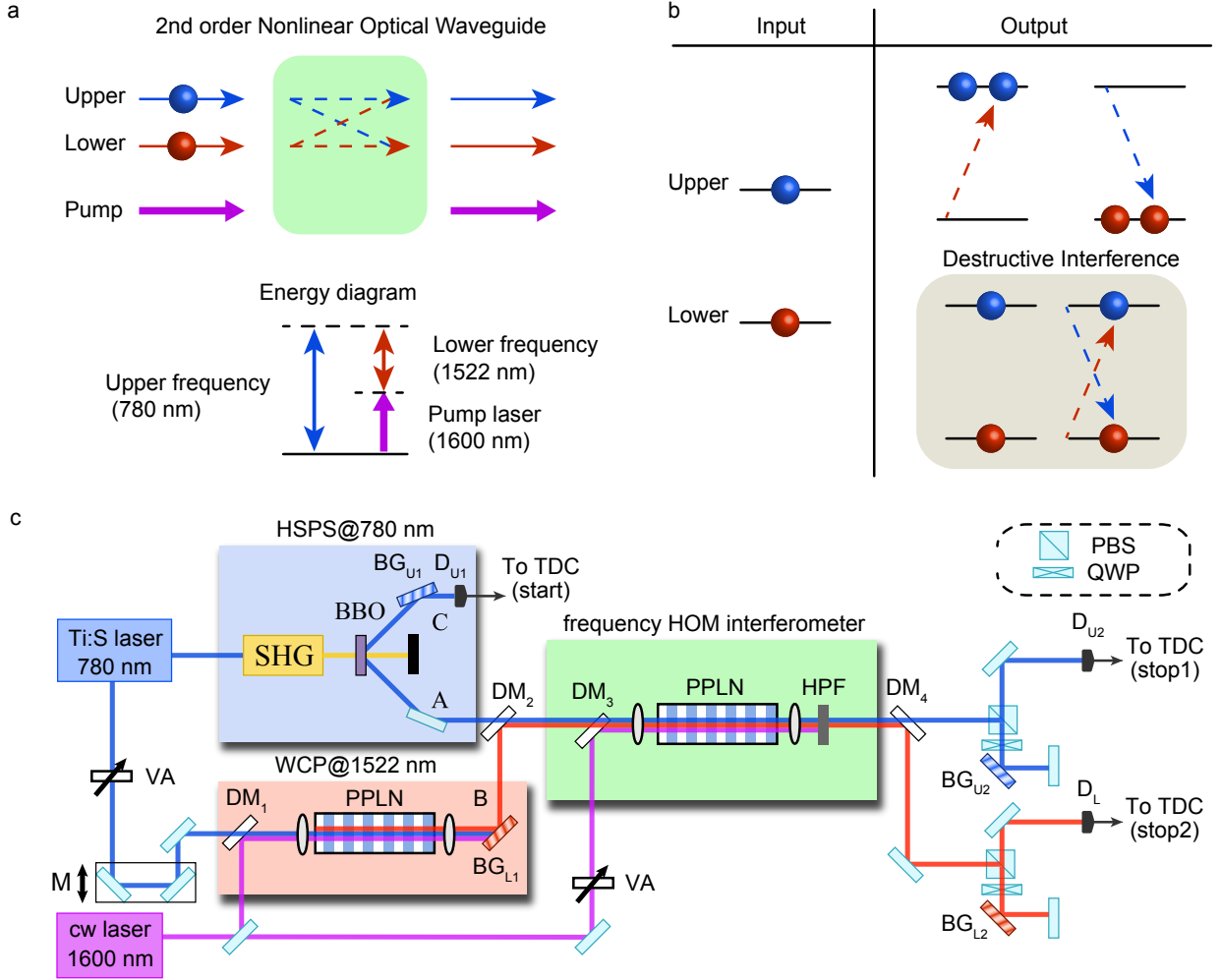
Figure 1 | Frequency-domain HOM interferometer. a, Frequency converter based on second-order nonlinear optical effect. It partially converts the wavelengths of the photons in a single spatial mode from/to 780 nm to/from 1522 nm via sum/difference frequency generation. b, Principle of the frequency-domain HOM effect. When a single photon in upper mode and another photon in the lower mode are mixed by the frequency converter, the single photon occupation events in the output disappear due to the destructive interference. c, The experimental setup of the frequency-domain HOM interference. In the experiment, the heralded single photon source (HSPS) at 780 nm and the weak coherent pulse (WCP) at 1522 nm are prepared to serve as two input photons to the frequency HOM interferometer. The two photons are combined by DM_2 to a single spatial mode and then go into the PPLN waveguide as the frequency-domain BS. The output light pulses are separated into two spatial modes by DM_4 for the photon detection of the two frequency modes.

light. By using Eq. (1), annihilation operators $\hat{a}_{s,\text{out}}$ and $\hat{a}_{i,\text{out}}$ of the signal and the idler modes coming out from the nonlinear optical medium are described by

$$\hat{a}_{s,\text{out}} = \cos(|g|\tau)\hat{a}_s - e^{i\phi}\sin(|g|\tau)\hat{a}_i \tag{2}$$

and

$$\hat{a}_{i,\text{out}} = e^{-i\phi}\sin(|g|\tau)\hat{a}_s + \cos(|g|\tau)\hat{a}_i, \tag{3}$$

where τ is the traveling time of the light pulses through the nonlinear optical medium. The probability of a photon staying in the same input frequency mode is given by $\cos^2(|g|\tau)$ and the transition probability of the photon from an input frequency mode to the other frequency mode is $\sin^2(|g|\tau)$. The process of the frequency conversion described in Eqs. (2) and (3) can be seen as a BS with two different frequency input (output) modes[27, 28, 29]. Borrowing the terminology of the spatial BS, we may regard $\cos^2(|g|\tau)$ and $\sin^2(|g|\tau)$ as the transmittance and the reflectance, respectively.

In the frequency converter, the transition probability can be adjusted by changing the pump power. When we choose the pump power such that $\cos^2(|g|\tau) = \sin^2(|g|\tau) = 1/2$, the frequency converter works as a half BS acting on the two frequency modes. In such a situation, when a single photon in upper mode and
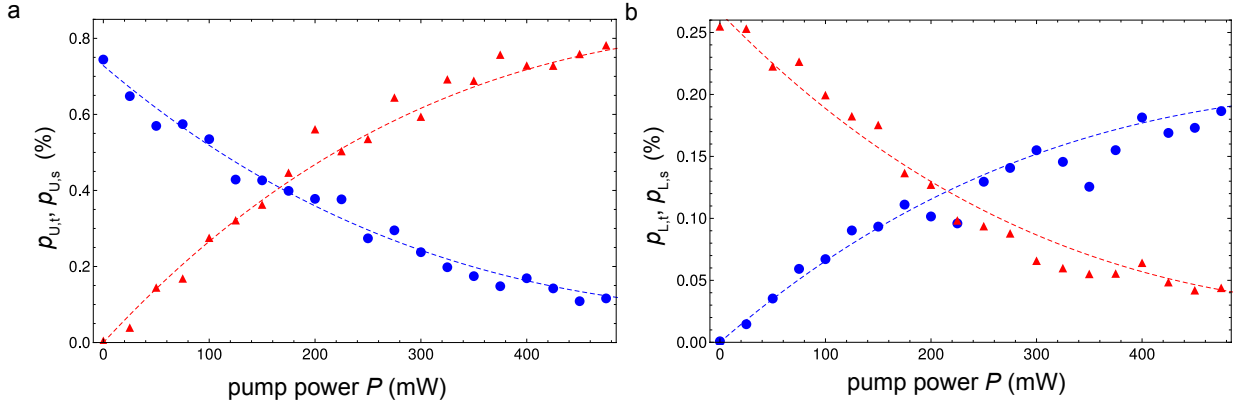
2

Figure 2 | The count rate vs. the pump power. a, The count rate of the transition/staying events per pulse $p_{U,t/s}$ (circle/triangle) for the heralded single photon at 780 nm when the heralding signal is detected at D_{V1}. b, The count rate of the transition/staying events per pulse $p_{L,t/s}$ (triangle/circle) for the coherent light pulse at 1522 nm. The dashed curves are obtained from our theoretical model with the observed values of $p_{U,t/s}$ and $p_{L,t/s}$ (see Supplementary material).
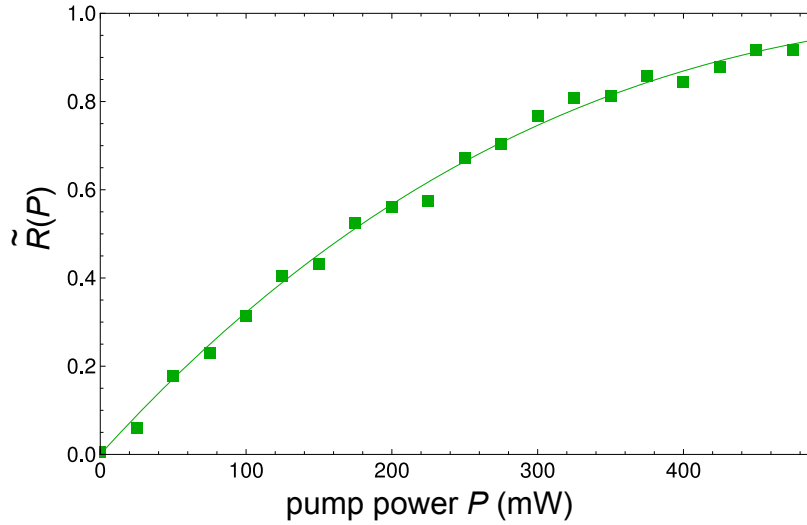

Figure 3 | The peak value of the internal transition probability. The curve is obtained by the best fit to $\widetilde{R}(P)$ with $A\sin^2(\sqrt{\eta P})$, where $A \approx 0.99$ and $\eta \approx 0.0036$ /mW.

another single photon in lower mode are injected to the frequency converter simultaneously, the two photons never come out from the different frequency modes but always come out in the same frequency modes. This phenomenon is a precise analog of the HOM interference in the frequency degree of freedom. One may wonder why unlike the conventional HOM experiments, two distinct bosons, which are distinguished by their frequencies, show the HOM interference. As shown in Fig. 1b, however, the crux of the HOM interference is the destructive interference between the output events. Therefore, what is really required in the process is the indistinguishability after the frequency conversion, which can be fulfilled by a suitable coherent property of the converter.

The experimental setup for the frequency-domain HOM interference by using the partial frequency converter is shown in Fig. 1c. We prepare a vertically(V) polarized heralded single photon at 780 nm in mode A and a V polarized weak coherent light at 1522 nm in mode B with an average photon number of 0.1 (see Method). The two light pulses are combined by a dichroic mirror (DM$_2$) and then focused on a type-0 quasi-phase-matched periodically-poled LiNbO$_3$ (PPLN) waveguide[24] for the frequency conversion. The time difference between the two light pulses is adjusted by mirrors (M) on a motorized stage. The V polarized cw pump laser at 1600 nm is combined with the two input light pulses by DM$_3$ and focused on the PPLN waveguide. The length of the PPLN crystal is 20 mm and the acceptable bandwidth is calculated to be $\Delta_{WG} \equiv 140$ GHz which corresponds to 0.28 nm for 780-nm light and 1.1 nm for 1522-nm light. The pump power is adjusted by a variable attenuator (VA) and determines the transition probability of the frequency converter.

After the frequency converter, the light pulses at 780 nm and 1522 nm are separated by DM$_4$ and Bragg
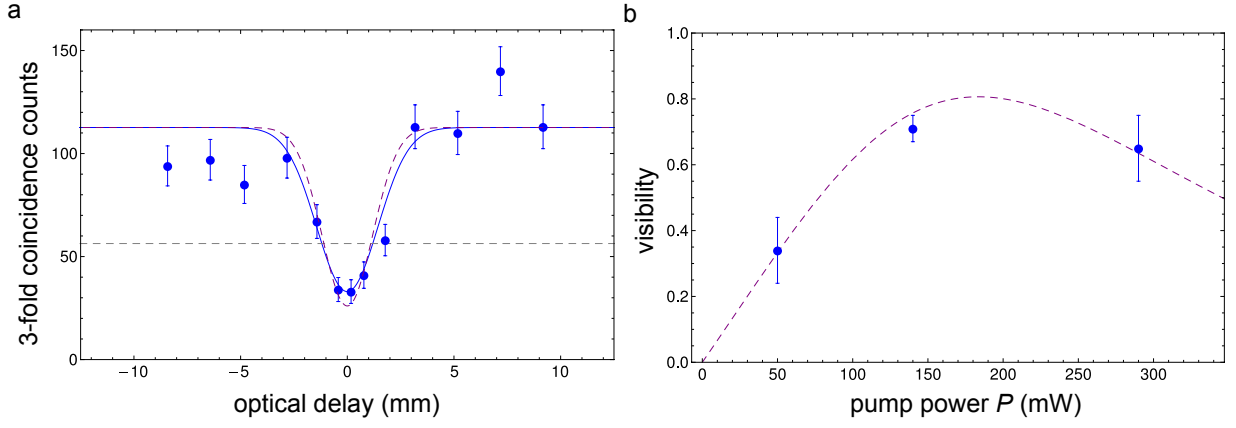
Figure 4 | Observed frequency-domain HOM interference. a, The observed HOM dip at 140-mW pump power. The circles represent the experimental threefold coincidence counts. The solid curve is the Gaussian fit to the experimental counts. The dashed curve is obtained from our theoretical model with the experimental parameters. The dashed horizontal line describes the half values of the maximum of the fitting result. b, The pump power dependence of the visibility. The circles are obtained from the experimental result. The dashed curve is obtained from our theoretical model with the experimental parameters.

gratings (BG$_{U2}$ and BG$_{L2}$ with bandwidths of $\Delta_U \equiv 99$ GHz and $\Delta_L \equiv 130$ GHz, respectively). They are then measured by an avalanche photodiode with the quantum efficiency of about 60% for 780-nm photons (D$_{U2}$) and by a superconducting single-photon detector (SSPD)[30] with the quantum efficiency of about 60% for the 1522-nm photons (D$_L$), respectively. In order to observe the HOM interference, we collect the threefold coincidence events among the three detectors D$_{U1}$,D$_{U2}$ and D$_L$. Note that SSPD with the quantum efficiency of about 10% for 780-nm photons (D$_{U1}$) is used for heralding the 780-nm input photon in mode A.

Before we demonstrate the frequency-domain HOM interference, we first measured the dependencies of the count rates of the transition/staying events $p_{U,t/s}$ and $p_{L,t/s}$ on the pump power for each of the upper and the lower input photons, respectively. The experimental result is shown in Figs. 2a and 2b. From the experimental result, we estimate the internal transition probability R of the frequency converter, which depends on the pump power P and the frequency of the input light, by constructing a theoretical model as follows.

We assume that the internal transition probability $R = R(P, \omega)$ is a Gaussian with the bandwidth of Δ_{WG} around the center of 780 nm/1522 nm for upper/lower input light, at which the peak value is given by $\widetilde{R}(P)$. We also assume that the optical circuit for the frequency-domain HOM after preparing the single photon and the coherent light pulse is decomposed to a lossless frequency converter with the transition probability $R(P, \omega)$ and the staying probability $T(P, \omega)(= 1 - R(P, \omega))$, two lossy media inducing the loss in each of the upper and the lower input modes, and two spectral filters acting on two output modes. The transmittances of the lossy media for the upper and the lower input light are denoted by $T_{in,U}$ and $T_{in,L}$, respectively, which describe the optical loss including the insertion loss to the frequency converter. The transmittances of the spectral filters for the upper and the lower output light are denoted by $T_{out,U}(\omega)$ and $T_{out,L}(\omega)$, respectively, which describe the optical loss including the BGs after the frequency converter and the quantum efficiencies of the detectors. We assume that the transmittance $T_{out,U/L}(\omega)$ is Gaussian with the peak value of $\widetilde{T}_{out,U/L}$. The bandwidths of $T_{out,U}(\omega)$ and $T_{out,L}(\omega)$ are calculated to be 70 GHz and 92 GHz, respectively, from the effect of using the BG$_{U2}$ and BG$_{L2}$ twice. For the input light pulses, we assume that the spectral shapes of the heralded single photon and the coherent light pulse are Gaussian with bandwidths of 740 GHz and 93 GHz, respectively. These are calculated by using the experimental parameters $\Delta_{WG}, \Delta_U, \Delta_L$ and the pulse width of the laser source $\Delta t \equiv 1.2$ ps. The emission rate of the photon pair from the SPDC is so small that the multiple-pair events are negligible. In addition, the heralded genuine single photon is assumed to be in a pure state because the narrow-band spectral filtering destroys the spectral correlation between the photon pair from SPDC. Under these assumptions, the initial state composed of the single photon and the coherent state is regarded as a pure state described by $|\phi\rangle = \hat{a}_U^\dagger \hat{D}_L(\alpha) |0\rangle$, where $\hat{a}_U^\dagger$ is a creation operator of the upper mode, $\hat{D}_L(\alpha)$ is a displacement operator with complex number α of the lower mode and $|0\rangle$ is the vacuum state for both modes. Based on the above theoretical model with the use of the observed values $p_{U,t}(P), p_{U,s}(P), p_{L,t}(P)$ and $p_{L,s}(P)$, we calculated the four parameters $\widetilde{R}(P)$, $T_{in,U}\widetilde{T}_{out,U}$, $T_{in,U}\widetilde{T}_{out,L}$ and

$|\alpha|^2 T_{\text{in,L}}/T_{\text{in,U}}$ for various values of P (see Supplementary material). The result of $\widetilde{R}(P)$ is shown in Fig. 3.

Next we demonstrated the frequency-domain HOM interference by using the frequency converter. We set a pump power to be 140 mW which results in the transition probability ~ 0.4 of the frequency converter according to Fig. 3. The experimental result of the dependency of the threefold coincidence counts on the optical delay is shown in Fig. 4a. The observed visibility of 0.71 ± 0.04 at the zero delay point was obtained by the best fit to the experimental data with a Gaussian. The full width at the half maximum was approximately 1.7 mm which corresponds to ~ 6 ps of a delay time. The high visibility clearly shows the nonclassical HOM interference between the two light pulses in a single spatial mode with different frequencies. We also measured the visibilities at the pump power 50 mW and 290 mW, which corresponds to the transition probabilities ~ 0.2 and ~ 0.7, respectively. The experimental result is shown in Fig. 4b. The observed visibilities are 0.34 ± 0.10 at 50 mW and 0.65 ± 0.10 at 290 mW.

In the following, we discuss the reasons for the degradation of the visibility. In our theoretical model, we can calculate the visibility by using the experimental parameters $\widetilde{R}(P_0)$, $T_{\text{in,U}}\widetilde{T}_{\text{out,U}}$, $T_{\text{in,U}}\widetilde{T}_{\text{out,L}}$ and $|\alpha|^2 T_{\text{in,L}}/T_{\text{in,U}}$ (see Supplementary material), the results of which are the dashed curves shown in Figs. 4a and 4b. These are in good agreement with the experimental results. Fig. 4b indicates that the highest visibility of 0.81 will be obtained when the pump power is 190 mW which corresponds to the transition probability ~ 0.5. In our theoretical model, main reasons for the degradation of the visibility comes from the input light pulses; the effect of the multiphoton components in the coherent light pulse at 1522 nm and the broad bandwidth of the heralded single photon at 780 nm. If we replace the coherent light pulse by a single photon and set $T_{\text{in,L}}$ to be equal to $T_{\text{in,U}}$, the visibility of the HOM interference is expected to be 0.95 at 190-mW pump power. If we use the single photon at 780 nm with the same bandwidth as that of the coherent light pulse, the visibility will be 0.93 at 190-mW pump power. If we take both improvement for the input light pulses, the visibility will be 0.98 at 190-mW pump power.

In conclusion, we have demonstrated the frequency-domain HOM interference between a heralded single photon at 780 nm and a weak laser light at 1522 nm in a single spatial mode by using the partial frequency converter based on the nonlinear optical effect. We also deduced that the performance of the demonstrated frequency domain HOM interferometer is almost ideal from the fact that the estimated visibility in the case of the ideal input photons is 0.98, which is close to unity. So far the spatial HOM interferometer has been exploited in a wide variety of quantum phenomena including a large scale quantum information processing. We thus believe that the frequency domain HOM interferometer will open up a novel frequency domain quantum interferometry and give a novel tool for exploiting quantum phenomena and a way of scaling up the quantum information processing with a large Hilbert space spanned by widely spreading frequency modes.

Methods

Preparation of the two input light pulses.. A light pulse from a mode-locked Ti:sapphire laser at 780 nm (pulse width: $\Delta t \equiv 1.2$ ps; repetition rate: 82 MHz) is used for the preparation. It is divided into two beams. One beam is used for preparing a heralded single photon at 780 nm. The beam is frequency doubled (wavelength: 390 nm; power: 200 mW) by second-harmonic generation (SHG), and then pumps a type-I phase-matched 1.5-mm-thick β-barium borate (BBO) crystal to generate a photon pair at 780 nm in modes A and C through the spontaneous parametric down conversion. The photon in mode C is measured by a SSPD denoted by D_{U1}, which prepares a heralded single photon in mode A. The spectral filtering of the photon in mode C is performed by a Bragg grating (BG_{U1}) with a bandwidth of $\Delta_{\text{U}} \equiv 99$ GHz which corresponds to 0.2 nm for 780-nm light.

The other beam from Ti:S laser is used for preparing the weak coherent light pulse at 1522 nm. The beam enters a difference frequency generation (DFG) module. In the DFG module, a V polarized cw pump laser at 1600 nm is combined with the input light pulse at 780 nm by a dichroic mirror (DM_1). They are focused on a type-0 quasi-phase-matched PPLN waveguide[24]. The length of the PPLN crystal is 20 mm and the acceptable bandwidth is calculated to be $\Delta_{\text{WG}} = 140$ GHz. After passing through the PPLN waveguide, the converted light at 1522 nm is extracted by BG_{L1} with a bandwidth of $\Delta_{\text{L}} \equiv 130$ GHz which corresponds to 1 nm for 1522-nm light. We adjust the average photon number of the coherent light pulse at 1522 nm to be ~ 0.1 by a variable attenuator (VA).

References

[1] Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. Phys. Rev. Lett. **59**, 2044-2046 (1987).

[2] Pan, J.-W. *et al.* Multiphoton entanglement and interferometry. Rev. Mod. Phys. **84**, 777-838 (2012).

[3] Santori, C., Fattal, D., Vuckovic, J., Solomon, G. S. & Yamamoto, Y. Indistinguishable photons from a single-photon device. Nature **419**, 594-597 (2002).

[4] Patel, R. B. *et al.* Two-photon interference of the emission from electrically tunable remote quantum dots. Nat. Photon. **4**, 632-635 (2010).

[5] Beugnon, J. *et al.* Quantum interference between two single photons emitted by independently trapped atoms. Nature **440**, 779-782 (2006).

[6] Maunz, P. *et al.* Quantum interference of photon pairs from two remote trapped atomic ions. Nat. Phys. **3**, 538-541 (2007).

[7] Sipahigil, A. *et al.* Quantum Interference of Single Photons from Remote Nitrogen-Vacancy Centers in Diamond. Phys. Rev. Lett. **108**, 143601 (14 2012).

[8] Sipahigil, A. *et al.* Indistinguishable Photons from Separated Silicon-Vacancy Centers in Diamond. Phys. Rev. Lett. **113**, 113602 (2014).

[9] Di Martino, G. *et al.* Observation of Quantum Interference in the Plasmonic Hong-Ou-Mandel Effect. Phys. Rev. Applied **1**, 034004 (2014).

[10] Fakonas, J. S., Lee, H., Kelaita, Y. A. & Atwater, H. A. Two-plasmon quantum interference. Nat. Photon. **8**, 317-320 (2014).

[11] Lopes, R. *et al.* Atomic Hong-Ou-Mandel experiment. Nature **520**, 66-68 (2015).

[12] Toyoda, K., Hiji, R., Noguchi, A. & Urabe, S. Hong-Ou-Mandel interference of two phonons in trapped ions. Nature **527**, 74-77 (2015).

[13] Kok, P. *et al.* Linear optical quantum computing with photonic qubits. Rev. Mod. Phys. **79**, 135-174 (2007).

[14] Barz, S. *et al.* Demonstration of Blind Quantum Computing. Science 335, 303-308 (2012).

[15] Spagnolo, N. *et al.* Experimental validation of photonic boson sampling. Nat. Photon. **8**, 615-620 (2014).

[16] Carolan, J. *et al.* Universal linear optics. Science **349**, 711-716 (2015).

[17] Tang, Y.-L. *et al.* Measurement-Device-Independent Quantum Key Distribution over 200 km. Phys. Rev. Lett. **113**, 190501 (2014).

[18] Guan, J.-Y. *et al.* Experimental Passive Round-Robin Differential Phase-Shift Quantum Key Distribution. Phys. Rev. Lett. **114**, 180502 (2015).

[19] Sangouard, N., Simon, C., de Riedmatten, H. & Gisin, N. Quantum repeaters based on atomic ensembles and linear optics. Rev. Mod. Phys. **83**, 33-80 (2011).

[20] Hofmann, J. *et al.* Heralded Entanglement Between Widely Separated Atoms. Science **337**, 72-75 (2012).

[21] Bao, X.-H. *et al.* Quantum teleportation between remote atomic-ensemble quantum memories. Proceedings of the National Academy of Sciences **109**, 20347-20351 (2012).

[22] Teich, M., Saleh, B., Wong, F. & Shapiro, J. Variations on the theme of quantum optical coherence tomography: a review. Quantum Information Processing **11**, 903-923. issn: 1570- 0755 (2012).

[23] Tanzilli, S. *et al.* A photonic quantum information interface. Nature **437**, 116-120 (2005).

[24] Ikuta, R. *et al.* Wide-band quantum interface for visible-to-telecommunication wavelength conversion. Nat. Commun. **2**, 537- (2011).

[25] Ikuta, R. *et al.* High-fidelity conversion of photonic quantum information to telecommunication wavelength with superconducting single-photon detectors. Phys. Rev. A **87**, 010301 (2013).

[26] Kumar, P. Quantum frequency conversion. Opt. Lett. **15**, 1476-1478 (1990).

[27] Giorgi, G., Mataloni, P. & De Martini, F. Frequency Hopping in Quantum Interferometry: Efficient

Up-Down Conversion for Qubits and Ebits. Phys. Rev. Lett. **90**, 027902 (2003).

[28] Raymer, M., van Enk, S., McKinstrie, C. & McGuinness, H. Interference of two photons of different color. Optics Communications **283**, 747 -752 (2010).

[29] Ikuta, R. *et al.* Observation of two output light pulses from a partial wavelength converter preserving phase of an input light at a single-photon level. Opt. Express **21**, 27865-27872 (2013).

[30] Miki, S., Yamashita, T., Terai, H. & Wang, Z. High performance fiber-coupled NbTiN superconducting nanowire single photon detectors with Gifford-McMahon cryocooler. Opt. Express **21**, 10208-10214 (2013).

Acknowledgements

Supplementary material

Estimation of the internal transition probability of the frequency converter. We estimated the internal transition probability of the frequency converter by constructing a theoretical model as follows. We assume that the optical circuit of the frequency-domain HOM interferometer is decomposed to a lossless frequency converter, two loss media inducing the loss in each of the upper and the lower input modes, and two spectral filters acting on each of two output modes. The transition probability and the staying probability of the lossless frequency converter are described by $R = R(P,\omega)$ and $T = 1 - R(P,\omega)$. The transmittances of the loss media and the spectral filters for the upper/lower light are denoted by $T_{\mathrm{in,U/L}}$ and $T_{\mathrm{out,U/L}}(\omega)$, respectively. As is described in the main text, we assume the state of the input light composed of the single photon and the coherent state is $|\phi\rangle = \hat{a}_{\mathrm{U}}^{\dagger}\hat{D}_{\mathrm{L}}(\alpha)|0\rangle$, where $\hat{a}_{\mathrm{U}}^{\dagger}$ is a creation operator of the upper mode, $\hat{D}_{\mathrm{L}}(\alpha)$ is a displacement operator with complex number α of the lower mode and $|0\rangle$ is the vacuum state for both modes. The spectral shapes of the heralded single photon and the coherent light pulse are denoted by $F_{\mathrm{U}}(\omega)$ and $F_{\mathrm{L}}(\omega)$, respectively. They are normalized as $\int d\omega F_{\mathrm{U}}(\omega) = \int d\omega F_{\mathrm{L}}(\omega) = 1$.

From the theoretical model, the observed values in Fig. 2 are described by

$$p_{\mathrm{U,t}}(P) = T_{\mathrm{in,U}} \int d\omega F_{\mathrm{U}}(\omega) R(P,\omega) T_{\mathrm{out,L}}(\omega), \tag{4}$$

$$p_{\mathrm{U,s}}(P) = T_{\mathrm{in,U}} \int d\omega F_{\mathrm{U}}(\omega)(1 - R(P,\omega)) T_{\mathrm{out,U}}(\omega), \tag{5}$$

$$p_{\mathrm{L,t}}(P) = |\alpha|^2 T_{\mathrm{in,L}} \int d\omega F_{\mathrm{L}}(\omega) R(P,\omega) T_{\mathrm{out,U}}(\omega), \tag{6}$$

$$p_{\mathrm{L,s}}(P) = |\alpha|^2 T_{\mathrm{in,L}} \int d\omega F_{\mathrm{L}}(\omega)(1 - R(P,\omega)) T_{\mathrm{out,L}}(\omega). \tag{7}$$

We assume that $F_{\mathrm{U}}(\omega)$, $F_{\mathrm{L}}(\omega)$ and $R(P,\omega)$ are Gaussian with bandwidths of $\Delta_{\mathrm{in,U}} = 740$ GHz, $\Delta_{\mathrm{in,L}} = 93$ GHz and $\Delta_{\mathrm{WG}} = 140$ GHz, respectively. The peak value of $R(P,\omega)$ is described by $\widetilde{R}(P)$. We also assume that $T_{\mathrm{out,U}}(\omega)$ and $T_{\mathrm{out,L}}(\omega)$ are Gaussian with the bandwidths of $\Delta_{\mathrm{out,U}} = 70$ GHz and $\Delta_{\mathrm{out,L}} = 92$ GHz around the center of 780 nm/1522 nm for the upper/lower input light, at which the peak values are given by $\widetilde{T}_{\mathrm{out,U}}$ and $\widetilde{T}_{\mathrm{out,L}}$. From the four observed values $p_{\mathrm{U,t}}(P), p_{\mathrm{U,s}}(P), p_{\mathrm{L,t}}(P)$ and $p_{\mathrm{L,s}}(P)$ in Fig. 2, we obtain the four parameters $\widetilde{R}(P), T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}, T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$ and $|\alpha|^2 T_{\mathrm{in,L}}/T_{\mathrm{in,U}}$ for various values of P. The four observed values are described by

$$p_{\mathrm{U,t}}(P) = T_{\mathrm{in,U}}\widetilde{R}(P)\widetilde{T}_{\mathrm{out,L}}\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,L}}^2}}, \tag{8}$$

$$p_{\mathrm{U,s}}(P) = T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}\left(\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2}} - \widetilde{R}(P)\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2}}\right), \tag{9}$$

$$p_{\mathrm{L,t}}(P) = |\alpha|^2 T_{\mathrm{in,L}}\widetilde{R}(P)\widetilde{T}_{\mathrm{out,U}}\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,U}}^2}}, \tag{10}$$

$$p_{\mathrm{L,s}}(P) = |\alpha|^2 T_{\mathrm{in,L}}\widetilde{T}_{\mathrm{out,L}}\left(\frac{1}{1 + \sqrt{\Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2}} - \widetilde{R}(P)\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2}}\right). \tag{11}$$

From the above equations, we obtain the quadratic equation $0 = A\widetilde{R}(P)^2 + B\widetilde{R}(P) + C$, where

$$A \equiv \frac{1}{\sqrt{(1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,L}}^2)(1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,U}}^2)}}$$
$$- \frac{1}{\sqrt{(1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2)(1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2)}} \frac{p_{\mathrm{U,t}}(P)p_{\mathrm{L,s}}(P)}{p_{\mathrm{U,s}}(P)p_{\mathrm{L,t}}(P)}, \quad (12)$$

$$B \equiv \left(\frac{1}{\sqrt{(1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2)(1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2)}} \right.$$
$$\left. + \frac{1}{\sqrt{(1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2)(1 + \Delta_{\mathrm{out,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{out,L}}^2/\Delta_{\mathrm{in,L}}^2)}} \right) \frac{p_{\mathrm{U,t}}(P)p_{\mathrm{L,s}}(P)}{p_{\mathrm{U,s}}(P)p_{\mathrm{L,t}}(P)}, \quad (13)$$

$$C \equiv - \frac{1}{\sqrt{(1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2)(1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2)}} \frac{p_{\mathrm{U,t}}(P)p_{\mathrm{L,s}}(P)}{p_{\mathrm{U,s}}(P)p_{\mathrm{L,t}}(P)}. \quad (14)$$

We see that $\widetilde{R}(P) = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$ is the physical solution which satisfies $0 \le \widetilde{R}(P) \le 1$. The result of $\widetilde{R}(P)$ is shown in Fig. 3. The other three parameters $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$ and $|\alpha|^2 T_{\mathrm{in,L}}/T_{\mathrm{in,U}}$ are described by

$$\frac{|\alpha|^2 T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} = \frac{p_{\mathrm{L,s}}(P)}{p_{\mathrm{U,t}}(P)} \frac{\widetilde{R}(P)}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,L}}^2}}$$
$$\times \left(\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2}} - \widetilde{R}(P) \frac{1}{\sqrt{1 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,L}}^2/\Delta_{\mathrm{out,L}}^2}} \right)^{-1}, \quad (15)$$

$$T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}} = p_{\mathrm{U,s}}(P) \left(\frac{1}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2}} - \widetilde{R}(P) \frac{1}{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,U}}^2}} \right)^{-1}, \quad (16)$$

$$T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}} = p_{\mathrm{U,t}}(P) \frac{\sqrt{1 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{WG}}^2 + \Delta_{\mathrm{in,U}}^2/\Delta_{\mathrm{out,L}}^2}}{\widetilde{R}(P)}. \quad (17)$$

From these equations, we calculated these parameters by using the observed values. The result is shown in Fig. S1. We note that these parameters should be independent of the pump power. The estimated values of $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$ and $|\alpha|^2 T_{\mathrm{in,L}}/T_{\mathrm{in,U}}$ take about constant values of 0.078, 0.081 and 0.047, respectively, which is consistent with our theoretical model. By using these estimated values, we obtain the dashed curves in Fig. 2. These curves are in good agreement with the experimental results.

Estimation of the degradation of the visibility. We estimated the degradation of the visibility by using the theoretical model as follows. The coincidence probability $p_c(\tau)$ on the time delay τ is described by

$$p_c(\tau) = 1 - p_{\mathrm{U0}}(\tau) - p_{\mathrm{L0}}(\tau) + p_{\mathrm{U0,L0}}(\tau), \quad (18)$$

where $p_{\mathrm{U0/L0}}(\tau)$ and $p_{\mathrm{U0,L0}}(\tau)$ represent probabilities where no photon is detected on the time delay τ in the upper/lower mode and in both modes, respectively. The visibility of the HOM interference is defined by $1 - p_c(0)/p_c(\infty)$. In our theoretical model, these probabilities are described by

$$p_{\mathrm{U0}}(\tau) = (1 - d_{\mathrm{U}}) \| \langle 0|_{\mathrm{U}} \hat{U}_{\mathrm{FBS}} |\phi\rangle_{\mathrm{U,L}} \otimes |0\rangle_{\mathrm{E}} \|^2, \quad (19)$$

$$p_{\mathrm{L0}}(\tau) = (1 - d_{\mathrm{L}}) \| \langle 0|_{\mathrm{L}} \hat{U}_{\mathrm{FBS}} |\phi\rangle_{\mathrm{U,L}} \otimes |0\rangle_{\mathrm{E}} \|^2, \quad (20)$$

$$p_{\mathrm{U0,L0}}(\tau) = (1 - d_{\mathrm{U}})(1 - d_{\mathrm{L}}) \| \langle 0|_{\mathrm{U,L}} \hat{U}_{\mathrm{FBS}} |\phi\rangle_{\mathrm{U,L}} \otimes |0\rangle_{\mathrm{E}} \|^2, \quad (21)$$

where $d_{\mathrm{U/L}}$ is the observed detection probability of the background noise, shown in Fig. S2. $|\phi\rangle$ is the initial state described by $\hat{a}_{\mathrm{U}}^\dagger \hat{D}_{\mathrm{L}}(\alpha) |0\rangle$. $|0\rangle_{\mathrm{E}}$ is the vacuum state in an ancillary system describing the input/output loss modes. $\hat{U}_{\mathrm{FBS}}$ is a unitary operator describing the frequency-domain BS including the input and output loss. We decompose $\hat{U}_{\mathrm{FBS}}$ to five lossless components, one of which describes lossless part of the frequency-domain BS and others describe the lossy media/spectral filters acting on input/output

9

modes. The actions of $\hat{U}_{\mathrm{FBS}}$ on the creation operators $\hat{a}_{\mathrm{U}}^{\dagger} = \int d\omega \sqrt{F_{\mathrm{U}}(\omega)}\hat{a}_{\mathrm{U},\omega}^{\dagger}$ in the input upper mode and $\hat{a}_{\mathrm{L}}^{\dagger} = \int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{L}}(\omega)}\hat{a}_{\mathrm{L},\omega}^{\dagger}$ in the input lower mode are written by

$$
\hat{U}_{\mathrm{FBS}}\hat{a}_{\mathrm{U}}^{\dagger}\hat{U}_{\mathrm{FBS}}^{\dagger} = \sqrt{R_{\mathrm{in,U}}}\int d\omega \sqrt{F_{\mathrm{U}}(\omega)}\hat{a}_{\mathrm{EU1},\omega}^{\dagger}
$$

$$
+\sqrt{T_{\mathrm{in,U}}}\int d\omega \sqrt{F_{\mathrm{U}}(\omega)T(P,\omega)}\left(\sqrt{T_{\mathrm{out,U}}(\omega)}\hat{a}_{\mathrm{U},\omega}^{\dagger} + \sqrt{R_{\mathrm{out,U}}(\omega)}\hat{a}_{\mathrm{EU2},\omega}^{\dagger}\right)
$$

$$
+\sqrt{T_{\mathrm{in,U}}}\int d\omega e^{-i\phi}\sqrt{F_{\mathrm{U}}(\omega)R(P,\omega)}\left(\sqrt{T_{\mathrm{out,L}}(\omega)}\hat{a}_{\mathrm{L},\omega}^{\dagger} + \sqrt{R_{\mathrm{out,L}}(\omega)}\hat{a}_{\mathrm{EL2},\omega}^{\dagger}\right), \quad (22)
$$

$$
\hat{U}_{\mathrm{FBS}}\hat{a}_{\mathrm{L}}^{\dagger}\hat{U}_{\mathrm{FBS}}^{\dagger} = \sqrt{R_{\mathrm{in,L}}}\int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{L}}(\omega)}\hat{a}_{\mathrm{EL1},\omega}^{\dagger}
$$

$$
+\sqrt{T_{\mathrm{in,L}}}\int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{L}}(\omega)T(P,\omega)}\left(\sqrt{T_{\mathrm{out,L}}(\omega)}\hat{a}_{\mathrm{L},\omega}^{\dagger} + \sqrt{R_{\mathrm{out,L}}(\omega)}\hat{a}_{\mathrm{EL2},\omega}^{\dagger}\right)
$$

$$
-\sqrt{T_{\mathrm{in,L}}}\int d\omega e^{-i\omega\tau+i\phi}\sqrt{F_{\mathrm{L}}(\omega)R(P,\omega)}\left(\sqrt{T_{\mathrm{out,U}}(\omega)}\hat{a}_{\mathrm{U},\omega}^{\dagger} + \sqrt{R_{\mathrm{out,U}}(\omega)}\hat{a}_{\mathrm{EU2},\omega}^{\dagger}\right), \quad (23)
$$

where $R_{\mathrm{in,U/L}} = 1 - T_{\mathrm{in,U/L}}, R_{\mathrm{out,U/L}}(\omega) = 1 - T_{\mathrm{out,U/L}}(\omega)$ and $[\hat{a}_{i,\omega}, \hat{a}_{j,\omega'}^{\dagger}] = \delta_{i,j}\delta(\omega - \omega')$ for $i,j =$ U,L,EU1,EU2,EL1,EL2 . From Eqs. (19) – (21), we have

$$
p_{\mathrm{U0,L0}}(\tau) = (1 - d_{\mathrm{U}})(1 - d_{\mathrm{L}})\left\|\sqrt{R_{\mathrm{in,U}}}\int d\omega \sqrt{F_{\mathrm{U}}(\omega)}\hat{a}_{\mathrm{EU1},\omega}^{\dagger}\langle 0|_{\mathrm{U,L}}\hat{U}_{\mathrm{FBS}}\hat{D}_{\mathrm{L}}(\alpha)|0\rangle_{\mathrm{U,L,E}}\right\|^{2}
$$

$$
+(1 - d_{\mathrm{U}})(1 - d_{\mathrm{L}})
$$

$$
\left\|\sqrt{T_{\mathrm{in,U}}}\int d\omega \sqrt{F_{\mathrm{U}}(\omega)R(P,\omega)}(\sqrt{R_{\mathrm{out,U}}(\omega)}\hat{a}_{\mathrm{EU2},\omega}^{\dagger} + e^{-i\phi}\sqrt{R_{\mathrm{out,L}}(\omega)}\hat{a}_{\mathrm{EL2},\omega}^{\dagger})\right.
$$

$$
\left.\times \langle 0|_{\mathrm{U,L}}\hat{U}_{\mathrm{FBS}}\hat{D}_{\mathrm{L}}(\alpha)|0\rangle_{\mathrm{U,L,E}}\right\|^{2} \quad (24)
$$

$$
= (1 - d_{\mathrm{U}})(1 - d_{\mathrm{L}})
$$

$$
\times \exp\left(-|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\int d\omega F_{\mathrm{L}}(\omega)(R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega) + T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega))\right)
$$

$$
\times \left(\int d\omega F_{\mathrm{U}}(\omega)\left(1 - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega) - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right)\right.
$$

$$
\left.+|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\left|\int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}(T_{\mathrm{out,U}}(\omega) - T_{\mathrm{out,L}}(\omega))\right|^{2}\right), \quad (25)
$$

$$
p_{\mathrm{U0}}(\tau) = (1 - d_{\mathrm{U}})\exp\left(-|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\int d\omega F_{\mathrm{L}}(\omega)R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega)\right)
$$

$$
\times \left(\int d\omega F_{\mathrm{U}}(\omega)(1 - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega))\right.
$$

$$
\left.+|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\left|\int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega)\right|^{2}\right), \quad (26)
$$

$$
p_{\mathrm{L0}}(\tau) = (1 - d_{\mathrm{L}})\exp\left(-|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\int d\omega F_{\mathrm{L}}(\omega)T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right)
$$

$$
\times \left(\int d\omega F_{\mathrm{U}}(\omega)(1 - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega))\right.
$$

$$
\left.+|\alpha|^{2}\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}}\left|\int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right|^{2}\right). \quad (27)
$$

By using the experimental parameters $\widetilde{R}(P)$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$ and $|\alpha|^{2}T_{\mathrm{in,L}}/T_{\mathrm{in,U}}$, we obtain the dashed curves in Figs. 4a and 4b from Eq. (18). Fig. 4b indicates that the highest visibility of 0.81 will be obtained at 190-mW pump power.

In our theoretical model, main reason for the degradation of the visibility comes from input light pulses; the effect of the multiphoton components in the coherent light pulse at 1522 nm and the broad bandwidth of $F_{\mathrm{U}}(\omega)$. When we replace the coherent light pulse by a single photon the visibility is calculated in a similar way with $|\phi\rangle = \hat{a}_{\mathrm{U}}^{\dagger}\hat{a}_{\mathrm{L}}^{\dagger}|0\rangle_{\mathrm{U,L}}$. In this case, the probabilities where no photon is detected in both modes $p_{\mathrm{U0,L0}}^{\mathrm{ph}}(\tau)$ and in the upper/lower mode $p_{\mathrm{U0/L0}}^{\mathrm{ph}}(\tau)$ are described by

$$
\begin{aligned}
p_{\mathrm{U0,L0}}^{\mathrm{ph}}(\tau) &\equiv (1-d_{\mathrm{U}})(1-d_{\mathrm{L}})\| \langle 0|_{\mathrm{U,L}}\, \hat{U}_{\mathrm{FBS}}\hat{a}_{\mathrm{U}}^{\dagger}\hat{a}_{\mathrm{L}}^{\dagger}|0\rangle_{\mathrm{U,L,E}} \|^2 \\
&= (1-d_{\mathrm{U}})(1-d_{\mathrm{L}}) \int d\omega F_{\mathrm{U}}(\omega)\left(1 - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega) - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right) \\
&\quad \times \frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \int d\omega F_{\mathrm{L}}(\omega)\left(1 - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega) - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right) \\
&\quad + (1-d_{\mathrm{U}})(1-d_{\mathrm{L}}) \\
&\quad \times \frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \left| \int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}(T_{\mathrm{out,U}}(\omega) - T_{\mathrm{out,L}}(\omega))\right|^2,
\end{aligned} \tag{28}
$$

$$
\begin{aligned}
p_{\mathrm{U0}}^{\mathrm{ph}}(\tau) &\equiv (1-d_{\mathrm{U}})\| \langle 0|_{\mathrm{U}}\, \hat{U}_{\mathrm{FBS}}\hat{a}_{\mathrm{U}}^{\dagger}\hat{a}_{\mathrm{L}}^{\dagger}|0\rangle_{\mathrm{U,L,E}} \|^2 \\
&= (1-d_{\mathrm{U}}) \int d\omega F_{\mathrm{U}}(\omega)\left(1 - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega)\right) \\
&\quad \times \frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \int d\omega F_{\mathrm{L}}(\omega)\left(1 - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega)\right) \\
&\quad + (1-d_{\mathrm{U}})\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \left| \int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}T_{\mathrm{out,U}}(\omega)\right|^2,
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
p_{\mathrm{L0}}^{\mathrm{ph}}(\tau) &\equiv (1-d_{\mathrm{L}})\| \langle 0|_{\mathrm{U,L}}\, \hat{U}_{\mathrm{FBS}}\hat{a}_{\mathrm{U}}^{\dagger}\hat{a}_{\mathrm{L}}^{\dagger}|0\rangle_{\mathrm{U,L,E}} \|^2 \\
&= (1-d_{\mathrm{L}}) \int d\omega F_{\mathrm{U}}(\omega)\left(1 - R(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right) \\
&\quad \times \frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \int d\omega F_{\mathrm{L}}(\omega)\left(1 - T(P,\omega)T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right) \\
&\quad + (1-d_{\mathrm{L}})\frac{T_{\mathrm{in,L}}}{T_{\mathrm{in,U}}} \left| \int d\omega e^{-i\omega\tau}\sqrt{F_{\mathrm{U}}(\omega)F_{\mathrm{L}}(\omega)T(P,\omega)R(P,\omega)}T_{\mathrm{in,U}}T_{\mathrm{out,L}}(\omega)\right|^2.
\end{aligned} \tag{30}
$$

When we adjust $T_{\mathrm{in,L}}$ such that $T_{\mathrm{in,L}}/T_{\mathrm{in,U}} = 1$ is satisfied, by using the experimental parameters $\widetilde{R}(P)$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}$ and $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$, we obtain the visibility of 0.95 at 190-mW pump power from Eqs. (28) – (30). In addition, if we narrow the bandwidth of $F_{\mathrm{U}}(\omega)$ from $\Delta_{\mathrm{in,U}}$ to $\Delta_{\mathrm{in,L}}$, the visibility will be 0.98 at 190-mW pump power from Eqs. (28) – (30). We note that if we take only the improvement of the bandwidth, the visibility will be 0.93 at 190-mW pump power from Eqs. (25) – (27).
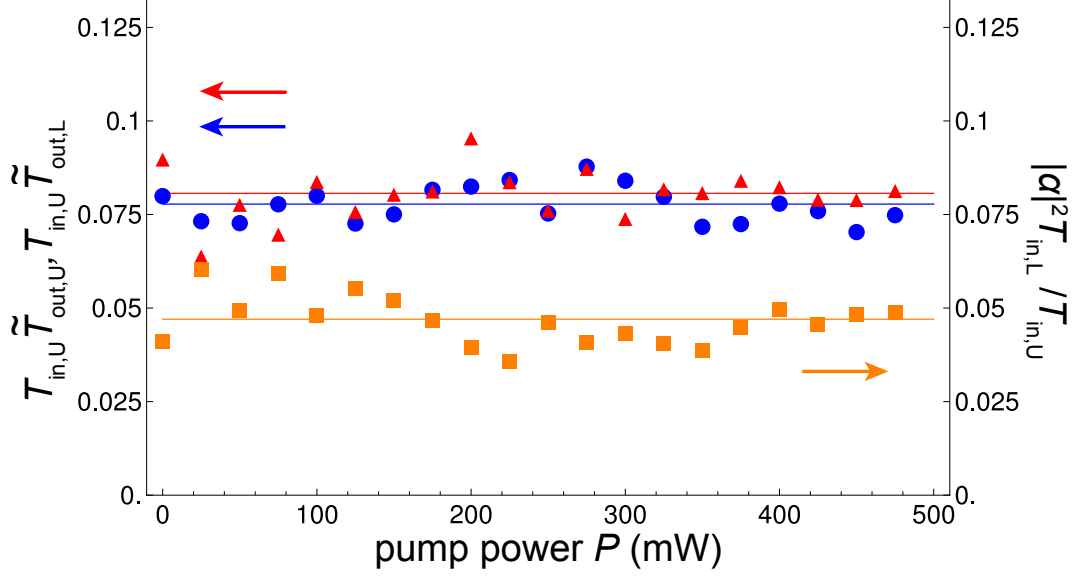
Figure S1 | The estimated values of the parameters except for the internal transition probability vs. the pump power. The circles, triangles and squares indicate the values of $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,U}}$, $T_{\mathrm{in,U}}\widetilde{T}_{\mathrm{out,L}}$ and $|\alpha|^2 T_{\mathrm{in,L}}/T_{\mathrm{in,U}}$, respectively. The horizontal lines represent the average values of these three parameters.
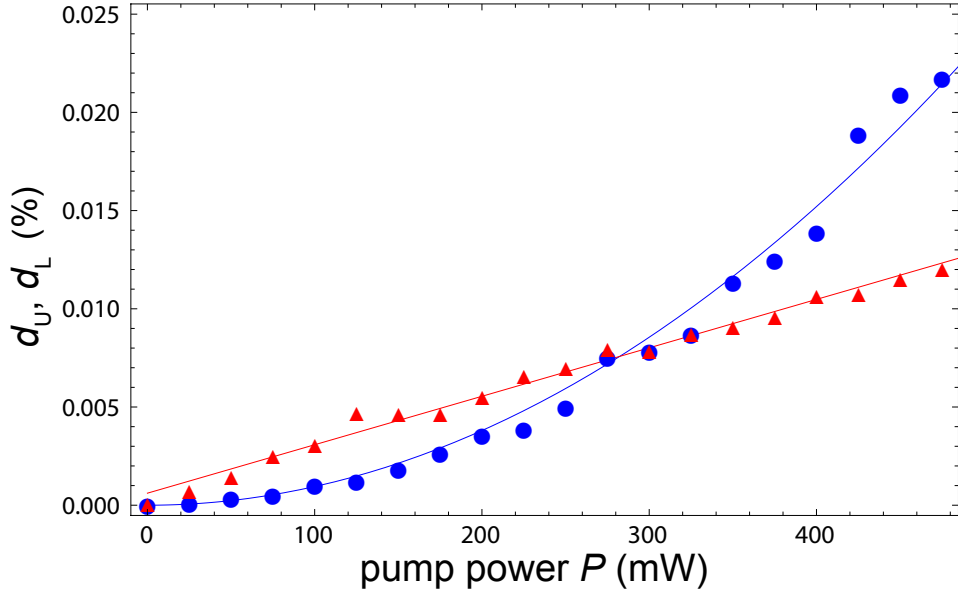

Figure S2 | The observed detection probability of the background noise. The circles/triangles indicate the observed detection probabilities of the background noise $d_{\mathrm{U/L}}(P)$. The curve for $d_{\mathrm{U}}(P)$ is obtained by the best fit to the observed values with $AP^2 + BP + C$, where $A \approx 9.5 \times 10^{-8}/\mathrm{mW}^2$, $B \approx 0.0/\mathrm{mW}$ and $C \approx 0.0$. The curve for $d_{\mathrm{L}}(P)$ is obtained by the best fit to the observed values with $DP + E$, where $D \approx 2.5 \times 10^{-5}/\mathrm{mW}$ and $E \approx 6.1 \times 10^{-4}$.