Improved heralded single-photon source using a photon-number-resolving superconducting nanowire detector: supplemental document

Here we detail the characterization of the photon-pair source, describe the experimental methods for data acquisition and analysis, and derive the theoretical model for our experiment using phase-space methods.

CONTENTS

1. PHOTON-PAIR SOURCE CHARACTERIZATION

First we study the joint spectral properties of the photon-pair source. The source produces photon pairs via spontaneous parametric down conversion (SPDC). The quantum state describing SPDC at the output of the periodically-poled lithium niobate (PPLN) waveguide is

$$|\Psi\rangle = A \int_0^\infty \int_0^\infty f(\omega_1,\omega_2)\hat{a}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2)d\omega_1 d\omega_2 \,|0\rangle \tag{S1}$$

where A is a constant prefactor that depends on the effective nonlinearity and interaction length, $\hat{a}(\omega_1)$ and $\hat{a}(\omega_2)$ are the daughter modes with frequency ω_1 and ω_2, and $f(\omega_1,\omega_2)$ is the joint spectral amplitude (JSA),

$$f(\omega_1,\omega_2) = \psi_{ph}(\omega_1,\omega_2) \times \psi_p(\omega_1,\omega_2), \tag{S2}$$

comprised of the phase-matching and pump envelope amplitudes $\psi_{ph}(\omega_1,\omega_2)$ and $\psi_p(\omega_1,\omega_2)$, respectively.

The joint spectral intensity (JSI) is,

$$|f(\omega_1,\omega_2)|^2 = |\psi_{ph}(\omega_1,\omega_2)|^2 \times |\psi_p(\omega_1,\omega_2)|^2. \tag{S3}$$

We model the phase-matching envelope intensity as,

$$|\psi_{ph}(\omega_1,\omega_2)|^2 = \mathrm{sinc}^2\left(\frac{\Delta kL}{2}\right), \tag{S4}$$

where $L = 1$ cm is the length of the waveguide and Δk is the phase-mismatch (see Fig. S1a). The phase mismatch for collinear quasi-phase-matching is [1],

$$\Delta k = 2\pi \left(\frac{n(\lambda_p)}{\lambda_p} - \frac{n(\lambda_1)}{\lambda_1} - \frac{n(\lambda_2)}{\lambda_2} - \Gamma \right), \qquad \Gamma = \frac{m}{\Lambda}, \tag{S5}$$

where $n_{p(1)(2)}$ is the pump (daughter 1) (daughter 2) index of refraction, $\lambda_{p(1)(2)} = \frac{2\pi c}{\omega_{p(1)(2)}}$ is the pump (daughter 1) (daughter 2) wavelength, m is an integer, and Λ is the poling period of the crystal. The index of refraction for light of wavelength λ in PPLN is [2],

$$n(\lambda) = \sqrt{1 + \frac{2.6734\lambda^2}{\lambda^2 - 0.01764} + \frac{1.2290\lambda^2}{\lambda^2 - 0.05914} + \frac{12.614\lambda^2}{\lambda^2 - 474.60}}. \tag{S6}$$

For $\lambda = 1540$ nm, $n(\lambda) = 2.21$, and for $\lambda = 770$ nm, $n(\lambda) = 2.26$.

We model the pump envelope intensity as,

$$|\psi_p(\omega_1, \omega_2)|^2 = \exp\left(-\frac{(\omega_p - \omega_1 - \omega_2)^2}{\sigma_p^2} \right), \tag{S7}$$

where $\omega_p = \frac{2\pi c}{770\text{ nm}}$ and $\sigma_p \sim \frac{2\pi}{100\text{ ps}} = 60$ GHz, which is subject to energy conservation condition $\omega_p = \omega_1 + \omega_2$ (see Fig. S1b).

To characterize the photon pair source, we perform a joint spectral measurement using tunable filters with 0.22 nm bandwidths on the daughter 1 and 2 paths. We measure coincidence detection events of the daughter 1 and 2 photons for various wavelengths. The measured JSI includes the detector response,

$$|f_{meas}(\omega_1, \omega_2)|^2 = |\psi_{ph}(\omega_1, \omega_2)|^2 \times |\psi_p(\omega_1, \omega_2)|^2 \times |\psi_d(\omega_1, \omega_2)|^2, \tag{S8}$$

where the third factor in Eq. S8 is the detector efficiency distribution (see Fig. S1c),

$$|\psi_d(\omega_1, \omega_2)|^2 = \exp\left(-\frac{(\lambda_1 - \lambda_d)^2 + (\lambda_2 - \lambda_d)^2}{\sigma_d^2} \right), \tag{S9}$$

which we model as a Gaussian centered at the optimal detection wavelength of $\lambda_d = 1550$nm with a spread of $\sigma_d = 53$ nm found from detector reflectivity data. The measured JSI including detector response is shown in Fig. S1d. The data are represented by circular markers with colors proportional to the corresponding number of coincidence detection events (intensity). The filled contour is the best fit of Eq. S8 to the data with Γ as a floating parameter.

The relevent JSI for the experiment includes the detector response as well as the filter response of the CWDM. The two output modes of the CWDM are centered at 1550 nm and 1530 nm with 13 nm bandwidths. We label the mode centered at 1550 nm as the idler mode and the mode centered at 1530 nm the signal mode.

The JSI for the experiment is modeled as,

$$|f_{exp}(\omega_1, \omega_2)|^2 = |f_{meas}(\omega_1, \omega_2)|^2 \times |\psi_f(\omega_1, \omega_2)|^2, \tag{S10}$$

where $|\psi_f(\omega_1, \omega_2)|^2$ is the filter response (see Fig. S1e),

$$\begin{aligned} |\psi_f(\omega_1, \omega_2)|^2 = &\exp\left(-\frac{(\lambda_1 - 1550\text{ nm})^2 + (\lambda_2 - 1530\text{ nm})^2}{\sigma_f^2} \right) \\ &+ \exp\left(-\frac{(\lambda_1 - 1530\text{ nm})^2 + (\lambda_2 - 1550\text{ nm})^2}{\sigma_f^2} \right). \end{aligned} \tag{S11}$$

The relevant JSI for the experiment is shown in Fig. S1f, which is found from Eq. S10 using the best fit parameters from the measured JSI.

We perform a Schmidt decomposition of the JSI for the experiment by evaluating the singular value decomposition of Fig. S1f [3]. The obtained eigenvalues, normalized by their sum, are shown in Fig. S2, corresponding to a Schmidt number of $K = 33.7$.

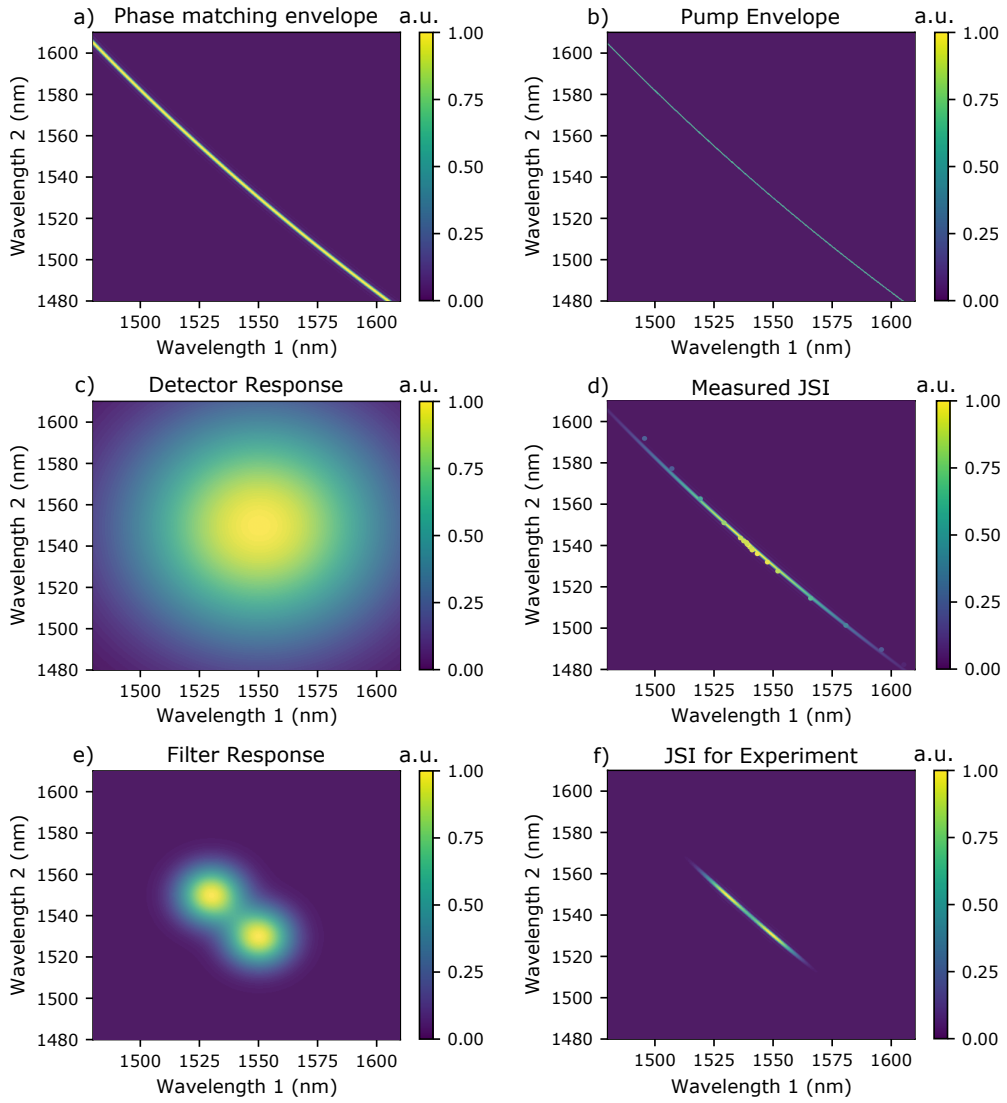

Fig. S1. a) Phase-matching envelope (Eq. S4), b) pump spectral envelope (Eq. S7), c) detector response (Eq. S9), d) measured JSI, e) filter response of CWDM (S12), and f) JSI for the experiment (Eq. S10). In d), the circles are the data and the filled contour plot is the fit of Eq. S8 to the data, with a best fit of $\Gamma = 583\ 1/\mathrm{m}$.

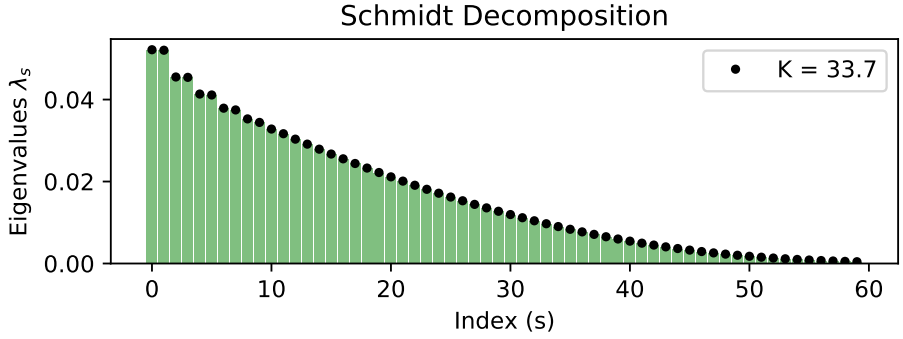

Fig. S2. Eigenvalues obtained from Schmidt decomposition of the JSI in Fig. S1f.

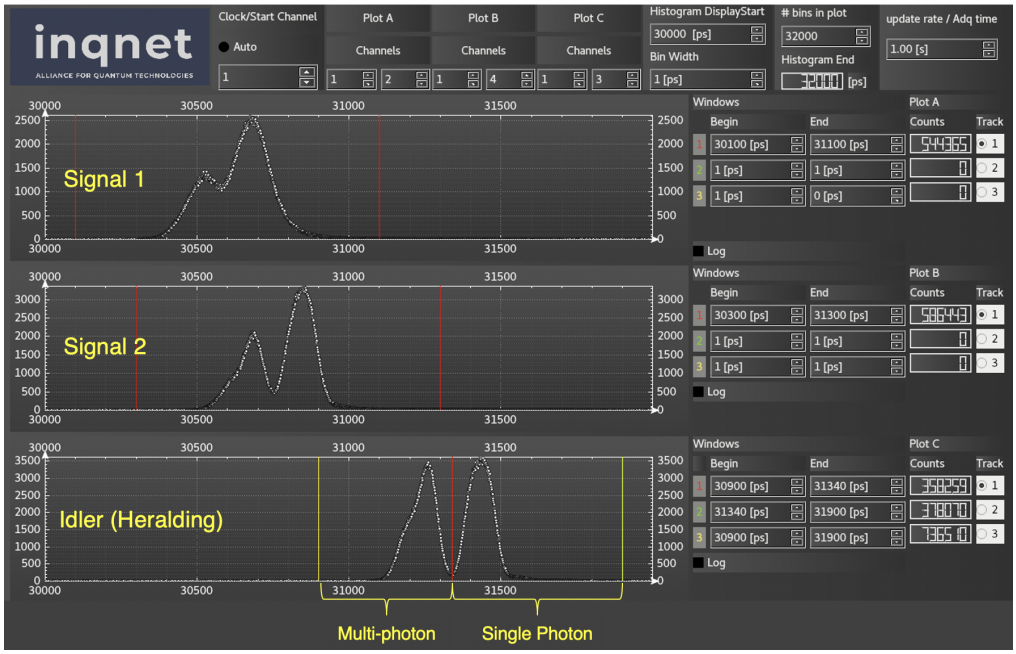

Fig. S3. Custom-made Graphical User Interface allows time-resolved detection of photons and real-time filtering of multi-photon events. The idler channel depicts a bimodal distribution of time tags relative to the clock for an acquisition time of 1 s. The left bin corresponds to the multi-photon events and the right bin corresponds to the single-photon events.

2. DATA ACQUISITION AND ANALYSIS

We developed a custom-made graphical user interface (GUI) for real-time analysis of photon arrival times and photon number discrimination, see Fig. S3. The recorded detection events in a time bin are the number of time-tags arriving in a temporal interval defined by the red and yellow markers, integrated over an acquisition time of 1 s. The markers enable user-defined time binning of photon arrival times and allows discrimination of different photon number events. A range of potential arrival times of photons in the signal paths are shown in the top two channels of the GUI, and the single and multi-photon events at the idler PNR detector are shown in the bottom channel of the GUI. Heralding of photons in the signal path conditioned on threshold (PNR) detection of photons in the idler path is performed by measuring single detection events, twofold coincidence events, and threefold coincidence events conditioned on the single- and multi-photon (single-photon) detection events at the idler detector.

3. SUPPLEMENTAL DATA

A. Single detection, twofold coincidence, and threefold coincidence events

The single detection, twofold and threefold coincidence detection probabilities are shown in Fig. S4. Detection rate is converted to probability by dividing by the experiment clock rate of 1 MHz. Experimental data is represented by circular markers and the fits from our theoretical model (described in Supplemental Sec. 3) are represented by the solid lines, using $\eta_i = 0.3280$, $\eta_{s_1} = 0.1802$, $\eta_{s_2} = 0.2210$, and $k = 2.71$.

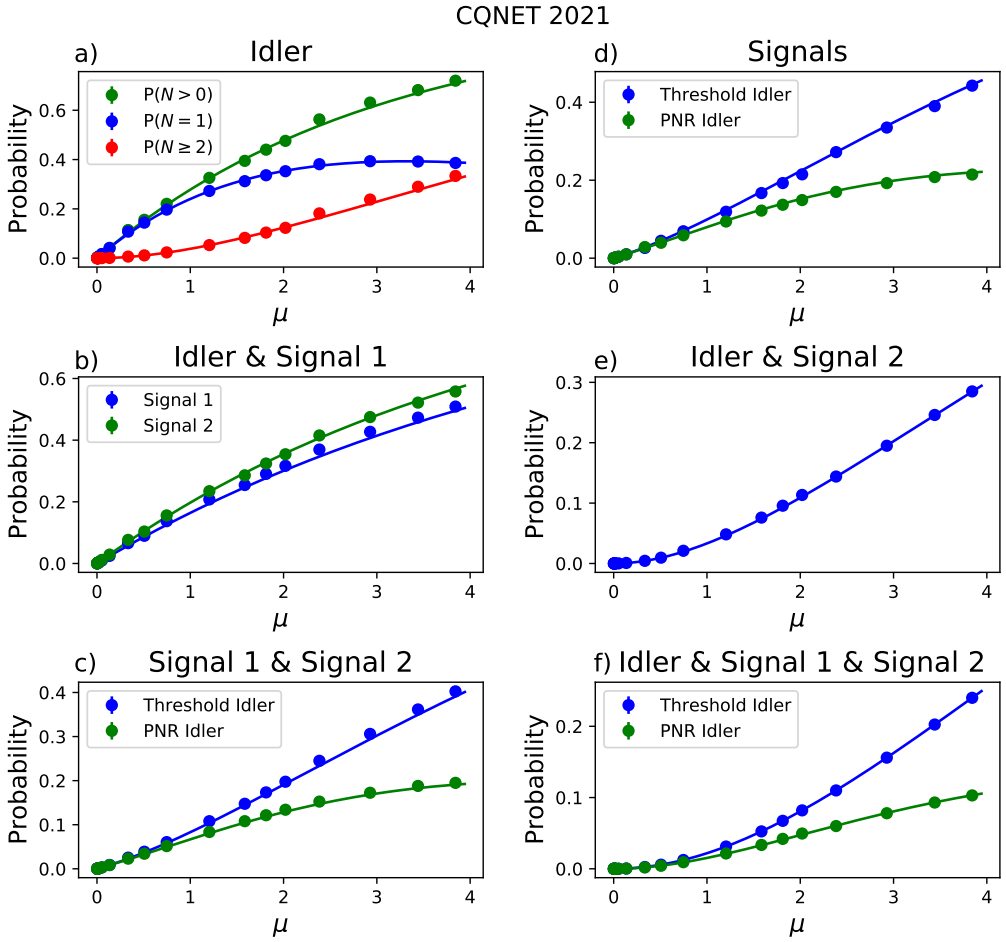

Fig. S4. a) Idler detection probabilities, b) signal 1 and signal 2 detection probabilities, c) signal 1 and idler twofold coincidence probabilities, d) signal 2 and idler twofold coincidence probabilities, e) signal 1 and signal 2 twofold coincidence probabilities, and f) threefold coincidence probabilities as a function of μ. In a), $P(N > 0)$ denotes the probability of at least one photon in the threshold configuration, $P(N = 1)$ denotes the probability of detection from the single-photon bin of the PNR configuration, and $P(N \geq 2)$ denotes the probability of detection from the multi-photon bin of the PNR configuration (see Fig. S3).

4. MODEL

The output of our photon-pair source at low mean photon number μ, neglecting loss, can be approximated as a two-mode squeezed state of the following form,

$$|\psi\rangle_{\mathrm{TMSV}} \approx \sqrt{1-\mu}\,|0_i 0_s\rangle + \sqrt{\mu}\,|1_i 1_s\rangle\,. \tag{S12}$$

In this limit, the probability of generating a single photon pair is given by μ. Correspondingly, the number of single, twofold, and threefold detection events are given by

$$C_i \approx R\eta_i\mu, \tag{S13}$$
$$C_{s_j} \approx R\eta_{s_j}\mu, \tag{S14}$$
$$C_{is_j} \approx R\eta_i\eta_{s_j}\mu, \tag{S15}$$
$$C_{is_1s_2} \approx R\eta_i\eta_{s_1}\eta_{s_2}\mu, \tag{S16}$$

where R is the clock rate of the experiment, $C_{is_1s_2}$ is the number of threefold coincidence detection events of photons in the idler and the two signal paths, C_i is the number of idler detection events,

and C_{is_j} is the number of idler and signal j twofold coincidence detection events, where $j = 1, 2$. We find estimates for the efficiencies from,

$$\eta_i \approx \frac{1}{2}\left(\frac{C_{is_1}}{C_{s_1}} + \frac{C_{is_2}}{C_{s_2}}\right), \tag{S17}$$

$$\eta_{s_1} \approx \frac{C_{is_1}}{C_i}, \tag{S18}$$

$$\eta_{s_2} \approx \frac{C_{is_2}}{C_i}, \tag{S19}$$

calculated using the data with $\mu < 6 \times 10^{-3}$. However, this not valid for large μ, where multi-photon contributions become non-negligible. To incorporate these effects, we use methods from the phase space formulation of quantum optics to derive analytical expressions for detection rates, twofold and threefold coincidence rates, and $g^2(0)$ as a function of μ. In particular, we determine the output characteristic function prior to detection, which in turn allows us to calculate the expectation values for all the measurements of interest. We include all major imperfections, such as coupling and detector inefficiencies as well as multi-photon effects. We note that the model can be extended to include dark counts, which are negligible for our experiment. In the following sections, we briefly outline the formalism used to model our experiment.

A. Phase space formalism

A characteristic function for an $N-$mode bosonic system is defined as,

$$\chi(\xi) = \mathrm{Tr}\left\{\hat{\rho}\exp(-i(\hat{x}_1, \hat{p}_1, \hat{x}_2, \hat{p}_2, \ldots \hat{x}_N, \hat{p}_N)\cdot\xi)\right\}, \tag{S20}$$

where $\hat{\rho}$ is the density matrix describing the state of the system, $\hat{x}_i$ and $\hat{p}_i$ are the canonically conjugate quadrature operators, and $\xi \in \mathbb{R}^{2N}$. The quadrature operators can be expressed in term of the bosonic creation and annihilation operators as,

$$\hat{x}_i = \frac{1}{\sqrt{2}}\left(\hat{a}_i^\dagger + \hat{a}_i\right) \qquad \hat{p}_i = \frac{i}{\sqrt{2}}\left(\hat{a}_i^\dagger - \hat{a}_i\right).$$

It is shown [4] that Eq. (S20) defines a unique mapping from the space of all possible quantum states to a space of functions over $\mathbb{R}^{2N}$, i.e. a quantum system is completely characterized by its characteristic function $\chi(\xi)$.

An important subclass of the possible quantum states is defined by the states whose characteristic function is given by a multivariate Gaussian function,

$$\chi(\xi) = \exp(-\xi^T\gamma\xi - id^T\xi), \tag{S21}$$

i.e. they are completely characterized by the mean vector d and covariance matrix γ, corresponding to the first and second moments. Representatives of this subclass include the vacuum state, coherent state, single- and two-mode squeezed states, as well as thermal states.

Relevant for our experiment is the non-degenerate output of an SPDC crystal, which can be described as a two-mode squeezed vacuum state whose covariance matrix is given by,

$$\gamma_{\mathrm{SPDC}}(\mu) = \begin{pmatrix} 1+2\mu & 0 & 2\sqrt{\mu(\mu+1)} & 0 \\ 0 & 1+2\mu & 0 & -2\sqrt{\mu(\mu+1)} \\ 2\sqrt{\mu(\mu+1)} & 0 & 1+2\mu & 0 \\ 0 & -2\sqrt{\mu(\mu+1)} & 0 & 1+2\mu \end{pmatrix},$$

where μ is the mean photon number per mode. This description is only valid for a narrow band SPDC source, where only one signal and one idler mode are present. If the source allows for multiple signal and idler modes, then the initial state must be modified to include all relevant Schmidt modes, which can be obtained through singular value decomposition of the JSI [3]. In this case, the covariance matrix of the system is given by a direct sum of the covariance matrices of the respective modes. For an SPDC source that supports N modes with Schmidt coefficients $\lambda_1, \lambda_2, \cdots \lambda_N$, the covariance matrix is,

$$\gamma = \gamma_{\mathrm{SPDC}}(\lambda_1\mu) \oplus \gamma_{\mathrm{SPDC}}(\lambda_2\mu) \oplus \ldots$$

where the sum runs over all relevant modes and we have $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_N$ and $\sum_i \lambda_i = 1$.

It is known that linear optics operations preserve the Gaussian nature of a state [4], i.e. they map Gaussian states onto Gaussian states, and hence can be described by a symplectic transformation S of the mean vector and covariance matrix,

$$d' = S^T d, \qquad \gamma' = S^T \gamma S.$$

For example, the transformation between the input modes $\hat{a}$, $\hat{b}$ and the output modes $\hat{a}'$, $\hat{b}'$ of a beamsplitter with transmittivity t is given by,

$$\hat{a}' = t\hat{a} + i\sqrt{1-t^2}\hat{b},$$
$$\hat{b}' = t\hat{b} + i\sqrt{1-t^2}\hat{a}.$$

From this we can find the symplectic transformation S that transforms the quadrature operators,

$$\begin{pmatrix} \hat{x}'_a \\ \hat{p}'_a \\ \hat{x}'_b \\ \hat{p}'_b \end{pmatrix} = \begin{pmatrix} t & 0 & 0 & -\sqrt{1-t^2} \\ 0 & t & \sqrt{1-t^2} & 0 \\ 0 & -\sqrt{1-t^2} & t & 0 \\ \sqrt{1-t^2} & 0 & 0 & t \end{pmatrix} \begin{pmatrix} \hat{x}_a \\ \hat{p}_a \\ \hat{x}_b \\ \hat{p}_b \end{pmatrix}$$
$$= S^T \begin{pmatrix} \hat{x}_a & \hat{p}_a & \hat{x}_b & \hat{p}_b \end{pmatrix}^T.$$

The propagation and coupling losses together with the detector inefficiencies can be modelled by combining the channel mode with a vacuum input on a beamsplitter with transmittivity η_{ch} and tracing out the reflected mode. Here η_{ch} corresponds to overall channel efficiency. Given that our setup consists of linear optics and that loss is modeled as an additional linear optic transformation, we are able to derive an overall symplectic transformation S_{total}, with which we calculate the characteristic function of the system up to detection,

$$\gamma_{\text{out}} = S_{\text{total}}^T \gamma_{\text{in}} S_{\text{total}}.$$

From the covariance matrix of the final Gaussian state, we can calculate several relevant experimental values such as rates and fidelities. For a measurement device with measurement operator $\hat{\Pi}$, the probability of detecting state $\hat{\rho}$ is,

$$Tr[\hat{\rho}\hat{\Pi}] = \left(\frac{1}{2\pi}\right)^N \int dx^{2N} \chi_\rho(x)\chi_\Pi(-x), \tag{S22}$$

where $\chi_\Pi(-x)$ is the characteristic function of the measurement operator and is defined in the same way as Eq.(S20) but with $\hat{\Pi}$ instead of $\hat{\rho}$.

For threshold detectors, which discriminate between zero and non-zero photons, the "on" and "off" measurement operators are,

$$\hat{\Pi}_{\text{off}} = |0\rangle\langle 0|,$$
$$\hat{\Pi}_{\text{on}} = \hat{I} - |0\rangle\langle 0|,$$

i.e. we can model the threshold detectors by projection onto the vacuum state. Since the vacuum state is a Gaussian state, the integrand in the Eq. (S22) is a multi-variant Gaussian function yielding,

$$Tr[\hat{\rho}\hat{\Pi}_{\text{off}}] = \frac{2}{\sqrt{\det(\gamma_{\text{red}} + I)}} e^{-d_{\text{red}}^T(\gamma_{\text{red}}+I_{\text{red}})^{-1}d_{\text{red}}}, \tag{S23}$$

where γ_{red} is the reduced covariance matrix and d_{red} is the reduced displacement vector obtained from γ and d by tracing out all but the measured modes.

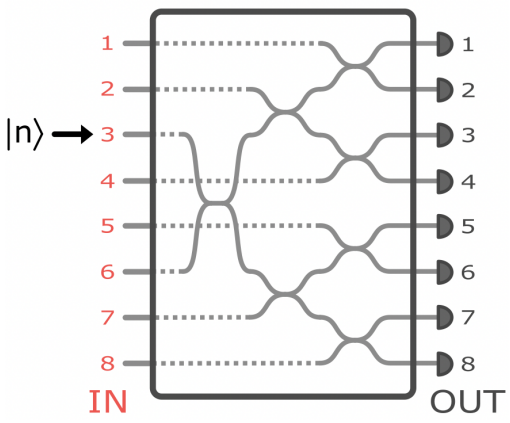

Fig. S5. $2N$-port beamsplitter for $N = 2^3$ realized as binary tree with threshold detectors at each output.

.

B. Photon-number-resolving detector

The characteristic functions for the measurement operators of the PNR detector are not Gaussian, so evaluating Eq. S22 is not straightforward. Therefore, we model the PNR detector as a 2N-port beamsplitter [5], which we realize as a binary tree of $50:50$ beamsplitters of "depth" k with $N = 2^k$ input and output ports (see Fig. S5).

To calculate the probability of coincidences across multiple modes of an N-mode state, we assume that we have a threshold detector at each of the N modes. Then the probability of detecting non-zero photons at m of the modes ("m-fold click") for an N-mode Gaussian state with covariance matrix γ_N is:

$$\sum_{\{m\}} Tr[\hat{\rho}^{\gamma_N}(\hat{\Pi}_{\text{on}}^{\otimes m} \otimes \hat{\Pi}_{\text{off}}^{\otimes(N-m)})] = \sum_{\{m\}} Tr[\hat{\rho}^{\gamma_N}((\hat{I} - |0\rangle\langle 0|)^{\otimes m} \otimes |0\rangle\langle 0|^{\otimes(N-m)})], \tag{S24}$$

where $\sum_{\{m\}}$ indicates a sum over the all possible choices of m output modes. This results in a linear combination of $\binom{N}{m}$ terms of the form,

$$Tr[\hat{\rho}^{\gamma_N}(\hat{I}^{\otimes k} \otimes |0\rangle\langle 0|^{\otimes(N-k)})] = Tr[\hat{\rho}_{\text{red}}^{\gamma_N} |0\rangle\langle 0|^{\otimes(N-k)}] \tag{S25}$$

which is a partial trace of $\hat{\rho}_N^{\gamma}$ over $N-k$ modes where $0 \leq k \leq N$. One useful property of Gaussian states is that the covariance matrix of the reduced state after a partial trace is simply the submatrix corresponding to the remaining system. It can be shown that for an N-mode system with covariance matrix γ_N and displacement vector d_N, the probability of measuring zero photons across the N modes is,

$$Tr[\hat{\rho}\hat{\Pi}_{\text{off}}^{\otimes N}] = Tr[\hat{\rho}\,|0\rangle\langle 0|^{\otimes N}] = \frac{2^N}{\sqrt{\det(\gamma_N + I_N)}} e^{-d_N^T(\gamma_N + I_N)^{-1} d_N}. \tag{S26}$$

So, (S25) simplifies to,

$$Tr[\hat{\rho}_{\text{red}}^{\gamma_N} |0\rangle\langle 0|^{\otimes(N-k)}] = \frac{2^{N-k}}{\sqrt{\det(\gamma_{N-k} + I_{N-k})}} e^{-d_{N-k}^T(\gamma_{N-k} + I_{N-k})^{-1} d_{N-k}}, \tag{S27}$$

where I_m is the m by m identity matrix, γ_{N-k} is the submatrix of γ_N and d_{N-k} is the sub vector of d_N corresponding to the remaining subsystem of $N-k$ modes. By knowing γ_N and d_N of the full N-mode system, we can find the m-fold click probability for arbitrary m, where $0 \leq m \leq N$.

For our experiment, we are interested in the single-photon detection probability P_1^N of the PNR detector. We model this as the probability that a single output mode of the $2N$-port beamsplitter is turned on:

$$P_1^N = NTr[\hat{\rho}^{\gamma'_N}(\hat{\Pi}_{\text{on}} \otimes \hat{\Pi}_{\text{off}}^{\otimes(N-1)})]$$
$$= \frac{N2^{N-1}}{\sqrt{\det(\gamma'_{N-1} + I_{N-1})}} e^{-{d'_{N-1}}^T(\gamma'_{N-1}+I_{N-1})^{-1}d'_{N-1}} - \frac{N2^N}{\sqrt{\det(\gamma'_N + I_N)}} e^{-{d'_N}^T(\gamma'_N+I_N)^{-1}d'_N}. \quad \text{(S28)}$$

The $2N$-port model can also be used to describe the photon number discrimination capability of the detector as discussed below.

B.1. Photon number discrimination

When a single-photon is sent to an input of the $2N$-port beamsplitter (with vacuum modes at each of the $N-1$ inputs), the action of the beamsplitter corresponding to unitary U_N splits the photon into an equal superposition of the N output modes. An arbitrary Fock state $|n\rangle$ transforms as,

$$U_N |n\rangle = \frac{1}{(\sqrt{N})^n} \sum_{j_1+j_2+...+j_N=n} \sqrt{\frac{n!}{j_1!j_2!...j_N!}} |j_1\rangle |j_2\rangle ... |j_N\rangle , \quad \text{(S29)}$$

so the joint probability of finding j_1 photons at output 1, j_2 photons at output 2, ..., and j_N photons at output N is,

$$P_n(j_1, j_2, ..., j_N) = \frac{1}{N^n} \frac{n!}{j_1!j_2!...j_N!}, \qquad \text{where} \quad \sum_{i=1}^{N} j_i = n. \quad \text{(S30)}$$

Then, the probability that n photons at a single input port trigger an m-fold click, i.e. a click at each of the m output ports, is

$$P_{m,n}^N = \frac{n!}{N^n} \sum_{j_1+j_2+...+j_N=n}^{(m)} \frac{1}{j_1!j_2!...j_N!} = \frac{m!}{N^n} \binom{N}{m} S(n,m), \quad \text{(S31)}$$

where $1 \leq m \leq n$, the notation (m) refers to the condition that m of $\{j_i\}$ are nonzero, and $S(n,m)$ is the Stirling number of the second kind. The Stirling number corresponds to the number of ways of partitioning a set of n elements into m non-empty sets [6].

As $N \to \infty$, the $2N$-port model approaches a PNR detector with perfect discrimination efficiency, such that the single-photon detection probability equals the single-photon probability of the input state.

$$\lim_{N\to\infty} P_1^N = \langle 1| \rho^{\gamma'_N} |1\rangle . \quad \text{(S32)}$$

For example, the probability of detecting a click at one output of a tree with depth k for an input thermal state with mean photon number μ is,

$$P_1^k = \frac{2^k\mu}{(1+\mu)\left(2^k + \left(2^k - 1\right)\mu\right)}, \quad \text{(S33)}$$

and for a coherent state with mean photon number $|\alpha|^2$ is,

$$P(1)_k = 2^k e^{-|\alpha|^2} \left(e^{|\alpha|^2/2^k} - 1\right). \quad \text{(S34)}$$

By taking the limit $k \to \infty$, we recover the single photon probabilities for a thermal state and coherent state.

$$\lim_{k\to\infty} \frac{2^k\mu}{(1+\mu)\left(2^k + \left(2^k - 1\right)\mu\right)} = \frac{\mu}{(1+\mu)^2} \quad \text{(S35)}$$

$$\lim_{k\to\infty} 2^k e^{-|\alpha|^2} \left(e^{|\alpha|^2/2^k} - 1\right) = e^{-|\alpha|^2}|\alpha|^2 \quad \text{(S36)}$$

B.2. POVM elements and counting statistics

The values of $P^N_{m,n}$ correspond to the matrix elements of the conditional probability matrix $\mathbf{C}$ defined in Ref. [7]. The rows correspond to the POVM elements of the measurement outcomes and the columns correspond to the Fock projection operators. The matrix for a threshold detector (a tree with $k = 0$) is,

$$\begin{array}{c} \\ \hat{\Pi}_{\text{off}} \\ \hat{\Pi}_{\text{on}} \end{array} \begin{array}{c} \begin{array}{ccccccccc} |0\rangle\langle 0| & |1\rangle\langle 1| & |2\rangle\langle 2| & |3\rangle\langle 3| & |4\rangle\langle 4| & |5\rangle\langle 5| & |6\rangle\langle 6| & \cdots \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots \end{pmatrix} \end{array},$$

whereas the matrix for an ideal PNR detector is the identity matrix,

$$\begin{array}{c} \\ \hat{\Pi}_0 \\ \hat{\Pi}_1 \\ \hat{\Pi}_2 \\ \hat{\Pi}_3 \\ \hat{\Pi}_4 \\ \hat{\Pi}_5 \\ \hat{\Pi}_6 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccccccc} |0\rangle\langle 0| & |1\rangle\langle 1| & |2\rangle\langle 2| & |3\rangle\langle 3| & |4\rangle\langle 4| & |5\rangle\langle 5| & |6\rangle\langle 6| & \cdots \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}.$$

For a detector with efficiency η, the probability that n photons trigger an m-fold click is,

$$P^N_{m,n}(\eta) = \sum_{j=0}^{n} P^N_{m,j} \binom{n}{j} \eta^j (1-\eta)^{n-j} = \sum_{j=0}^{n} C_{m,j} L_{j,n} = (\mathbf{C} \cdot \mathbf{L})_{m,n}\,, \tag{S37}$$

where $\mathbf{L}$ is the loss matrix with matrix elements,

$$L_{j,n} = \binom{n}{j} \eta^j (1-\eta)^{n-j}. \tag{S38}$$

The matrix corresponding to $\mathbf{C} \cdot \mathbf{L}$ for a tree with $k = 2.71$ and $\eta = 0.71$ is,

$$\begin{array}{c} \\ \hat{\Pi}_0 \\ \hat{\Pi}_1 \\ \hat{\Pi}_2 \\ \hat{\Pi}_3 \\ \hat{\Pi}_4 \\ \hat{\Pi}_5 \\ \hat{\Pi}_6 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccccccc} |0\rangle\langle 0| & |1\rangle\langle 1| & |2\rangle\langle 2| & |3\rangle\langle 3| & |4\rangle\langle 4| & |5\rangle\langle 5| & |6\rangle\langle 6| & \cdots \end{array} \\ \begin{pmatrix} 1 & 0.29 & 0.084 & 0.024 & 0.007 & 0.002 & 0.001 & \cdots \\ 0 & 0.71 & 0.489 & 0.255 & 0.119 & 0.052 & 0.022 & \cdots \\ 0 & 0 & 0.427 & 0.511 & 0.412 & 0.280 & 0.174 & \cdots \\ 0 & 0 & 0 & 0.211 & 0.381 & 0.438 & 0.408 & \cdots \\ 0 & 0 & 0 & 0 & 0.081 & 0.205 & 0.317 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0.022 & 0.075 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.004 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}.$$

The counting statistics $p(n)$ can be related to the input photon number distribution $\varrho(n)$ by

$$p_m = \sum_n \sum_{j=0}^{n} C_{m,j} L_{j,n} \varrho_n \tag{S39}$$

where $p_m = p(m)$ and $\varrho_n = \varrho(n)$ following the notation from Eq. (9) of Ref. [7]. In matrix notation,

$$\vec{p} = \mathbf{C} \cdot \mathbf{L}\vec{\varrho}. \tag{S40}$$

The transpose of the matrix $(\mathbf{C} \cdot \mathbf{L})^T$ is matrix $\mathbf{B}$ from Ref. [8], which relates probabilities and density matrices as,

$$\vec{p} = \mathbf{B}\hat{\rho}. \tag{S41}$$

B.3. Single photon discrimination efficiency

The element of positive-operator value measure corresponding to the single photon outcome for a phase independent PNR detector can be described by

$$\hat{\Pi}_1 = \sum_{n=0}^{\infty} c_n^1 \left|n\right\rangle \left\langle n\right|, \tag{S42}$$

where c_n^1 are the matrix elements corresponding to the representation of the operator in the photon number basis. The single photon outcome corresponds to the second row of the conditional probability matrix $\mathbf{C}$ defined in the previous section. The single photon outcome for a threshold detector is the "on" outcome,

$$\hat{\Pi}_{\text{on}} = \sum_{n=1}^{\infty} \left|n\right\rangle \left\langle n\right|, \tag{S43}$$

and single photon outcome for an ideal PNR detector is,

$$\hat{\Pi}_1 = \left|1\right\rangle \left\langle 1\right|. \tag{S44}$$

For the detector used in the experiment with $\eta = 0.71$ and $k = 2.71$, the single photon outcome is,

$$\begin{aligned}\hat{\Pi}_1^{exp} = {} & 0.710 \left|1\right\rangle \left\langle 1\right| + 0.489 \left|2\right\rangle \left\langle 2\right| + 0.255 \left|3\right\rangle \left\langle 3\right| \\ & + 0.119 \left|4\right\rangle \left\langle 4\right| + 0.052 \left|5\right\rangle \left\langle 5\right| + 0.022 \left|6\right\rangle \left\langle 6\right| + \cdots .\end{aligned} \tag{S45}$$

The single photon discrimination efficiency is defined in the main text as,

$$\eta_{PNR} = \frac{P(1|1)}{\sum_{n=0}^{\infty} P(1|n)} = \frac{c_1^1}{\sum_{n=0}^{\infty} c_n^1}, \tag{S46}$$

where c_n^1 is equal to $P(1|n)$, the probability of registering one photon given n incident photons. The single photon discrimination efficiency is equivalent to unity minus the trace norm between the ideal $\hat{\Pi}_1^{ideal}$ and the normalized $\hat{\Pi}_1^{exp}$,

$$\eta_{PNR} = 1 - \frac{1}{2} \sum_n \left| \frac{c_n^{exp}}{\sum_n c_n^{exp}} - \frac{c_n^{ideal}}{\sum_n c_n^{ideal}} \right| = \frac{c_1^{exp}}{\sum_n c_n^{exp}}, \tag{S47}$$

which is the weight of $\left|1\right\rangle \left\langle 1\right|$ contribution in $\hat{\Pi}_1^{exp}$. For a threshold detector, η_{PNR} approaches zero and for an ideal PNR detector, $\eta_{PNR} = 1$. For our detector, we calculate $\eta_{PNR} = 0.427$.

C. Analytical expressions

For a $2N$-port beamsplitter realized as a finite depth binary tree, we derive expressions for single detection, two-fold coincidence, and three-fold coincidence probabilities as a function of the efficiencies and tree depth k, where $N = 2^k$. The equations reduce to the threshold detection case for $k = 0$. In the following, we use $\hat{\Pi}_{\text{off},m}$ and $\hat{\Pi}_{\text{on},m}$ to denote the measurement operators for a threshold detector at mode m,

$$\hat{\Pi}_{\text{off},m} = \left|0\right\rangle \left\langle 0\right|_m, \tag{S48}$$

$$\hat{\Pi}_{\text{on},m} = \hat{I}_m - \left|0\right\rangle \left\langle 0\right|_m. \tag{S49}$$

For the PNR detector, we use $\hat{\Pi}_{\text{on},m} \otimes \hat{\Pi}_{\text{off}}^{\otimes N-1}$ to denote an "on" measurement outcome for a detector at the mth output and "off" measurement outcomes for the detectors at the remaining $N-1$ outputs of the tree.

C.1. Idler detection probability

The idler detection probability is calculated as

$$P_i(\mu,\eta_i,k) = N\mathrm{Tr}\left[\rho\left(I_{s_1}\otimes I_{s_2}\otimes\hat{\Pi}_{\mathrm{on},m}\otimes\hat{\Pi}_{\mathrm{off}}^{\otimes N-1}\right)\right], \tag{S50}$$

and is evaluated to

$$P_i(\mu,\eta_i,k) = 2^k\left(\prod_s\frac{2^k}{2^k+(2^k-1)\lambda_s\mu\eta_i}-\prod_s\frac{1}{1+\lambda_s\mu\eta_i}\right), \tag{S51}$$

where λ_s are the Schmidt coefficients obtained from the joint spectral intensity of the experiment. The products in the expressions run over all Schmidt coefficients. The Schmidt coefficients for our source are shown in Fig. S2.

C.2. Idler & signal twofold coincidence probability

The idler and signal twofold coincidence probabilities are calculated as

$$P_{is_1}(\mu,\eta_i,\eta_{s_1},k) = N\mathrm{Tr}\left[\rho\left(\hat{\Pi}_{\mathrm{on},s_1}\otimes I_{s_2}\otimes\hat{\Pi}_{\mathrm{on},m}\otimes\hat{\Pi}_{\mathrm{off}}^{\otimes N-1}\right)\right], \tag{S52}$$

$$P_{is_2}(\mu,\eta_i,\eta_{s_2},k) = N\mathrm{Tr}\left[\rho\left(I_{s_1}\otimes\hat{\Pi}_{\mathrm{on},s_2}\otimes\hat{\Pi}_{\mathrm{on},m}\otimes\hat{\Pi}_{\mathrm{off}}^{\otimes N-1}\right)\right]. \tag{S53}$$

The idler and signal j twofold coincidence probability where $j=1,2$ is evaluated to

$$\begin{aligned}P_{is_j}(\mu,\eta_i,\eta_{s_j},k) = 2^k\Bigg(&\prod_s\frac{2^k}{2^k+(2^k-1)\lambda_s\mu} \\ &-\prod_s\frac{2^{k+1}}{\lambda_s\mu\eta_{s_j}(2^k-(2^k-1)\eta_i)+2(2^k+(2^k-1)\lambda_s\mu\eta_i)} \\ &-\prod_s\frac{1}{1+\lambda_s\mu\eta_i}+\prod_s\frac{2}{2+2\lambda_s\mu\eta_i+\eta_{s_j}+\lambda_s\mu(1-\eta_i)}\Bigg).\end{aligned} \tag{S54}$$

C.3. Idler & signal 1 & signal 2 threefold coincidence probability

The idler, signal 1, and signal 2 threefold coincidence probability is calculated as

$$P_{i,s_1,s_2}(\mu,\eta_{s_1},\eta_{s_2},\eta_i,k) = N\mathrm{Tr}\left[\rho\left(\hat{\Pi}_{\mathrm{on},s_1}\otimes\hat{\Pi}_{\mathrm{on},s_1}\otimes\hat{\Pi}_{\mathrm{on},m}\otimes\hat{\Pi}_{\mathrm{off}}^{\otimes N-1}\right)\right], \tag{S55}$$

and is evaluated to

$$P_{i,s_1,s_2}(\mu,\eta_{s_1},\eta_{s_2},\eta_i,k) = 2^k\Bigg(\prod_s\frac{2^k}{(2^k+(2^k-1)\lambda_s\mu\eta_i)} \tag{S56}$$

$$-\prod_s\frac{2^{k+1}}{\lambda_s\mu\eta_{s_1}(2^k-(2^k-1)\eta_i)+2(2^k+(2^k-1)\lambda_s\mu\eta_i)} \tag{S57}$$

$$\begin{aligned}&-\prod_s\frac{2^{k+1}}{\lambda_s\mu\eta_{s_2}(2^k-(2^k-1)\eta_i)+2(2^k+(2^k-1)\lambda_s\mu\eta_i)} \\ &+\prod_s\frac{2^{k+1}}{\lambda_s\mu(\eta_{s_1}+\eta_{s_2})(2^k-(2^k-1)\eta_i)+2(2^k+(2^k-1)\lambda_s\mu\eta_i)} \\ &-\prod_s\frac{1}{1+\lambda_s\mu\eta_i}+\prod_s\frac{2}{2+2\lambda_s\mu\eta_i+\eta_{s_1}\lambda_s\mu(1-\eta_i)}\end{aligned} \tag{S58}$$

$$+\prod_s\frac{2}{2+2\lambda_s\mu\eta_i+\eta_{s_2}\lambda_s\mu(1-\eta_i)}$$

$$-\prod_s\frac{2}{2+2\mu\lambda_s\eta_i+(\eta_{s_1}+\eta_{s_2})\mu\lambda_s(1-\eta_i)}\Bigg).$$

With these expressions, we derive analytical expressions for the second order correlation function $g^{(2)}(0)$,

$$g^{(2)}(0) = \frac{P_{i,s_1,s_2}P_i}{P_{i,s_1}P_{i,s_2}}. \tag{S59}$$

REFERENCES

1. F. Laudenbach, H. Hübel, M. Hentschel, P. Walther, and A. Poppe, "Modelling parametric down-conversion yielding spectrally pure photon pairs," Opt. express **24**, 2712–2727 (2016).
2. D. E. Zelmon, D. L. Small, and D. Jundt, "Infrared corrected sellmeier coefficients for congruently grown lithium niobate and 5 mol.% magnesium oxide–doped lithium niobate," JOSA B **14**, 3319–3322 (1997).
3. K. Zielnicki, K. Garay-Palmett, D. Cruz-Delgado, H. Cruz-Ramirez, M. F. O'Boyle, B. Fang, V. O. Lorenz, A. B. U'Ren, and P. G. Kwiat, "Joint spectral characterization of photon-pair sources," J. Mod. Opt. **65**, 1141–1160 (2018).
4. C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," Rev. Mod. Phys. **84**, 621 (2012).
5. H. Paul, P. Törmä, T. Kiss, and I. Jex, "Photon chopping: new way to measure the quantum state of light," Phys. review letters **76**, 2464 (1996).
6. E. W. Weisstein, "Stirling number of the second kind," From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/StirlingNumberoftheSecondKind.html.
7. D. Achilles, C. Silberhorn, C. Sliwa, K. Banaszek, I. A. Walmsley, M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, "Photon-number-resolving detection using time-multiplexing," J. Mod. Opt. **51**, 1499–1515 (2004).
8. A. Feito, J. Lundeen, H. Coldenstrodt-Ronge, J. Eisert, M. B. Plenio, and I. A. Walmsley, "Measuring measurement: theory and practice," New J. Phys. **11**, 093038 (2009).