Hybrid Detectors

Ivo Pietro Degiovanni

I.N.RI.M., Strada delle Cacce 91, 10135 Torino, Italia

Sergey Polyakov and Alan Migdall

Joint Quantum Institute and National Institute of Standards and Technology,

100 Bureau Dr, Stop 8410, Gaithersburg, MD, 20899 USA

Hendrik B. Coldenstrodt-Ronge and Ian A. Walmsley

University of Oxford, Clarendon Laboratory,

Parks Road, Oxford, OX1 3PU, United Kingdom

Franco N. C. Wong

Massachusetts Institute of Technology,

Research Laboratory of Electronics, Cambridge, MA 02139, U.S. A.

(Dated: July 31, 2013)

PACS numbers:

I. INTRODUCTION

We present an overview of efforts to improve photon-counting detection systems through the use of hybrid detection techniques such as spatial- and time-multiplexing of conventional detectors, and frequency up-conversion. We review the basic operation for these methods and illustrate their utility in a number of applications showing new or improved capabilities compared with conventional methods.

Photon-number-resolving (PNR) detectors are a critically important tool for many fields of optical science and technology [1] such as quantum metrology [2–4] quantum imaging [5], quantum information [6–8], and foundations of quantum mechanics [9]. Unfortunately, most conventional single-photon detectors can only distinguish between zero photons detected ("no-click") and one or more photons detected ("click"). What is generally meant by detectors with inherent PNR capability are those able to produce a signal proportional to the number of absorbed photons. However, there are currently few types of detectors with this intrinsic PNR ability [1], with the most promising type being the transition edge sensor (TES) [10] (discussed in Chap. 6). However, these PNR detectors require advanced cryogenic equipment and are hardly accessible or convenient for average laboratories or for end users. While future developments may increase the accessibility of these detectors, other technologies that promise more cost-effective PNR capability are being actively pursued as well. In this review, we examine the use of spatial [11–13] or temporal [14–17] multiplexing using conventional click/no-click detectors with no intrinsic PNR capability. In addition, multiplexing can also be used to address other deficiencies of detectors (such as accessing high order correlations of light [18, 19]).

Frequency up-conversion techniques address a different photon-counting issue: the challenges associated with the steep tradeoffs among efficiency, maximum count rates, dark count rates, etc. at the longer wavelengths beyond 1 μm. Silicon single-photon avalanche diodes (SPADs) are well suited for continuous-wave (CW) operation with low dark counts, low afterpulsing, and reasonably high efficiency of greater than 50% at visible and near-infrared (near-IR) wavelengths. At the longer wavelengths, recent advances in InGaAs detector technologies (see Chap. 4) have closed the gap in these performance metrics (detection efficiencies up to 50% have been achieved [20]), but must still be operated in gated mode. Based on sum-frequency generation, up-conversion uses a strong pump to change the signal wavelength to the visible or near-IR for detection using standard Si SPADs. Interestingly, using short pump pulses in the up-conversion process allows optical sampling of the signal photons to yield sub-picosecond temporal profiles of the signal.

In Section II, we survey different types of spatially multiplexed detectors and compare their advantages and disadvantages in terms of important detector characteristics such as PNR capability, dark count rates, and afterpulsing. A spatially multiplexed detection system uses several conventional detectors appropriately connected with passive beam splitters and/or active optical switches to provide measurement enhancement, particularly PNR capability and shorter effective dead times. A simple theory for the operation of a spatially multiplexed detector is described, and the detection statistics for both pulsed and continuous-wave photon sources are formulated. We also present a scheme using a tree of actively controlled optical switches that can be superior to a passive tree configuration in achieving higher count rates, and we verify its operation in a proof-of-concept experiment.

In Section III, we discuss time-multiplexed detection (TMD), which uses beam splitters and time-delay loops to redistribute the photons in a signal pulse among multiple time slots. An advantage of TMD is that it reduces the number of detectors required for a multiplexing scheme. We present a theoretical description of TMD and discuss an experimental implementation. Finally we discuss phase-sensitive TMD and show proof-of-principle data obtained from such a system.

In Section IV, we review several experimental implementations of up-conversion-based single-photon counting including CW and pulsed modes of operation, and bulk and waveguide configurations. The basic theory of operation is described, and the special case in which the input quantum state is preserved is singled out as particularly useful for quantum information processing. Excessive background counts in the up-conversion process can be a problem in photon-counting applications and methods to mitigate this issue are discussed. Optical sampling using ultrashort pump pulses in up-conversion is shown to yield temporal information about single and entangled photons that would not have been possible with conventional detector technologies.

II. SPACE-MULTIPLEXED DETECTORS

A. Introduction

In this section we review the theory and practice of the spatially multiplexed detection approach and its application as a PNR detector. We present a simple general theory describing the operation of a spatially multiplexed detector. This theory models random and passive distribution of the photons towards each click/no-click (non-PNR) detector, for both pulsed and continuous photon sources. We discuss the various detectors developed in a spatially multiplexed form to date, such as the Si-photomultiplier, superconducting nanowire single-photon detector (SNSPD), or SPAD, and present the strategies behind them along with the technologies used to implement them. We compare the advantages and disadvantages of the different technologies in terms of important detector characteristics like PNR capability, as well as, dark count rates, afterpulsing, etc.

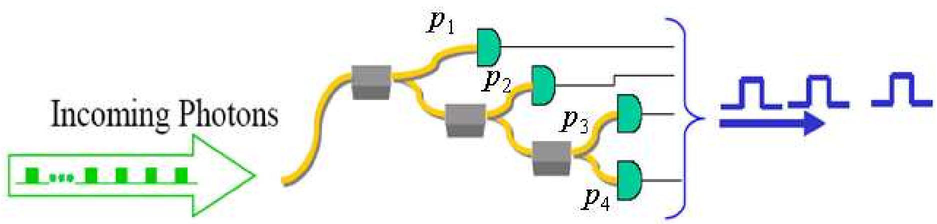

FIG. 1: A passive spatially-multiplexed detection assembly.

In its simplest form, a spatially multiplexed detection system is an assembly of detectors connected to an input optical field through a beam-splitter tree, see Fig. 1. This implementation will be referred to as a passive multiplexing scheme. A beam-splitter tree can be replaced with optical switches for better control over the multiplexing process and this approach will also be discussed using a scheme where the photons are distributed to the detectors through actively controlled optical switches. Active switching yields another interesting application of the spatially multiplexed detectors: the reduction of the effective dead time. It turns out that properly designed control of optical switches offers an advantage over a simple passive tree configuration in achieving high count rates (i.e., in reducing the effective dead time). We will summarize both theoretical considerations and proof-of principle experiments of this active-control approach.

B. Theory of Operation

In this section we present some basic statistical models describing detection by spatially multiplexed detectors. Because in practice there are distinct differences in the detection of light from pulsed versus CW sources, we discuss these two cases separately, with the theory for a pulsed source presented in Section II B 1, followed by CW sources in Section II B 2.

1. *Spatially Multiplexed PNR Detector with a Pulsed Source*

Consider a state of N photons. The most obvious way to represent the distribution of photons from that state onto a spatially multiplexed array of $\mathcal{N}$ click/no-click detectors is through a multinomial distribution [21]

$$P(n_1, n_2, ..., n_{\mathcal{N}}|N, p_1, p_2, ..., p_{\mathcal{N}}) = \delta_{N,\sum_{i=1}^{\mathcal{N}} n_i}\ \delta_{1,\sum_{i=1}^{\mathcal{N}} p_i}\ N! \prod_{i=1}^{\mathcal{N}} \frac{p_i^{n_i}}{n_i!}\ , \tag{1}$$

where each photon has a probability p_i to impinge on the i-th click/no-click detector, n_i is the number of photons impinging on the i-th click/no-click detector, and δ_i is the Kronecker delta function. The conditional probability that n_i incident photons will produce a click on the i-th detector is $p_i(\text{click}|n_i) = 1 - p_i(\text{no-click}|n_i)$ where $p_i(\text{no-click}|n_i) = (1 - \eta_i)^{n_i}$ and η_i is the detection efficiency of the i-th detector (including the optical losses of the multiplexing system).

The probability that just the two detectors j and k click, while the others do not click when N photons are incident on the spatially multiplexed detector is

$$\begin{aligned}\mathcal{P}(\text{click}_j, \text{click}_k|N) \ = & \sum_{n_1,..n_j,...,n_k,...,n_{\mathcal{N}}} P(n_1, ..n_j, ..., n_k, ..., n_{\mathcal{N}}|N, p_1, ..., p_j, ..., p_k, ..., p_{\mathcal{N}}) \times \\ & p_j(\text{click}|n_j) p_k(\text{click}|n_k) \prod_{i=1 (i \neq j,k)}^{\mathcal{N}} p_i(\text{no-click}|n_i).\end{aligned} \tag{2}$$

Thus, the probability of obtaining two clicks when N photons impinge on the detector array can be derived by summing over all the possible values of j and k:

$$Q_{\text{click}}(2|N) = \sum_{j=k+1}^{\mathcal{N}} \sum_{k=1}^{\mathcal{N}-1} \mathcal{P}(\text{click}_j, \text{click}_k|N). \tag{3}$$

An analogous argument holds for the probability $Q_{\text{click}}(k|N)$ of obtaining k clicks given N photons in the input field (with the obvious condition that $k \leq N$). Thus, the positive-operator-valued-measure (POVM) representing the detection process for a state with N photons is $\widehat{Q}_{\text{click}}(k) = \sum_{N=k}^{\infty} Q_{\text{click}}(k|N)|N\rangle\langle N|$.

The probability of observing k clicks per pulse is then

$$\mathrm{P}_{\text{click}}(k) = \mathrm{Tr}[\widehat{\rho}\ \widehat{Q}_{\text{click}}(k)] = \sum_{N=k}^{\infty} \rho_{N,N}\ Q_{\text{click}}(k|N)\ , \tag{4}$$

where $\widehat{\rho}$ is the density matrix representing the quantum state emitted by the pulsed source and $\rho_{N,N}$ is the probability that the pulse contains N photons (the diagonal element of the density matrix $\widehat{\rho}$).

The situation is greatly simplified when the photons are distributed with equal probability among the click/no-click detectors (which, incidentally, is also the most efficient solution). In this case $p_i = 1/\mathcal{N}$, and all these detectors have the same efficiency $\eta_i = \eta$. Note that we use η instead of η_{DE} for each component detector's efficiency, η_{DE} is reserved to denote the instrument's *overall* detection efficiency. In this case the probability of having k clicks in the presence of N photons simplifies to

$$\begin{aligned} Q_{\text{click}}(k|N) \ = \ & \frac{\mathcal{N}!}{k!(\mathcal{N}-k)!} \sum_{n_1,..,n_{\mathcal{N}}} \frac{N!}{\prod_{i=1}^{\mathcal{N}} n_i!} \left(\frac{1}{\mathcal{N}}\right)^N \left[\prod_{j=k+1}^{\mathcal{N}} (1-\eta)^{n_j}\right] \times \\ & \left\{\prod_{j=1}^{k} [1-(1-\eta)^{n_j}]\right\} \delta_{N,\sum_{i=1}^{\mathcal{N}} n_i}\ . \end{aligned} \tag{5}$$

An example of this probability function is shown in Fig. 2.

So far we have considered the case of perfect single-photon detectors. However, real detectors can have distorting effects, such as dead time, although if the dead time of the individual detectors is less than the inverse of the repetition rate of the source it will not be an issue. Similarly, if the maximum number of photons in a state is much smaller than the

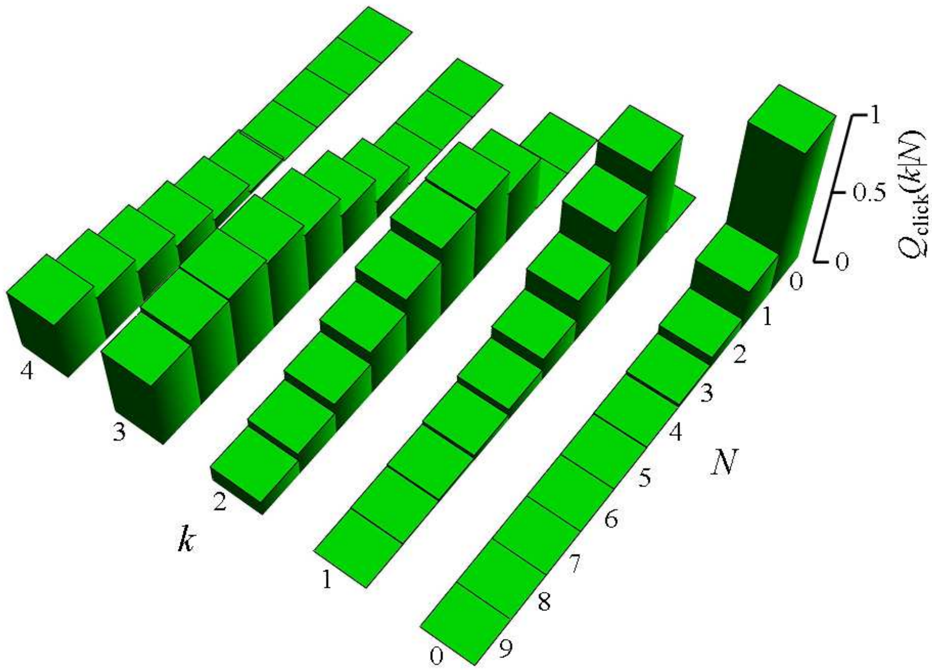

FIG. 2: Probability $Q_{\mathrm{click}}(k|N)$ of obtaining k clicks in the presence of N photons given by Eq. (5), and evaluated for a detection system of $\mathcal{N} = 4$ detectors with $\eta = 0.7$.

number of detectors, dead-time effects may be negligible. Otherwise the statistical model must include the effects of dead time. Such treatment is beyond the scope of this book, but the investigation of some specific cases can be found in Ref. [22].

2. Spatially Multiplexed PNR Detector with a CW Source

The most traditional and well established radiometric measurements involving macroscopic light levels and analog detectors are performed with continuous, stabilized light sources tied to radiometric scales. Operating any such radiometrically-tied sources at photon-counting levels remains an outstanding technological challenge. The most commonly available photon-counting detectors exhibit nonlinearity due to dead time, limited PNR ability, etc. This nonlinear behavior is also strongly influenced by the temporal statistics of the emission source, e.g., dead-time losses are different for sub-Poissonian and super-Poissonian light fields [23].

In this section we consider the simplest case: a Poissonian photon source [21]. The probability density function of the time interval Δ between the emission of two subsequent photons is $f_1(\Delta) = \mu \exp(-\mu\Delta)$, where μ is the mean photon rate of the emitted photons.

In this case, the probability of having m photons in the time interval T can be written in terms of the cumulative probability distribution function $F_m(T)$ as the difference $F_m(T) - F_{m+1}(T)$ [24, 25]. $F_m(T)$ represents the probability that m photons impinge on a detector during a time interval from $t = 0$ to $t = T$:

$$F_m(T) = \int_0^T f_m(t)\mathrm{d}t \ , \tag{6}$$

where

$$f_m(t) = \mu(\mu t)^{m-1} \exp(-\mu t)/(m-1)! \tag{7}$$

is the probability density function corresponding to an m photon Poissonian process, i.e., the convolution of m density functions $f_1(\Delta)$ describing m individual photons each arriving at some time during the interval T. Thus, one finds that the probability of having exactly m photons in the time interval T with a Poissonian photon source is:

$$F_m(T) - F_{m+1}(T) = \frac{(\mu T)^m}{m!} \exp(-\mu T) = \mathrm{Poisson}(m|\mu T), \tag{8}$$

which is the well known Poissonian distribution with mean photon number μT.

We now describe the detection of photons from such a source by a spatially multiplexed detector. The first step is to find the distribution of photons for each single click/no-click detector. Following Section II B 1, this can be calculated from the discrete convolution of the multinomial distribution of Eq. (1) and the Poissonian distribution,

$$\sum_{m=0}^{\infty} P(n_1, n_2, ..., n_{\mathcal{N}}|m, p_1, p_2, ..., p_{\mathcal{N}})\mathrm{Poisson}(m|\mu T) = \prod_{i=1}^{\mathcal{N}} \mathrm{Poisson}(n_i|p_i\mu T) \ . \tag{9}$$

Thus, we get $\mathcal{N}$ independent Poissonian processes, one for each of the detectors, and the mean photon number impinging on the i-th detector in the time interval T is $p_i\mu T$.

This result can be used for both coherent and multimode thermal light. Indeed, in the limit of a measurement time T much longer than the correlation time of thermal light, the Poisson distribution of photons in Eq. (8) asymptotically holds for thermal light when the mean number of photons per mode of the thermal source is much smaller than one. Because for a Poissonian process the counting may be considered as random (i.e., the events are equally distributed in time [21]), the photons impinge one at the time on the i-th detector.

For a CW source, the distortion of the photon statistics caused by the i-th detector's dead time $t_{\rm dead}$ can be neglected only when the mean time between two subsequent incoming photons $(p_i\mu)^{-1} \gg t_{\rm dead}$ (where $p_i\mu$ is the photon rate impinging on the i-th detector). When this condition is satisfied other detector characteristics, such as dark counts or dead time, may become important and should be included in the model. An example of such a more inclusive model can be found in Ref. [26]. For simplicity here, we only include dead time in our model.

For negligible dead time (e.g., $p_i\mu t_{\rm dead} \ll 10^{-2}$) and a detection efficiency of the i-th detector, η_i, the probability distribution of counted events, $\mathcal{P}(m_i)$, can be written as a discrete convolution between the binomial distribution representing the counting process and the Poissonian distribution of the incoming photon

$$\mathcal{P}(m_i) = \sum_{n_i=m_i}^{\infty} \begin{pmatrix} n_i \\ m_i \end{pmatrix} \eta_i^{m_i}(1-\eta_i)^{n_i-m_i}\text{Poisson}(n_i|p_i\mu T) = \text{Poisson}(m_i|\eta_i p_i\mu T)\ . \qquad (10)$$

The effect of a less-than-unity detection efficiency when measuring a Poissonian state is merely an attenuation of the input, with no other changes in detection statistics, see Fig. 3.

For a CW source neglecting the dead times is clearly too restrictive and all click/no-click detectors always exhibit some dead time. Since the time interval between subsequent arrivals of photons can be arbitrarily small in a Poissonian process with any mean number, dead time necessarily distorts the distribution of photon counts.

Due to differences in detector's electronics, it is usual to distinguish between two main types of dead time, i.e. non-paralyzable and paralyzable [27]. In the first, a dead time $t_{\rm dead}$ follows the detection of a photon, and all photons arriving within this fixed time interval are simply ignored. In the second, an incoming photon produces a dead time of $t_{\rm dead}$ if the detector was ready, or extends the dead time by the same interval $t_{\rm dead}$ if the detector was already dead. In both cases, all photons that arrive during the dead time are not counted.

For the case of paralyzable dead time, the distortion of the count statistics have been fully discussed in the literature [28–32]. The probability of counting a photon corresponds to the probability that nothing triggered the detector during the time interval $t_{\rm dead}$, prior to

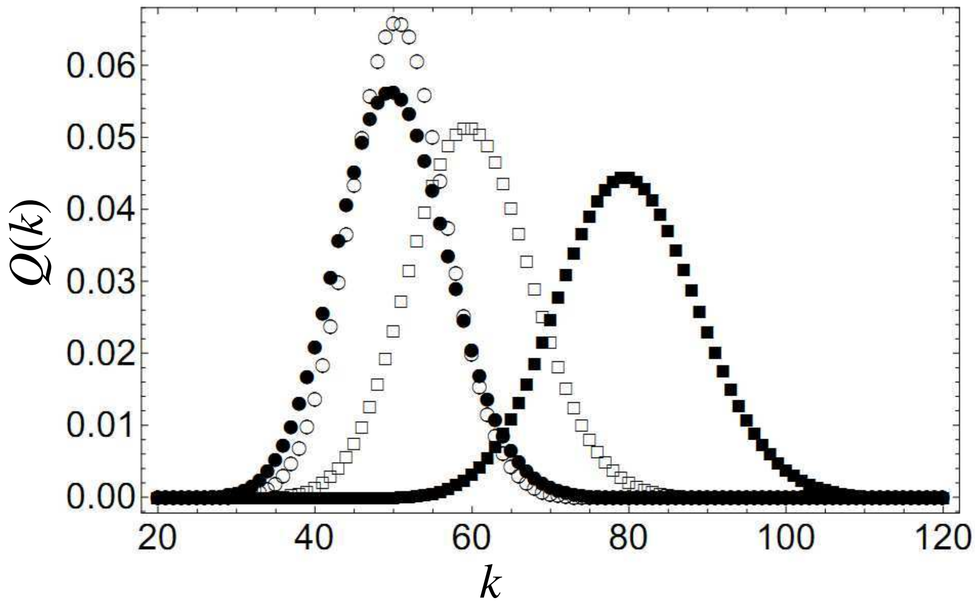

FIG. 3: Distortion of incoming photon statistics induced by the detection efficiency and dead time. The incoming state with Poissonian statistics Poisson$(k|\mu T)$ with $\mu T = 80$ (filled squares) gets effectively attenuated due to a finite quantum efficiency $\eta = 0.7$ (open squares), see Eq. (10). The output distribution remains Poissonian with Poisson$(m|\eta\mu T)$. For the detectors with an paralyzable dead time t_{dead}, the output statistics is governed by Eq. (11), which gives a distribution Poisson$(k|\eta\mu T \exp(-\eta\mu t_{\mathrm{dead}}))$, i.e., it also remains Poissonian with a new mean value of $\eta\mu T \exp(-\eta\mu t_{\mathrm{dead}})$ (filled circles). For the detectors with a non-paralyzable dead time t_{dead}, the output distribution is no longer Poissonian, but rather sub-Poissonian, see Eq. (13) (open circles).

the arrival of the photon in question. Based on Eq. (10), this probability can be expressed as $\pi = \exp(-\eta_i p_i \mu t_{\mathrm{dead}})$. Thus, similar to Eq. (10), the probability of detecting k_i photons by the i-th detector is simply

$$\mathcal{Q}(k_i) = \sum_{m_i=k_i}^{\infty} \begin{pmatrix} m_i \\ k_i \end{pmatrix} \pi^{k_i}(1-\pi)^{m_i-k_i}\,\mathcal{P}(m_i) = \mathrm{Poisson}(k_i|\pi\eta_i p_i \mu T). \tag{11}$$

We use this result to estimate the true photon rate of a source for a given count rate measured by a multiplexed detector. As seen in Fig. 3, an paralyzable dead time also does not distort the Poissonian character of the input field.

For detectors with non-paralyzable dead time [33, 34], a similar expression for the probability can be obtained by using a typical argument of renewal theory [24, 25]. The detection circuit of such a detector is blocked for a fixed time t_{dead} after which it is able to detect a subsequent photon. This means that each time interval between adjacent photon arrivals is split into two parts. First, a fixed time interval t_{dead} after each successful detection during which the detector is inhibited (dead). Second, the remaining time (i.e., the interval between the end of the dead time t_{dead} and the succeeding detection). This gives rise to a Markovian stochastic process, in contrast to a Poissonian process that has no memory effects [21, 24]. We treat this time interval between the two detections as the sum of the two time intervals introduced above.

Considering a measurement duration T, the probability of registering m counts given a fixed dead time t_{dead} is equal to $\Theta(T-mt_{\mathrm{dead}})$, where Θ is a Heaviside function. The related probability density function is then $h_m(t) = \delta(t-mt_{\mathrm{dead}})$ where δ is the Dirac delta function [35]. The probability density distribution of m counted photons by the i-th detector is the convolution of the two probability density functions f_m (from Eq. (7)) and h_m,

$$\begin{aligned} g_m(t) &= \int f_m(t_1)h_m(t_2)\delta(t-t_1-t_2)\mathrm{d}t_1\mathrm{d}t_2 \\ &= \eta_i p_i \mu \left[\eta_i p_i \mu \left(t - mt_{\mathrm{dead}}\right)\right]^{m-1} \frac{\exp\left[-\eta_i p_i \mu \left(t - mt_{\mathrm{dead}}\right)\right]}{(m-1)!}\Theta(t-mt_{\mathrm{dead}}). \end{aligned} \tag{12}$$

Following the same approach that led to Eq. (8), the probability of having exactly k_i photons counted in the time interval T in the presence of a Poissonian source and a detector with a non-paralyzable dead time is

$$\mathcal{Q}(k_i) = \frac{\gamma\left[k_i; \eta_i p_i \mu \left(T - k_i t_{\mathrm{dead}}\right)\right]}{(k_i-1)!} - \frac{\gamma\left[k_i+1; \eta_i p_i \mu \left(T - (k_i+1)\, t_{\mathrm{dead}}\right)\right]}{(k_i)!}\ , \tag{13}$$

where γ is the lower incomplete Gamma function [36]. Note that in this case of non-paralyzable dead times the shape of the statistics for the detected signal differs from the Poissonian shape of the input, see Fig. 3.

It can be shown [33] that the first moment of the probability distribution in Eq. (13) (i.e., the mean number of photon counts) $\mathcal{E}(k_i) = \sum_{k_i} k_i \mathcal{Q}(k_i)$, in the asymptotic limit of $T \gg t_{\mathrm{dead}}, (\eta_i p_i \mu)^{-1}$, corresponds to the well known formula

$$\mathcal{E}(k_i) \simeq \frac{\eta_i p_i \mu T}{1+\eta_i p_i \mu t_{\mathrm{dead}}}, \tag{14}$$

and analogously it can be demonstrated that the variance is,

$$\mathcal{E}(k_i^2) - \mathcal{E}(k_i)^2 \simeq \frac{\mathcal{E}(k_i)}{(1 + \eta_i p_i \mu t_{\text{dead}})^2} \ . \tag{15}$$

Thus, the average total number of counts measured by a spatially multiplexed detection system is

$$\mathcal{M} = \sum_{i=1}^{\mathcal{N}} \mathcal{E}(k_i) \ . \tag{16}$$

When the photons are distributed with equal probability among the click/no-click detectors, $p_i = 1/\mathcal{N}$, and all these detectors have the same efficiency $\eta_i = \eta$, Eq. (16) simplifies to

$$\mathcal{M} = \frac{\eta \mu T}{1 + \eta \mu t_{\text{dead}}/\mathcal{N}} \ . \tag{17}$$

Note that Eq. (17) can also be interpreted as a reduced effective dead time (with a reduction factor $1/\mathcal{N}$). Therefore, a multiplexed system not only enables PNR capabilities from non-PNR detectors, but also can improve the dead-time properties of a detector.

It turns out that when PNR capability is not required, dead-time properties can be reduced beyond this $1/\mathcal{N}$ factor. We discuss this in the next section.

3. *Dead Time Reduction*

As we have seen, spatial multiplexing of detectors can not only enable PNR capabilities, but can also reduce the effective dead time, allowing for higher detection rates. Particularly, we saw that the effective dead time of a multiplexed system is reduced relative to a single detector by a factor of $\mathcal{N}$, the number of the detectors in the passive "detector tree" arrangement.

There exists a method to improve detection rates in a more efficient way than just randomly sending the photons towards the elements of the click/no-click detector array. This method requires a means to actively monitor the state of each detector to check if it is ready to register a photon or it is dead, and an optical switch to route subsequent incoming photons to a detector that is known to be ready [37–40], as shown in Fig. 4. We analyze this strategy analytically and numerically, and show that this scheme allows the $\mathcal{N}$ detector system to be operated at a detection rate significantly higher than $\mathcal{N}$ times the detection rate of an individual detector, while reducing the overall dead time.

The system's switching operation could be sequential, with each detector firing in order (the control system only switches to the next detector if the previous has fired), or it could be set up to switch the input to any live detector regardless of whether the previous detector had just fired. This latter implementation may allow for optimum use of an array of detectors where each detector may have a different dead time or when the switching time of the system is not negligible. In the simple model discussed here, we assume that all the detectors have the same dead time and switch transition time. The switch transition time includes any latency or other possible delays. While we include switch latency as part of the switch transition time, rather than as a separate parameter, we point out that latency may affect the choice of what firing order is used depending on what detection characteristic is most important for the particular application. For example, an application might have a repetitive pulsed photon source where the time to sense a detection is longer than the pulse period while the switch time by itself is less than the pulse period. In that case the detection system might benefit from operating in a mode where the input is immediately switched to another live detector regardless of whether the previous detector fired. This would reduce the effect of the long latency although at a cost of a decreased likelihood of having at least one live detector available.

The switching strategy for the optical switch discussed here consists of simply re-routing photons to the next detector in the sequence of $\mathcal{N}$-detectors after the previous detector fires. This is the simplest implementation and is all that is required when the optical switching time is not a large fraction of each individual click/no-click detector's dead time. Even when the switching time is non-negligible, our assessment shows an advantage from using this scheme versus passive detector-tree schemes.

The relevant figure of merit in this context is a dead-time fraction (DTF), defined as the ratio of missed- to incident-events. A good device-independent benchmark for comparing different detection systems at high photon detection rates is the rate of incoming photons that results in a DTF of 10%, $R_{\mathrm{DTF}=10\%}$. This is a practical limit for detector operation in many real-world applications.

We analytically estimate the DTF from the mean total count rate of the overall detector pool and effective dead time for each detector (which depends on their position in the switching system). We consider a Poissonian source, as described above, and a pool of

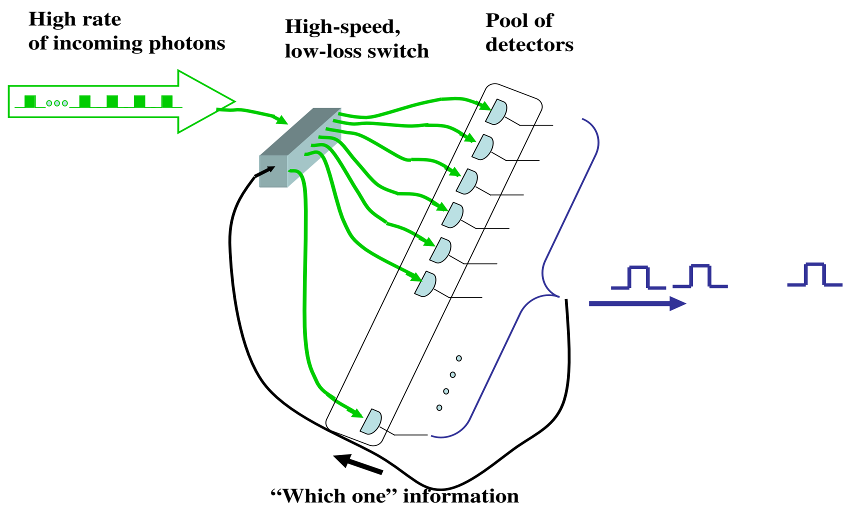

FIG. 4: A pool of detectors and a fast switch are used to register a high rate of incoming photons. Incoming photons are switched to a ready detector. If that detector fires, it is switched out of the ready pool until recovery. If it does not fire, that detector remains active.

identical detectors (with both equal detection efficiencies η and equal non-paralyzable dead times $t_{\rm dead}$). From Eq. (17) we find the DTF for a detector tree to be

$$\mathrm{DTF} = \frac{\eta\mu T - \mathcal{M}}{\eta\mu T} = 1 - \frac{1}{1+\eta\mu\frac{t_{\rm dead}}{\mathcal{N}}} \ . \tag{18}$$

Analogously, for detection system with an array of the same $\mathcal{N}$ detectors, and active multiplexing, an overall or "effective" dead time $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ can be introduced, treating the whole system as a single detection unit:

$$\mathrm{DTF} = 1 - \frac{1}{1+\eta\mu\mathcal{T}_{\mathrm{dead}(\mathcal{N})}} \ . \tag{19}$$

Therefore, the task reduces to calculating the effective dead time of the system. Because the optical switch only switches photons to a new detector after a count is registered, the effective dead time is given by the statistical contribution of the switching time, $t_{\rm s}$ and the single-detector dead time, $t_{\rm dead}$, governed by the two cases illustrated in Fig. 5: (a) $\mathcal{N}$ events are counted in a time interval longer than $t_{\rm dead} - t_{\rm s}$, or (b) they occur in a time interval shorter than $t_{\rm dead} - t_{\rm s}$. In the latter case, a photon is switched back to a detector that is

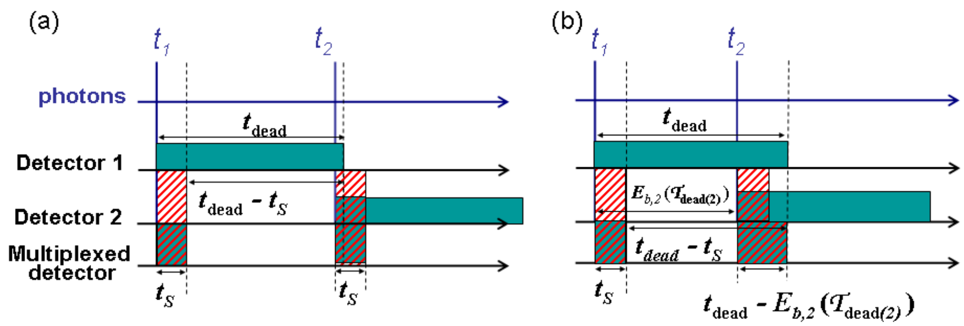

FIG. 5: Detection of two consecutive photons by a multiplexed detector with $\mathcal{N}=2$. The effective dead time $\mathcal{T}_{\mathrm{dead}(\in)}$ is given by the statistical contribution of the two possible scenarios. (a) Full saturation is avoided: after a click of detector 1, the detector 2 clicks after a time interval greater than $t_{\mathrm{dead}}-t_s$. By the time the detection assembly recovers from switching time t_s, the detector 1 will be ready to accept a new photon. (b) Full saturation of the assembly: after the click of detector 1, the detector 2 clicks after a time interval smaller than $t_{\mathrm{dead}}-t_s$. In addition to switching dead time t_s, the assembly will remain fully saturated, until the first detector recovers.

still dead, and the entire detector assembly saturates. Due to this saturation, the assembly dead time is longer than t_{s} by the remaining dead time of an individual detector. We write the effective dead time for an $\mathcal{N}$-detector assembly as:

$$\mathcal{T}_{\mathrm{dead}(\mathcal{N})}=p_{a,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})t_{\mathrm{s}}+p_{b,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})[t_{\mathrm{dead}}-E_{b,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})]\ , \tag{20}$$

where

$$p_{a,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})=\int_{t_{\mathrm{dead}}-t_{\mathrm{s}}}^{+\infty} g_{\mathcal{N}}(\Delta t|\eta\mu T,\mathcal{T}_{\mathrm{dead}(\mathcal{N})})\mathrm{d}\Delta t\ , \tag{21}$$

and

$$p_{b,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})=\int_{0}^{t_{\mathrm{dead}}-t_{\mathrm{s}}} g_{\mathcal{N}}(\Delta t|\eta\mu T,\mathcal{T}_{\mathrm{dead}(\mathcal{N})})\mathrm{d}\Delta t\ , \tag{22}$$

are the probabilities that case (a) or (b) occurs for $g_{\mathcal{N}}(\Delta t|\eta\mu T,\mathcal{T}_{\mathrm{dead}(\mathcal{N})})$, which is the probability density distribution of the time interval Δt, between a count and the $(\mathcal{N}-1)$-th preceding one. Note that the dependence of the above probabilities on $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ requires solving an integral equation to obtain $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$. The mean time interval between a count and a $(\mathcal{N}-1)$-th preceding count when case (b) occurs is given by:

$$E_{b,\mathcal{N}}(\mathcal{T}_{\mathrm{dead}(\mathcal{N})})=\frac{\int_0^{t_{\mathrm{dead}}-t_{\mathrm{s}}}\Delta t g_{\mathcal{N}}(\Delta t|\eta\mu T,\mathcal{T}_{\mathrm{dead}(\mathcal{N})})\mathrm{d}\Delta t}{\int_0^{t_{\mathrm{dead}}-t_{\mathrm{s}}} g_{\mathcal{N}}(\Delta t|\eta\mu T,\mathcal{T}_{\mathrm{dead}(\mathcal{N})})\mathrm{d}\Delta t}. \tag{23}$$

Note that Eq. (12) allows writing an expression for the probability density distribution $g_{\mathcal{N}}(\Delta t|\eta\mu T, \mathcal{T}_{\mathrm{dead}(\mathcal{N})})$, assuming a Poissonian input and introducing a constant effective dead time $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ [34]:

$$\begin{aligned} g_{\mathcal{N}}(\Delta t|\eta\mu T, \mathcal{T}_{\mathrm{dead}(\mathcal{N})}) \;=\; & \frac{(\eta\mu)^{\mathcal{N}-1}[\Delta t - (\mathcal{N}-1)\mathcal{T}_{\mathrm{dead}(\mathcal{N})}]^{\mathcal{N}-2}}{(\mathcal{N}-2)!} \\ & \cdot \exp\{-\eta\mu[\Delta t - (\mathcal{N}-1)\mathcal{T}_{\mathrm{dead}(\mathcal{N})}]\}\Theta[\Delta t - (\mathcal{N}-1)\mathcal{T}_{\mathrm{dead}(\mathcal{N})}] \;. \end{aligned} \tag{24}$$

An analytical formula for $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ exists only for $\mathcal{N} = 2$ detectors:

$$\mathcal{T}_{\mathrm{dead}(2)} = \frac{t_{\mathrm{dead}}}{2} - \frac{1 + 2\ W\left[\frac{(2\ t_{\mathrm{s}} - t_{\mathrm{dead}})\eta\mu - 1}{2}\right]}{2\ \eta\mu}\ , \tag{25}$$

where W is the principal value of the Lambert W-function [41]. For more detectors we use numerical methods to determine $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$.

Interestingly, while neglecting the dynamic nature of the dead time and introducing a constant effective dead time $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ seems to be a restrictive assumption, the results obtained with this approach are in excellent agreement with experimental results, as well as Monte-Carlo simulations in all the regimes considered [38]. This is not surprising, because such treatment merely swaps the order of integration when computing the averages.

Figure 6(a) shows the dead-time fraction versus the incoming photon rate for systems with $\mathcal{N} = 1$ to 5 detectors with single-detector dead times of 1 μs and switching times equal to 1% and 10% of the single-detector dead time. For $t_{\mathrm{s}} = 0.01\,t_{\mathrm{dead}}$ the effect of switching time on the system is negligible, while for $t_{\mathrm{s}} = 0.1\,t_{\mathrm{dead}}$, the multiplexed scheme shows much less increase of the $R_{\mathrm{DTF}=10\%}$ points with increasing detector number. Fig. 6(b) compares the analytic theory with the Monte-Carlo results, showing good agreement with the simulation for all switching times t_{s}. Fig. 6(c) compares the active multiplexed scheme just described with a passive scheme (detector/beam splitter tree configuration) for $t_{\mathrm{s}} = 0.1t_{\mathrm{dead}}$. As judged by the $R_{\mathrm{DTF}=10\%}$ points, the active multiplexed scheme surpasses the passive arrangement for relatively few detectors, $\mathcal{N} = 4, 5$.

Figure 7(a) shows the mean effective dead time $\mathcal{T}_{\mathrm{dead}(\mathcal{N})}$ for $\mathcal{N}$ up to 5, versus the mean incident photon rate $(\eta\mu)$, for $t_{\mathrm{s}} = 0.1\,t_{\mathrm{dead}}$ and $0.01\,t_{\mathrm{dead}}$. The effective dead time clearly satisfies the condition $t_{\mathrm{s}} \leq \mathcal{T}_{\mathrm{dead}(\mathcal{N})} \leq t_{\mathrm{d}}/\mathcal{N}$. We see that the maximum effective dead time of the multiplexed scheme coincides with the detector-tree dead time. This means that for

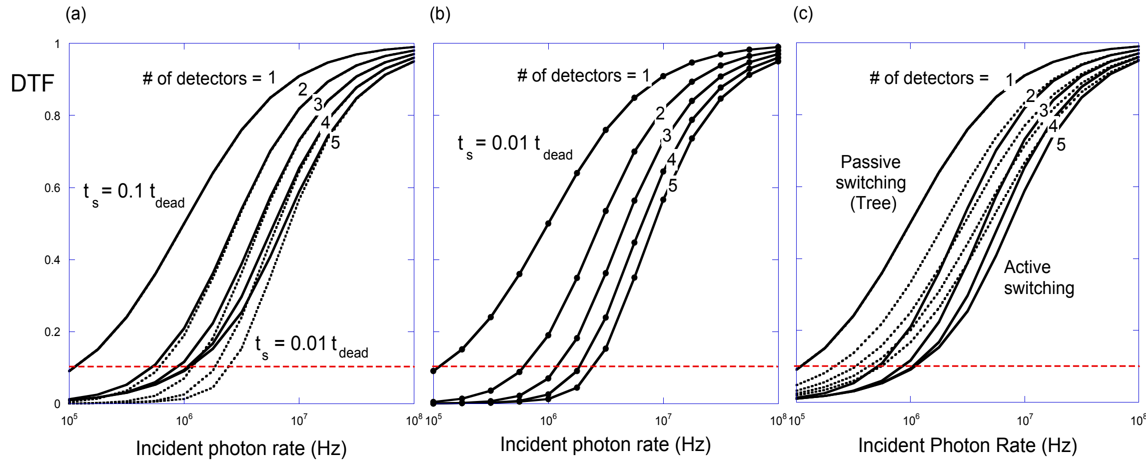

FIG. 6: DTF versus the incoming photon rate for $\mathcal{N}$=1 to 5 detectors with t_{dead}=1 μs. (a) Analytically determined DTF for actively switched system with $t_{\text{s}} = 0.01\, t_{\text{dead}}$ (dotted lines) and $t_{\text{s}} = 0.1\, t_{\text{dead}}$ (solid lines). (b) DTF determined analytically (solid lines) compared to DTF determined using Monte-Carlo simulations (points) for $t_{\text{s}} = 0.01\, t_{\text{dead}}$. (c) DTFs with $t_{\text{s}} = 0.1\, t_{\text{dead}}$ for an actively switched scheme (solid lines) compared to a passive scheme (dotted lines). Horizontal dashed line indicates DTF=10%, our benchmark level for practical detector operation.

an optical switch with $t_{\text{s}} < t_{\text{dead}}/N$, our scheme surpasses what is possible with a passive scheme. Figure 7(b) shows the ratio of the mean count rate for the multiplexed scheme to the count rate of a single detector, versus the mean incoming photon rate. We see that, as expected for high count rates, the maximum gain is $\mathcal{N}$-times the rate that would be obtained by a single detector.

Figure 8 shows $R_{\text{DTF}=10\%}$ versus the number of detectors for the active switching system at several switching times. The $t_{\text{s}} = 0.001\, t_{\text{dead}}$ result differs little from the case when the switching time is neglected. Up to $t_{\text{s}} = 0.02\, t_{\text{dead}}$ the results show significant advantage of the active-switch scheme over a passive beamsplitter tree for all numbers of detectors shown. Above $t_{\text{s}} = 0.2\, t_{\text{dead}}$ the advantage of the active system is significantly reduced until ultimately its figure of merit falls below that of the passive scheme for just a few detectors.

C. Experimental Implementations of Space-Multiplexed Detectors

1. Detector Tree Arrangements

A spatially-separated detector arrangement where detectors are connected by a series of beam splitters is the simplest example of a space-multiplexed detector. The first experiment

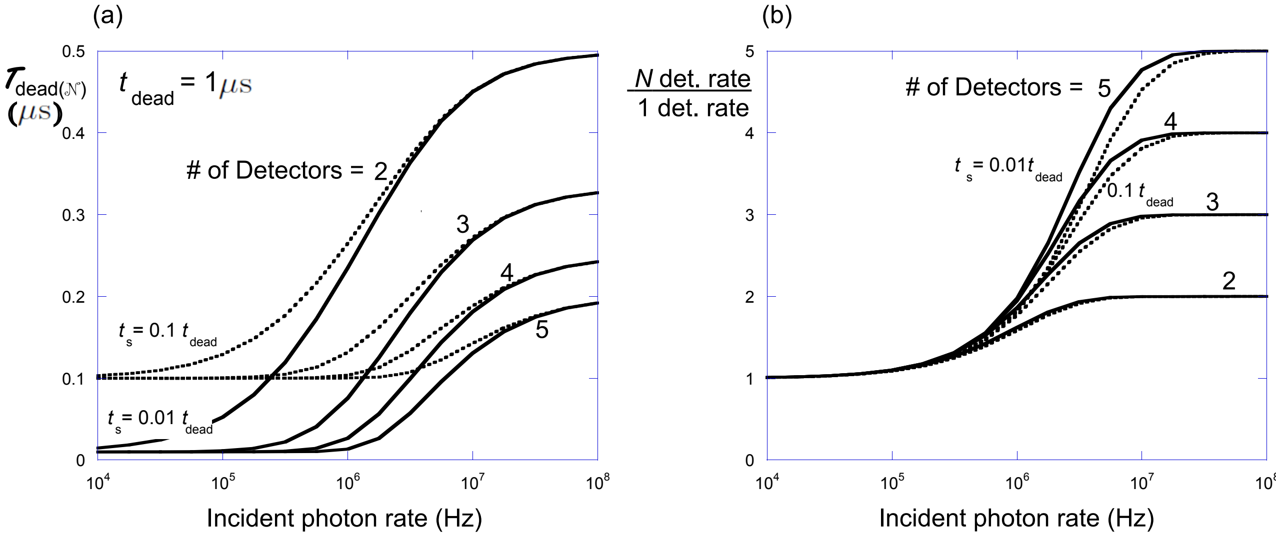

FIG. 7: (a) Plots of $\mathcal{T}_{\text{dead}(\mathcal{N})}$ versus incident photon rate for systems of 2 to 5 detectors (all with $t_{\text{dead}} = 1\ \mu\text{s}$) with $t_s = 0.1\, t_{\text{dead}}$ (dotted lines) and $t_s = 0.01\, t_{\text{dead}}$ (solid lines). Note that the system dead time saturates at $t_{\text{dead}}/\mathcal{N}$ for high photon rates, while for low photon rates it is limited by t_s. (b) Ratio of the number of photons counted by multiplexed systems to those counted by a single detector, all for $t_{\text{dead}}= 1\ \mu$s, and $t_s = 0.1\, t_{\text{dead}}$ (dotted lines), $t_s = 0.01\, t_{\text{dead}}$ (solid lines). The ratio limit for a high incoming photon rate is $\mathcal{N}$.

of this kind, where the two detectors were used to characterize multi-photon character of an input state of light, was that by Robert Hanbury Brown and Richard Q. Twiss in 1956 [42]. They use that system to show the difference in the photon number statistics of single-mode thermal light from that of multi-mode thermal light. Similar setups have been used in numerous experiments since, most notably to characterize second-order correlation properties of nonclassical states of light (with two-detector arrangements) and their multi-order correlation properties (with three or more detectors connected via a beam splitter tree). The underlying theory and experimental techniques of this measurement are described in Chapter 2.

Because a detector tree arrangement reduces the dead time of a detection system, multiplexing can also be used for counting photons at rates higher than that of an individual detector [39]. One example where this technique may be useful is in high-speed quantum communications, although in applications where security is the paramount concern, the subtleties of the detection system, such as detection efficiency variation with count rate, are often critical.

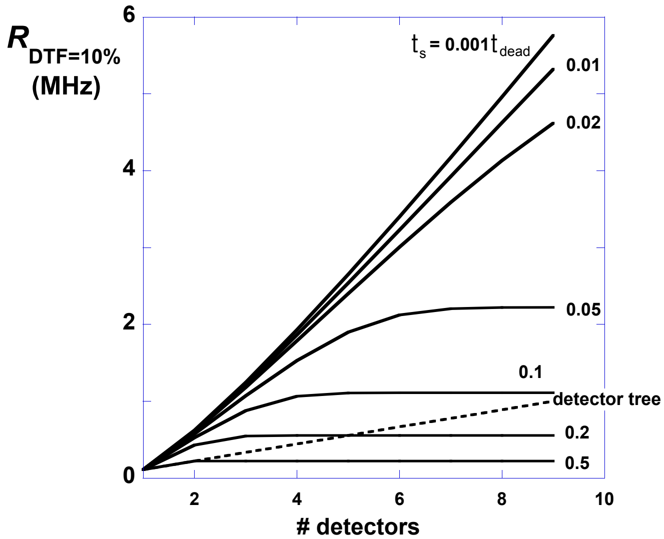

FIG. 8: Incoming rate giving a DTF=10% ($R_{\mathrm{DTF=10\%}}$) versus the number of detectors, for t_s = 0.001, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5 of t_dead (solid lines). The detector tree configuration result is shown for comparison (dashed line).

2. Photon-Number Resolving Detector Arrays

As integration techniques develop, integration of many detectors into a single device becomes possible. New technology is allowing the integration of hundreds of click/no-click detectors, making it possible to achieve significant photon-number resolving capacity, and making the effects of the dead time of each individual detector negligibly small.

One of the practical implementations of spatially-multiplexed detectors uses a microlens array to image light from a fiber on a two-dimensional SPAD array [43–45]. This is a simple and effective approach to achieving photon-number resolution with non-photon resolving detectors by having the optical mode geometrically split among the detectors in an array. Each SPAD's output can be either read individually [43] or summed to give a single output pulse with amplitude proportional to the number of detected photons [44, 45]. The drawbacks of these approaches usually include low overall detection efficiency, which means only a lower limit estimate of the photon number in an input pulse is possible, and uneven splitting of the input mode between the pixels so that saturation levels of different regions of a 2D array can vary significantly. In addition, there may be crosstalk between adjacent pixels, where the firing of one pixel causes one or more of its neighbor to fire, see Chapter 4.

SNSPDs can also be used in array formats. For example, the meanders of SNSPDs can be interleaved in such a way that each covers an equal part of the input mode. However, the number of multiplexed SNSPD devices used to detect the output of one mode has been low, with implementations so far reporting less than 10 separate individual meanders. A significant advantage of this arrangement is that it offers the potential of even faster operation speed than the already-fast single SNSPD, because the inductance of the individual wires is much lower than the longer single-wire meander of the original SNSPD, whose recovery time is inductance limited [46]. On the other hand, the need for cryogenic equipment to operate these detectors is a significant drawback to their use.

As with SPADs, there are two possible implementations for multiplexing. The first of these arrangements, the parallel-SNSPD [13], uses nanowires connected electrically in parallel. The currents through the parallel wires are summed so that the single analog output signal is proportional to the number of wires that have gone normal due to absorption of incident photons. This scheme was demonstrated with niobium nitride (NbN) nanowires 100 nm wide with a capability of counting up to four photons, with a dark-count rate of 0.15 Hz and a repetition rate of 80 MHz. A drawback of this parallel nanowire arrangement is that the analog output signals corresponding to different photon number states are not always well-separated, leading to an uncertainty in interpreting the output. The demonstrated parallel-SNSPD performs well relative to other photon-number-resolving detectors in regards to dark-count rate and maximum count rate.

The second scheme also runs parallel wires, but does so as completely separate detectors with individual outputs, thus the result is a digital output, i.e., the number of output pulses gives the number of photons detected. This scheme was demonstrated in a system of four separate wires [47]. In a recent publication [48], a system detection efficiency as high as 76% is reported. Therefore, this detector arrangement offers a combination of high detection efficiency, high temporal resolution, photon-number resolution, as well as reduced dead time through multiplexing, although the additional thermal loading of many output lines may be a challenge for such cryogenic detectors.

3. *Dead Time Fraction Reduction via Active Detector Multiplexing*

As shown from theory, active multiplexing of the detectors can boost detection rates while maintaining low saturation of the assembly by a factor exceeding the number of detectors

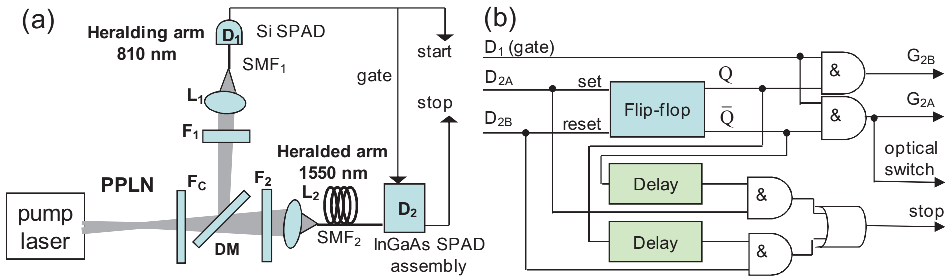

FIG. 9: (a) Experimental setup for measuring DTF of multiplexed detectors. (b) Electronics schematic for detection dead-time reduction.

in the assembly. This was verified in a proof of principle experiment with an assembly of just two detectors and a fast optical switch [39, 40].

The experimental setup, presented in Fig. 9(a), is built around a parametric down-conversion crystal that produces photon pairs at two different frequencies. The photon at 810 nm is detected by a silicon SPAD D_1 (with a dead time of 50 ns, that is negligible compared to the dead time of the infrared detectors in the system under test). The detection of an 810 nm photon heralds a 1550 nm photon in the signal arm, where different detector arrangements were tested. Two InGaAs detectors ($D_{2\mathrm{A}}$, $D_{2\mathrm{B}}$) connected through a fast optical switch were used to implement the multiplexing arrangements. The real-time logic for the active detector arrangements was built using a Field Programmable Gate Array (FPGA) (for some background on this, see for example [49]). For dead-time reduction as discussed earlier in this chapter, the design for two detectors was built around an asynchronous Set-Reset Flip-Flop (RS-Trigger), see Fig. 9(b). Note that even though FPGAs usually operate synchronously, asynchronous codes of low complexity can be implemented, although extensive testing of the asynchronous gate solutions is required. In particular, devices built using asynchronous gates may exhibit dead time. Fortunately, when SPADs outputs are used, single-photon detector dead time is usually longer than that of the FPGA logic. Another complication is the need to carefully consider timing (latency) and duration of the single-photon detector's output, which may differ from one detector to the other (even if the detectors are from the same batch) and be ready to modify the logical circuit accordingly.

Several detector configurations were compared: (i) a single detector, (ii) a detector tree arrangement, and (iii) a multiplexed arrangement that was designed to reduce detection dead time, as discussed above (Fig. 10). The count rate for configuration (iii) is the highest for all values of DTF. In particular, for our chosen threshold of $DTF = 10\%$ we see a heralding count rate 2.3 times higher for the actively multiplexed scheme (iii) as compared to a single detector. Note that this improvement factor is achieved with just two detectors in the assembly.

Another interesting feature of active multiplexing is that other properties of detectors can be improved along with the dead-time reduction. Because the output of only one of the detectors is monitored at any point of the protocol, the overall dark count rate is just that of a single detector instead of scaling linearly with the number of the detectors, as happens with passive multiplexing arrangements [39]. Another benefit of active multiplexing is that the afterpulse probability of an active arrangement will be always lower than that of a single detector or a detector tree and will depend on count rate [39].

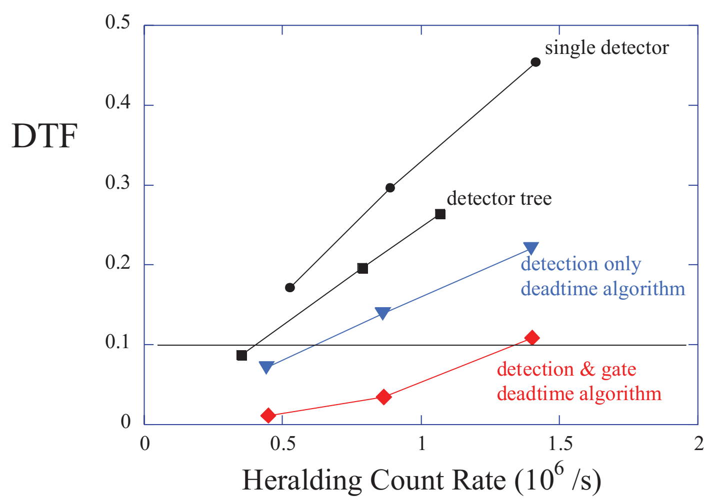

FIG. 10: Measured DTF vs. heralding (D_1) count rate for a single detector, a detector tree, and multiplexed detector arrays with different dead-time reduction algorithms: one considering only dead time arising from a detection as described in this chapter, and one that includes dead time due to detection of a photon and dead time due to gate electronics even if no detection occurs. Such gate electronics is specifically required for infrared SPADS [40]. Horizontal line is the benchmark level $DTF = 10\%$.

While analyzing the performance limitations of active multiplexing, it was found that the maximum count rate increase is limited by a feature often found in gated InGaAs detectors that is not commonly appreciated. In addition to detection dead time, these detectors have a dead time of its gate input (i.e., in the case when a detector is gated on but no photon is detected, the gate circuitry requires some time to be ready to respond to a subsequent gate), thus a strategy for suppressing the impact of this gate dead time was developed, requiring only a modification of the FPGA firmware. A measurement using this modification with a two-detector multiplexed assembly in configuration (iii) resulted in a dead-time reduction factor of nearly five when compared with a single InGaAs detector (Fig. 10). The details of this modification are beyond the scope of this chapter, but can be found in [40].

III. TIME-MULTIPLEXED DETECTORS

A. Introduction

In a time-multiplexed detector the pulse of photons impinging on the detector is split into several temporal modes as depicted on the left in Fig. 11. The resulting temporal modes can subsequently be detected with a non-photon-number-resolving, or click/no-click detector, such as a standard SPAD. A simple, however very powerful time-multiplexing detection scheme was suggested by Banaszek and Walmsley [14] in 2003 and is depicted on the right of Fig. 11. It is centered around a fiber loop used to store the light pulse under investigation. An electro-optic switch (S) can be used to couple the light pulse to be measured into the storage loop. A highly asymmetric coupler (C) is used to keep most of the pulse in the storage loop in each pass, but each time less than one photon per pulse leaves the cavity. A single-photon detector can be used to detect the output of this coupler and reconstruct the number of photons in the input pulse.

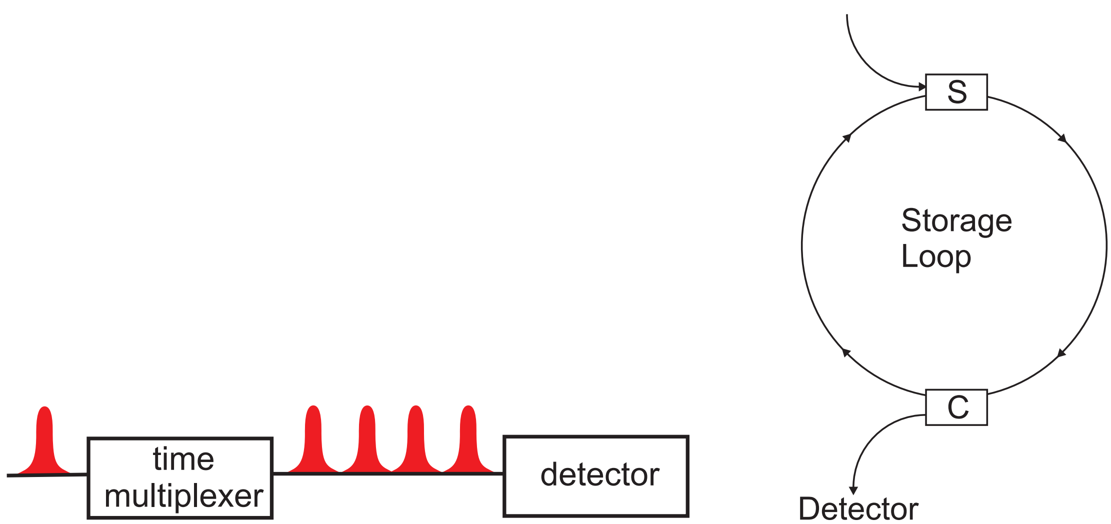

FIG. 11: Left: Basic principle of time-multiplexed detection. Right: simple storage loop TMD [14].

Depending on the fiber delay and source used, this time-multiplexed detector requires a fast switch to couple the pulse into the cavity and to ensure only a single pulse is transmitted into the cavity. The strength of the output coupler and the time required between two measurements depend on the number of photons in the input pulse. When a pulse is coupled into the loop while a residual amount from the previous pulse is still present the overlap can prevent the correct estimation of the incoming photon statistics. In the following subsection we show an approximation of this active scheme using only passive splitters, though at the expense of limited photon-number resolving capabilities.

B. Fiber-Loop Detectors

Figure 12 shows the schematic of a time-multiplexed detector based on fiber loops. An incoming pulse (I) is split at a fiber beam splitter with a 50:50 splitting ratio resulting in two spatial output modes (II). One of these outputs is sent through additional optical fiber which realizes a delay of ΔT between them (III). At a second fiber beam splitter the two temporal modes are recombined and split again, which leads to two temporal modes in each spatial mode (IV). In a second iteration more delay $2\Delta T$ is added to one of the two spatial modes shifting the two pulses yet again (V). The next recombination beam splitter finally produces four temporal modes in two spatial outputs, which are detected by two single-photon detectors, for example SPADs.

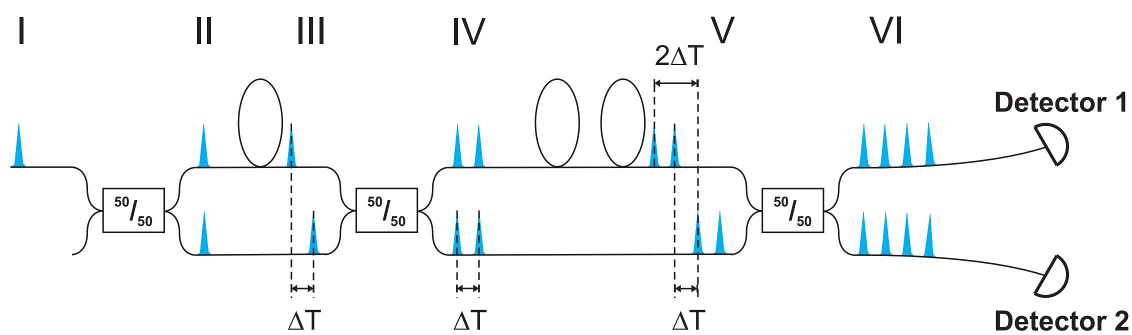

FIG. 12: Time-multiplexed detector based on two fiber loops.

When using non-photon-number-resolving detectors, the photon-number resolution of this network is constrained by the eight output modes of the beam splitter network. In principle more delay and recombination stages could be added to the detector, increasing the number of output modes and thus the level of photon-number resolution. However this reduces the maximum possible repetition rate of the detector. If the mode that acquired the highest delay in the network could, in principle, overlap or interfere with the leading, undelayed mode of the next pulse entering the detector, the photon numbers in each pulse cannot be retrieved accurately. The maximum possible repetition rate of the detector is thus given by

$$f = \frac{1}{\sum \Delta T} \,. \tag{26}$$

The time delays must be chosen according to the photon detectors used, taking into account the dead time of the detector, as well as its afterpulsing probability. Current off-the-shelf SPADs have dead times of around 50 ns; however, due to afterpulsing, a base delay of 100 ns is typical. For two delay stages this sums to a maximum delay of 300 ns, which after 200 ns is added to separate two consecutive pulse trains, allowing a maximum detector repetition rate of 2 MHz. Data acquisition can be controlled by an FPGA, reading out the number of clicks from the detector for an incoming pulse. Time gating can be implemented with the FPGA to reduce effects from dark counts and afterpulsing.

The measured number of clicks, however, does not correspond directly to the number of photons in the input pulse. In Fig. 13 the effects of losses and the beam splitter network are illustrated. Coupling into, as well as at splices within the fiber network, and beam splitter imperfections are sources of loss, in addition to the intrinsic losses of the detector inefficiency. These may lead to only k clicks from an incoming pulse with n photons. A second effect is the distribution of the photons into the spatio-temporal modes, the convolution effect. Several photons can end up in the same mode, causing only one click event. For example, in Fig. 13 (right) two out of the three incoming photons end up in the same mode, causing only two clicks. Both these effects make the use of TMDs challenging for single-shot measurements. For ensemble measurements the photon number statistics can be determined.

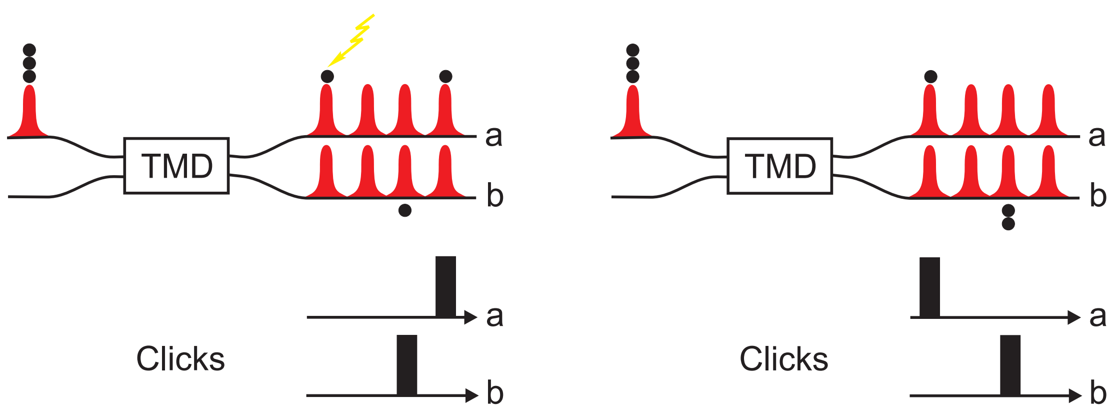

FIG. 13: Left: Influence of losses (the first photon in channel a is lost). Right: Influence of the fiber loop network.

We consider the photon number statistics $\vec{\sigma}$ of the input states, which are given by the diagonal elements of the density matrix $\hat{\rho}$ in the photon number basis. Most photon-number resolving detectors cannot resolve phase information of the incoming state, making this a valid model. In Subsection III C we will introduce the concept of weak-homodyne detection for extracting phase information from the measurement. Similar to photon-number statistics, we define the click statistics measured with the TMD $\vec{p}$, which are simply the probabilities p_n of measuring n clicks. The click statistics $\vec{p}$ for a set of input photon number statistics $\vec{\sigma}$ is given by [17]:

$$\vec{p} = \mathbf{C} \cdot \mathbf{L} \cdot \vec{\sigma}\,, \tag{27}$$

where the matrices $\mathbf{L}$ and $\mathbf{C}$ represent the effects of losses and convolution of several photons into a single bin, respectively. A single beam splitter with reflectivity ϵ_{loss} in front of the detector can be used to model all losses with a single loss parameter [50, 51]. The loss matrix contains the probabilities of k out of l photons being transmitted as

$$\mathbf{L}_{kl} = \begin{cases} \binom{l}{k} \left(1-\epsilon_{\text{loss}}\right)^{k} \epsilon_{\text{loss}}^{l-k} & \text{if: } l \geqslant k, \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

where $\binom{l}{k}$ is the binomial coefficient.

The convolution matrix is best calculated by considering all possible routes photons could travel through the beam-splitter network [17, 52]. The probabilities for these photons ending up in a certain spatio-temporal mode after the fiber network is given by a set of N probabilities $b_1, ..., b_N$, where N is the number of spatio-temporal modes of the TMD. b_j then denotes the probability of a photon ending up in the j$^{\text{th}}$ bin, which are a characteristic of a given TMD. Several simplifications, for example, for perfect (50:50) splitting ratios [15] exist, but the most general form in which to calculate the convolution matrix are the probabilities of l incident photons distributed into k spatio-temporal modes. These k clicks can be caused by all possible combinations of k out of the N modes of the TMD, which can be described by the k-tuples $c = (c_1, c_2, ..., c_k)$ with $c_i \in [1, N] \cap \mathbb{N}$ and $c_1 \neq c_2 \neq ... \neq c_k$.

The convolution matrix is then given by

$$\mathbf{C}_{kl} = \begin{cases} 0 & k > l \\ \sum_c b_{c_1} b_{c_2} ... b_{c_l} & k = l \\ \sum_c \left[\sum_d \frac{1}{\mathrm{id}(d)!} \prod_{j=1}^{k} \binom{l-\sum_{i=0}^{j-1} d_i}{d_j} b_{c_1}^{d_1} b_{c_2}^{d_2} ... b_{c_k}^{d_k} \right] & k < l \end{cases} , \tag{29}$$

where the first two cases are trivial. In the third case more than one photon can end up in a single mode. l photons can be distributed into k bins according to the k-tuples $d = (d_1, d_2, ..., d_k)$ with $d_i \in [1, n-k+1] \cap \mathbb{N}$ and $d_1 \geqslant d_2 \geqslant ... \geqslant d_k$, with $\sum_k d_k = n$ and the definition $d_0 = 0$. $\prod_{j=1}^{k} \binom{l-\sum_{i=0}^{j-1} d_i}{d_j}$ accounts for the different ways the photons which are distributed into a single bin can be chosen and $\mathrm{id}(d)!$ denotes the number of permutations with bins filled with the same number of photons to prevent overestimation by the binomial coefficient. Several illustrative examples for use of this formula are given in the literature [52].

The splitting ratios b_j required to calculate the convolution matrix of a TMD can be measured with bright light techniques [53] or photon counting [17]. The convolution matrix can also be determined by performing detector tomography, which is introduced in Chap. 9. Once measured, the convolution matrix for a given fiber network remains fixed, the loss matrix, however, may still change, for example, due to variations in coupling efficiency. Calibration techniques, introduced in Chap. 8, are therefore crucial for characterizing a TMD detector.

Many applications and experiments require the use of two detectors, such as the measurement of coincidences in down-conversion experiments (c.f. Chap. 12). In these experiments we can define the joint photon number statistics [54] $\boldsymbol{\sigma}$ of the two modes entering the two detectors. The entries of this matrix $(\boldsymbol{\sigma})_{m,n}$ are the probabilities of simultaneously having m photons in mode 1 and n photons in mode 2. Similar to the click statistics, we define the joint click statistics $(\mathbf{p})_{m,n}$ of getting m clicks in mode 1 simultaneously with n clicks in mode 2. Equation (27) then reads

$$\mathbf{p} = \mathbf{C}_1 \cdot \mathbf{L}(\eta_1) \cdot \boldsymbol{\sigma} \cdot \mathbf{L}^{\mathrm{T}}(\eta_2) \cdot \mathbf{C}_2^{\mathrm{T}}, \tag{30}$$

where matrix transposition is denoted by the superscript $^{\mathrm{T}}$ and subscripts indicate the properties of detectors 1 and detector 2, respectively. Figure 14 show examples of reconstructed joint photon number statistics.

The direct inversion of Eqs. (27) and (30) is possible; however, for low efficiencies in particular, it may lead to unphysical results such as negative probabilities. The use of optimization algorithms, such as least squares or maximum likelihood in the reconstruction incorporates the constraint of obtaining physical states and enables state reconstruction even for overall efficiencies as low as 5% [54].

Background processes such as fluorescence can mix with the process under investigation and obscure the input photon number distribution. Assuming a background photon number distribution $\boldsymbol{\sigma}_{\mathrm{Bg}}$, the photon number distribution entering the detector $\boldsymbol{\sigma}_{\mathrm{Det}}$ is given by the convolution [54]

$$\boldsymbol{\sigma}_{\mathrm{Det}} = \boldsymbol{\sigma} * \boldsymbol{\sigma}_{\mathrm{Bg}} \,, \tag{31}$$

where $\boldsymbol{\sigma}$ denotes the joint photon number statistics of the processes under investigation. However, the joint click statistics $\mathbf{p}_{\mathrm{Det}}$, which are the entity measured by the detector, cannot be written as a direct convolution of background click statistics $\mathbf{p}_{\mathrm{Bg}}$ and clicks caused by the process under investigation $\mathbf{p}$. Photons originating in the actual process and the background are mixed in the fiber network, and then detected by SPADs that saturate at one photon. This effective interaction is described by the convolution matrix in Eqs. (27) and (30).

With the convolution matrices $\mathbf{C}_{1(2)}$ for the detector in mode 1 (mode 2) we define the reduced joint click statistics $\mathbf{r}_{\mathrm{Det}}$[54], which can be obtained by applying the inverse convolution matrix to Eq. (30)

$$\mathbf{r}_{\mathrm{Det}} = \mathbf{C}_1^{-1} \cdot \mathbf{p}_{\mathrm{Det}} \cdot \left(\mathbf{C}_2^{T}\right)^{-1} = \mathbf{L}\left(\eta_1\right) \cdot \boldsymbol{\sigma}_{\mathrm{Det}} \cdot \mathbf{L}^{\mathrm{T}}\left(\eta_2\right). \tag{32}$$

These reduced statistics can be assumed independent processes without interaction and we can rewrite Eq. (30) as

$$\mathbf{r}_{\mathrm{Det}} = \mathbf{r} * \mathbf{r}_{\mathrm{Bg}} = \left[\mathbf{C}_1^{-1} \cdot \mathbf{p} \cdot \left(\mathbf{C}_2^{T}\right)^{-1}\right] * \mathbf{r}_{\mathrm{Bg}} \,. \tag{33}$$

Using the convolution theorem, the Fourier transform $\mathcal{F}$, and element by element matrix division, we find the desired reduced joint click statistics of the process under investigation

[54]

$$\mathbf{p} = \mathbf{C}_1 \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\left\{\mathbf{r}_{\mathrm{Det}}\right\}}{\mathcal{F}\left\{\mathbf{r}_{\mathrm{Bg}}\right\}} \right\} \mathbf{C}_2^{\mathrm{T}} , \tag{34}$$

which can then be used to reconstruct its joint photon number statistics $\boldsymbol{\sigma}$. The concept of reduced statistics remains valid with only one detector mode, described by Eq. (27). Figure 14 shows the effects of background on the measurement using the joint statistics of the two modes produced by spontaneous parametric down-conversion. This process, introduced in much greater detail in Chap. 12, produces photons in pairs – its joint statistics should therefore be diagonal. In the left of Fig. 14 the data shows significant off-diagonal components, indicating background contributions. After measuring the background, Eq. (34) was used to determine the reduced joint statistics and plot the reconstructed state on the right of Fig. 14, showing the success of background subtraction.

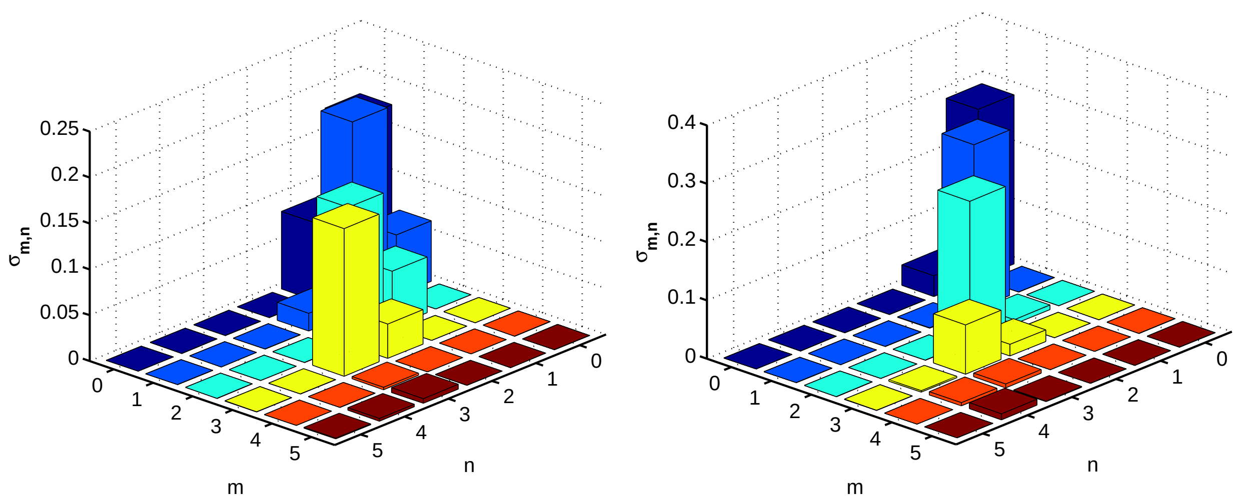

FIG. 14: Joint statistics without (Left) and with (Right) background subtraction [54].

C. Weak Homodyne Detection

Detectors with PNR capability usually work in the photon number basis without any phase reference and thus only provide access to the diagonal elements of the density matrix in the number basis. However in many quantum experiments, for example entanglement witnessing [55, 56], more information about the state is required. A standard way to do this is full state reconstruction by means of strong-field homodyne tomography [57, 58],

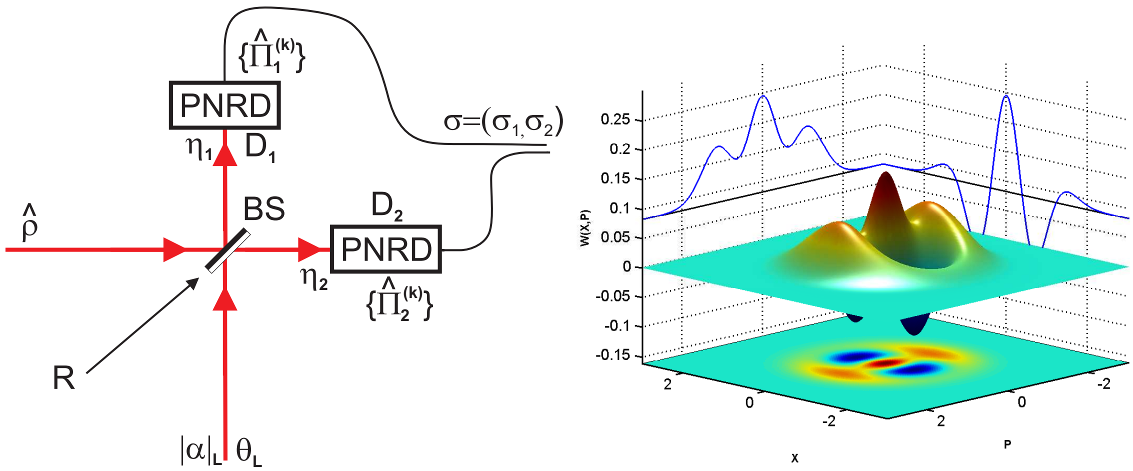

FIG. 15: Left: Schematic of a phase-sensitive PNRD. Right: Example of the exotic POVM elements [59] achievable with a weak homodyne detector, illustrated by its Wigner function.

which mixes a local oscillator (LO) with the signal under investigation and subsequently takes measurements for different phase settings between the LO and the state. On the other hand, detectors with PNR capabilities cannot be used in strong-field homodyne detection as they do not resolve the number of photons usually present in the strong LO and might even be damaged by the high LO field strengths.

Based on the same concept, the weak-field homodyne detector mixes the probe state with a LO beam. However the LO strength is comparable to the signal to be detected and thus compatible with PNR detectors (PNRDs). Weak-field homodyne detectors have been suggested and analyzed theoretically [59], and also demonstrated in recent experiments [59–62]. In Fig. 15 we show the principle of a weak-field homodyne detector: an input state $\hat{\rho}$ is mixed with a weak local oscillator at a beam splitter of reflectivity R. Both outputs of the beam splitter are subsequently detected by PNRDs **D1** and **D2** and the click distribution $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)$ recorded. We describe the detectors using their POVM sets $\left\{\hat{\Pi}_1^{(k)}\right\}$ and $\left\{\hat{\Pi}_2^{(k)}\right\}$. Further parameters required for a full description of the detector are the amplitude $|\alpha_{\mathrm{LO}}|$ and phase Θ_{LO} of the LO.

Analysis and rigorous mathematical treatment of the detector reveals the POVM elements for obtaining joint clicks $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)$

$$\hat{\Pi}^{(\boldsymbol{\sigma})}_{\gamma,\mathrm{theo}} = \sum_{k,l=0}^{M} \left[\mathbf{C_1} \cdot \mathbf{L_1}\right]_{\sigma_1 k} \left[\mathbf{C_2} \cdot \mathbf{L_2}\right]_{\sigma_2 l} \hat{\Pi}^{(k,l)}_{\gamma,\mathrm{ideal}}\,, \tag{35}$$

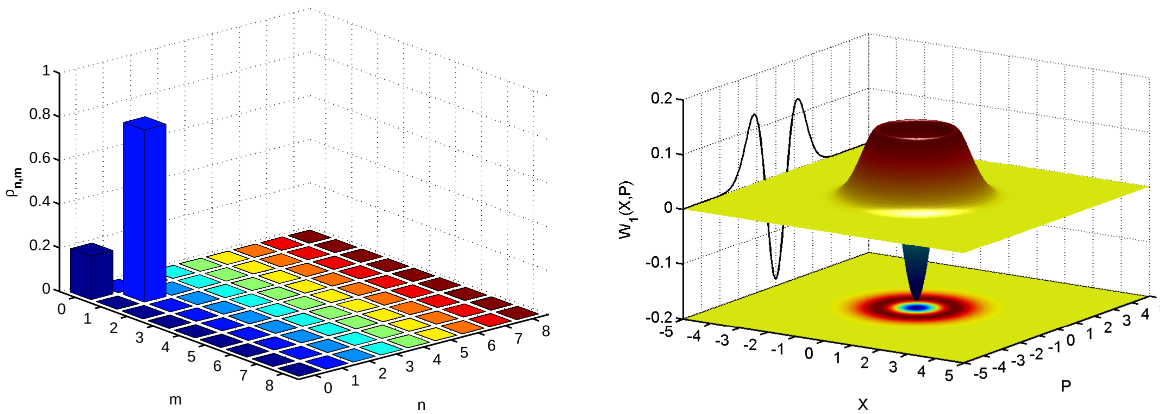

FIG. 16: Results with the phase sensitive TMD. Left: Joint photon number statistics of the heralded single-photon state Right: Wigner function of the single-photon state.

where we use the loss and convolution matrices for the TMD and $\hat{\Pi}^{(k,l)}_{\gamma,\text{ideal}}$ is the POVM element of a perfect photon number resolving detector, with the unitary transformation $\hat{U}_{BS}$ describing the beam splitter

$$\hat{\Pi}^{(k,l)}_{\gamma,\text{ideal}} = \left(\hat{\mathbb{1}}_{in} \otimes |\alpha\rangle_L \langle\alpha|_L\right) \hat{U}^{\dagger}_{BS} \left(|k\rangle_1 |l\rangle_2 \otimes \langle k|_1 \langle l|_2\right) \hat{U}_{BS} \,. \tag{36}$$

In Fig. 16 we show the reconstruction of a heralded single-photon state, produces by spontaneous parametric down-conversion with a phase-sensitive TMD.

IV. UP-CONVERSION DETECTORS

A. Introduction

Optical frequency up-conversion at the single-photon level [63–65] enables an alterative method for detecting single photons at wavelengths longer than 1 μm. Based on sum-frequency mixing in a nonlinear crystal, followed by detection in the visible or near-IR range, the technique provides a convenient way to detect single photons at wavelengths for which the performance of inexpensive or commercially available single-photon counters is lacking. This is especially true for wavelengths greater than 1 μm where Si SPADs are no longer suitable. With a strong pump at an appropriate wavelength, the long-wavelength photon can be frequency translated to the visible or near-IR for detection by Si SPADs that are easier to use and have more favorable characteristics than InGaAs SPADs. For telecom-band photons, up-conversion and subsequent detection by Si SPADs can lead to significantly improved performance with higher detection efficiency, CW operation without gating, and lower dark counts compared with InGaAs SPADs. Even as this performance gap has narrowed somewhat in recent years with steady improvements of InGaAs SPAD operation (See Chap. 4 for details), detection via up-conversion remains a viable option. Up-conversion can also be utilized to provide unique single-photon measurement capabilities. In ultrafast up-conversion [66, 67], ultrashort pump pulses are used to time-sample single-photon pulses and yield a single-photon temporal profile with sub-picosecond resolution that is orders of magnitude better than any state-of-the-art single-photon detector technologies.

Up-conversion of telecom-band photons at 1.3 μm and 1.55 μm for efficient detection is well suited to various tasks in optical and quantum communications. Communications-based applications include optical time-domain reflectometry [68], quantum key distribution [69–71], and photon-counting pulse-position modulation communication [72]. Up-conversion serves another important purpose in quantum optical information processing. Unity-efficient up-conversion preserves the quantum state of the input photon so that the output photon is identical to that of the input photon except it is at a different wavelength. Therefore, up-conversion offers wavelength freedom of choice in quantum optical information processing. For example, two frequency-nondegenerate photons can be made to produce a Hong-Ou-Mandel (HOM) interference signature [73] if one of the photons is frequency translated to match the wavelength of the other photon, or if both photons are up-converted to the same frequency [74]. A more practical application in a multi-wavelength quantum network [75] is to generate the entanglement at a convenient wavelength for high flux and high quality entanglement, followed by frequency translation to the desired wavelength without losing the entanglement [76].

B. Theory of Single-Photon Up-conversion

Consider the generation of an up-converted photon at frequency ω_2 from an input photon at frequency ω_1 through interaction with a strong monochromatic pump field at frequency ω_p in a second-order ($\chi^{(2)}$) nonlinear crystal. Energy conservation constrains these fields' frequencies to satisfy $\omega_1 + \omega_p = \omega_2$, and momentum conservation at the photon level requires that their wave vectors obey $\vec{k}_1 + \vec{k}_p = \vec{k}_2$. Assuming perfect phase matching and operation well within the phase-matching bandwidth, the input and up-converted output fields are related by a simple beam splitter relationship [77]. In particular, the output at ω_2 is given by

$$\hat{A}_2(L) = \cos(\kappa|A_p|L)\hat{A}_2(0) + i\frac{A_p}{|A_p|}\sin(\kappa|A_p|L)\hat{A}_1(0)\,, \tag{37}$$

where $\hat{A}_i(x)$ is the quantum field operator for frequency ω_i at the input ($x = 0$) or output ($x = L$) of the crystal, A_p is the classical pump field, κ is the nonlinear coupling coefficient, and L is the length of the crystal.

For vacuum input at ω_2 the up-converted output power, in photon units, is proportional to the input power

$$\langle \hat{N}_2(L)\rangle = \langle \hat{A}_2^\dagger(L)\hat{A}_2(L)\rangle = \sin^2(\kappa|A_p|L)\langle \hat{N}_1(0)\rangle\,. \tag{38}$$

The up-conversion efficiency is given by $\sin^2(\kappa|A_p|L)$ and reaches a maximum 100% when the pump field $|A_p| = A_{\pi/2} = \pi/2\kappa L$ (at a pump power $P_{\pi/2}$). We note that for $|A_p| = A_{\pi/2}$ Eq. (37) shows that

$$\hat{A}_2(L) = i\frac{A_p}{|A_p|}\hat{A}_1(0)\,. \tag{39}$$

Except for an unimportant absolute phase, the output field operator will be in the same quantum state as the input field operator, and hence represents the ideal case of converting a photon from one frequency to another frequency without altering its quantum state. Quantum-state frequency conversion also applies to the field amplitude and was first demonstrated in squeezed states of light by Huang and Kumar [78]. Such state-preserving conversion is useful in many quantum information processing applications. For example, entangled photons can be generated at a convenient wavelength with high purity and then converted to a different wavelength that is optimized for a specific application. With efficient state-preserving frequency conversion a quantum key distribution (QKD) network can operate seamlessly between free-space operation at 800 nm and fiber-based system at 1.55 μm. We note that for nonzero detuning from perfect phase matching the quantum-state frequency translation is imperfect, even when $|A_p| = A_{\pi/2}$, because of dispersion [77].

For single-photon detection, frequency translation is primarily used to shift the photon wavelength to a region where good detector performance is readily available. Frequency conversion is not restricted to sum-frequency generation as we have discussed so far. Difference-frequency generation can also be used for frequency translation, satisfying the relation $\omega_1 - \omega_p = \omega_2$. In this case, the input photon has the highest frequency ω_1 and is converted to a lower frequency ω_2. For example, a visible photon can be mixed with a strong near-IR (say 800 nm) pump for conversion to the telecom wavelength, which can then be transmitted in a low-loss single-mode optical fiber to a remote location for subsequent detection by a SNSPD or TES, or for frequency up-conversion back to the visible, thus allowing states of visible photons to be transmitted over long distances. In difference-frequency conversion, the input quantum state can also be maintained at the output under appropriate pump and phase-matching conditions [79, 80]. In general, frequency conversion works well for single-photon detection as long as the strong pump does not have the highest frequency, in which case the parametric down-conversion process would dominate and result in a large background of spontaneously emitted "noise" photons.

C. Up-conversion Techniques

Single-photon frequency up-conversion can be implemented in several ways depending on the choice of nonlinear material, pump source, and the type of application. Below we briefly discuss several successful experimental demonstrations to illustrate the advantages and disadvantages of specific choices such as continuous-wave or pulsed pumping, and nonlinear bulk crystal or nonlinear waveguide.

Figure 17 shows the setup of one of the earliest single-photon up-conversion experiment by Albota and Wong [63]. Weak laser light at 1560 nm was converted to the visible at 633 nm using a 4 cm long bulk periodically poled lithium niobate (PPLN) crystal inside an enhancement ring cavity for the 1064 nm pump. The cavity, which was single-pass for both the input and output up-converted photons, produced a circulating pump power of over 25 W at 1064 nm to reach near-unity conversion efficiency. With properly coated cavity mirrors and PPLN crystal there was very little loss for the input and output light, which is important for achieving a high overall detection efficiency that was primarily limited, in this case, by the spectral and spatial filters and by the Si SPAD detection efficiency. Intracavity up-conversion has also been demonstrated by placing the up-conversion crystal inside the pump laser cavity, which removes the need for phase locking the cavity to an external field [81].

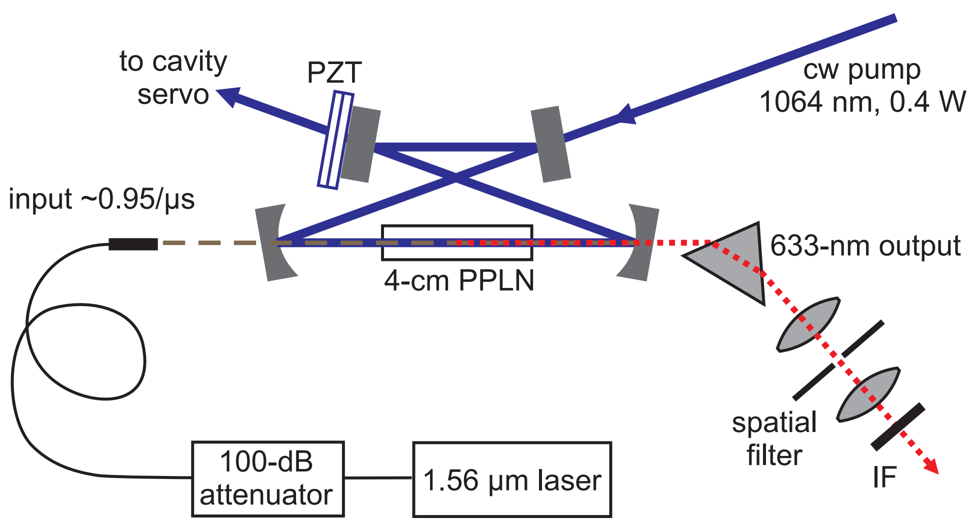

FIG. 17: Experimental setup for cw single-photon up-conversion of 1.56-μm light using a bulk PPLN crystal and a pump enhancement cavity to reduce the required input pump power [63]. Spectral filtering with a dispersing prism and a 10 nm interference filter (IF) centered at 633 nm and spatial filtering with a pinhole are used to reduce pump-induced background counts. Adapted from [63], Fig. 1.

Figure 18 shows the results of the single-photon up-conversion measurements that are in good agreement with the functional power dependence of Eq. (37). The up-conversion efficiency reaches 90%, and is limited by insufficient circulating pump power. It also plots the number of detected background photons (measured without input) and up-converted photons (measured difference with and without input) showing that the amount of noise photons can be significant and would adversely affect photon counting applications such as QKD. The background photons were thought to have originated from a two-step cascaded process: pump-induced fluorescence and non-phase-matched parametric fluorescence that are in the same spatial mode and at the same wavelength λ_1 as the input, followed by efficient up-conversion from λ_1 to λ_2. The parametric fluorescence can be suppressed if the pump wavelength is chosen to be longer than the input wavelength, $\lambda_p > \lambda_1$, such that the energy of a pump photon is lower than the energy of the input photon.

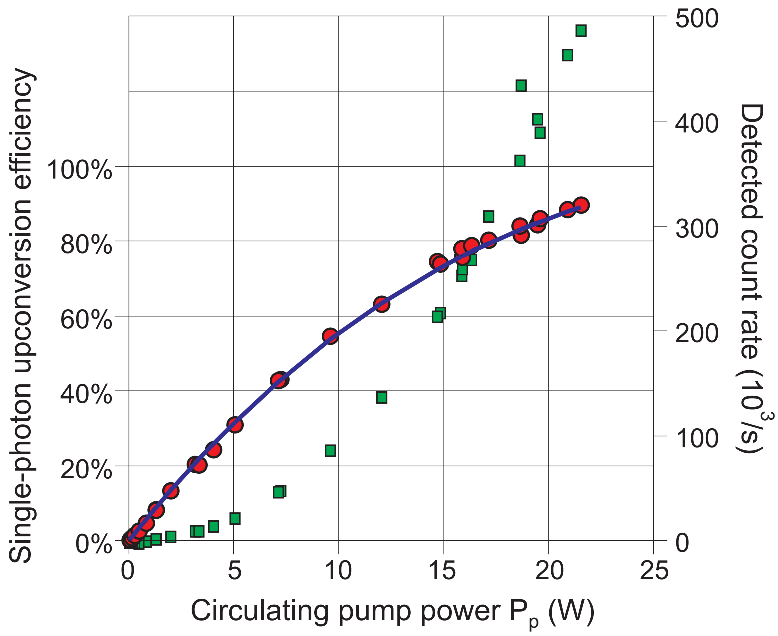

FIG. 18: Intrinsic single-photon up-conversion efficiency (filled circles) of Ref. [63] using Fig. 17 setup as a function of circulating pump power, in accord with the functional form (solid curve) of Eq. (37). Up-converted signal counts (filled circles) and background counts (filled squares) are also plotted (right axis) indicating the significant background noise at high up-conversion efficiency. At the highest pump level, the overall system detection efficiency is 34%. Adapted from [63], Fig. 3.

The idea of using a pump wavelength longer than the input wavelength was first implemented by Langrock *et al.* [82]. Moreover, in view of the high pump power required for a bulk nonlinear crystal, a significantly more efficient nonlinear waveguide was used as the up-conversion medium [65], thus reducing the required pump power. Figure 19 shows the up-conversion results with a fiber-coupled PPLN waveguide and an input signal at 1.32 μm and a pump at 1.55 μm [82]. Two improvements are evident: only $\approx$100 mW of pump was required to achieve maximum intrinsic up-conversion efficiency, and a dark count rate of ≈1.6 × 10^4/s was measured at that pump level. The use of a longer pump wavelength suppresses the dark count rate by a factor of 50 compared with the same waveguide setup but with the pump and signal wavelengths interchanged. The dark count rate can be further reduced with additional filtering to enable up-conversion-based applications such as 1310 nm QKD [71].

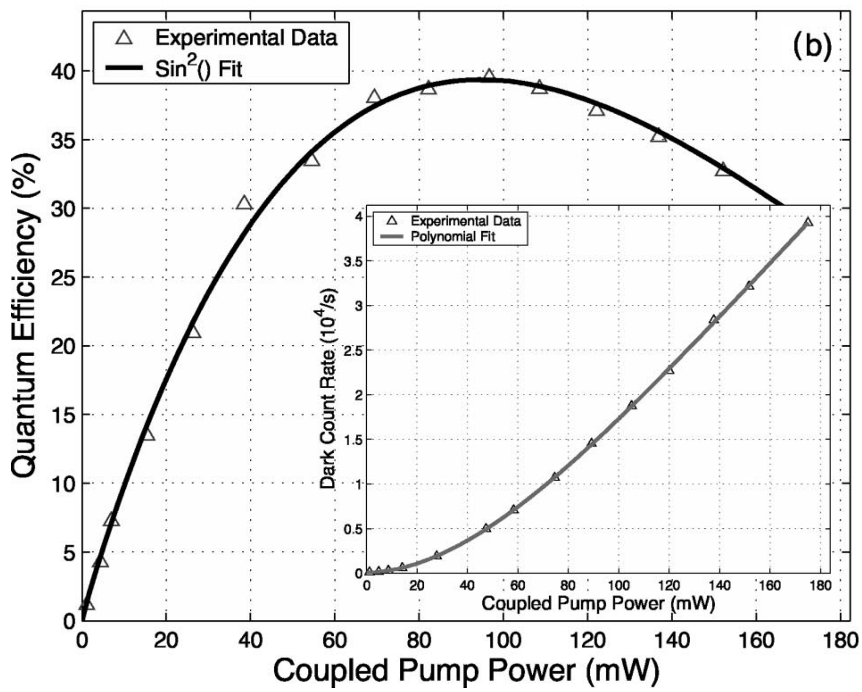

FIG. 19: System detection efficiency and dark counts measurements made with a fiber-coupled PPLN waveguide with a 1.55 μm pump and an input signal wavelength of 1.32 μm [82]. At ≈100 mW of pump, unity intrinsic up-conversion efficiency is reached with significantly lower dark counts than results shown in Fig. 18 that uses a pump wavelength shorter than the signal wavelength. Reproduced from [82], Fig. 2(b).

In designing an up-conversion detection system, ultimately it is the overall system detection efficiency that should be maximized. One can relate the overall detection efficiency η_{DE} to various processes by $\eta_{DE} = \eta_u\eta_c\eta_f\eta_p\eta_s$, where η_u is the intrinsic up-conversion efficiency, η_c is the in- and out-coupling efficiency through the nonlinear medium, η_f is the filtering efficiency for spatial or spectral filtering that is needed to block background photons, η_p is the propagation efficiency through the up-conversion system other than those we have specifically noted, and η_s is the Si SPAD detection efficiency. It is straightforward to find a judicious combination of a highly nonlinear crystal, adequate pump power, and up-conversion configuration that produces near-unity intrinsic up-conversion efficiency. Extra efforts are usually needed to reduce other losses to achieve high system efficiency. For example, high-efficiency spectral filtering can be obtained using a reflective Bragg grating instead of a more lossy interference filter. Waveguide-based up-converters are convenient and require low pump powers, but coupling and waveguide losses can be a problem especially for free-space propagating signal photons. It should be noted that for certain quantum information processing tasks in which coincidences of two or more photons are measured, some amount of background counts can be tolerated because they are unlikely to be correlated with the two or more detected outputs.

For $\chi^{(2)}$ nonlinear crystals the up-conversion process is polarization specific and therefore the polarization of the input photon must be set accordingly. It is useful to be able to up-convert a photon with unknown polarization efficiently [83] and, better yet, retain its polarization in the output state after up-conversion which is essential for coherent up-conversion of polarization qubits. Figure 20 shows a simple two-up-converter scheme that can accomplish polarization-preserving up-conversion. A polarization beam splitter separates the incoming photon into orthogonally polarized components that are up-converted by the two nonlinear crystals and then recombined at a second polarization beam splitter for output. For any polarization-preserving up-conversion scheme it is important to maintain the relative phase between the two components by stabilizing the two paths interferometrically. Polarization-preserving up-conversion has been implemented using a single nonlinear crystal with a double-pass configuration [77]. In the first pass it up-converts one polarization component and in the return pass, after the light undergoes a 90° rotation for both input and output wavelengths, the orthogonal component is up-converted and then combined with the other polarization component for output. At the single-photon level, up-conversion of a polarization-qubit while retaining the entanglement property with another photon has been demonstrated by Ramelow *et al.* [84].

D. Pulsed Up-conversion

For cw sources of single photons and bi-photons the arrival times of the photons are random and the up-conversion apparatus must be continuously on, as in the examples discussed in the previous subsection. For pulsed sources, one can take advantage of the periodic nature of the photon arrival times by using a pulsed pump that is synchronized with the

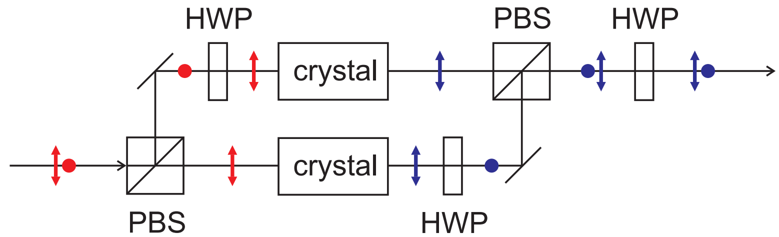

FIG. 20: Schematic of a polarization-preserving up-converter. Input with an arbitrary polarization (in red) is separated by a polarization beam splitter (PBS) into horizontal and vertical polarizations. The vertical component in the upper arm is rotated into horizontal polarization before up-conversion. The lower-arm up-converted horizontal component is rotated into vertical polarization for recombination with the upper-arm horizontal component at the second PBS. A final output half-wave plate (HWP) restores the output state (in blue) to its original input polarization state.

photon arrival times. To optimize the up-conversion efficiency the peak power of the pulsed pump should be equal to $P_{\pi/2}$, which can be easily obtained with a low average-power pump without the need for an enhancement cavity. The pulse width of the pump should also be larger than the input photon coherence time and the temporal jitters of the system.

Kwiat *et al.* first demonstrated pulsed up-conversion using a pulsed 1064 nm pump with an attenuated cw laser source at 1550 nm [64]. True synchronized pulsed up-conversion was later demonstrated using ≈600 ps pump pulses at 1064 nm and 200 ps signal pulses at 1550 nm showing near-unity intrinsic up-conversion efficiency [85]. Pulsed up-conversion using two or more spectrally and temporally distinct pump pulses can be utilized to increase the single-photon detection rate by spectrally separating the up-converted pulses and sending them to different single-photon detectors, thus overcoming the dead time of a single SPAD [86]. For pulsed up-conversion it is more appropriate to consider background count probability per pump pulse. A background count of 3×10^{-4} per 1 ns measurement time (comparable to pump pulse duration) was measured [64], which is comparable to that obtained in cw measurements [63] when normalized to 1 second. The same strategy of using a longer-wavelength pump to reduce the background counts was adopted by Xu *et al.* [71] in pulsed up-conversion of 1.32 μm photons showing a low dark-count rate of $\approx 2.2 \times 10^3$/s for a system detection efficiency of 20%. When the overall detection efficiency and the duty cycle of the pulsed system are taken into account, the dark-count rates for pulsed [71] and cw [82] up-conversion using a 1.55 μm pump are similar.

E. Ultrafast Up-conversion

We can take advantage of the temporal degree of freedom in pulsed up-conversion in a different way. Instead of employing a pump pulse that is wider than the input photon's pulse width to ensure maximum temporal coverage and hence efficiency, the pump pulse can be much shorter to reveal temporal information of the incoming photons that is absent from typical up-conversion measurements and with a time resolution that is far better than any currently available single-photon counters. The idea is to use sub-picosecond pump pulses to optically sample the input photons which have typical coherence times of 1 ps. The input photon is up-converted only if the sub-ps pump pulse is present, and by scanning the arrival time of the pump pulse relative to the input photon, the temporal profile of the input photon can be mapped.

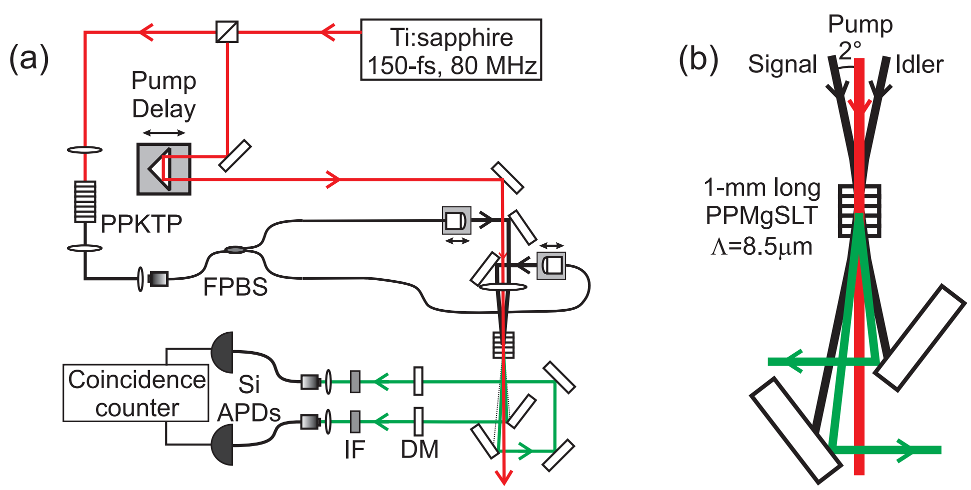

FIG. 21: Schematic of sub-picosecond up-conversion experiments for measuring spatial variation in generation efficiencies along the length of a nonlinear crystal and for two-photon joint temporal measurements. Reproduced from [66], Fig. 1.

Kuzucu *et al.* utilized this time-resolved measurement technique in the setup of Fig. 21 to map the spatial variation of SPDC generation probability in a PPKTP crystal with a resolution of ≈1 mm [66]. A 790 nm mode-locked Ti:sapphire laser with 150 fs pulses was used to pump a 1 mm long periodically poled MgO-doped stoichiometric lithium tantalate (PPMgSLT) crystal that served as the up-conversion nonlinear medium. The type-0 phase-matched PPMgSLT crystal length was much shorter than the typical 40-mm length used in other up-conversion experiments because it is necessary to have a large phase-matching bandwidth (which is inversely proportional to the crystal length) for the short signal and pump pulses. The high peak power of the 150 fs pump pulse provided adequate up-conversion efficiency for the short crystal. To reduce timing jitter between the pump and the SPDC outputs, the same pump pulse was used to drive both the PPKTP down-conversion and the PPMgSLT up-conversion processes. Three optical delay lines provided separate timing control for the pump and the two input IR beams to the up-converter. Figure 21(b) shows that the two independent up-converters are implemented with a single crystal and a single pump by arranging the two IR inputs centered at 1580 nm from the PPKTP crystal in a noncollinear configuration. The three beams are in a non-planar geometry to avoid coincident detection of any non-phase-matched down-converted photons that could be generated in PPMgSLT by the strong pump.

Because the pump and the SPDC outputs propagated at different group velocities in PPKTP, the location of the photon-pair emission in the crystal can be inferred from their arrival times relative to the pump pulse, with no (maximum) difference if they were emitted at the exit (entrance) facet of the crystal. The spatial resolution was set by the pulse width of the pump. Figure 22 displays the temporal profiles of signal and idler outputs from the PPKTP crystal, clearly showing that the generation probabilities in the front and back halves of the crystal are different. The time-resolved technique enables one to monitor the quality of the crystal's periodic grating structure and can be utilized as a diagnostic tool in the poling process.

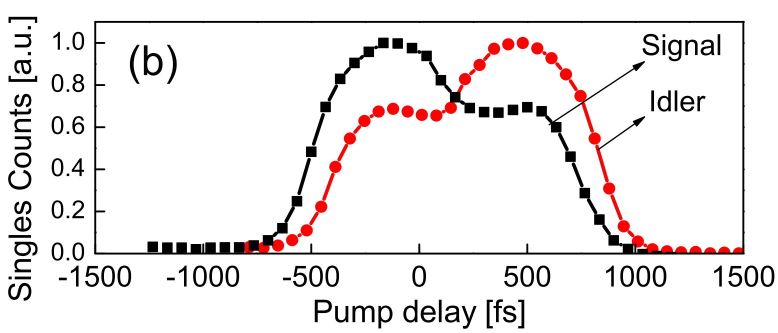

FIG. 22: Normalized singles histogram for signal and idler outputs of PPKTP as a function of the pump delay. Signal and idler profiles are mirror images of each other because under extended phase matching conditions the signal-pump and idler-pump time delays are equal with opposite signs [67]. Reproduced from [66], Fig. 2(b).

Ultrafast up-conversion can be used as a new tool for characterizing two-photon entanglement. Tunable narrowband spectral filters are traditionally used for making joint spectral measurements of two correlated photons, but they can be quite lossy for very narrow bandwidths. Time-frequency Fourier duality suggests that one can learn as much from joint temporal measurements by optical sampling based on ultrafast up-conversion, whose timing resolution is easily adjustable and can be as fine as a few femtoseconds.

Kuzucu *et al.* applied the ultrafast up-conversion technique to directly measure the time anticorrelation characteristics of coincident-frequency entangled photon pairs generated by ultrafast SPDC in type-II phase-matched PPKTP under extended phase-matching conditions [87]. These specially phase-matched SPDC signal and idler photons are positively correlated in their frequencies and therefore, by Fourier duality, they should be anticorrelated in the time domain (relative to the pump excitation pulse for the PPKTP down-converter). That is, if the signal photon arrives a certain amount of time after the pump pulse, then the idler photon would arrive the same amount of time before the pump pulse. Using the same setup of Fig. 21 a background-free two-photon-coincidence temporal profile was obtained for a pump bandwidth of 6 nm as a function of the signal and idler delay times [67], as shown in Fig. 23(a). It verifies that the arrival times of the two photons were indeed anticorrelated. Moreover, the joint temporal measurement capability makes it possible to monitor the result when one modifies the temporal correlation characteristics. Figure 23(b) shows a joint temporal density of a two-photon state that is nearly temporally unentangled when the pump bandwidth was reduced from 6 nm in Fig. 23(a) to 1.1 nm in (b) [67].

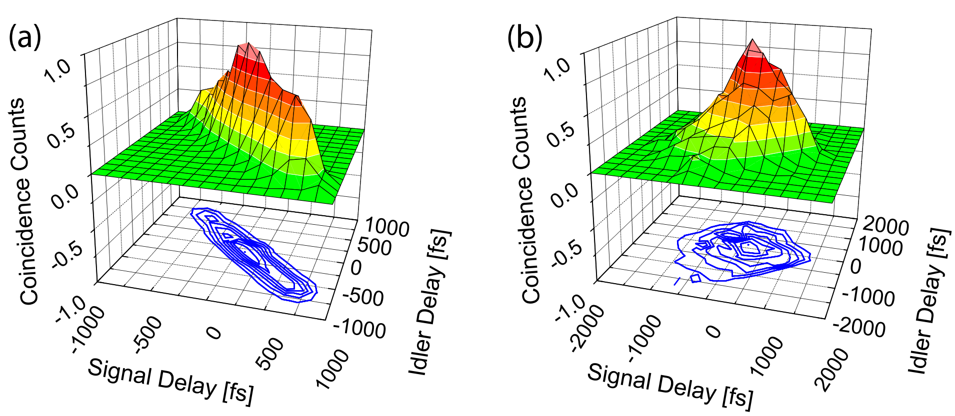

FIG. 23: Two-photon joint temporal density for pump 3-dB bandwidth of (a) 6 nm and (b) 1.1 nm as measured by ultrafast up-conversion with 150 fs pulses [67]. Reproduced from [67], Fig. 4.

V. CONCLUSION

This chapter has covered a variety of hybrid detectors aimed at improving and expanding the capabilities of photon-counting measurements beyond those of a SPAD. Multiplexing in space involving multiple SPADs, or in time with delay loops and a single SPAD, makes it possible to use conventional detectors to achieve a certain level of photon-number-resolving capability essential to many applications. Spatial multiplexing is well suited to overcoming single-SPAD limitations in maximum count rates, afterpulsing, and dead times. Multiplexing also enables new capabilities such as the weak-field homodyne detector based on the phase-sensitive time-multiplexed detection system.

Frequency up-conversion removes the need to operate at wavelengths within the range of sensitivity of Si SPADs. Instead, the photon can be translated from another wavelength of interest, such as the telecom band for low-loss transmission to a remote location via optical fibers. One particularly novel capability is sub-picosecond up-conversion that serves to optically sample the longer-duration photons to probe their rich temporal characteristics.

The future for hybrid detectors is bright because the demands of newly developed applications for high-performance photon-counting measurement capabilities cannot be met by existing single-photon detectors in a compact and convenient package. These hybrid detection techniques are expected to be further refined and utilized in an expanding range of applications, some of which have been developed during the time this book is being written. Hybrid measurement capabilities will form part of an essential detection toolbox for the growing fields of quantum information processing and low-light detection and imaging.

[1] R. H. Hadfield, Nature Photon. **3,** 696 (2009); C. Silberhorn, Contemp. Phys. **48,** 143 (2007) *and refs. therein.*

[2] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96,** 010401 (2006).

[3] Y. Gao, P. M. Anisimov, C. F. Wildfeuer, J. Luine, H. Lee, and J. P. Dowling , J. Opt. Soc. Am. B **27,** A170 (2010).

[4] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photon. **5,** 222 (2011).

[5] G. Brida, M. Genovese, and I. Ruo Berchera, Nature Photon. **4,** 227 (2010).

[6] E. Knill, R. Laflamme, and G. J. Milburn, Nature **409,** 46 (2001).

[7] L. M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Nature **414,** 413 (2001).

[8] T. Langer and G. Lenhart, New J. Phys. **11,** 055051 (2009); J. L. O'Brien, A. Furusawa, and J. Vuckovic, Nature Photon. **3,** 687 (2009); N. Gisin and R. Thew, Nature Photon. **1,** 165 (2007) *and refs. therein.*

[9] M. Genovese, Phys. Rep. **413,** 319 (2005) *and refs. therein.*

[10] K. D. Irwin and G. C. Hilton, *Cryogenic Particle Detection*, Edited by C. Enss, Topics Appl. Phys., Vol. 99, (Springer-Verlag, Berlin, 2005), 63.

[11] H. Paul, P. Törmä, T. Kiss, and I. Jex, Phys. Rev. Lett. **76,** 2464 (1996).

[12] L. A. Jiang, E. A. Dauler, and J. T. Chang, Phys. Rev. A **75,** 062325 (2007).

[13] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva, G. Gol'tsman, K. G. Lagoudakis, M. Benkhaoul, F. Lvy, and A. Fiore, Nat. Phot. **2,** 302 (2008).

[14] K. Banaszek and I. A. Walmsley, Opt. Lett. **28,** 52 (2003).

[15] M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, Phys. Rev. A **68,** 043814 (2003).

[16] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, and I. A. Walmsley, Opt. Lett. **28,** 2387 (2003).

[17] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, I. A. Walmsley, M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, J. Mod. Optics **51,** 1499 (2004).

[18] M. Avenhaus, K. Laiho, M. V. Chekhova, and C. Silberhorn, Phys. Rev. Lett. **104,** 063602 (2010).

[19] E. A. Goldschmidt, F. Piacentini, I. R. Berchera, S. V. Polyakov, S. Peters, S. Kueck, G.

Brida, I. P. Degiovanni, A. Migdall, and M. Genovese, Phys. Rev. A **88,** 013822 (2013).

[20] A. Restelli, J. C. Bienfang, and A. L. Migdall, Appl. Phys. Lett. **102,** 141104 (2013).

[21] A. Papoulis, *Probability, Random variables, and Stochastic process*, (McGraw-Hill-International Edition, 1991).

[22] I. P. Degiovanni, S. Polyakov, A. Migdall, private communication.

[23] Rodney Loudon, *The Quantum Theory of Light* (Oxford Science Publications, 2000).

[24] D. R. Cox and H. D. Miller, *The theory of stochastic processes*, (Chapman and Hall Ed., 1978).

[25] D. R. Cox, *Renewal Theory*, (Chapman and Hall Ed., 1982).

[26] S. Castelletto, I. P. Degiovanni, and M. L. Rastello, Phys. Rev. A **67,** 022305 (2003).

[27] J. W. Müller, Nucl. Instr. and Meth. in Physics Reserach **A301,** 543 (1991).

[28] J. Libert, Nucl. Instr. and Meth. in Physics Research **126,** 589 (1975).

[29] J. Libert, Nucl. Instr. and Meth. in Physics Research **136,** 563 (1976).

[30] J. W. Müller, Report BIPM-89/1.

[31] J. W. Müller, *Applied modulo counting*, *Advanced Mathematical Tools in Metrology*, (World Scientific, 1994, p.133).

[32] F. Esposito, N. Spinelli, R. Velotta and V. Berardi, Rev. Sci. Instrum. **62,** 2822 (1991).

[33] F. Carloni, A. Corberi, M. Marseguerra and C. M. Porceddu, Nucl. Instr. and Meth. in Physics Reserach **78,** 70 (1970).

[34] S. Castelletto, I. P. Degiovanni and M. L. Rastello, Metrologia **37,** 613 (2000).

[35] I. M. Gel'fand and G. E. Shilov, *Generalized functions*, (vol.1, Academic Press, 1964).

[36] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, (Dover Publications Inc., 1965).

[37] S. A. Castelletto, I. P. Degiovanni, V. Schettini, and A. L. Migdall, J. Mod. Opt., **54,** 337, (2007).

[38] V. Schettini, S. V. Polyakov, I. P. Degiovanni, G. Brida, S. Castelletto, and A. L. Migdall, IEEE Quant. Electron. Sel. Topics. **13,** 978 (2007).

[39] G. Brida, I. P. Degiovanni, V. Schettini, S. V. Polyakov, and A. Migdall, J. Mod. Opt. **56,** 405 (2009).

[40] G. Brida, I. P. Degiovanni, F. Piacentini, V. Schettini, S. V. Polyakov, and A. Migdall, Rev. Sci. Instr. **80,** 116103 (2009).

[41] http://mathworld.wolfram.com/LambertW-Function.html.

[42] R. Hanbury Brown and R. Q. Twiss, Nature **177,** 27 (1956).

[43] L. A. Jiang, E. A. Dauler, and J. T. Chang, Phys. Rev. A **75,** 062325 (2007).

[44] D. A. Kalashnikov, S.-H. Tan, M. V. Chekhova and L. A. Krivitsky, Opt. Express **19,** 9352 (2011).

[45] http://jp.hamamatsu.com/resources/products/ssd/pdf/s10362-11_series_kapd1022e05.pdf (2009).

[46] A. J. Kerman, E. A. Dauler, W. E. Keicher, J. K. W. Yang, K. K. Berggren, G. Gol'tsman, and B. Voronov, Appl. Phys. Lett. **88,** 111116 (2006).

[47] E. A. Dauler, A. J. Kerman, B. S. Robinson, J. K. W. Yang, B. Voronov, G. Goltsman, S. A. Hamilton, and K. K. Berggren, J. Mod. Opt. **56,** 364 (2009).

[48] D. Rosenberg, A. J. Kerman, R. J. Molnar, and E. A. Dauler, Opt. Exp. **21,** 1440 (2013).

[49] http://www.nist.gov/pml/div684/grp03/multicoincidence.cfm.

[50] C. Silberhorn, D. Achilles, A. B. U'Ren, K. Banaszek, and I. A. Walmsley, "Characterization and preparation of higher photon number states," in *Proceedings of 7th International Conference on Quantum Communication, Measurement and Computing,* (S. M. Barnett, E. Andersson, J. Jefferes, P. Ohberg, and O. Hirota, Eds.) pp. 342–345 (2004).

[51] D. Achilles, C. Silberhorn, and I. A. Walmsley, Phys. Rev. Lett. **94,** 043602 (2006).

[52] H. B. Coldenstrodt-Ronge, J. S. Lundeen, K. L. Pregnell, A. Feito, B. J. Smith, W. Mauerer, C. Silberhorn, J. Eisert, M. B. Plenio, and I. A. Walmsley, J. Mod. Opt. **56,** 432 (2009).

[53] H. B. Coldenstrodt-Ronge, "Photon number resolved characterization of quantum states," Diploma Thesis, Max Planck Research Group, Institute of Optics, Information and Photonics, University Erlangen-Nuremberg, Germany, March 2006.

[54] A. P. Worsley, H. B. Coldenstrodt-Ronge, J. S. Lundeen, P. J. Mosley, B. J. Smith, G. Puentes, N. Thomas-Peter, and I. A. Walmsley, Opt. Express **17,** 4397 (2009).

[55] J. Eisert, F. G. S. L. Brandão and K. M. R. Audenaert, New J. Phys. **9,** 46 (2007).

[56] G. Puentes, A. Datta, A. Feito, J. Eisert, M. B. Plenio, and I. A. Walmsley, New J. Phys. **12,** 033042 (2010).

[57] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70,** 1244 (1993).

[58] U. Leonhardt, *Measuring the quantum state of light*, Cambridge University Press, 2005.

[59] G. Puentes, J. S. Lundeen, M. P. A. Branderhorst, H. B. Coldenstrodt-Ronge, B. J. Smith, and I. A. Walmsley, Phys. Rev. Lett. **102,** 080404 (2009).

[60] K. Laiho, K. N. Cassemiro, D. Gross, and C. Silberhorn, Phys. Rev. Lett. **105,** 253603 (2010).

[61] L. Zhang, H. B. Coldenstrodt-Ronge, A. Datta, G. Puentes, J. S. Lundeen, X.-M. Jin, B. J. Smith, M. B. Plenio, and I. A. Walmsley, Nature Photon. **6**, 364 (2012).

[62] L. Zhang, A. Datta, H. B. Coldenstrodt-Ronge, X.-M. Jin, J. Eisert, M. B. Plenio, and I. A. Walmsley, New J. Phys. **14,** 115005 (2012).

[63] M. A. Albota and F. N. C. Wong, Opt. Lett. **29,** 1449 (2004).

[64] A. P. Vandevender and P. G. Kwiat, J. Mod. Opt. **51,** 1433 (2004).

[65] R. V. Roussev, C. Langrock, J. R. Kurz and M. M. Fejer, Opt. Lett. **29,** 1518 (2004).

[66] O. Kuzucu, F. N. C. Wong, S. Kurimura, and S. Tovstonog, Opt. Lett. **33,** 2257 (2008).

[67] O. Kuzucu, F. N. C. Wong, S. Kurimura, and S. Tovstonog, Phys. Rev. Lett. **101,** 153602 (2008).

[68] M. Legré, R. Thew, H. Zbinden, and N. Gisin, Opt. Express **15,** 8237 (2007).

[69] R. T. Thew, S. Tanzilli, L. Krainer, S. C. Zeller, A. Rochas, I. Rech, S. Cova, H. Zbinden, and N. Gisin, New J. Phys. **8,** 32 (2006).

[70] H. Takesue, T. Honjo, and H. Kamada, Jpn. J. Appl. Phys. **45,** 5757 (2006).

[71] H. Xu, L. Ma, A. Mink, B. Hershman, and X. Tang, Opt. Express **15,** 7247 (2007).

[72] M. A. Albota and B. S. Robinson, Opt. Lett. **35,** 2627 (2010).

[73] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59,** 2044 (1987).

[74] H. Takesue, Phys. Rev. Lett. **101,** 173901 (2008).

[75] J. H. Shapiro, New J. Phys. **4,** 47.1 (2002).

[76] S. Tanzilli, W. Tittel, M. Halder, O. Alibart, P. Baldi, N. Gisin, and H. Zbinden, Nature **437,** 116 (2005).

[77] M. A. Albota, F. N. C. Wong, and J. H. Shapiro, J. Opt. Soc. Am. B **23,** 918 (2006).

[78] J. Huang and P. R. Kumar, Phys. Rev. Lett. **68,** 2153 (1992).

[79] Y. Ding and Z. Y. Ou, Opt. Lett. **35,** 2591 (2010).

[80] N. Curtz, R. Thew, C. Simon, N. Gisin, and H. Zbinden, Opt. Express **18,** 22099 (2010).

[81] H. Pan and H. Zeng, Opt. Lett. **31,** 793 (2006).

[82] C. Langrock, E. Diamanti, R. V. Roussev, Y. Yamamoto, M. M. Fejer, and H. Takesue, Opt. Lett. **30,** 1725 (2005).

[83] H. Takesue, E. Diamanti, C. Langrock, M. M. Fejer, and Y. Yamamoto, Opt. Express **14,** 13067 (2006).

[84] S. Ramelow, A. Fedrizzi, A. Poppe, N. K. Langford, and A. Zeilinger, Phys. Rev. A **85,** 013845 (2012).

[85] A. P. VanDevender and P. G. Kwiat, J. Opt. Soc. Am. B **24,** 295 (2007).

[86] L. Ma, J. C. Bienfang, O. Slattery, and X. Tang, Opt. Express **19,** 5470 (2011).

[87] O. Kuzucu, M. Fiorentino, M. A. Albota, F. N. C. Wong, and F. X. Kärtner, Phys. Rev. Lett. **94,** 083601 (2005).