Post-processing with linear optics for improving the quality of single-photon sources

Dominic W Berry\textsuperscript{1}, Stefan Scheel\textsuperscript{2}, Casey R Myers\textsuperscript{3}, Barry C Sanders\textsuperscript{1,4}, Peter L Knight\textsuperscript{2} and Raymond Laflamme\textsuperscript{3,5}

\textsuperscript{1} Department of Physics, Australian Centre for Quantum Computer Technology, Macquarie University, Sydney, New South Wales 2109, Australia
\textsuperscript{2} Quantum Optics and Laser Science, Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2BW, UK
\textsuperscript{3} Institute for Quantum Computing, University of Waterloo, Ontario N2L 3G1, Canada
\textsuperscript{4} Institute for Quantum Information Science, University of Calgary, Alberta T2N 1N4, Canada
\textsuperscript{5} Perimeter Institute for Theoretical Physics, 35 King Street N., Waterloo, Ontario N2J 2W9, Canada

E-mail: berry@ics.mq.edu.au

Abstract. Triggered single-photon sources produce the vacuum state with non-negligible probability, but produce a much smaller multiphoton component. It is therefore reasonable to approximate the output of these photon sources as a mixture of the vacuum and single-photon states. We show that it is impossible to increase the probability for a single photon using linear optics and photodetection on fewer than four modes. This impossibility is due to the incoherence of the inputs; if the inputs were pure-state superpositions, it would be possible to obtain a perfect single-photon output. In the more general case, a chain of beam splitters can be used to increase the probability for a single photon, but at the expense of adding an additional multiphoton component. This improvement is robust against detector inefficiencies, but is degraded by distinguishable photons, dark counts or multiphoton components in the input.
One of the most promising methods for quantum information processing is linear optics and photodetection. Linear optics and photodetection may be used for provably secure quantum communication [1] as well as quantum computation [2]. An important requirement for these schemes is the ability to produce a single photon on demand [2, 3], yet generating high-fidelity single-photon states is challenging. The traditional method for generating single photons involves photodetection on one output mode from a non-degenerate parametric down-conversion process to post-select a single photon in the correlated mode [4, 5]. This method has the drawback that the time of the photon emission is not controlled. More recently, triggered photon sources have been developed, including molecules [6], quantum wells [7], colour centres [8], atoms in high-Q cavities [9] and quantum dots [10]. These sources have a significant vacuum contribution, but the multiphoton contribution may be made very small [11].

In this study, we consider triggered sources because it is desirable to generate the photons at the same time in order to perform interference. For the majority of this study, we approximate these sources by taking the multiphoton probability to be zero. That is, we consider an idealized single-mode single-photon source, which may be represented by the density operator

\[ \hat{\rho}_p = (1 - p)|0\rangle\langle 0| + p|1\rangle\langle 1|. \number{1} \]

Here \( p \) is the probability for a single photon, and is also called the efficiency. Increasing the efficiency is important because many quantum optics experiments, especially those concerned with linear optical quantum computation, require high-efficiency sources. We also derive results for states with a coherent superposition of the vacuum and a single photon; however, this form of state is not produced by single-photon sources.

Much effort is directed towards improving sources, but here we consider the problem of performing post-processing to obtain higher efficiency. Ideally, this post-processing should also maintain a zero multiphoton contribution (although a very small multiphoton contribution...
would also be acceptable). A promising method of post-processing is linear optics and photodetection. As mentioned above, linear optics and photodetection can be used to perform quantum computation [2] and optical controlled-NOT gates have recently been demonstrated [12]; however, there are also no-go theorems for linear optics [13].

In a recent publication [14], we investigated the possibility of improving the single-photon efficiency $p$ with linear optical elements and photodetection. Here we present a number of new results that clarify the limitations inherent in this method as well as reviewing the results presented in [14]. In particular, we analyse the effects of various experimental limitations, present a scheme that gives perfect results for inputs with a coherent superposition of zero and one photon, and show connections between some difficult unsolved problems.

We begin by showing that it is not possible to obtain an improvement for the simple case of a beam splitter in section 2. In section 3 we show that, if we allow a coherent superposition of zero and one photon, rather than an incoherent mixture, it is possible to obtain a perfect single-photon state. We then proceed to the case of a multimode interferometer with incoherent inputs in section 4. In sections 5 and 6 we give an expanded discussion of the limitations on the improvement that is possible to obtain, and the method to obtain an improvement. Section 7 gives a detailed discussion of the impact of various experimental problems on this method. We give further discussion of no-go theorems for post-processing in section 8. Then, in section 9, we show that there are deep connections between the unsolved problems for post-processing. Lastly, we give a discussion of how the theory is changed by allowing multiphoton contributions in the inputs in section 10, and we conclude in section 11.

2. Beam splitter

To begin with, we consider the simplest case of two copies of the quantum state (1) combined on a single beam splitter. The initial state can be written in the form

$$\hat{\rho}_{in}^{(2)} = (1 - p_1)(1 - p_2)|00\rangle\langle00| + p_1(1 - p_2)|10\rangle\langle10| + p_2(1 - p_1)|01\rangle\langle01| + p_1p_2|11\rangle\langle11|. \quad (2)$$

Each photonic mode operator gets transformed by the beam splitter in the following way:

$$\hat{a}_{1}^\dagger \mapsto \Lambda_{11} \hat{a}_{1}^\dagger + \Lambda_{21} \hat{a}_{2}^\dagger, \quad \hat{a}_{2}^\dagger \mapsto \Lambda_{12} \hat{a}_{1}^\dagger + \Lambda_{22} \hat{a}_{2}^\dagger, \quad (3)$$

where $\Lambda$ is a $2 \times 2$ unitary matrix [15]. We have two options: to project onto the vacuum or onto the single-photon Fock state (projecting onto two photons results in vacuum output in the unmeasured mode). After vacuum projection in mode 2, we end up with an unnormalized state of the form

$$\hat{\rho}_{out}^{(2)} \propto |0\rangle\langle0| + \left( \frac{p_1}{1 - p_1}|\Lambda_{11}|^2 + \frac{p_2}{1 - p_2}|\Lambda_{12}|^2 \right)|1\rangle\langle1| + \text{two-photon term}. \quad (4)$$

Thus the ratio between the probabilities for one and zero photons is just a weighted average of $p_1/(1 - p_1)$ and $p_2/(1 - p_2)$ and cannot exceed either of these. That is, it is not possible to improve the ratio between the probabilities for obtaining one and zero photons. This implies that it is not possible to improve the absolute probability of obtaining one photon.
In the case of a single-photon detection, the resulting unnormalized state is

$$\hat{\rho}_{\text{out}}^{(2)} \propto \left( \frac{1-p_1}{p_1} |\Lambda_{22}|^2 + \frac{1-p_2}{p_2} |\Lambda_{21}|^2 \right) |0\rangle \langle 0| + |\text{per } \Lambda|^2 |1\rangle \langle 1|,$$

(5)

where \( \text{per } \Lambda \) is the permanent [16] of the beam splitter matrix \( \Lambda \). Since the absolute value of the permanent of a unitary matrix is bounded from above by unity, and the term in brackets is a weighted sum of terms \((1-p_i)/p_i\), we do not find any improvement for this case either. Hence, there is no improvement in the probability for a single photon if zero, one or two photons are detected. These results demonstrate that for mixed-state inputs it is impossible to obtain an improvement in the single-photon probability using a beam splitter.

3. Pure-state inputs

It is possible to obtain an improvement using a beam splitter if the inputs are in pure-state superpositions of zero and one photon, instead of incoherent mixtures. Consider two input modes that are each in the state \( \alpha |0\rangle + \beta |1\rangle \). The initial state may be written as

$$|\psi\rangle_{\text{in}}^{(2)} = [\alpha^2 + \alpha \beta (a_1^\dagger + a_2^\dagger) + \beta^2 a_1^\dagger a_2^\dagger] |00\rangle.$$

(6)

Applying the beam splitter transformation (3) gives

$$|\psi\rangle_{\text{trans}}^{(2)} = [\alpha^2 + \alpha \beta (\Lambda_{11} + \Lambda_{12}) a_1^\dagger + (\Lambda_{21} + \Lambda_{22}) a_2^\dagger]$$

$$+ \beta^2 [\Lambda_{11} \Lambda_{12} (a_1^\dagger)^2 + \Lambda_{21} \Lambda_{22} (a_2^\dagger)^2 + (\text{per } \Lambda) a_1^\dagger a_2^\dagger)] |00\rangle.$$

(7)

Conditioning on detection of zero photons in mode 2 gives the output state

$$|\psi\rangle_{\text{out}}^{(2)} \propto \alpha^2 |0\rangle + \alpha \beta (\Lambda_{11} + \Lambda_{12}) |1\rangle + \sqrt{2} \beta^2 \Lambda_{11} \Lambda_{12} |2\rangle.$$

(8)

It is easily seen that this output state may have a higher probability for a single photon. For example, if the initial state is close to the vacuum state (i.e. \( \alpha \gg \beta \)), then an improvement by a factor of 2 may be obtained by using \( \Lambda_{11} = \Lambda_{12} = 1/\sqrt{2} \).

In fact, it is possible to further process this output state to obtain a perfect single-photon state. If we combine this state with mode 3, which is also assumed to be prepared in state \( \alpha |0\rangle + \beta |1\rangle \), then the total state may be represented as

$$|\psi\rangle_{\text{in}}^{(3)} \propto [\alpha^2 + \alpha \beta (\Lambda_{11} + \Lambda_{12}) a_1^\dagger + \beta^2 \Lambda_{11} \Lambda_{12} (a_1^\dagger)^2] (\alpha + \beta a_3^\dagger) |00\rangle.$$

(9)

Applying the beam splitter transformation (3) (except using a prime to distinguish this beam splitter from the previous one) gives

$$|\psi\rangle_{\text{trans}}^{(3)} \propto [\alpha^2 + \alpha \beta (\Lambda'_{11} + \Lambda'_{12}) (\Lambda'_{11} a_1^\dagger + \Lambda'_{31} a_3^\dagger)$$

$$+ \beta^2 \Lambda'_{11} \Lambda'_{12} (\Lambda'_{11} a_1^\dagger + \Lambda'_{31} a_3^\dagger)^2] [\alpha + \beta (\Lambda'_{13} a_1^\dagger + \Lambda'_{33} a_3^\dagger)] |00\rangle.$$

(10)
Conditioning on detection of two photons gives

$$|\psi^{(3)}\rangle_{\text{out}} \propto \sqrt{2} \beta^2 \Lambda_3' \cdot \alpha[(\Lambda_{11} + \Lambda_{12})\Lambda_{33} + \Lambda_{11} \Lambda_{12} \Lambda_{31}']|0\rangle + \beta\Lambda_{11} \Lambda_{12}(2\Lambda_{11}' \Lambda_{33} + \Lambda_{31}' \Lambda_{13}')|1\rangle.$$  

(11)

To make this equation more clear, we use

$$\Lambda = \begin{bmatrix} e^{i\phi} \cos \theta & -\sin \theta \\ \sin \theta & e^{-i\phi} \cos \theta \end{bmatrix}, \quad \Lambda' = \begin{bmatrix} e^{i\phi'} \cos \theta' & -\sin \theta' \\ \sin \theta' & e^{-i\phi'} \cos \theta' \end{bmatrix}.$$  

(12)

The condition that the output state is a pure single-photon state then becomes

$$(\cos \theta - e^{-i\phi} \sin \theta) e^{-i\phi'} \cos \theta' - \sin \theta \cos \theta \sin \theta' = 0.$$  

(13)

This equation may be satisfied by taking $\theta'$ and $\phi'$ to be

$$\theta' = \arctan \left( \frac{|\cos \theta - e^{-i\phi} \sin \theta|}{\sin \theta \cos \theta} \right), \quad \phi' = \arg(\cos \theta - e^{-i\phi} \sin \theta).$$  

(14)

That is, regardless of the characteristics of the initial beam splitter (provided $\sin \theta$ and $\cos \theta$ are non-zero), it is possible to obtain a perfect single-photon output.

Another issue is the probability for obtaining the desired pattern of detection results. Using the unnormalized expression for $|\psi^{(3)}_{\text{out}}\rangle$ above, this probability is given by

$$P = 2|\beta|^6 \sin^2 \theta \cos^2 \theta \sin^2 \theta' (2 \cos^2 \theta' - \sin^2 \theta')^2$$

$$= |\beta|^6 \frac{\frac{1}{2} \sin^2 2\theta(1 - \cos \phi \sin 2\theta) \left( \frac{1}{4} \sin^2 2\theta - 1 + \cos \phi \sin 2\theta \right)}{\left( \frac{1}{4} \sin^2 2\theta + 1 - \cos \phi \sin 2\theta \right)^3}.$$  

(15)

In the second line we have used the expression (14) for $\theta'$. This probability is plotted in figure 1 for the range 0 to $\pi$ in $\theta$ and $\phi$ ($P$ is periodic with period $\pi$ in these variables). There are four maxima in this range, for $(\theta, \phi) = (\pi/4, \pi), (\pi/4, \acos(13/14)), (3\pi/4, 0)$ and $(3\pi/4, \acos(-13/14))$. The exact values of $\phi$ of $\acos(13/14)$ and $\acos(-13/14)$ are not obvious from the plot but are straightforward to obtain analytically. The two maxima $(\pi/4, \pi)$ and $(3\pi/4, 0)$ correspond to the same beam splitter, so there are only three maxima that correspond to distinct beam splitters. Each of these maxima is exactly the same height,

$$P_{\max} = 16|\beta|^6/81.$$  

(16)

One factor that distinguishes the maxima is the sensitivity to the parameters. Clearly, the maxima at $(\pi/4, \acos(13/14))$ and $(3\pi/4, \acos(-13/14))$ are far more sensitive to the values of $\theta$ and $\phi$, and it is therefore better to use the beam splitter corresponding to $(\pi/4, \pi)$ and $(3\pi/4, 0)$. For the second beam splitter, the appropriate parameters are $\theta' = \acos(1/3)$ and $\phi' = 0$. That is, the best result is obtained by using a 50/50 beam splitter followed by a beam splitter with a reflectivity of 1/9.

Thus we see that if the inputs to an interferometer are in pure superposition states, it is possible to obtain a perfect single-photon output for three modes. In contrast, if the inputs to the interferometer are incoherent superpositions of Fock states, it is impossible to obtain an
improvement in the single-photon probability for three modes [14]. These results imply that it is the incoherence in the inputs that prevents an improvement in the single-photon probability. It would be interesting to determine the degree of decoherence that is sufficient to prevent an improvement in the single-photon efficiency. However, this is a difficult problem, which we leave to further investigation.

4. Multimode incoherent inputs

Although it is possible to obtain perfect single-photon states from pure superposition states, this method cannot be applied to current experiments, as single-photon sources do not produce pure superposition states. Real single-photon sources produce an incoherent combination of Fock states; therefore, we consider input states of this form for the remainder of this paper. In the multimode case, we start with a supply of \(N\) mixed states of the form (1). For additional generality, we allow the different inputs to have different probabilities for a single photon, \(p_i\), and we denote the maximum of these probabilities by \(p_{\text{max}}\). The initial input state may be described by

\[
\hat{\rho}_\text{in}^{(N)} = \bigotimes_{i=1}^N [(1 - p_i)|0\rangle\langle 0| + p_i|1\rangle\langle 1|]
\]

\[
= \sum_s P_s \left( \prod_i (\hat{a}_i^\dagger)^{s_i} |0\rangle \langle 0| \prod_i (\hat{a}_i)^{s_i} \right),
\]

where \(P_s = \prod_i p_i^{s_i} (1 - p_i)^{1-s_i}\), and the vector \(s = (s_1, \ldots, s_N)^T\) \((s_j = 0, 1)\) gives the photon numbers in the inputs. The quantity \(P_s\) is the probability of obtaining this combination of input photon numbers.

Figure 1. The probability of obtaining the desired detection results as a function of the beam splitter parameters \(\theta\) and \(\phi\) for the first beam splitter.
Figure 2. Schematic set-up of the network. We assume \( N \) incoming modes prepared in the state (17) with different \( p_i \). The photon number is measured in output modes 2 to \( N \), and we wish to improve the probability for a single photon in mode 1.

This input is then passed through a passive interferometer which consists of beam splitters, mirrors and phase shifters (see figure 2). Each of these elements preserves total photon number from input to output under ideal conditions. No energy is required to operate these optical elements, hence the term passive (these are also known as linear optical elements). More generally, polarization transforming elements can be included, but here we are concerned only with a scalar field treatment; in fact, polarization effects could be included by doubling the number of channels and treating the two polarizations in a mode as two separate channels.

Classically, the field amplitude of channel \( i \) would be represented by the complex number \( a_i \). The set of all field amplitudes for the \( N \)-channel interferometer is given by the vector \( a = (a_1, \ldots, a_N)^T \). The passive interferometer transforms the input amplitudes to the output amplitudes via the matrix transformation \( a \mapsto \Lambda \hat{a} \) with \( \Lambda \in U(N) \), where \( U(N) \) is the set of all \( N \times N \) unitary matrices. Quantization of the field is obtained by the replacement of \( a \) by the vector annihilation operator \( \hat{a} \), and the interferometer transforms the operators according to \( \hat{a} \mapsto \Lambda \hat{a} \). [15] This transformation of the operators yields

\[
\hat{\rho}_{\text{trans}}^{(N)} = \sum_x P_x \left[ \prod_i \left( \sum_k \Lambda_{ki} \hat{a}_k \right)^{x_i} |0\rangle \langle 0| \prod_i \left( \sum_k \Lambda_{ki}^* \hat{a}_k \right)^{x_i*} \right].
\] (18)

In the completely general case, we could perform photodetections on \( N - N_1 \) of the modes, and use the remaining \( N_1 \) modes as single-photon sources if the desired combination of detection
results is obtained. However, this generality is not needed here because we are concerned with the maximum improvement in the photon statistics in a single mode. In order to fix notation, we denote by mode 1 that mode for which we want to improve the statistics, and label the other modes where photodetections have not been performed as modes 2 to $N_1$. The reduced density matrix in mode 1 is then identical to what would be obtained if photodetections were performed on modes 2 to $N_1$ (as well as $N_1+1$ to $N$), and the results of these photodetections discarded. Therefore, the probability for a single photon will be a weighted average of the single-photon probabilities for the different combinations of detections in modes 2 to $N_1$. Hence, the maximum single-photon probability in mode 1 will be obtained for some combination of detections in modes 2 to $N_1$. For this reason, we consider the state in mode 1 conditioned on photodetections in the other $N-1$ modes. As our aim is to determine the best results possible using linear optics and photodetection, we also assume that the photodetectors perform perfect photon counting measurements (imperfect detection is discussed later).

Before determining the conditional output state, we introduce some additional notation. The total number of photons detected is $D$, and the maximum possible number of photons input to the interferometer is $M$. As some of the $p_i$ may be equal to zero, $M$ may be less than $N$; $M$ is equal to the number of non-zero values of $p_i$. For $j>1$, $n_j$ is the number of photons detected in mode $j$, and $n_1$ is the photon number in mode 1 (the output mode). We use the notation $\Sigma_n = \sum_i n_i$ (so $\Sigma_n = D + n_1$) and $\Sigma_i = \sum_j s_i$. In addition, we define the set $\Phi_s = \{i \mid s_i = 1\}$, and let $\Upsilon_s$ be the set that consists of all vectors comprising elements of $\Phi_s$.

The conditional state in mode 1 after photodetection in modes 2 to $N$ is

$$\hat{\rho}_{\text{out}}^{(N)} = \sum_{n_1=0}^{N} c_{n_1} |n_1\rangle \langle n_1|.$$  \hspace{1cm} (19)

Each coefficient $c_{n_1}$ is given by

$$c_{n_1} = K |n\rangle \hat{\rho}_{\text{trans}}^{(N)} |n\rangle,$$  \hspace{1cm} (20)

where $|n\rangle$ is a tensor product of number states in each of the output modes and the normalization constant $K$ is equal to

$$K = \left[ \sum_{n_1=0}^{N} \langle n | \hat{\rho}_{\text{trans}}^{(N)} |n\rangle \right]^{-1}. $$  \hspace{1cm} (21)

Evaluating $c_{n_1}$ gives

$$c_{n_1} = \frac{K'}{n_1!} \sum_{\sigma} P_s |S_{\sigma,n}\rangle^2, $$  \hspace{1cm} (22)

where $K' = K / \prod_{j=2}^{N} n_j!$, and

$$S_{\sigma,n} = \sum_{\Sigma_1} (\Lambda_{1,\sigma_1} \cdots \Lambda_{1,\sigma_{n_1}}) \cdots (\Lambda_{N,\sigma_{N-1}} \cdots \Lambda_{N,\sigma_{N+1}}). $$  \hspace{1cm} (23)
This quantity may alternatively be expressed using permanents as

$$S_{s,n} = \text{per}(A[n, s]).$$

(24)

Here the notation $A[n, s]$ is used to indicate that the $i$th column of $A$ is repeated $s_i$ times and the $j$th row is repeated $n_j$ times.

Two figures of merit for a single-mode field of the form $\sum_i q_i |i\rangle \langle i|$ are

$$R = \frac{q_1}{q_0}, \quad G = \frac{q_2}{q_1} \frac{q_1}{q_0}.$$  

(25)

We use the subscript ‘out’ to indicate the output field and ‘in’ to indicate the input field. For the output field we simply have $q_i = c_i$. For the input field, we have $G_{\text{in}} = 0$ as the two-photon component is assumed to be negligible. For simplicity, we define $R_{\text{in}}$ to be the maximum input ratio $p_{\text{max}}/(1 - p_{\text{max}})$.

The figure of merit $G$ characterizes the two-photon contribution and is equal to $1/2$ for Poisson photon statistics. If the multiphoton component in the output is zero, then comparing $R_{\text{in}}$ and $R_{\text{out}}$ immediately tells us if there is an improvement in the probability for a single photon. Even if the multiphoton component is non-zero, using $R_{\text{out}}$ has the following advantages:

1. The common constant $K'$ cancels, so it is possible to evaluate $R_{\text{out}}$ analytically.
2. If $R_{\text{out}} \leq R_{\text{in}}$, then it is clear that $c_1 \leq p_{\text{max}}$. Thus we can determine those cases where there is no improvement.
3. For $p_{\text{max}} \ll 1$, $c_0 \approx 1$ and $R_{\text{in}} \approx p_{\text{max}}$. Therefore the improvement in $R$ is approximately the same as the improvement in the single-photon probability over $p_{\text{max}}$.

An alternative measure of the multiphoton contributions is given by how sub-Poissonian the field is. That is, we may define the measure

$$\Pi = \frac{\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2}{\langle \hat{n} \rangle}.$$  

(26)

For a sub-Poissonian field, $\Pi < 1$. States of the form (1) are sub-Poissonian with $\Pi = 1 - p$. We take $\Pi_{\text{in}}$ to be the minimum value in the inputs, $1 - p_{\text{max}}$, and $\Pi_{\text{out}}$ is simply the value for the output mode. If an output $c_1$ greater than $p_{\text{max}}$ is obtained, while maintaining a multiphoton contribution that is zero or very small, then it is clear that $\Pi_{\text{out}}$ will be smaller than $\Pi_{\text{in}}$. On the other hand, if the output has multiphoton contributions similar to those for a Poisson distribution, $\Pi_{\text{out}}$ will be closer to 1.

5. Limit on improvement

Ideally, we wish to obtain an improvement in the figure of merit $R$, while maintaining a value of $G$ that is zero, or at least small with respect to $1/2$ (the value for a Poisson distribution). This is a difficult task so, for simplicity, we begin by focusing on improving $R$. As was shown in [14], there is an upper limit on how far $R$ can be increased. Here we show this result in more detail.

New Journal of Physics 6 (2004) 93 (http://www.njp.org/)
First we consider the expression for $c_0$:

$$c_0 = K' \sum_{s, \Sigma_2 = D} P_s |S_{s, n^0}|^2. \quad (27)$$

Here $n_i^0 = 0$ and $n_i^0$ is the combination of detection results for $j > 1$. A simple way of re-expressing this summation is

$$c_0 = \frac{K'}{N - D} \sum_{s, \Sigma_2 = D + 1} \sum_{k, \Sigma_k = 1} P_s |S_{s', n^0}|^2,$$

where $s'_k = s_k$ except for $s_k^j = 0$. That is, we consider combinations of input photons $s$ with one too many photons, then remove one of these photons to obtain the correct number of input photons. This expression for the sum gives each term $N - D$ times, so it is necessary to divide by $N - D$ to obtain the correct result. Specifically, each alternative $s_k$ may be obtained from an $s$ which is identical, except one of the zeros of $s_k$ is replaced with a one. As each $s_k$ has $N - D$ zeros, there are $N - D$ possible alternative $s$ that give the same $s_k$.

If some of the inputs to the interferometer are simply vacuum states (i.e. some of the $p_i$ are zero), it is possible to express the summation in a more efficient way. First note that those terms in the sum in equation (27) with $P_s = 0$ do not contribute to the sum; therefore we may restrict to terms with $P_s \neq 0$.

$$c_0 = K' \sum_{s, P_s \neq 0, \Sigma_2 = D} P_s |S_{s, n^0}|^2. \quad (29)$$

We may re-express this equation as

$$c_0 = \frac{K'}{M - D} \sum_{s, P_s \neq 0, \Sigma_2 = D + 1} \sum_{k, \Sigma_k = 1} P_s |S_k|^2,$$

where $S_k = S_{s', n^0}$. Here the dividing factor required is only $M - D$.

Recall that the maximum total number of photons is $M$, so there are $N - M$ inputs with $p_i = 0$. Each alternative $s_k^j$ still has $N - D$ zeros, but some of these zeros will correspond to inputs with $p_i = 0$. Because we are restricting to terms with $P_s \neq 0$, for all $i$ with $p_i = 0$, $s_i = 0$ and therefore $s_k^j$ must be equal to zero. Thus all $N - M$ of the inputs with $p_i = 0$ must correspond to zeros of $s_k^j$, and so there will only be $M - D$ zeros of $s_k^j$ that correspond to non-zero $p_i$. As in the previous case, each $s_k^j$ may be obtained from an $s$ which is identical, except one of the zeros of $s_k^j$ is replaced with a one. However, because we have restricted to terms with $P_s \neq 0$, the zero that is replaced with a one must be for an $i$ with non-zero $p_i$. Hence there are only $M - D$ alternative $s$ that lead to the same $s_k^j$, and the redundancy in this case is only $M - D$. That is why a dividing factor of $M - D$ is required in this case.

Simplifying equation (30), we obtain

$$c_0 = \frac{K'}{M - D} \sum_{s, P_s \neq 0, \Sigma_2 = D + 1} P_s \sum_{k, \Sigma_k = 1} \frac{1 - p_k}{p_k} |S_k|^2. \quad (31)$$
Since the sum is limited to terms where $P_s \neq 0$, $p_k$ is non-zero, and therefore the ratio $(1 - p_k)/p_k$ does not diverge. Using $p_k \ll p_{\text{max}}$, we obtain the inequality

$$c_0 \geq \frac{K'/R_{\text{in}}}{M - D} \sum_{s; \Sigma_0 = D+1} P_s \sum_{k; s_1 = 1} |S_k|^2. \tag{32}$$

We now allow terms with $P_s = 0$, because they do not contribute to the sum.

It is also possible to obtain an inequality for $c_1$. The probability $c_1$ is given by

$$c_1 = K' \sum_{s; \Sigma_0 = D+1} P_s |S_{s,n'}|^2. \tag{33}$$

In this case, the notation $n^i$ means $n^i_1 = 1$, and $n^i_j$ is the combination of detection results for $j > 1$. We may express $S_{s,n'}$ as

$$S_{s,n'} = \sum_{\sigma \in Y_s} \Lambda_1,\sigma_1 (\Lambda_2,\sigma_2 \cdots \Lambda_2,\sigma_{n^i-1}) \cdots (\Lambda_N,\sigma_N,\sigma_N \cdots \Lambda_N,\sigma_N)$$

$$= \sum_{k; s_1 = 1} \Lambda_{1k} \sum_{\sigma \in Y_k} (\Lambda_2,\sigma_1 \cdots \Lambda_2,\sigma_{n^i-1}) \cdots (\Lambda_N,\sigma_N,\sigma_N \cdots \Lambda_N,\sigma_N)$$

$$= \sum_{k; s_1 = 1} \Lambda_{1k} S_{k,n^0}$$

$$= \sum_{k; s_1 = 1} \Lambda_{1k} S_k. \tag{34}$$

Therefore, we may re-express the equation for $c_1$ as

$$c_1 = K' \sum_{s; \Sigma_0 = D+1} P_s \left| \sum_{k; s_1 = 1} \Lambda_{1k} S_k \right|^2. \tag{35}$$

This then gives the inequality

$$c_1 \leq K' \sum_{s; \Sigma_0 = D+1} P_s \sum_{k; s_1 = 1} |S_k|^2. \tag{36}$$

Combining equations (32) and (36), we can see that

$$R_{\text{out}} = \frac{c_1}{c_0} \leq R_{\text{in}} (M - D). \tag{37}$$

This yields an upper limit on the ratio between the probabilities for one and zero photons. This result allows one to draw three main conclusions:

1. As $M \leq N$ and $D \geq 0$, the improvement in $R$ can never be greater than $N$. There is no known scheme that saturates this upper bound, but there is a scheme known that achieves an improvement of approximately $N/4$ [14].

*New Journal of Physics* 6 (2004) 93 (http://www.njp.org/)
2. If the number of photons detected is one less than the maximum input number, then $M - D = 1$, and there cannot be an improvement.\footnote{This no-go theorem has also been shown by E Knill (personal communication) for the case that $M = N$.} This case is important because it is the most straightforward way of eliminating the possibility of two or more photons in the output mode. We have not proven that it is impossible to obtain an improvement while eliminating the multiphoton component, but if such a scheme is possible it cannot eliminate the multiphoton component by detecting one less than the maximum input number of photons.

3. It is impossible to obtain a single photon with unit probability if $p_{\text{max}} < 1$. If $c_1 = 1$ were obtained, then $R_{\text{out}}$ would be infinite; from equation (37), this is clearly not possible unless $R_{\text{in}}$ is infinite (which would correspond to $p_{\text{max}} = 1$).

6. Method for improvement

We have previously shown that it is possible to obtain an improvement in the probability for a single photon [14]. Here we review this method, giving more motivation for this scheme, and propose a simple realization using a line of beam splitters. In order to obtain a value for the ratio $R_{\text{out}}$ that is close to the upper limit, we require the two inequalities (32) and (36) to be as close to equality as possible. We may achieve equality in the first case (32) by taking all non-zero $p_i$ equal to $p_{\text{max}}$. To obtain equality in equation (36), we would require $s_k = 1$ whenever $\Lambda_{1k}$ is non-zero, and $S_k$ to be proportional to $\Lambda_{1k}^*$ for those values of $k$ where $s_k = 1$. These conditions would need to be satisfied for all $s$ that give non-zero $P_s$. The first condition is a problem, because it cannot be satisfied unless $D = M - 1$. To see this, note that this condition is equivalent to requiring that $s_k = 0$ implies $\Lambda_{1k} = 0$, for all $s$ that give $P_s > 0$. If $P_s > 0$, then $s_k = 0$ for all $k$ such that $p_k = 0$. Therefore, $p_k = 0$ implies $\Lambda_{1k} = 0$. Now if $D + 1 < M$, then it must be the case that $s_k = 0$ for some $k$ such that $p_k > 0$. In addition, for arbitrary $k$ such that $p_k > 0$, there will be an $s$ with $P_s > 0$ such that $s_k = 0$. Therefore, for the first condition to be satisfied, it would be necessary for $\Lambda_{1k}$ to be equal to zero for all $k$. This is clearly not possible, because $\Lambda$ is a unitary matrix.

On the other hand, if $D + 1 = M$, then the only $s$ giving $P_s > 0$ is that with $s_k = 1$ for $p_k > 0$, and $s_k = 0$ for $p_k = 0$. Therefore, it is possible for the first condition to be satisfied, by choosing a $\Lambda$ with $\Lambda_{1k} = 0$ for all $k$ such that $p_k = 0$. However, the case with $D = M - 1$ is unimportant, because it is not possible to obtain an improvement in $R$. Hence, we see that, in any case where it is possible to obtain an improvement, it is not possible to obtain equality in equation (37).

On the other hand, we can determine a scheme that gives $S_k \propto \Lambda_{1k}^*$. This condition can be satisfied using the interferometer with matrix elements

\begin{align*}
\Lambda_{11} &= -\epsilon, \\
\Lambda_{21} &= \sqrt{1 - \epsilon^2}, \\
\Lambda_{1i} &= \sqrt{(1 - \epsilon^2)/(N - 1)}, \\
\Lambda_{2i} &= \epsilon/\sqrt{N - 1}
\end{align*}

for $i > 1$ (the values of $\Lambda_{ij}$ for $i > 2$ do not enter into the analysis). Here $\epsilon$ is a small number, and we ignore terms of order $\epsilon$ or higher. Now let $p_i = p_{\text{max}}$, and consider the measurement
record where zero photons are detected in modes 3 to \( N \), so the number of photons detected in mode 2 is \( D \).

This scheme is the same as in [14], except input modes 1 and 2 have been swapped. Expressing the scheme in this form allows us to determine a simple realization using beam splitters. This scheme may be performed using the chain of beam splitters shown in figure 3. The first \( N - 2 \) beam splitters (those on the right) result in an output beam with equal contributions from \( N - 1 \) of the inputs. This equal combination is achieved by decreasing the reflectivities from \( \frac{1}{2} \) for the first beam splitter to \( \frac{1}{N-1} \) for beam splitter \( N - 2 \). The last beam splitter has the low reflectivity \( \epsilon^2 \). With appropriate phase shifts, these beam splitters give the overall interferometer described by \( \Lambda \) in equation (38).

To determine \( c_{n_1} \), note first that \( \Lambda_{21} \gg \Lambda_{2i} \) for \( i > 1 \), so we may ignore those terms in the sum for \( S_{k,n} \) where \( \Lambda_{21} \) does not appear. Each term has magnitude \( \Lambda_{12}^{n_1} \Lambda_{21} \Lambda_{22}^{D-1} \), and there are \( D(D+n_1-1)! \) such terms. Therefore, provided \( s_1 = 1 \),

\[
S_{k,n} \approx D(D+n_1-1)!\Lambda_{12}^{n_1} \Lambda_{21} \Lambda_{22}^{D-1}. \tag{39}
\]

If \( s_1 = 0 \), then \( S_{k,n} \) is of order \( \epsilon \) times this value.

Recall that \( S_k = S_{s',n'} \), where \( s' \) is equal to \( s \), except for \( s'_k = 0 \) and \( n'_1 = 0 \). (We do not consider \( k \) where \( s'_k = 0 \).) The result for \( k > 1 \) may be obtained by replacing \( n_1 \) with 0 and \( D \) with \( \Sigma_s - 1 \) in equation (39), giving

\[
S_k \approx (\Sigma_s - 1)(\Sigma_s - 2)!\Lambda_{21} \Lambda_{22}^{\Sigma_s-2}. \tag{40}
\]

For \( k = 1 \), we simply obtain \( S_1 \) of order \( \epsilon \) times this value. Similarly, \( \Lambda_{1k} \) is constant for \( k > 1 \), and of order \( \epsilon \) for \( k = 1 \). Thus this scheme gives \( S_k \propto \Lambda_{1k}^* \), as claimed above.

---

\[^{7}\text{We use } \Lambda_{12} \text{ and } \Lambda_{22} \text{ to indicate the values of } \Lambda_{1i} \text{ and } \Lambda_{2i} \text{ for } i > 1.\]
In order to determine \( c_{n_1} \), note that there are \( \binom{N-1}{D+n_1-1} \) different combinations of inputs such that \( \Sigma = D + n_1 \) and \( s_1 = 1 \). Combining this expression with equation (39), we have

\[
c_{n_1} \approx \frac{K'}{n_1!} p_{\max}^{D+n_1} (1 - p_{\max})^{N-D-n_1} \frac{(N-1)!D^2(D+n_1-1)!}{(N-D-n_1)!} \Lambda_{12}^{2n_1} \Lambda_{21}^2 \Lambda_{22}^{2D-2}.
\]

We have combined those factors that do not depend on \( n_1 \) into a new constant \( K'' \), and used \( \Lambda_{12} \approx 1/\sqrt{N - 1} \). Using equation (41) gives

\[
R_{\text{out}} \approx R_{\text{in}} \frac{D(N-D)}{N-1}.
\]

The maximum improvement in \( R \) is obtained for \( D = \lceil N/2 \rceil \), where \( R_{\text{out}} \approx R_{\text{in}} [N^2/4]/(N-1) \). The multiplicative factor \( [N^2/4]/(N-1) \) is larger than 1 for all \( N \geq 4 \). Thus we find that, provided there are at least 4 modes, we may obtain an improvement in \( R_{\text{out}} \). For \( p_{\max} \ll 1 \), \( c_1 \approx p_{\max} [N^2/4]/(N-1) \). For large \( N \), the probability of a single photon increases approximately as \( N/4 \), but does not achieve the upper bound of \( N \).

Although we find an improvement in the measure \( R \), the two-photon contribution is not negligible. Using the measure \( G \), we find

\[
G_{\text{out}} = \frac{c_2/c_1}{c_1/c_0} \approx \frac{(D+1)(N-D-1)}{2D(N-D)}.
\]

For \( D = \lceil N/2 \rceil \), this measure is close to 1/2, so the two-photon component is similar to that for a Poisson distribution. By taking \( D = N - 2 \), it is possible to obtain an improvement in \( R \) of about a factor of 2, with a value of \( G_{\text{out}} \) about half that for a Poisson distribution. However, this two-photon contribution is still much greater than for good single-photon sources [11].

The multiphoton contributions are especially important for larger \( p_{\max} \). Although the improvement in \( R \) is independent of \( p_{\max} \), the multiphoton component means that improvements in \( c_1 \) are obtained only for values of \( p_{\max} \) below 1/2. That is, this method can only be used to obtain improvements in the probability of a single photon up to 1/2, but not to make the probability of a single photon arbitrarily close to 1.

This scheme also performs poorly when evaluated via the measure \( \Pi \). It is more difficult to evaluate this scheme using this measure, and we do not have a simple solution for \( \Pi \). However, numerically we find that \( \Pi_{\text{out}} \) is close to 1 for \( D = \lceil N/2 \rceil \), again indicating that the output state is close to Poissonian. For \( D = N - 2 \), \( \Pi_{\text{out}} \) is closer to \( \Pi_{\text{in}} (=1-p_{\max}) \), but for no \( D \) is a value of \( \Pi_{\text{out}} \) less than \( \Pi_{\text{in}} \) obtained.

Nevertheless, this scheme does give an improvement in \( R \). It can be expected that this improvement is close to the maximum possible, because this scheme satisfies \( S_k \propto \Lambda_{1k}^* \). However, note that it satisfies the other condition for optimality, i.e. that \( \Lambda_{1k} \neq 0 \) implies \( s_k = 1 \), fairly poorly. As shown above, this condition cannot be satisfied completely (unless there is no improvement in \( R \)), and there does not appear to be any method of satisfying it better than the method we have described above. Extensive numerical searches have failed to find any scheme that gives a better improvement in \( R \) than the above scheme, strongly indicating that it is optimal for increasing \( R \).
7. Experimental limitations

In practice, there will be a number of limitations to using this method for improving the probability of a single photon. The main ones are the following:

1. The photons produced by different sources are not completely indistinguishable.
2. Real photodetectors do not give perfect photon counting measurements. Most photodetectors can only distinguish between the vacuum state and a state with one or more photons.
3. Real photodetectors have limited efficiency and dark counts.
4. The desired combination of detection results will occur with low probability.
5. Real sources have a finite multiphoton component.

The first point is probably the most important for experimental realizations, because it is very difficult to produce photons from different sources that are indistinguishable. The above scheme will not work if the photons are distinguishable, and it is not possible to obtain an improvement via linear optics and photodetection for distinguishable photons.

This result may be shown in the following way. For the combination of input photons \( s \), the probability of obtaining the combination of detections \( n \) is

\[
P(s, n) = \frac{1}{\prod_{j=1}^{N} n_j!} S_{s,n}^{(2)},
\]

(44)

where

\[
S_{s,n}^{(2)} = \sum_{\sigma \in Y_s} |(\Lambda_1, \sigma_1 \cdots \Lambda_1, \sigma_n) \cdots (\Lambda_N, \sigma_{\Sigma_1-N \cdots \Sigma_1-1} \cdots \Lambda_N, \sigma_{\Sigma_1})|^2.
\]

(45)

That is, the probability is the sum over the probabilities for obtaining the output photons from the different input modes. The total probability for \( n_1 \) photons in output 1 is therefore

\[
c_{n_1} = \frac{K}{n_1!} \sum_{s; \Sigma_1 = \Sigma_a} P_s S_{s,n}^{(2)}.
\]

(46)

To consider the improvement possible, consider the expression for \( c_1 \). In an analogous way to the case for indistinguishable inputs, we may write

\[
c_1 = K \sum_{s; \Sigma_1 = D+1} P_s \sum_{k; \Sigma_2 = 1} |\Lambda_{1k}|^2 S_{s^k,n^0}^{(2)}.
\]

(47)

Exchanging the order of the summations gives

\[
c_1 = K \sum_k |\Lambda_{1k}|^2 \frac{P_k}{1 - P_k} \sum_{s; \Sigma_2 = D+1} P_s S_{s^k,n^0}^{(2)}
\]

\[\leq K \frac{p_{\text{max}}}{1 - p_{\text{max}}} \sum_k |\Lambda_{1k}|^2 \sum_{s; \Sigma_1 = D} P_s S_{s,n^0}^{(2)}
\]

\[= R_{\text{im}} c_0.
\]

(48)
Therefore, regardless of the interferometer and combination of detections, the conditional probability for a single photon in the output state cannot be larger than $p_{\text{max}}$.

There is a subtlety in this result because it is theoretically possible to determine which input the photon came from. This means that output port 1 contains $N$ modes, one for each of the input ports. The number of photons in mode 1 is the sum over the photon number in each of these modes. If the available photodetectors are unable to distinguish between the modes, then photodetection would yield the same result as for the same number of photons in a single mode.

Given a detector with a limited capability to distinguish the modes, the outputs from $N$ single-photon sources could be manipulated to be observed as the same mode by this detector. Then, if $p_i = p_{\text{max}}$ for each of these modes, the probability for a total photon number of 1 is $Np_{\text{max}}(1 - p_{\text{max}})^{N-1}$. Hence, for small $p_{\text{max}}$, the probability for a single photon is improved by a factor of $N$. However, this method does not increase the probability for a single photon in a single mode, which is necessary for quantum information applications.

For $N - 2$ of the detectors, point 2 will not be a problem. The reason for this is that we are conditioning on detection of zero photons at these detectors. It is only the detector on mode 2 that is required to perform a photon counting measurement. Even for this detector, it is not necessary to determine the exact photon number. That is because the probability of a single photon will be increased for any number of photons from 2 to $N - 2$. Therefore, if the detector can register that the photon number is in this range, rather than the exact photon number, it will be sufficient to produce an improvement in the single-photon probability.

Alternatively, an improvement can also be achieved using detection that simply verifies that there is more than one photon. For example, the visible light photon counter [17] can do this task with high efficiency. Even though the possibilities of $N - 1$ or $N$ photons have not been eliminated, they have lower probability, and will not contribute significantly to the conditional photon probabilities in the output mode.

Limited efficiency is not a severe problem for small $p_{\text{max}}$, because the probability for the vacuum is relatively large. Dark counts will not be a problem for the first $N - 2$ detectors, because we are conditioning on vacuum detection at these detectors. Dark counts will merely slightly reduce the probability of obtaining the desired combination of detection results. On the other hand, dark counts will be a problem for detector 2, as we are conditioning on detection of more than one photon at this detector.

Point 4 will always be a problem, because the above scheme is only effective for small $\epsilon$. The probability for this combination of detection results becomes very small in the limit of small $\epsilon$. If larger values of $\epsilon$ are used, then the final probability for a single photon becomes smaller. Thus there is a trade-off involved. Point 5 is more complicated, and it is not clear how important this problem is without performing direct calculations.

To estimate the relative importance of each of the above problems, we have calculated the final conditional probability for a single photon taking each of these points into account. In figure 4 we have plotted the conditional probability for a single photon versus the probability for obtaining the desired combination of detection results. These curves are parametrized by $\epsilon$; that is, both probabilities were calculated for a range of values of $\epsilon$. In general, as $\epsilon$ is decreased, the probability for obtaining the desired detection results decreases, and the final conditional probability for a single photon increases. The particular example we have shown is of a four-mode interferometer where $p_{\text{max}} = 0.2$ for the inputs.

To perform these calculations, the density matrix was left unnormalized. The trace of the density matrix at the end of the calculation then gives the probability for that combination of
Figure 4. The final probability for a single photon versus the probability for obtaining the appropriate detection results for a four-mode interferometer and $p_{\text{max}} = 0.2$. The black line is the ideal case, the light blue line is that for the case where $D = 3$ and $D = 4$ are also allowed, and the green line is that if, in addition, the photodetectors have 90% efficiency. The dark blue line is that including all of these experimental limitations, plus 0.1% chance of dark counts at detector 2, and the red line is that taking into account these experimental limitations, as well as allowing 0.1% probability for two photons in the inputs (without dark counts). The dashed black line is for photons that are 90% indistinguishable, with ideal photodetectors and without multiphoton inputs.

detection results. For the case of partially distinguishable photons, the single-photon states were taken to be a coherent superposition of a single photon in the correct input mode (that interferes with the other photons) and a single photon in an additional mode (that does not interfere). For these calculations the photons were assumed to be 90% indistinguishable.

For the calculations with imperfect detectors the input photons were assumed to be indistinguishable. For those cases where the detector at mode 2 does not register the exact photon number, the density matrices for $D = 2$, 3 and 4 were simply added. To take account of finite efficiency detectors and dark counts, the density matrices for the other detection results were multiplied by constant factors, and added to the density matrix for the desired detection result. It was assumed that the inefficient detectors register single-, two- and three-photon states as vacuum with probabilities of 10, 1 and 0.1%, respectively. For the detector on mode 2 it was assumed that a single-photon state is registered as two or more photons with 0.1% probability, and the vacuum state is registered as two or more photons with 0.0001% probability (due to the lower probability of simultaneous dark counts). The appropriate equations to use for the case with multiphoton inputs are derived in section 10.

For perfect sources and detectors, the final probability for a single photon is above the initial probability of 20% when the probability for obtaining the detection results is below about 0.7%. Thus, in order to obtain the desired detection results, the experiment needs to be repeated.
roughly 200 times, which is not unreasonable. If the photons are only 90% indistinguishable, then it is still possible to obtain an improvement in the probability for a single photon. The final probability for a single photon is only reduced by about 1%, so the results are only moderately sensitive to photon distinguishability.

If we consider a final detector that cannot distinguish between two, three or four photons, the results are almost identical to those for the ideal case, so this problem is relatively trivial. Even photodetectors with finite efficiency do not greatly affect the results. If the first two photodetectors have 90% efficiency, the final probability for a single photon is reduced by about 0.3%. The greatest problems are dark counts at detector 2, and multiphoton components at the inputs. If a dark count rate of 0.1% is allowed, then the maximum single-photon probability is reduced below 0.23%. For small values of $\epsilon$ the single-photon probability drops significantly, rather than approaching the maximum value. The results are similar if a two-photon probability of 0.1% is allowed in the inputs (while the vacuum probability is decreased by 0.1%). The single-photon probability again drops for small values of $\epsilon$, and the maximum single-photon probability is less than 0.22%. For two-photon probabilities of 0.4% or more it is not possible to obtain any increase in the single-photon probability. Thus we see that the final single-photon probability is most sensitive to two-photon components in the input and dark counts at detector 2. The results are also moderately sensitive to the distinguishability of the photons.

8. No-go theorems

In this section, we prove a number of no-go theorems for post-processing via linear optics and photodetection. Note that one limitation of the scheme given in section 6 is that it only gives improvements in $R$ for four or more modes. In fact, it is impossible to obtain improvements for fewer than four modes [14]. This result may be shown in the following way. First consider the case $D = 0$. Then there is only one term in the sum for $c_0$ and $c_0 = K'P_0$. The expression for $c_1$ becomes

$$c_1 = K' \sum_{k=1}^{N} \frac{p_k}{1 - p_k} P_0 |\Lambda_{1k}|^2$$

$$\leq K' R_{in} \sum_{k=1}^{N} P_0 |\Lambda_{1k}|^2$$

$$= K' R_{in} P_0$$

$$= c_0 R_{in}.$$

Thus we have shown that $R_{out} \leq R_{in}$, so $c_1 \leq p_{max}$. Hence, there can be no improvement in the photon statistics if zero photons are detected.

This result can also be shown in a more intuitive way as follows. First note that an arbitrary $U(N)$ interferometer can be obtained using a line of $N - 1$ beam splitters followed by a $U(N - 1)$ interferometer (figure 5). This is an immediate consequence of the algorithmic construction of arbitrary $U(N)$ interferometers from beam splitters [18]. If the $N - 1$ modes upon which the
A $U(N)$ interferometer can be represented by a $U(N - 1)$ interferometer preceded by $N - 1$ beam splitters. This figure shows the example for $N = 5$.

$U(N - 1)$ interferometer acts are those that are measured, then we may omit the $U(N - 1)$ interferometer entirely (because detecting zero photons at the output of this interferometer is identical to detecting zero photons at the input). Thus, this case may be reduced to the case of a line of beam splitters where zero photons are detected at each stage.

The case of a line of beam splitters with vacuum detection may be deduced from the case for a single beam splitter. As was shown above, with a single beam splitter there is no improvement in the ratio between the probabilities for one and zero photons. It is easily seen that the same result holds if there are non-zero probabilities for photon numbers larger than 1 in the inputs (corresponding to photon numbers larger than one in the output).

Thus, if we have a line of beam splitters, the ratio between the probabilities for one and zero photons in the output cannot be increased above the maximum of that for the inputs. This result implies that the probability of one photon in the output can never exceed $p_{\text{max}}$. This result holds for a line of beam splitters, and therefore for an arbitrary $U(N)$ interferometer.

We can also obtain a similar result for the case $D = 1$, provided all the input $p_i$ are equal. If the single photon is detected in mode $m$, then

$$c_0 = K' \sum_k \frac{p_{\text{max}}}{1 - p_{\text{max}}} P_0 |\Lambda_{mk}|^2 = K' R_{\text{in}} P_0.$$  \hspace{1cm} (50)
The value of \( c_1 \) is given by

\[
\begin{align*}
    c_1 &= \frac{1}{2} K' \sum_{k} \sum_{l,l \neq k} R_{in}^2 P_0 |\Lambda_{1l} \Lambda_{mk} + \Lambda_{1k} \Lambda_{ml}|^2 \\
    &\leq \frac{1}{2} K' R_{in}^2 P_0 \sum_{k,l} |\Lambda_{1l} \Lambda_{mk} + \Lambda_{1k} \Lambda_{ml}|^2 \\
    &= \frac{1}{2} K' R_{in}^2 P_0 \left[ \sum_{l} |\Lambda_{1l}|^2 \sum_{k} |\Lambda_{mk}|^2 + \sum_{l} |\Lambda_{1l}|^2 \sum_{l} |\Lambda_{ml}|^2 \right] \\
    &\quad + \sum_{l} \Lambda_{1l} \Lambda_{mk}^* \sum_{k} \Lambda_{mk}^* \Lambda_{1k} + \sum_{l} \Lambda_{1l} \Lambda_{ml}^* \sum_{k} \Lambda_{ml}^* \Lambda_{1k} \\
    &= K' R_{in}^2 P_0. 
\end{align*}
\]

In the last line we have used the fact that \( \Lambda_{1k} \) and \( \Lambda_{mk} \) are orthonormal. Thus we again find \( R_{out} \leq R_{in} \), so \( c_1 \leq p_{\max} \).

These results can be used for an alternative proof that no improvement is possible for the case of a single beam splitter. We have shown that detecting zero photons does not give an improvement, and if one photon is detected, then we must have \( M - D = 1 \) or \( 0 \), so there again can be no improvement.

We can also eliminate the case of a three-mode interferometer, although the reasoning is not as straightforward. First note that an input with probability \( p_i \) of a photon can be obtained by randomly selecting between a source with efficiency \( p_{\max} \) and the vacuum. That is, with probability \( q = p_i / p_{\max} \) we use the source with efficiency \( p_{\max} \), and with probability \( 1 - q \) we use the vacuum state. If we discard the information about which source was used, this is obviously equivalent to a source with efficiency \( p_i \). Hence, the value of \( c_1 \) for the source with efficiency \( p_i \) is the weighted average of the values of \( c_1 \) for the cases where the efficiency is \( p_{\max} \) and zero. Thus, the maximum \( c_1 \) must be obtained with all of the non-zero \( p_i \) equal to \( p_{\max} \).

Therefore, in considering the three-mode interferometer, we can let all the non-zero \( p_i \) be equal to \( p_{\max} \). If all the \( p_i \) are non-zero, then we can use the result showing that there is no improvement with one photon detected and all \( p_i \) equal. The other alternatives for detection have \( D = 0 \) or \( M - D \leq 1 \), so there can be no improvements in these cases either. If one or more of the \( p_i \) are zero, then the only detection alternatives are with \( D = 0 \) or \( M - D \leq 1 \). Thus, we have shown that there can never be an increase in the probability of a single photon if less than four modes are used.

9. Unsolved problems

Although it seems that we were able to answer most of the relevant questions concerning the possibility of improving the efficiency of single-photon sources, there are, in fact, still a number
of open questions we would now like to address. The two main unsolved problems for this post-processing are:

1. Is it possible to increase the probability for a single photon, regardless of the value of $p_{\text{max}}$?
2. Can the single-photon probability be increased without adding a multiphoton component?

At this time the indications are that the answer to both these questions is no. We have performed numerical searches for interferometers that give improvements for $p \geq 1/2$. These searches have been unsuccessful, indicating that it is not possible to obtain an improvement for $p_{\text{max}} \geq 1/2$. We have not been able to prove this assertion; however, we can show that there are various implications if there is any value of $p_{\text{max}}$ such that it is impossible to obtain an improvement.

First note that it is sufficient to use $p_i = p_{\text{max}}$ in the input modes. As discussed above, for a given interferometer the maximum improvement will always be obtained with all of the non-zero $p_i$ equal to $p_{\text{max}}$. It is possible to obtain a vacuum state from inputs with efficiency $p_{\text{max}}$, simply by using a beam splitter and conditioning on detection of two photons at one of the outputs. Therefore, if there is an interferometer that achieves a certain result using inputs with $p_i = 0$ or $p_{\text{max}}$, there will always be another (expanded) interferometer that achieves the same result with $p_i = p_{\text{max}}$.

Using this simplification, the expression for $c_{n_1}$ simplifies to $c_{n_1} = K'' d_{n_1} R_{\text{in}}^{n_1}/n_1!$, where

$$d_{n_1} = \sum_{s; \Sigma_s = \Sigma_n} |S_{s,n}|^2. \quad (52)$$

The values of $d_{n_1}$ are independent of $p_{\text{max}}$, and depend only on the interferometer and combination of detection results. There is an improvement in the probability of a single photon if

$$d_1 > d_0 + \sum_{n_1=2}^{N} d_{n_1} R_{\text{in}}^{n_1}/n_1!. \quad (53)$$

Let $p_0$ be a value of $p_{\text{max}}$ such that there is an improvement in the probability of a single photon, and let the corresponding value of $R_{\text{in}}$ be $R_0$. Then there exists an interferometer and combination of detection results such that

$$d_1 > d_0 + \sum_{n_1=2}^{N} d_{n_1} R_0^{n_1}/n_1!. \quad (54)$$

Since each of the $d_{n_1}$ are positive, and the right-hand side is increasing as a function of $R_{\text{in}}$, we find that (53) is satisfied for all $0 < p_{\text{max}} \leq p_0$. Thus we find that, for any value of $p_{\text{max}}$ such that there is an improvement, there is an improvement for all smaller values of $p_{\text{max}}$. In turn this result implies that, if there is no improvement for $p_{\text{max}} = p_0$, then there can be no improvement for larger values of $p_{\text{max}}$.

It is also possible to show that, if it were possible to obtain an improvement with no multiphoton contribution, there would be no value of $p_{\text{max}}$ for which we cannot obtain an improvement. However, this requires a more detailed analysis of the interferometer and detection results.

---

*Provided $p_{\text{max}}$ is non-zero and less than 1. These restrictions on $p_{\text{max}}$ are implied in the following text.*

*New Journal of Physics* 6 (2004) 93 (http://www.njp.org/)
improvement. To show this, note that zero multiphoton contribution implies that \( d_{n_1} = 0 \) for \( n_1 \geq 2 \). Therefore, if this improvement is possible for \( p_{\text{max}} = p_0 \), then equation (54) becomes simply

\[
d_1 > d_0.
\]

(55)

Similarly, the condition to obtain an improvement for any other value of \( p_{\text{max}} \) is simply \( d_1 > d_0 \), which is automatically satisfied. In addition, because \( d_{n_1} = 0 \) for \( n_1 \geq 2 \), \( c_{n_1} = 0 \) for \( n_1 \geq 2 \), for any \( p_{\text{max}} \).

Thus, if it is possible to obtain an improvement for some value of \( p_{\text{max}} \) while maintaining zero multiphoton contribution, then it will be possible to obtain an improvement for all values of \( p_{\text{max}} \). In addition, the relative improvement in \( R \) is independent of \( p_{\text{max}} \). To see this, note that

\[
\frac{R_{\text{out}}}{R_{\text{in}}} = \frac{c_1/c_0}{R_{\text{in}}} = \frac{d_1}{d_0}.
\]

(56)

A further implication is that it would be possible to obtain an output state that is arbitrarily close to the pure single-photon state. As the output from the interferometer has no multiphoton contribution, outputs from \( N \) of these interferometers may be used as the input to another, thus increasing \( R \) by a factor of \( (d_1/d_0)^2 \). Further iterations may be used to increase \( R \) by a factor of \( d_1/d_0 \) to any arbitrary power, thus obtaining a final probability for a single photon that is arbitrarily close to 1.

At this stage there is no known scheme that can give an improvement in the probability for a single photon while maintaining zero multiphoton contribution. As shown above, if this were possible for any value of \( p_{\text{max}} \), then it would be possible for all values of \( p_{\text{max}} \). As increasing the single-photon probability without maintaining zero multiphoton component is a less difficult problem, if it were possible to obtain an improvement while maintaining zero multiphoton component, for all values of \( p_{\text{max}} \) there would be an enormous range of schemes that give improvements without the constraint on the multiphoton component. Such a wide range of schemes would be relatively easy to find numerically; the fact that numerical searches have failed to find any scheme that gives an improvement in the single-photon probability for \( p_{\text{max}} \geq 1/2 \) therefore implies that it is extremely unlikely that there is any scheme that gives an improvement while maintaining zero multiphoton contribution. Nevertheless, these numerical results are not sufficient to rule out this possibility.

10. Multiphoton inputs

The majority of this study is based upon inputs from photon sources that have zero multiphoton contribution. It is also possible to derive results for inputs with non-zero probabilities for two or more photons, but this case is more difficult. The simplest case is for a beam splitter with multiphoton inputs. Let us denote the probability for \( m \) photons in input mode \( i \) by \( p_{im} \). Then the input state may be written as

\[
\rho_{\text{in}}^{(N)} = \sum_{k,l} p_{1k} p_{2l} |kl\rangle \langle kl| = \sum_{k,l} \frac{p_{1k} p_{2l}}{k!l!} (a_1^k a_2^l)^\dagger |00\rangle \langle 00| (a_1^k a_2^l)^\dagger.
\]

(57)
The beam splitter transformation (3) gives

$$\hat{\rho}^{(N)}_{\text{trans}} = \sum_{k,l} \frac{p_{1k} p_{2l}}{k! l!} (\Lambda_{11} a_1^\dagger + \Lambda_{21} a_2^\dagger)^k (\Lambda_{12} a_1^\dagger + \Lambda_{22}^* a_2^\dagger)^l |00\rangle \langle 00| (\Lambda_{11}^* a_1 + \Lambda_{21}^* a_2)^k (\Lambda_{12}^* a_1 + \Lambda_{22}^* a_2)^l.$$

(58)

Expanding in a series and conditioning upon detection of $D$ photons in mode 2 gives

$$\hat{\rho}^{(N)}_{\text{out}} = K \sum_{k,l} p_{1k} p_{2l} k! l! D! (k+l-D)!$$

\begin{align*}
&\times \sum_{m=\max(D-k,0)}^{\min(D,l)} \frac{\Lambda_{11}^{l-D+m} \Lambda_{12}^{D-m} \Lambda_{21}^{l-m} \Lambda_{22}^{m}}{(D-m)! m! (k-D+m)! (l-m)!} |k+l-D\rangle \langle k+l-D|
&= KD! \left| \frac{\Lambda_{12}}{\Lambda_{22}} \right|^{2D} \sum_{k,l} p_{1k} p_{2l} k! l! D! (k+l-D)! |\Lambda_{22}|^{-2k} |\Lambda_{12}|^{-2l}
&\times \sum_{m=\max(D-k,0)}^{\min(D,l)} \frac{(-1)^m |\Lambda_{22}/\Lambda_{12}|^{2m}}{(D-m)! m! (k-D+m)! (l-m)!} |k+l-D\rangle \langle k+l-D|,
\end{align*}

(59)

where $K$ is a normalization constant. If $K$ is omitted, the trace gives the probability for this detection result. This is the expression used to calculate the numerical results in section 7.

The transformation for a multimode interferometer is also straightforward to determine. For this case the input state may be written as

$$\hat{\rho}^{(N)}_{\text{in}} = \bigotimes_{i=1}^{N} \left[ \sum_{m} p_{\text{in}} |m\rangle \langle m| \right]$$

$$= \sum_{s} P_s \prod_{i} s_i \left( \prod_{i} (a_i^\dagger)^{s_i} |0\rangle \langle 0| \prod_{i} (a_i)^{s_i} \right),$$

(60)

where $P_s = \prod_{i} (p_{i,s_i})^{s_i}$. From this point on we use the abbreviated notation $P'_s = P_s / \prod_{i} s_i!$. Note that $s_i$ may take any integer value $\geq 0$, rather than simply 0 and 1. Now applying the interferometer transformation $a^\dagger \mapsto A \Lambda a^\dagger$ gives

$$\hat{\rho}^{(N)}_{\text{trans}} = \sum_{s} P'_s \left[ \prod_{i} \left( \sum_{k} \Lambda_{ki} a_k^\dagger \right)^{s_i} |0\rangle \langle 0| \prod_{i} \left( \sum_{k} \Lambda_{ki}^* a_k \right)^{s_i} \right].$$

(61)

After detection on modes 2 to $N$, the final state obtained is

$$\hat{\rho}^{(N)}_{\text{out}} = \sum_{n_1=0}^{N} c_{n_1} |n_1\rangle \langle n_1|,$$

(62)
with

\[ c_{n_1} = \frac{K'}{n_1!} \sum_{s : \Sigma s = \Sigma n} P_s' |S_{s,n}|^2, \]

where \( S_{s,n} = \text{per}(A[n, s]) \) and \( K' \) is a normalization constant. This result is very similar to the case for inputs with no multiphoton contribution. The only difference is that values of \( s_i \) larger than 1 are now permitted, and there is the additional dividing factor of \( \prod s_i! \). We do not use this result in this paper, but it is a useful general form.

Only one of the no-go theorems still applies for the case where multiphoton inputs are allowed. For the case of a beam splitter, the multiphoton contributions in the input give multiphoton contributions in the output. Therefore, if zero photons are detected, there can be no improvement in the ratio between the probability for one photon to the probability for zero photons. In the multimode case where zero photons are detected, it is possible to decompose the interferometer as in figure 5, then omit the \( U(N - 1) \) interferometer. At each beam splitter in the chain the ratio between the probabilities for one and zero photons is not increased, so the final ratio cannot be above the maximum for the inputs. Nevertheless, the result in this case is not as strong as in the case where the inputs have no multiphoton contribution. Proving that the ratio between the probabilities for one and zero photons has not increased does not prove that the absolute probability for a single photon has not increased. The problem is that it is possible, in principle, for the multiphoton component to be decreased sufficiently that the probability for a single photon is increased.

If the inputs have multiphoton contributions, it is not impossible to obtain a perfect single-photon output state. In particular, consider a state \( \rho_0 = \sum q_i |i\rangle \langle i| \), such that \( q_D = 0, q_{D+1} > 0 \), and \( q_i = 0 \) for \( i > D + 1 \). If this state is combined with the vacuum at a beam splitter, and \( D \) photons are detected, then the output state will be a pure single-photon state. This result also demonstrates that there is no initial value for the single-photon probability that cannot be improved upon.

The drawback to these results is that states such as \( \rho_0 \) would be very difficult to produce in the laboratory. It is probable that there is some measure of the quality of the state that cannot be improved upon using linear optics and photodetection. However, it is difficult to determine what measure this would be. For example, it is clear that \( \Pi \) can be improved, because \( \rho_0 \) can be heavily super-Poissonian. The same considerations rule out other simple possibilities such as the entropy. Finding a measure that is non-increasing under linear optics and photodetection is a promising direction for future research.

11. Conclusions

Triggered single-photon sources produce an incoherent mixture of zero and one photons, with much smaller probabilities for two or more photons. Provided the multiphoton contributions in the inputs may be ignored, we have shown that it is possible to significantly increase the probability for a single photon by using post-processing via linear optics and photodetection. This method has the drawback that it produces a significant multiphoton component that is comparable with that for the Poisson distribution.

We have shown that there are severe limitations on what post-processing can be performed. In particular, there is an upper limit on the increase in the probability for a single photon. This
upper limit cannot be achieved, but the method we have found for increasing the probability for a single photon gives the same scaling with the number of modes. It is probable that this method achieves the maximum increase in the probability for a single photon. This result is indicated numerically, but has not been proven.

In addition, it is impossible to obtain an increase in the probability for a single photon using a single beam splitter. Alternatively, if zero photons, or one less than the maximum input photon number are detected, it is again impossible to obtain an improvement in the probability for a single photon. In the restricted case that all the inputs are identical, it is impossible to obtain an improvement in the probability for a single photon if one photon is detected. These no-go theorems are sufficient to prove that at least a four-mode interferometer is required to obtain an improvement.

Another important no-go theorem is that it is impossible to obtain a perfect single-photon output with imperfect inputs. It must be emphasized that this no-go theorem is only for mixed states with no multiphoton components. If we relax these constraints, by considering pure superposition states of zero and one photon, then it is possible to obtain a pure single-photon output. Alternatively, some multiphoton states can be processed to yield a perfect single-photon output. However, it must be emphasized that it is not probable that pure input states, or the appropriate multiphoton states, can be produced experimentally.

There are also a number of important unsolved problems. It is currently unknown whether there is an upper limit (less than 1) to the initial probability for a single photon such that it is possible to obtain an improvement. It is also unknown if it is possible to obtain an improvement in the probability for a single photon while maintaining zero multiphoton contribution. If such a scheme were possible it would be very significant, because it would be possible to obtain a state arbitrarily close to a single photon state from arbitrarily poor input states.

Acknowledgments

RL acknowledges valuable discussions with E Knill in the early stages of this work. This work was funded in parts by the UK Engineering and Physical Sciences Research Council. SS enjoys a Feodor-Lynen fellowship of the Alexander von Humboldt foundation. RL thanks the NSA, NSERC and MITACS for support. BCS and RL thank the CIAR program on Quantum Information Processing. This research has also been supported by an Australian Department of Education Science and Training Innovation Access Program Grant to support collaboration in the European Fifth Framework project QUPRODIS and by Alberta’s informatics Circle of Research Excellence (iCORE).

References

[1] Bennett C H and Brassard G 1984 Proc. IEEE Int. Conf. on Computers, Systems and Signal Processing, Bangalore, India (New York: IEEE)
Bennett C H, Brassard G and Ekert A K 1992 Sci. Am. 267 50
[2] Knill E, Laflamme R and Milburn G J 2001 Nature 409 46
Scheel S, Nemoto K, Munro W J and Knight P L 2003 Phys. Rev. A 68 032310
[3] Brassard G, Lütkenhaus N, Mor T and Sanders B C 2000 Phys. Rev. Lett. 85 1330
Brassard G and Crépeau C 1996 SIGACT News 27 13
[4] Hong C K and Mandel L 1986 Phys. Rev. Lett. 56 58
Sergienko A V, Atatüre M, Walton Z, Jaeger G, Saleh B E A and Teich M C 1999 Phys. Rev. A 60 R2622
Jennewein T, Simon C, Weis H, Weinfurter H and Zeilinger A 2000 Phys. Rev. Lett. 84 4729
[5] Waks E, Diamanti E and Yamamoto Y 2003 Preprint quant-ph/0308055
[6] DeMartini F, DiGiuseppe G and Marocco M 1996 Phys. Rev. Lett. 76 900
Brunel C, Lounis B, Tamarat P and Orrit M 1999 Phys. Rev. Lett. 83 2722
Lounis B and Moerner W E 2000 Nature 407 491
[7] Kim J, Benson O, Kan H and Yamamoto Y 1999 Nature 397 500
[8] Kurtz C, Mayer S, Zarda P and Weinfurter H 2000 Phys. Rev. Lett. 85 290
Brouvi R, Beveratos A, Poizat J P and Grangier P 2000 Opt. Lett. 25 1294
Beveratos A, Kühn S, Brouvi R, Gacoin T, Poizat J-P and Grangier P 2002 Eur. Phys. J. D 18 191
[9] Kuhn A, Henrich M, Bondo T and Rempe G 1999 Appl. Phys. B 69 373
Varcoe B T H, Brattke S, Weidinger M and Walther H 2000 Nature 403 743
Kuhn A, Henrich M and Rempe G 2002 Phys. Rev. Lett. 89 067901
[10] Michler P, Kiraz A, Becher C, Schoenfeld W V, Petroff P M, Zhang L, Hu E and Imamoglu A 2000 Science 290 2282
Santori C, Pelton M, Solomon G, Dale Y and Yamamoto Y 2001 Phys. Rev. Lett. 86 1502
Zwiller V, Blom H, Jonsson P, Panev N, Jeppesen S, Tsegaye T, Goobar E, Pistol M E, Samuelson L and Björk G 2001 Appl. Phys. Lett. 78 2476
Moreau E, Robert I, Gerad J M, Abram I, Manin I and Thierry-Mieg V 2001 Appl. Phys. Lett. 79 2865
Yuan Z, Kardynal B E, Stevenson G R, Shields A J, Lobo C J, Cooper K, Beattie N S, Ritchie D A and Pepper M 2002 Science 295 102
Gérard J-M and Gayral B 1999 J. Lightwave Technol. 17 2089
Santori C, Fattal D, Vuckovic J, Solomon G S and Yamamoto Y 2002 Nature 419 594
[11] Vučković J, Fattal D, Santorini C and Solomon G S 2003 Appl. Phys. Lett. 82 3596
[12] Ralph T C, Langford N K, Bell T B and White A G 2002 Phys. Rev. A 65 062324
O’Brien J L, Pryde G J, White A G, Ralph T C and Branning D 2003 Nature 426 264
[13] Bartlett S D and Sanders B C 2002 Phys. Rev. Lett. 89 207903
Eisert J, Scheel S and Plenio M B 2002 Phys. Rev. Lett. 89 137903
[14] Berry D W, Scheel S, Sanders B C and Knight P L 2004 Phys. Rev. A 69 031806(R)
[15] Vogel W, Wallentowitz S and Welsch D-G 2001 Quantum Optics: An Introduction (Berlin: Wiley-VCH)
[16] Minc H 1978 Permanents (Reading, MA: Addison-Wesley)
[17] Waks E, Inoue K, Diamanti E and Yamamoto Y 2003 Preprint quant-ph/0308054
[18] Reck M, Zeilinger A, Bernstein H J and Bertani P 1994 Phys. Rev. Lett. 73 58