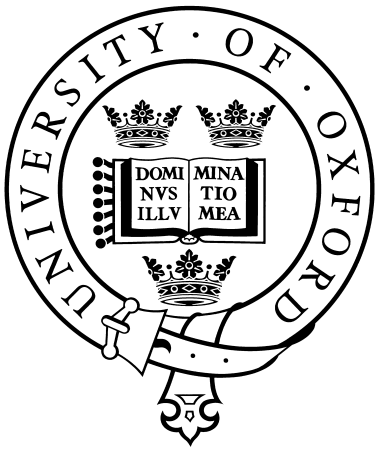
Measurement of the joint statistics of a single photon source generating pure quantum states

Andrew P. Worsley
Hertford College



M.Phys. Project Report

Trinity Term 2008

Supervised by
Dr. Jeff S. Lundeen
Prof. Ian A. Walmsley

Institute for Experimental Photonics
Clarendon Laboratory
University of Oxford
United Kingdom

Contents

Acknowledgements

Thanks must go to Peter Mosley and Hendrik Coldenstrodt-Ronge for their invaluable expertise in the lab, and without whom this project would have taken significantly longer! Thanks also to Jeff, Ian and the rest of the Ultrafast group for making my time in the Clarendon an interesting and enjoyable one.

Abstract

We measure the joint statistics of photons created by parametric down conversion in a type-II phasematched potassium dihydrogen phosphate (KDP) crystal pumped by an amplified mode-locked Titanium:Sapphire laser at 830nm. Two time multiplexed detectors (TMDs) are used, and the counting is performed by field programmable gate array (FPGA) electronics. The photon statistics of the source are reconstructed and fitted to a general multithermal distribution to determine the state purity.

1 Introduction

1.1 Background

Research in quantum information processing (QIP) seeks to exploit quantum mechanical effects for applications such as computing and communication [1]. Quantum systems have been shown to have significant advantages over their classical counterparts, such as more efficient computational algorithms [2] and theoretically ideal cryptography [3].

Classical information is stored as bits that have a defined state of either 1 or 0. To process these bits, logical operations are carried out on them by applying electronic gates such as AND and OR [4]. Quantum information is carried in the state of quantum bits, or *qubits* [1]. These can be treated mathematically as objects that, under a particular operator, have eigenstates of $|0\rangle$ and $|1\rangle$, but generally exist in a superposition $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ (where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$). Quantum logic gates can be implemented by applying quantum mechanical operators to individual qubits or systems of qubits.

One physical system that can be used as a qubit is a photon. It has been shown that all the gates necessary for optical quantum computation can be constructed by using only optical beam splitters (e.g. partially silvered mirrors), phase shifters (e.g. wave plates), single photon sources and photodetectors [5]. This is known as linear optics quantum computation (LOQC). A necessary component of LOQC is a beam splitter at which non-classical interference of two photons can occur. For error-free computation, the interfering photons must be indistinguishable and in a pure quantum state [6]. This requires all the properties of the two photons (such as spatial mode, time of arrival, frequency and polarisation) to be well controlled. A major challenge for quantum computation is the construction of a source that produces such photons on demand [7].

There are several methods for producing pulses of light containing a single photon [8, 9, 10], but parametric down conversion (PDC), further explained in §1.2, is the most successful. It is a simple and reliable process that produces pairs of photons simultaneously. This is useful because one photon from the pair can be detected in order to indicate the presence of the other: a technique known as *heralding*. The heralded photon can then be used in QIP applications [11].

A problem with PDC sources is that the photons generally exhibit pairwise quantum correlations. LOQC requires pure states, but the operation of detecting the heralding photon projects the heralded photon into a mixed state that is useless for QIP. Recently however, Ian Walmsley's group at the Clarendon Laboratory developed a photon source that minimises these correlations and achieves a heralded state purity of over 95% [12, 13], measured using the Hong-Ou-Mandel (HOMI) interference technique [6].

The *photon statistics* of the number of pairs created by PDC depend on the purity of this state, further explained in §1.2 and §1.3. This report presents work in measuring the statistics of the source, and attempts to verify the form of the pure state.

1.2 Parametric down conversion

The simplest model of PDC considers a pump photon that enters a nonlinear crystal and decays to two daughter photons [14]. The process is spontaneous because it is stimulated by random quantum fluctuations of the vacuum. However, the details [15] are beyond the scope of this report. It suffices to say that PDC is observed in crystals with a nonlinear $\chi^{(2)}$ electric suscep-

tibility that provides the coupling between the higher frequency pump photon and the lower frequency daughter photons. The opposite process is sum frequency generation[1], where two photons of lower frequency couple to one photon of higher frequency.

'Parametric' means that the crystal is left unchanged by the process. Energy and momentum are therefore conserved between pump and daughter photons. In this report, we consider collinear down conversion, where the output fields are created with momenta parallel to that of the pump field, thus we need only look at momentum conservation in one dimension, yielding:

$$\omega_p = \omega_1 + \omega_2, \quad (1)$$

$$k(\omega_p) = k(\omega_1) + k(\omega_2), \quad (2)$$

where the index p corresponds to the pump photon, and the indices 1 and 2 correspond to the two daughter photons.

Given that there is a large array PDC emitting sites within a crystal, Equation 2 tells us that they must oscillate with the correct phase differences in order that they constructively interfere to produce the output field. Equations 1 and 2 are therefore called the *phasematching* equations. Equation 1 is satisfied through conservation of energy, whilst Equation 2 is satisfied using the birefringent properties of a nonlinear crystal: by cutting and rotating the crystal correctly, the refractive index n seen by each photon can be changed. Since $k = \frac{n(\omega)\omega}{c}$, this adjusts the wave vectors accordingly. We will consider type-II phasematching, shown in Figure 1, with the pump photon in horizontal polarisation (H) and the two daughter photons orthogonally polarised: one in H, and one in vertical polarisation (V).

[1]The $\chi^{(2)}$ nonlinear polarisation $\mathbf{P}_{NL}$ in a crystal is related to the applied electric field $\mathbf{E}$ by $\mathbf{P}_{NL} = \epsilon_0 \chi^{(2)} \mathbf{E}^2$ [16]. An applied electric field comprising two frequencies, $\mathbf{E}_{\omega_1}(t) = \frac{E_1}{2}(e^{i\omega_1 t} + \text{c.c.})$ and $\mathbf{E}_{\omega_2}(t) = \frac{E_2}{2}(e^{i\omega_2 t} + \text{c.c.})$ enters in $\mathbf{P}_{NL}(t)$ as $(\mathbf{E}_{\omega_1}(t) + \mathbf{E}_{\omega_2}(t))^2$. This produces a term (amongst others) in $e^{i(\omega_1+\omega_2)t}$ representing an oscillation at the sum of the two applied frequencies. Indeed, if $\omega_1 = \omega_2$, we see second harmonic generation (SHG).

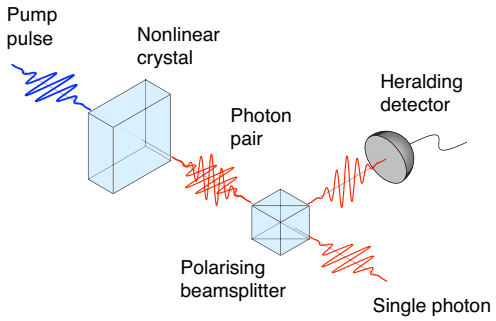


Figure 1: Schematic of type-II phasematched parametric down conversion with a heralding setup. Reproduced from [12].

In general, the phasematching equations have a broad continuum of solutions. The output state is thus a sum over all the possible spectral modes in which the equations are satisfied, and the source is a 'multimode' down converter. Therefore, as mentioned in §1.1, each detection of a heralding photon projects the system into a different state[2]. Since the projection is probabilistic, we obtain a mixed state for the heralded photon.

The idea behind the source developed in the Walmsley group is to constrain the solution of the phasematching equations to one spectral mode by exploiting certain properties of a potassium dihydrogen phosphate (KDP) crystal. Thus, when one photon is detected, the other always collapses into the same state. This removes any randomness from the ensemble of heralded photons and produces a pure state. This is the purity to which we refer in this report.

1.3 Photon number statistics

The laser pump pulse incident on a nonlinear crystal contains many pump photons. Thus, when the crystal is prepared properly, any of these photons may undergo PDC to create a

[2]Note that the system is not projected into a definite frequency state because the pump photons have a broadband spectrum as a result of being in an ultrafast ($\Delta t \approx 50$fs) laser pulse used to control their timing precisely. The Fourier transform result for Gaussian pulses [16] shows $\Delta\omega\Delta t \approx \pi$. Thus, these short pulses must have a broadband spectrum ($\Delta\lambda \approx 20$nm @ 830nm).

daughter pair. There is a probability distribution for the number of generated pairs n known as the photon statistics. For a general multimode down converter emitting in s modes with equal weight, the distribution is *multithermal* [17] and given by:

$$\tilde{\rho}_m(n,s) = \frac{\Gamma(n+s)}{n!\Gamma(s)} \left(1+\frac{\bar{n}}{s}\right)^{-s} \left(1+\frac{s}{\bar{n}}\right)^{-n}, \tag{3}$$

where $\tilde{\rho}_m(n,s)$ is the probability of generating n pairs and $\bar{n}$ is the mean number of pairs generated.

A pure source (as discussed at the end of §1.2) emitting in one mode has a *thermal* (or Bose-Einstein) distribution [18], given by Equation 3 with $s = 1$:

$$\tilde{\rho}_t(n) = \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}}. \tag{4}$$

Contrastingly, a completely multimode down converter is effectively an ensemble of an infinite number of independent pure down conversion emitters. This makes pair generation an entirely random process, hence the photon statistics are in a Poissonian distribution, given by the limit of Equation 3 as $s \to \infty$:

$$\tilde{\rho}_p(n) = \mathrm{e}^{-\bar{n}} \frac{\bar{n}^n}{n!}. \tag{5}$$

1.4 Previous work

Previously, the joint statistics of various PDC sources have been measured with, for example, cryogenically cooled detectors [19] and time multiplexed detectors [20]. However, these experiments only measured the statistics of up to two or three pairs of photons. There has also been some work towards measuring a thermal state [21], but a verification of the form of a pure PDC state has yet to be made. This work seeks to measure the thermal statistics of up to eight photon pairs.

2 Time multiplexed detection

2.1 Photodetection and the avalanche photodiode (APD)

There are several instruments available for detecting single photons [22, 23, 24], and most rely upon an incident photon taking the detecting medium beyond some kind of critical point, thereby causing a transition in the material that produces a macroscopic voltage. In an APD, this transition is between the insulating and conducting states of a semiconductor junction, which in our case is made of silicon. The APD has the advantages of a high quantum efficiency (up to 60%), room temperature operation and a low dark count rate (signals when there is no incident photon) [25]. We operate the APDs in the high gain Geiger mode, in which the junction is reverse biased to just above breakdown voltage. One or more incident photons create carriers in the depletion layer that then cause an 'avalanche' of more carriers. The resulting large current that then flows produces a detectable voltage pulse. The current is removed quickly after the detection by a 'quenching circuit' [26]. A detection event signal is commonly referred to as a *click*. Both the signal pulse time and the detector *dead time* (a time during which photons cannot be detected as the voltage builds back up) are of the order of tens of nanoseconds. This provides the additional advantage of high detection rates, so lasers can be run at high repetition rates and data can be collected quickly.

One drawback is that the amplitude of the pulse is independent of the number of incident photons, and thus they cannot be used directly as photon number resolving detectors. To overcome this, a technique known as *multiplexing* is employed.

2.2 Multiplexing

Multiplexing involves splitting an incident optical pulse into more modes than the mean number of photons in the pulse. Since photons in separate modes can be detected separately, the aim is to have less than one photon per mode on average, such that a click from a mode indicates the

presence of roughly one photon in that mode. The modes can be separated in space (spatial multiplexing) and detected by different APDs [27], or separated in time (time multiplexing) to produce a train of pulses which can be detected by one APD [28] as long as the pulse separation is greater than the detector dead time.

The arrangement is usually such that each photon has a roughly equal probability of ending up in any one of these modes, so as long as the number of modes available is greater than the number of incident photons, it is likely that the number of clicks is representative of the number of photons in the incident field. The relationship between the number of clicks and the number of incident photons is determined statistically, and discussed in §2.4.

Time multiplexing has the advantage that it is more readily scalable to detecting higher photon numbers, which requires splitting the incident pulse into more modes. In spatial multiplexing, adding another mode requires an extra detector. However, only a single APD is required to count an arbitrary number of temporal modes. Much of the following theory of time multiplexed detection was developed in the Walmsley lab [28] for the use of the TMDs in this experiment.

2.3 Time multiplexed detector (TMD) construction

Figure 2 shows how the time multiplexed detector splits up a pulse. At each of the first two beam splitters, a photon can be coupled into either the long or short fibre. The fibre loops are of approximate length L and the short fibres can be considered of negligible length. Table 1 lists the possible delays that can be introduced.

2.4 Convolution

The probabilistic nature of photon propagation in the fibre arrangement necessarily means that even with a lossless TMD, the *click statistics* (the probability distribution of the number of clicks) are not those of the photon statistics of the *input state* (the state before the TMD). There are calculable probabilities for the number of clicks k in the detector given a certain number of incident photons n. For example, two photons incident on the detector can either end up in different modes (giving two clicks) or the same mode (giving one click). In general, all that is required to detect k clicks is that the number of incident photons $n \geq k$. Thus, the probability p_k of detecting k clicks is a sum over the probabilities $C(k|n)$ of n photons ending up in k modes:

$$p_k = \sum_n C(k|n)\tilde{\rho}_n, \tag{6}$$

where $\tilde{\rho}_n$ is the probability that n photons are incident on the TMD.

We can now construct a *convolution matrix* $\mathbf{C}$, with elements $C_{kn} = C(k|n)$, that relates a vector $\tilde{\boldsymbol{\rho}}$ containing the photon statistics to a vector $\mathbf{p}$ containing the click statistics:

$$\mathbf{p} = \mathbf{C} \cdot \tilde{\boldsymbol{\rho}}. \tag{7}$$

Figure 3 gives a visualisation of the convolution matrix. It shows how a high incident photon

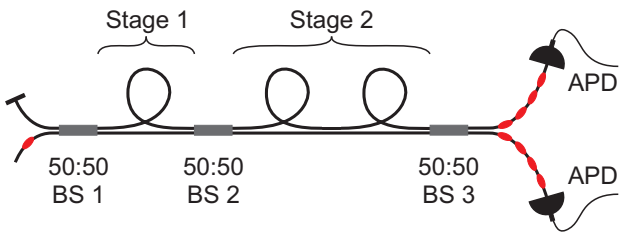


Figure 2: Schematic of a TMD. In a perfect efficiency model, the fibre loops and beam splitters (BS) in this two-stage TMD place the input state from the left into a superposition of eight states. On detection, an incident photon can be found in either of the two spatial modes, one for each avalanche photodiode (APD), and in one of four temporal modes, making a total of eight possible modes.

Stage 1	Stage 2	Time
Short	Short	0
Long	Short	t
Short	Long	2t
Long	Long	3t

Table 1: The time delay of a photon taking four different paths in the TMD as a result of the choice between a long and short fibre at each of the two beam splitters, BS 1 and BS 2. The time $t = L/c$ is the delay introduced by one fibre loop of length L.

number means that a large convolution matrix is required to estimate the input photon statistics. However, in this experiment, the average photon numbers after loss (to be discussed in §2.5) are low and so a 9×9 truncated convolution matrix is sufficient. The details of the calculation of $\mathbf{C}$ are given in Appendix A.

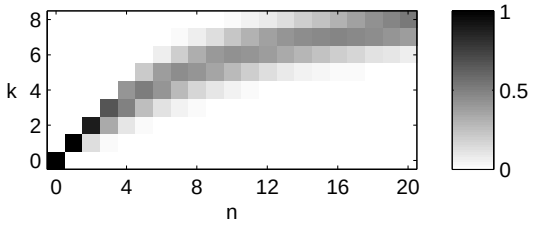


Figure 3: The probabilities C_{kn} of detecting k clicks when n photons are incident on a lossless TMD with 8 equally weighted modes. At low n, k is likely to equal n. When $n \gg 8$, the detector begins to saturate - it becomes more likely that all 8 modes will be filled, and thus 8 clicks will be detected. Notice that $C_{kn} = 0$ for $k > n$ since it is impossible to see more clicks than there are incident photons. The problem is analogous to randomly distributing n balls among 8 bins and finding the probability that k bins are occupied.

2.5 Loss

Photons are lost from the incident beam through many processes including: imperfect coupling of beams to fibres and fibres to each other, detector quantum efficiencies less than unity and scattering from optical components in the beam path. The analysis of loss can be simplified by noting that its effect is solely to remove photons randomly. It can therefore be modelled as a single beam splitter placed before the input to a lossless TMD. The beam splitter reflects photons out of the system at random. The probability that a photon gets removed is called the *loss parameter*, l, and is equivalent to the reflection coefficient of the beam splitter. The efficiency of the system, $\eta = 1 - l$, is equivalent to the transmission coefficient.

The parameter l generates a set of probabilities $L_{n'n}$ relating n photons incident on the beam splitter to n' photons being transmitted. In a similar way to the convolution matrix of §2.4, we can use these probabilities to construct the *loss matrix* $\mathbf{L}$ which acts on the input state $\tilde{\boldsymbol{\rho}}$ to generate a *reduced distribution* of photon statistics $\tilde{\boldsymbol{\rho}}' = \mathbf{L} \cdot \tilde{\boldsymbol{\rho}}$. This represents the photons incident upon the lossless TMD. Equation 7 is thus modified to:

$$\begin{aligned} \mathbf{p} &= \mathbf{C} \cdot \tilde{\boldsymbol{\rho}}' \\ \mathbf{p} &= \mathbf{C} \cdot \mathbf{L} \cdot \tilde{\boldsymbol{\rho}}. \end{aligned} \tag{8}$$

The details of the calculation of the loss matrix are given in Appendix B.

3 Experimental details

3.1 Setup

Figure 4 shows the laser set up. As explained in §1.2, the KDP crystal produces two orthogonally polarised down conversion photons. After the strong blue pump light is removed, these two photons can be separated by a polarising beam splitter that transmits H (detected by TMD 1) and reflects V (detected by TMD 2). This setup allows simultaneous counting of the number of photons in H and the number in V on each pulse.

The APDs had dead times of $\sim$ 60ns, so the temporal separation of TMD modes was made to be $\sim$ 100ns by making the fibre loops the relevant lengths. The precise timings of modes in each TMD are given in Appendix D. The latest temporal mode from a TMD is $\sim 0.4\mu$s after the laser pulse. This was noted to be well within the laser pulse separation of 4μs at 250kHz repetition rate, thus the laser pulses do not interfere with each other. However, this also means that for the majority of the time (between laser pulses) we do not want the electronics to count. Selecting the time window during which we expect a laser pulse, and hence want to count photons, is called *time gating*.

The photodiode PD produces a signal when there is a laser pulse, providing a trigger for the FPGA. We programmed the FPGA to use this trigger, along with the known TMD mode timings, to calculate when to expect photons. It generates a train of eight narrow time gating pulses, one for each of the four time windows for each TMD. On photon detection, the APDs produce

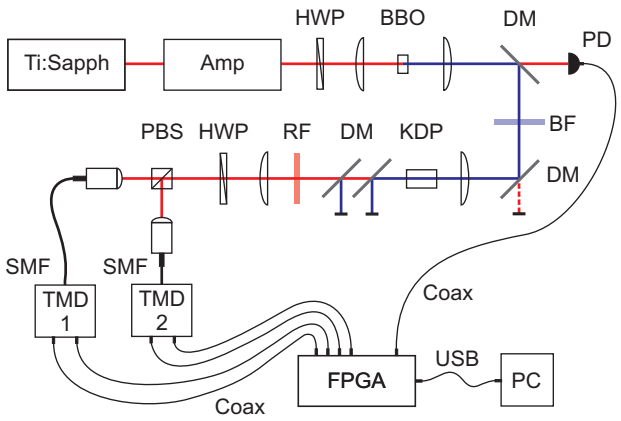


Figure 4: The experimental setup for taking joint photon number statistics. The Titanium:Sapphire (Ti:Sapph) 830nm modelocked laser produces 80MHz pulses of duration $\sim$ 50fs. These are coupled into a regenerative optical amplifier (Amp) that produces greater intensity pulses at a rate of $\sim$ 250kHz. The beam passes through a half wave plate (HWP) into the second harmonic generation (SHG) setup, where it is focused into a beta-barium borate (BBO) crystal. The emerging beam comprises the remaining fundamental red, some of which is collected by a photodiode (PD), and 415nm blue from SHG, which is reflected by dichroic mirrors (DM) and passed through a blue filter (BF) before being focused into the KDP crystal for PDC. The remaining blue pump beam is removed by more dichroic mirrors and a red filter (RF). The orthogonally polarised red photon pairs pass through another HWP before being separated by a polarising beam splitter (PBS). Each of the split beams is coupled into a single mode fibre (SMF) which goes to a TMD. The TMD signals and PD trigger signal are all connected to field programmable gate array (FPGA) electronics by coaxial cable (Coax), and the FPGA data is sent to a personal computer (PC) via universal serial bus (USB) cable.

pulses which are $\sim$ 25ns long. If the rising edge of an APD pulse is coincident with the relevant time gating pulse in the FPGA, then that APD click will be counted. Any APD pulses which occur outside the time gating windows (e.g. from background photons or *after pulsing*[3]) are not counted. The time gating windows in this experiment were 4ns long.

The amplifier is used for for two reasons. Firstly, it is used to reduce the 80MHz repetition rate of the Ti:Sapph laser, which is much too high for the TMD. A TMD receiving pulses every 10ns would mean more than one pulse would be in the detector at any one time, and they would interfere. Secondly, an amplified pump pulse with greater intensity increases the average photon number of the down conversion, making the distributions given by Equations 3, 4 and 5 more distinguishable. This allows us to reduce the error when fitting the reconstructed photon statistics to a multithermal state, and thus reduce the error on measuring the number of modes s.

Another degree of freedom for the photons not yet discussed is that of their transverse spatial profile in the laser beam, known as the *transverse mode* [16]. The down conversion beams from PDC may comprise several transverse modes, so to further restrict our photons to a single mode, we couple these beams into single mode fibre (SMF) before they go to the TMDs. SMF can carry only one transverse mode: the lowest Gaussian mode, known as TEM_{00}.

3.2 Data acquisition

The data was taken using Altera Cyclone FPGA electronics in communication with Labview 8.0 'Virtual Instrument' (VI) software on a PC. Each standard run was 40s long, which at a laser repetition rate of 250kHz allowed counting over approximately 10^7 pulses to reduce random error on the click statistics. Whilst the laser is active, a run is started from the VI, which starts the FPGA creating a 2-D histogram, with the number of clicks in each PDC beam on each axis. At the end of the run, the FPGA uploads the histogram to the VI. The laser power was adjusted by using a rotatable half wave plate and polarising beam splitter after the laser aperture (not shown).

4 State reconstruction

4.1 Theory

4.1.1 Joint statistics

Joint statistics give the probability of finding m photons in one beam at the same time as find-

[3]Sometimes APDs produce a a second signal shortly after a detection event. This is caused by the release of charge caught in the depletion layer, and is known as 'after pulsing'.

ing n in the other. In a joint statistics measurement, the photon statistics and click statistics are now represented by the matrices $\boldsymbol{\rho}$ and $\mathbf{P}$ respectively. The elements of $\boldsymbol{\rho}$ are the probabilities ρ_{mn}, where m is the number of photons generated in one beam and n is the number of photons generated in the other. For a down conversion source producing pairs of photons in H and V polarisations, we would expect to find $\rho_{mn} \neq 0$ only for $m = n$, i.e. a diagonal matrix. The elements of $\mathbf{P}$ are the corresponding probabilities P_{kl} of detecting k clicks in TMD 1 and l clicks in TMD 2, and are given by

$$P_{kl} = \sum_{m,q} \sum_{n,r} C_{kq} L_{qm} \rho_{mn} L_{nr} C_{rl}. \qquad (9)$$

We can write this as a matrix equation analagous Equation 8:

$$\mathbf{P} = \mathbf{C}_1 \cdot \mathbf{L}_1 \cdot \boldsymbol{\rho} \cdot \mathbf{L}_2^{\mathrm{T}} \cdot \mathbf{C}_2^{\mathrm{T}}, \qquad (10)$$

where $\mathbf{C}_i$ and $\mathbf{L}_i$ are the convolution and loss matrices for the i^{th} TMD. These matrices make up the detectors' characteristic transformation of the input state to the click statistics.

4.1.2 Convolution matrices

The $\mathbf{C}_i$ matrices can be calculated using the *mode weightings* - the probabilities that an incident photon appears in each of the modes of a TMD. This is done using a pulsed light source with a very high photon number ($\bar{n} \gg 8$) such that the TMD is saturated, and all 8 modes contain many photons. Using a fast photodiode, we can measure the intensity of light in each mode, which is proportional to the probability of a photon appearing in that mode. The mode weightings for our TMDs are given in Appendix A.

4.1.3 Loss matrices

The $\mathbf{L}_i$ matrices can be calculated if we find the loss parameter l_i for each TMD. We developed a generalisation of the Klyshko method [29] that exploits the experimental data itself.

It is difficult to determine the efficiency of a low photon number photodetector because it needs to be tested with a weak source. We should ideally like to send in a definite number of photons and see how many of them are detected. However, it is the nature of current weak sources to produce various numbers of photons with certain probabilities upon each pump pulse. This is a problem because, for example, when the source produces zero photons on a particular pulse, this is indistinguishable from one photon being produced, but being 'missed' by the photodetector.

Klyshko's scheme uses a two-photon light source to calibrate a photodetector PD1. For clarity of explanation, assume that upon a pump pulse this source produces either no pairs or a single pair of photons, but we do not know with what probabilities. One photon from the pair is sent to PD1 and the other is sent to a second photodetector PD2. All real photodetectors have quantum efficiencies less than unity, so for each pulse there are four unique outcomes: either both PDs click, only PD1 clicks, only PD2 clicks, or neither PD clicks. Now consider solely the subset of events in which we see a click in PD2. In this subset, we know that the source definitely produced one photon pair. Thus, there are now two unique events: a click in PD1 or no click in PD1. By counting over many pulses, the ratio of the number of each these two events tends to a measure of the efficiency of PD1, which can be called the *Klyshko efficiency*. Similarly, by using the subset of events in which we see a click at PD1, the efficiency of PD2 may also be calculated. Such a scheme has been tested experimentally using parametric down conversion [30].

The generalised method extends this for detectors that can resolve photon number, like TMDs. As before, the joint photon statistics are assumed to be diagonal (i.e. the source produces pairs), but the corresponding set of probabilities, ρ_{nn}, are left free.

We use an iterative optimisation to test the parameter space defined by the loss parameters of each TMD, l_1 and l_2 (where $0 \leq l_i \leq 1$). For each tested point in parameter space (l_1, l_2), a least squares method (with the constraints $\rho_{nn} \geq 0$) is used to find a best fit for $\boldsymbol{\rho}$. This inevitably leaves a residual matrix that is the difference between the left hand side and right and side

of Equation 10, representing the difference between the measured click statistics and the click statistics generated from the best fit. The magnitude of this residual is characterised by calculating its *Frobineus norm*[4]. The point (l_1, l_2) whose residual has the minimum Frobineus norm is then taken to be the best estimate of the losses. Since the problem appears to be convex[5], we move around the parameter space using Newton's method [31]. A MATLAB® program incorporating the 'lsqnonneg' function was written for determining the loss parameters. Further details are given in Appendix B.

We would expect that the loss parameters of the detectors are independent of the average incident photon number. So by using joint statistics data sets from different pump powers and performing this optimisation for each one, we can obtain a measurement of the loss parameters of the TMDs, and hence construct their respective loss matrices $\mathbf{L}_i$.

4.2 Reconstruction in practice

4.2.1 Optimisation of the state

Having calculated $\mathbf{C}_i$ and $\mathbf{L}_i$, it would appear that to find $\boldsymbol{\rho}$, one need only invert Equation 10:

$$\boldsymbol{\rho} = \mathbf{L}_1^{-1} \cdot \mathbf{C}_1^{-1} \cdot \mathbf{P} \cdot (\mathbf{C}_2^{\mathrm{T}})^{-1} \cdot (\mathbf{L}_2^{\mathrm{T}})^{-1}. \qquad (11)$$

However, an inversion of Equation 10 is difficult because with typical loss parameters, the loss matrices are almost singular (see Appendix B, Table 6). Consequently, their inverses blow up, which can result in unphysical photon statistics i.e. negative probabilities or probabilities that sum to greater than unity [32]. To avoid this problem, we use a least squares method to optimise $\boldsymbol{\rho}$ within the physical constraints that its elements satisfy $0 \leq \rho_{mn} \leq 1$. A MATLAB® program incorporating the 'lsqlin' function was written for this purpose.

[4]The Frobineus norm of a matrix is the square root of the sum of the squares of all its elements: $\mathrm{Norm}(\mathbf{A}) = \sqrt{\mathbf{A}\mathbf{A}^{\mathrm{T}}}$.

[5]Optimisation problems seek to minimise a function $f(x_1, x_2, ..., x_n)$. In this case, $f(l_1, l_2)$ is the magnitude of the residual of the least squares fit. A convex optimisation problem is one in which f has only one local minimum, which is thus the global minimum.

4.2.2 Background noise

Any photons that reach the detector and are not from the PDC source are considered to be background. Possible background sources are dark counts in the APDs, and photons from ambient light sources entering the TMD fibres. However, these are suppressed to some extent by time gating (see §3.1). Background that is not suppressed by time gating is that which arrives as a consequence of the laser pulse itself. There are some residual pump pulse photons that pass through the filters, and red photons from glass components that fluoresce slightly when subjected to strong blue light.

Just as loss processes randomly remove photons from a detector, similarly background noise randomly adds photons. Consider the case when one pair of photons is generated (ideal input state element ρ_{11}), and they go to TMD 1 and TMD 2 respectively. Ideally, this would appear in the click statistics as a diagonal P_{11} event. If a background photon is also detected at TMD 1, this changes the click event to P_{21}. Notice though that it is also possible to get a P_{21} event from two pair generation ρ_{22} followed by a loss of one photon in TMD 2.

This illustrates the general case that background in one detector very hard to distinguish from loss in the other. In a Klyshko loss calculation, when there is significant background in one detector, this will produce a higher loss parameter for the other detector. We see in §5 that this has consequences for our results. Dealing with background is discussed further in Appendix C.

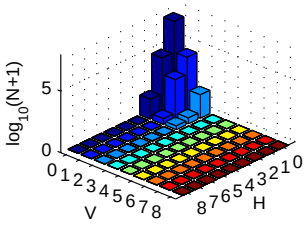
5 Results

5.1 Clicks

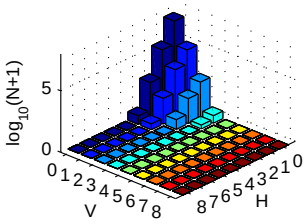
Figure 5 shows two examples of click histogram data. The mean detected photon number is much less than the number of modes in the TMD, and thus the detector was far from saturation. Indeed, the highest order significant events were those involving three clicks being detected in either TMD. The analysis is therefore not affected by the truncation of the convolution matrices to

9×9. Study of Figure 3 and its row $k = 3$ shows that there is little need to consider convolutions of pairs above about $n = 8$ since the probability of a higher order pair contributing to three clicks in the TMD is very low.

The P_{00} element is by far the largest in click data at all powers. This reflects primarily the low detector efficiency, but also the rarity of down conversion events: on most pulses, no photons reach the APDs. A diagonal element P_{nn} is always much smaller than the corresponding P_{ij} elements where either $i = n$ and $j < n$, or $j = n$ and $i < n$. This too is due to low detector efficiency, but also to background photons.



(a) 1mW, $\bar{n} = 0.003$



(b) 65mW, $\bar{n} = 0.124$

Figure 5: Examples of the click histograms from the FPGA. N_{mn} is the number of events where there were m clicks in H (TMD 1) and n clicks in V (TMD 2). The total number of events (i.e. laser pump pulses) was $\sum_{m,n} N_{mn} = 10\ 267\ 277$. $\bar{n}$ is the mean detected photon number.

Since such a large part of the form of the data is the random removal of photons from the down conversion beams due to low detector efficiency, the click statistics at a first glance look like those of a random Poissonian source. However, we seek a state buried within these statistics that is diagonal and non-Poissonian. This means that a precise determination of the efficiencies is critical in order to compensate for them, and this is attempted in §5.2.

5.2 Efficiency determination

Figure 6 shows the result of the efficiency determination using the generalised Klyshko method described in §4.1.3.

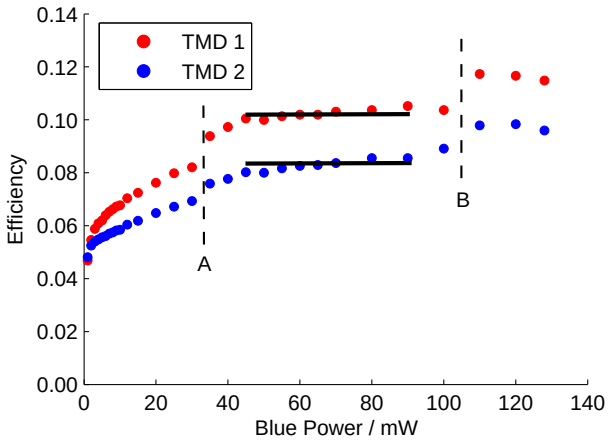


Figure 6: The efficiencies of the TMDs calculated with a generalised Klyshko method, as a function of the blue pump power to the KDP crystal. See text for details.

We would expect the efficiency of a TMD to be independent of the number of incident photons, and thus constant for different blue powers. However, the data points do not show this. The Klyshko efficiencies increase with power from zero, up to a small discontinuity indicated by Line A. There is an approximate plateau between 45mW and 100mW, and then there is an apparent discontinuity (indicated by Line B) to powers of 110mW and above, which give a significantly higher efficiency.

At lower powers, one explanation could be that random background (from ambient light and scattering in the system) was more significant and had the effect of adding to off-diagonal elements in the state (see §4.2.2), which looks like a higher loss and hence a lower efficiency. The discontinuity at 30mW could be due to the digi-

tal power meter used to measure the blue beam power: this is the power at which a change of scale must be made.

An explanation of the behaviour at high powers could be that above ~ 90mW, the BBO crystal was observed to exhibit other nonlinear effects: visible red and yellow light was emitted, possibly as a result of the high peak intensity of the pulses thermally damaging the crystal. An increase in the red power leaving the BBO could have meant that more red photons were transmitted through the blue filtering, since the filters are not perfect, and red background was reaching the TMDs. Furthermore, this background would have been synchronised with the laser pulses, so would not be suppressed by the time gating. Alternatively, the form of the blue beam mode from the BBO could have been so different that the down conversion light was more efficiently coupled into the single mode fibres, and so the detector efficiency was actually higher.

Considering the apparently anomalous behaviours at low and high power, for the purpose of a state reconstruction, the most reliable efficiency values were taken to be between 45mW and 90mW. These were averaged to obtain:

$$\begin{aligned} \eta_1 = 0.102(2) \quad &\Rightarrow \quad l_1 = 0.898(2), \\ \eta_2 = 0.083(2) \quad &\Rightarrow \quad l_2 = 0.917(2). \end{aligned}$$

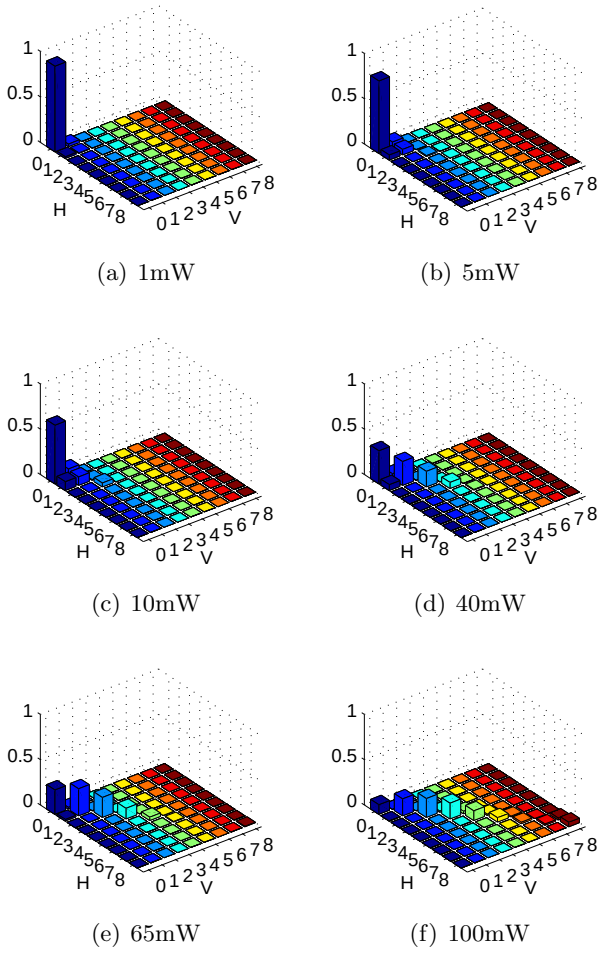
5.3 Reconstruction

Along with the convolution matrices calculated from the mode weightings of §4.1.2, the TMD efficiencies calculated in §5.2 were used to perform a reconstruction of the detector input states at each blue pump power, as described in §4.2.1. A selection of the reconstructed states is shown in Figure 7.

The physicality of the states was checked by ensuring that the sum of the probabilities, S, for each state is unity (since this was not an optimisation constraint), where

$$S = \sum_{mn} \rho_{mn}. \tag{12}$$

For our states, the mean of the respective S values $\bar{S} = 1.0002 \pm 0.0004$.



(a) 1mW (b) 5mW

(c) 10mW (d) 40mW

(e) 65mW (f) 100mW

Figure 7: The reconstructed input states for various powers of blue laser. Each bar shows the probability of the corresponding number of photons in each PDC beam from the KDP crystal. If the reconstructed state was purely that of PDC, we would expect non-zero values on only the diagonal elements of the state. However, background photons and error on the TMD efficiencies mean that there are some off-diagonal elements.

The reconstructions produced approximately diagonal states. We can use the following as a measure of how diagonal a state is:

$$D = \frac{\mathrm{Tr}(\boldsymbol{\rho})}{S}, \tag{13}$$

where $D = 1$ for an entirely diagonal state, and $D = 0$ for a state with no diagonal elements. For our reconstructed states, $\bar{D} = 0.87 \pm 0.08$.

5.4 Multithermal fit

Recall the photon number distributions of §1.3. If the state is pure, we would expect that a reconstruction fitted to a diagonal multithermal distribution (with the mean photon number $\bar{n}$ and number of modes s as free parameters) would yield the same $\bar{n}$ as the reconstruction, and show $s = 1$. Again, a MATLAB® program was written to perform this fit.

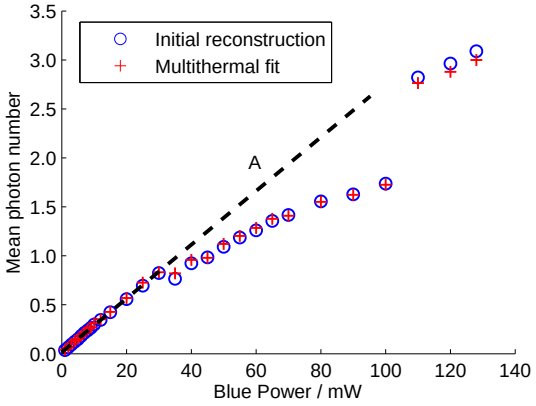


Figure 8: The mean photon number produced by the state reconstruction, and by a fit to a multithermal distribution. Line A shows a linear increase of $\bar{n}$ with power if the gradient at low powers was continued. Subsequent means lie below this line, indicating that the mean is increasing less than linearly with power. See text for details.

Figure 8 shows $\bar{n}$ for the reconstructed state and the multithermal fit. The multithermal fit follows the reconstruction very closely. This is what we would expect since the mean photon number in the click statistics is related to the mean number of input state photons solely by the loss parameters, which we have fixed. A good reconstruction must conserve probability, which will show in the correct reconstructed mean photon number.

The theory of PDC [33] shows that we expect $\bar{n} \propto \sinh gP$, where P is the pump power and g is a constant. This constant is not known for our down converter, thus we do not know the regime of the sinh function we are in. The most we can do is assess whether $\bar{n}$ is increasing at least linearly (as it would if gP was small). Line A illustrates that the mean photon number is increasing less than linearly if we consider the continuous regime of $0 - 100$mW. One explanation could be that an increase in blue power from the BBO may not be causing a proportional increase in the power coupled into the crystal due to a reduction of the beam quality. Alternatively, the effects of background in the click statistics and unreliable Klyshko efficiencies may simply be distorting the mean photon number too much.

The apparent discontinuity in mean photon number above 100mW corresponds to the discontinuity in the Klyshko efficiencies. It is likely that this increase appeared because the mean efficiencies used to reconstruct all the states were much lower than the calculated Klyshko efficiencies for powers > 100mW, as discussed in §5.2. If a reconstruction is performed assuming an efficiency lower than that of the system, it overestimates the loss of photons, thus producing an artificially high original mean photon number.

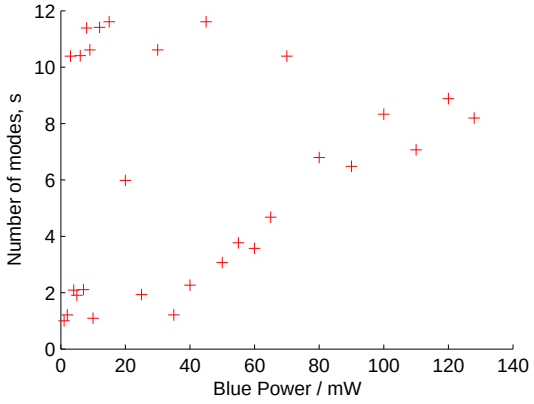


Figure 9: The number of modes in which the down converter emitted, as produced by the multithermal fit. See text for discussion.

Figure 9 shows the s values generated for each multithermal fit. An increase in power should not change the number of modes into which the down converter emits [33]. If we look for a mean value of s, however, we get $\bar{s} = 6 \pm 4$, which is clearly a very large error. This indicates that the reconstructions were not precise enough to provide resolution of the s parameter in a multithermal distribution. The effect on the distribution from a change in s is very small, thus significant error in the reconstructed state (due to errors on loss parameters and background) has meant that

a reliable measurement of s is not possible. Alternatively, it may be that the assumption of a multithermal fit is questionable. The weighting of the modes into which a down converter emits is determined by the spectrum of solutions to the phasematching equations (Equations 1 and 2). This spectrum does not in general assign equal weights to each mode, whereas a multithermal distribution does. However, more sensitive parameters such as mode weightings are likely to be undetectable with such a large error in s.

6 Conclusion

The convex nature of the loss parameter optimisation problem, together with a reasonably large subrange of pump powers yielding almost constant efficiencies, is good evidence that the efficiency of a photon number resolving detector can be found using a parametric down conversion source and a generalisation of the Klyshko method. It has also been shown that using these efficiencies for the TMDs in this experiment, joint photon statistics close to those of a PDC source can be reconstructed from the click statistics data, yielding physical states that are almost diagonal.

However, errors on the efficiency calculation and the reconstruction were too high to find a precise form for the photon statistics. Consequently, the fitting of a multithermal distribution was not reliable enough to determine the purity or otherwise of the source.

For the success of future experiments, it is firstly essential that the purity of the source is established with a HOMI interference dip (see §1.1) prior to a data run in order that s is as close to unity as possible. Ideally, the purity should be verified across the power range. An assumption of this work was that the form of the pump beam remained the same at all powers, however the observation of a change in colour of the SHG light at higher powers has made this assumption questionable. This suggests that the form of the pump beam modes at different powers should be investigated.

Secondly, ambient background light should be reduced as much as possible by lowering the intensity of laboratory lighting to the minimum level and shielding the TMDs from scattered laser light. The main cause of the errors in the analysis is likely to be a lack of compensation for background photons, which led to the apparently variable efficiency and errors in the reconstruction. In order to measure background that is synchronised with the laser pulse, at each blue power the background photon distribution should be measured by stopping the down conversion, but allowing the beam to propagate through the system as it otherwise would. This could be achieved by rotating the polarisation of KDP pump beam such that it is no longer phasematched in the crystal. A second set of click statistics would then be formed for the background. We have developed a theoretical framework for accommodating this background data, which is given in Appendix C.

References

[1] M A Nielsen and I L Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[2] P W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26:1484–1509, 1997.

[3] C H Bennett and G Brassard. Quantum cryptography: Public key distribution and coin tossing. *Proc. Int. Conf. on Computer Systems and Signal Processing (Bangalore)*, pages 175–179, 1984.

[4] D V Bugg. *Electronics: Circuits, Amplifiers and Gates*. Taylor & Francis, 1991.

[5] E Knill, R Laflamme, and G J Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, January 2001.

[6] C K Hong, Z Y Ou, and L Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59(18):2044–2046, November 1987.

[7] P Grangier, B Sanders, and J Vuckovic. Focus on single photons on demand. *New Journal of Physics*, 6(E04), 2004.

[8] C W Chou, S V Polyakov, A Kuzmich, and H J Kimble. Single-photon generation from stored excitation in an atomic ensemble. *Physical Review Letters*, 92(213601), 2004.

[9] J Rarity, J Fulconis, J Duligall, W Wadsworth, and P Russell. Photonic crystal fiber source of correlated photon pairs. *Optics Express*, 13:534–544, 2005.

[10] Q Lin, F Yaman, and G P Agrawal. Photon-pair generation in optical fibers through four-wave mixing: role of Raman scattering and pump polarization. *Physical Review A*, 75(023803), 2007.

[11] T B Pittman, B C Jacobs, and J D Franson. Heralding single photons from pulsed parametric down-conversion. *Optics Communications*, 246:545–550, 2005.

[12] P J Mosley. Generation of heralded single photons in pure quantum states. PhD Thesis, University of Oxford, 2007.

[13] P J Mosley, J S Lundeen, B J Smith, P Wasylczyk, A B U'Ren, C Silberhorn, and I A Walmsley. Heralded generation of ultrafast single photons in pure quantum states. *Physical Review Letters*, 100(133601), April 2008.

[14] M Fox. *Quantum Optics: An Introduction*. Oxford University Press, 2006.

[15] R Loudon. *The Quantum Theory of Light*. Oxford University Press, 1973.

[16] C C Davis. *Lasers and Electro-Optics: Fundamentals and Engineering*. Cambridge University Press, 1996.

[17] L Mandel. Fluctuations of photon beams: the distribution of the photo-electrons. *Proc. Phys. Soc.*, 74:233–243, 1959.

[18] A Luis and J Perina. Zeno effect in parametric down-conversion. *Physical Review Letters*, 76(23):4340–4343, June 1996.

[19] E Waks, B C Sanders, E Diamanti, and Y Yamamoto. Highly nonclassical photon statistics in parametric down-conversion. *Physical Review A*, 73(033814), 2006.

[20] M Avenhaus, H B Coldenstrodt-Ronge, K Laiho, W Mauerer, I A Walmsley, and C Silberhorn. Photon number statistics of multimode parametric down conversion. Submitted to PRL, 2008.

[21] F Paleari, A Andreoni, G Zambra, and M Bondani. Thermal photon statistics in spontaneous parametric downconversion. *Optics Express*, 12(13):2816–2824, 2004.

[22] J Kim, S Takeuchi, and Y Yamamoto. Multiphoton detection using visible light. *Applied Physics Letters*, 74(7), February 1999.

[23] B Cabrera, R M Clarke, P Colling, A J Miller, S Nam, and R W Romani. Detection of single infrared, optical, and ultraviolet photons using superconducting transition edge sensors. *Applied Physics Letters*, 73(6), August 1998.

[24] S Somani, S Kasapi, K Wilsher, W Lo, R Sobolewski, and G Goltsman. New photon detector for device analysis: superconducting single-photon detector based on a hot electron effect. *Journal of Vacuum Science & Technology B*, 19(6):2766–2769, November 2001.

[25] S Cova, M Ghioni, A L Lacaita, C Samori, and F Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Applied Optics*, 35(12):1956–1976, 1996.

[26] R G W Brown, R Jones, J G Rarity, and K D Ridley. Characterization of silicon avalanche photodiodes for photon-correlation measurements. 2: active quenching. *Applied Optics*, 26:2383–2389, 1987.

[27] P Kok and S L Braunstein. Detection devices in entanglement-based optical state preparation. *Physical Review A*, 63(3):033812, February 2001.

[28] D Achilles, C Silberhorn, C Sliwa, K Banaszek, I A Walmsley, M J Fitch, B C Jacobs, T B Pittman, and J D Franson. Photon-number-resolving detection using time-multiplexing. *Journal of Modern Optics*, 51(9-10):1499–1515, 2004.

[29] D N Klyshko. Use of two-photon light for absolute calibration of photoelectric detectors. *Sov. J. Quantum Electron.*, 10(9):1112–1117, 1980.

[30] J G Rarity, K D Ridley, and P R Tapster. Absolute measurement of quantum efficiency using parametric down conversion. *Applied Optics*, 26(21):4616–4618, November 1987.

[31] J Nocedal and S J Wright. *Numerical Optimization*. Springer-Verlag, 1999.

[32] N Thomas-Peter. Measurement of joint photon number statistics using two time multiplexed detectors. M.Phys. Project Report, University of Oxford, 2007.

[33] A Luis and J Perina. SU(2) coherent states in parametric down-conversion. *Physical Review A*, 53(3):1886–1893, March 1996.

[34] S S Wilks. *Mathematical Statistics*. Princeton University Press, 1943.

[35] G B Price. Distributions derived from the multinomial expansion. *The American Mathematical Monthly*, 53(2):59–74, February 1946.

[36] M L Boas. *Mathematical Methods in the Physical Sciences*. Wiley, 1983.

Appendices

A Convolution matrices

Outline

We want to know the number of modes k that will be occupied given n photons incident on the TMD. Therefore, we require the elements C_{kn} of the convolution matrix.

The calculation of a convolution matrix is most easily illustrated by example. Consider a TMD which has four possible output modes labeled $i = \{1, 2, 3, 4\}$. Let the probability that a photon ends up in the i^{th} mode be λ_i, and the set of λ_i be known as the mode weighting. Consider the case where eight photons are incident on the detector. The question is: what are the probabilities for {1,2,3,4} of the four modes being occupied? For example, the simplest case is that of a single mode being occupied by all the photons. Since there are four modes, the probability is simply the sum of the probabilities of all eight photons appearing in each mode:

$$C_{18} = \lambda_1^8 + \lambda_2^8 + \lambda_3^8 + \lambda_4^8. \tag{14}$$

We can consider Equation 15, which uses the fact that the probability of a photon appearing in any mode must be equal to unity:

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 1 \\ \Rightarrow (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)^8 &= 1. \end{aligned} \tag{15}$$

By the multinomial theorem[6], the left-hand side of Equation 15 can be expanded [34, 35] as:

$$\sum_{k_1,k_2,k_3,k_4} \frac{8!}{k_1!k_2!k_3!k_4!} \lambda_1^{k_1} \lambda_2^{k_2} \lambda_3^{k_3} \lambda_4^{k_4}. \tag{16}$$

This gives terms in various powers of λ, which are displayed in Table 2. So, for example, to calculate the probability of three modes being occupied given eight incident photons, we sum the five corresponding terms in the relevant powers of λ:

$$\begin{aligned} C_{38} = p(3|8) = \sum_{i\neq j\neq k}^{4} \Bigg(&\frac{8!}{6!1!1!} \lambda_i^6 \lambda_j \lambda_k \left\{\frac{1}{2!}\right\} + \frac{8!}{5!2!1!} \lambda_i^5 \lambda_j^2 \lambda_k + \frac{8!}{4!3!1!} \lambda_i^4 \lambda_j^3 \lambda_k \ldots \\ &\ldots + \frac{8!}{4!2!2!} \lambda_i^4 \lambda_j^2 \lambda_k^2 \left\{\frac{1}{2!}\right\} + \frac{8!}{3!3!2!} \lambda_i^3 \lambda_j^3 \lambda_k^2 \left\{\frac{1}{2!}\right\} \Bigg), \end{aligned} \tag{17}$$

where the factors $\{\frac{1}{q!}\}$ are to prevent multiple counting where there are q values of λ raised to the same power. For example, under the summation, the product $\lambda_j^2 \lambda_k^2$ produces two terms with $\lambda_1^2 \lambda_2^2$ and $\lambda_2^2 \lambda_1^2$, which are identical. Thus, given the set of λ_i for a TMD, all C_{kn} values can be calculated.

[6] The multinomial theorem is an extension of the binomial theorem:

$$(x_1 + x_2 + .. + x_m)^n = \sum_{k_1,k_2,..,k_m} \frac{n!}{k_1!k_2!..k_m!} x_1^{k_1} x_2^{k_2} ..x_m^{k_m},$$

where the sum is taken over all integer sequences of k_1 to k_m, and $\sum_i^m k_i = n$.

Term In	Modes Occupied
λ_i^8	1
$\lambda_i^7\lambda_j$	2
$\lambda_i^6\lambda_j^2$	
$\lambda_i^5\lambda_j^3$	
$\lambda_i^4\lambda_j^4$	
$\lambda_i^6\lambda_j\lambda_k$	3
$\lambda_i^5\lambda_j^2\lambda_k$	
$\lambda_i^4\lambda_j^3\lambda_k$	
$\lambda_i^4\lambda_j^2\lambda_k^2$	
$\lambda_i^3\lambda_j^3\lambda_k^2$	
$\lambda_i^5\lambda_j\lambda_k\lambda_l$	4
$\lambda_i^4\lambda_j^2\lambda_k\lambda_l$	
$\lambda_i^3\lambda_j^2\lambda_k^2\lambda_l$	
$\lambda_i^2\lambda_j^2\lambda_k^2\lambda_l^2$	

Table 2: The various powers of λ in which terms from the multinomial expansion of Equation 15 appear. The corresponding terms represent the probabilities of a certain number of modes in the TMD being occupied.

Calculation

The mode weightings for each TMD are shown in Table 3, along with their respective 9×9 convolution matrices in Table 4 and Table 5.

TMD 1	TMD 2
0.1278	0.1482
0.1238	0.1350
0.1201	0.1023
0.1086	0.1132
0.1324	0.1612
0.1281	0.1250
0.1338	0.1128
0.1255	0.1021

Table 3: The mode weightings for each TMD. Note that they do not sum to unity due to rounding errors.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.128 & 0.017 & 0.002 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0 & 0 & 0.872 & 0.334 & 0.101 & 0.028 & 0.008 & 0.002 & 0.000 \\ 0 & 0 & 0 & 0.649 & 0.496 & 0.265 & 0.123 & 0.054 & 0.020 \\ 0 & 0 & 0 & 0 & 0.401 & 0.509 & 0.422 & 0.291 & 0.182 \\ 0 & 0 & 0 & 0 & 0 & 0.198 & 0.375 & 0.443 & 0.423 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.073 & 0.192 & 0.308 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.018 & 0.063 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.002 \end{pmatrix}$$

Table 4: A convolution matrix for TMD 1.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.125 & 0.016 & 0.002 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0 & 0 & 0.875 & 0.329 & 0.096 & 0.026 & 0.007 & 0.002 & 0.000 \\ 0 & 0 & 0 & 0.655 & 0.493 & 0.258 & 0.116 & 0.049 & 0.020 \\ 0 & 0 & 0 & 0 & 0.409 & 0.512 & 0.417 & 0.282 & 0.172 \\ 0 & 0 & 0 & 0 & 0 & 0.204 & 0.383 & 0.448 & 0.421 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.076 & 0.201 & 0.318 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.019 & 0.067 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.002 \end{pmatrix}$$

Table 5: A convolution matrix for TMD 2.

B Loss in a TMD

Loss matrices

Given the loss parameter l for a TMD, the probabilities $p(n'|n)$ of n' photons being transmitted to an ideal TMD given n incident photons can be computed, forming the elements $L_{n'n}$ of a matrix $\mathbf{L}$. The problem is one of n trials with the same probability l each time, thus the distribution is binomial:

$$L_{n'n} = \left(\begin{array}{c} n \\ n' \end{array} \right) l^{n-n'}(1-l)^{n'}. \tag{18}$$

The 9×9 matrix generated when $l = 0.9$ is shown in Table 6.

$$\left(\begin{array}{ccccccccc} 1 & 0.900 & 0.810 & 0.729 & 0.656 & 0.591 & 0.531 & 0.478 & 0.431 \\ 0 & 0.100 & 0.180 & 0.243 & 0.292 & 0.328 & 0.354 & 0.372 & 0.383 \\ 0 & 0 & 0.001 & 0.027 & 0.049 & 0.073 & 0.098 & 0.124 & 0.149 \\ 0 & 0 & 0 & 0.001 & 0.004 & 0.008 & 0.015 & 0.023 & 0.033 \\ 0 & 0 & 0 & 0 & 0.000 & 0.000 & 0.001 & 0.003 & 0.005 \\ 0 & 0 & 0 & 0 & 0 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.000 & 0.000 & 0.000 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.000 & 0.000 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.000 \end{array} \right)$$

Table 6: A loss matrix with $l = 0.9$. Note that $L_{n'n} = 0$ for $n' > n$ since it is impossible for loss processes to add photons to the field.

Optimisations in the Klyshko method

Many readily available optimisation algorithms (for instance, in MATLAB®) take a problem in the *affine* form $\mathbf{y} = \mathbf{Mx}$, where the vector $\mathbf{y}$ and the matrix $\mathbf{M}$ are known, and the vector $\mathbf{x}$ is to be found such that the residual norm, $|\mathbf{r}| = |\mathbf{y} - \mathbf{Mx}|$ is minimised. In the loss optimisation problem, it is possible to reduce Equation 10 to such a form, however $\mathbf{y}$ alone is known (representing the click statistics), along with only assumed *forms* for $\mathbf{M}(l_1, l_2)$ (representing the convolution matrices and loss matrices) and $\mathbf{x}$ (representing the photon statistics).

The problem is thus one optimisation problem within another. There are many possibilities for $\mathbf{M}(l_1, l_2)$, each of which produces a certain $|\mathbf{r}|$ when optimising $\mathbf{x}$. The MATLAB® function 'lsqnonneg' was used to find $|\mathbf{r}|$ for a given $\mathbf{M}(l_1, l_2)$. However, in order to find the best $\mathbf{M}(l_1, l_2)$, another optimisation routine was developed. For a simulation of click statistics for a thermal distribution with $\bar{n} = 1$ and $(l_1, l_2) = (0.55, 0.45)$, Figure 10 shows $|\mathbf{r}|$ for various $\mathbf{M}(l_1, l_2)$.

Similar graphs are obtained for simulations with higher loss parameters, and with measured click statistics (although the form is less pronounced).

We can consider the residual norm to be a function, $f(l_1, l_2) = |\mathbf{r}|$, of the loss parameters. In order to find the minimum of this function, we can employ Newton's method. This provides a way of iteratively moving towards a stationary point of a function. For clarity, we consider a function of one variable, $f(x)$. For finding a minimum, the method uses an initial guess x_0 and then iterates with

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}. \tag{19}$$

where x_{n+1} is closer to the minimum. This method relies on the function being twice differentiable, and uses these derivatives at x_n to calculate how to move toward x_{n+1}. Thus, if we seek the minimum

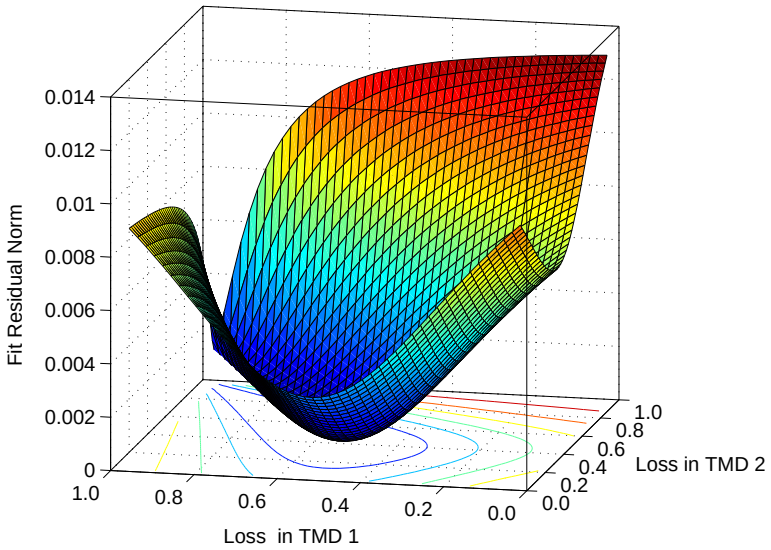


Figure 10: The value of the residual norm, $|\mathbf{r}|$, is an indication of the closeness of the fit to the measured statistics. The minimum residual norm is the point where l_1 and l_2 are optimal. The single minimum point of this function demonstrates that the problem is convex.

of the entire function (the global minimum), the derivatives must guide the iteration toward that minimum. If the function has multiple minima, Newton's method may find a local minimum which is not the global minimum, in which case some prior knowledge of the function must be used in order to choose x_0 sufficiently close to the global minimum.

For our optimisation problem, Figure 10 shows us that the $f(l_1, l_2)$ is indeed twice differentiable (being very smooth), and does not have multiple minima: its single minimum means that the problem is convex, and so the solution found by Newton's method should not depend on the choice of x_0.

In a multidimensional problem with $f(x_1, x_2, ..., x_m)$, Equation 19 becomes vectorised:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma[\mathbf{H}(f(\mathbf{x}_n))]^{-1}\boldsymbol{\nabla} f(\mathbf{x}_n). \tag{20}$$

where $\mathbf{x}_i$ is the vector position in parameter space, $\mathbf{H}$ is the *Hessian matrix* of second derivatives and γ is a parameter which can be used to optimise the step length between iterations. The straight implementation of Newton's method as performed in this work was very crude, and required much adjusting of iteration parameters in order to find a solution. However, there are refinements which would improve an algorithm that incorporates them [31].

C Background in a TMD

Combining probability distributions: more convolution

Consider the input state $\tilde{\boldsymbol{\rho}}$ to a single TMD. This vector represents the discrete statistics of the photons incident on the TMD. We know some of these photons are from the source photon statistics with distribution $\mathbf{a}$, and that there are extra photons from a background distribution $\mathbf{b}$. We should like to know how $\mathbf{a}$ and $\mathbf{b}$ combine to give $\tilde{\boldsymbol{\rho}}$ so that finding $\tilde{\boldsymbol{\rho}}$ and $\mathbf{b}$ would allow us to retrieve $\mathbf{a}$.

For two discrete random variables $X, Y \geq 0$ with probability distributions $f(X)$ and $g(Y)$, the probability distribution $h(X+Y)$ is the discrete convolution[7] of f and g [34]:

$$h(X+Y) = (f * g)(m) = \sum_{p=0}^{m} f(p)g(m-p). \tag{21}$$

Since the number of photons incident on the TMD is the sum of the numbers of photons from both the source and the background, we can conclude that the source distribution convolves with the background distribution in this way.

Equation 21 represents a sum over the photon numbers p in the range $0 \ldots 8$. However, vector elements are conventionally labeled $1 \ldots 9$, especially in mathematics software. Therefore, for convenience, we change the base to where photon number 0 is at vector element 1 and form a vector equation:

$$\begin{aligned} \tilde{\rho}_m &= \sum_{p=1}^{m} a_p b_{(m+1-p)} \\ \tilde{\boldsymbol{\rho}} &= \mathbf{a} * \mathbf{b}. \end{aligned} \tag{22}$$

The same form of convolution occurs in two-dimensions to produce the matrix $\boldsymbol{\rho}$ representing photon statistics from two TMDs, so we can extend Equation 22 to

$$\begin{aligned} \rho_{m,n} &= \sum_{p=1}^{m}\sum_{q=1}^{n} A_{p,q} B_{(m+1-p),(n+1-q)} \\ \boldsymbol{\rho} &= \mathbf{A} * \mathbf{B}. \end{aligned} \tag{23}$$

where $\mathbf{B}$ is the matrix of background photon statistics and $\mathbf{A}$ is the matrix of photon statistics from the source, which we should like to find.

Deconvolution

The convolution theorem [36] can be applied to Equation 23 to give:

$$\mathscr{F}\{\boldsymbol{\rho}\} = \mathscr{F}\{\mathbf{A}\}\mathscr{F}\{\mathbf{B}\}, \tag{24}$$

where the Fourier transforms are two-dimensional and discrete, and the matrix multiplication is element-wise. Thus, we can recover $\mathbf{A}$ by:

$$\mathbf{A} = \mathscr{F}^{-1}\left\{\frac{\mathscr{F}\{\boldsymbol{\rho}\}}{\mathscr{F}\{\mathbf{B}\}}\right\}, \tag{25}$$

where the division of the Fourier transform matrices is element-wise.

[7] Not to be confused with the convolution matrices introduced in §2.4.

Finding the loss parameters and input states

Finding the input states $\boldsymbol{\rho}$ and $\mathbf{B}$ presents a problem, since we can only measure their respective click statistics $\mathbf{P}$ and $\mathbf{P}_B$: their reconstruction requires knowledge of the loss parameters. Ideally we should like to remove the background from $\mathbf{P}$ to obtain the source click statistics $\mathbf{P}_A$ before we begin finding the loss parameters, making our assumption of a diagonal state in the Klyshko method a more robust one. We can use Equation 10 to find expressions for the click statistics from each of the photon statistics distributions:

$$\begin{aligned} \mathbf{P} &= \mathbf{C}_1 \cdot \mathbf{L}_1 \cdot \boldsymbol{\rho} \cdot \mathbf{L}_2^{\mathrm{T}} \cdot \mathbf{C}_2^{\mathrm{T}} \\ \mathbf{P}_A &= \mathbf{C}_1 \cdot \mathbf{L}_1 \cdot \mathbf{A} \cdot \mathbf{L}_2^{\mathrm{T}} \cdot \mathbf{C}_2^{\mathrm{T}} \\ \mathbf{P}_B &= \mathbf{C}_1 \cdot \mathbf{L}_1 \cdot \mathbf{B} \cdot \mathbf{L}_2^{\mathrm{T}} \cdot \mathbf{C}_2^{\mathrm{T}}. \end{aligned} \tag{26}$$

Our rationale for convolving the source and background photon distributions was the following: the convolved distribution is that of the sum of the photon numbers from the source and background. In general, the click distributions do not convolve in this way, since the number of clicks in a TMD is in general not the sum of the clicks from the source and the background. This is due to the convolution (now in the sense of §2.4) of photons that occurs in the TMD. For example, a photon from the background may end up in the same mode as a photon from the source, thus what would have been a single click in each of the separate distributions $\mathbf{P}_B$ and $\mathbf{P}_A$ is now also a single click in $\mathbf{P}$. Hence, we must note:

$$\mathbf{P} \neq \mathbf{P}_A * \mathbf{P}_B. \tag{27}$$

However, we can still remove the background without knowledge of the loss parameters by considering the *reduced photon statistics* (see §2.5), which are the statistics after the input state has undergone loss:

$$\begin{aligned} \boldsymbol{\rho}' &= \mathbf{L}_1 \cdot \boldsymbol{\rho} \cdot \mathbf{L}_2^{\mathrm{T}} \\ \mathbf{A}' &= \mathbf{L}_1 \cdot \mathbf{A} \cdot \mathbf{L}_2^{\mathrm{T}} \\ \mathbf{B}' &= \mathbf{L}_1 \cdot \mathbf{B} \cdot \mathbf{L}_2^{\mathrm{T}}. \end{aligned} \tag{28}$$

By Equations 26, the reduced photon statistics of Equations 28 can also be written as

$$\begin{aligned} \boldsymbol{\rho}' &= \mathbf{C}_1^{-1} \cdot \mathbf{P} \cdot (\mathbf{C}_2^{\mathrm{T}})^{-1} \\ \mathbf{A}' &= \mathbf{C}_1^{-1} \cdot \mathbf{P}_A \cdot (\mathbf{C}_2^{\mathrm{T}})^{-1} \\ \mathbf{B}' &= \mathbf{C}_1^{-1} \cdot \mathbf{P}_B \cdot (\mathbf{C}_2^{\mathrm{T}})^{-1}. \end{aligned} \tag{29}$$

Equations 29 rely on the existence of the inverses $\mathbf{C}_1^{-1}$ and $(\mathbf{C}_2^{\mathrm{T}})^{-1}$. Investigation of typical TMD convolution matrices, for example in Appendix A, reveals that these inverses do exist. In contrast to the click statistics, the reduced photon statistics distributions do convolve since it is still valid to sum photon numbers even after losses. We can therefore write:

$$\boldsymbol{\rho}' = \mathbf{A}' * \mathbf{B}'. \tag{30}$$

Hence, we can apply the deconvolution of Equation 25 with

$$\begin{aligned} \boldsymbol{\rho} &\rightarrow \boldsymbol{\rho}' \\ \mathbf{A} &\rightarrow \mathbf{A}' \\ \mathbf{B} &\rightarrow \mathbf{B}' \end{aligned}$$

and calculate $\mathbf{A}'$ completely. By Equations 29, we can then find $\mathbf{P}_A = \mathbf{C}_1 \cdot \mathbf{A}' \cdot \mathbf{C}_2^{\mathrm{T}}$. Now instead of performing the loss determination and state reconstructions of §4 with Equation 10, we can use:

$$\mathbf{P}_A = \mathbf{C}_1 \cdot \mathbf{L}_1 \cdot \mathbf{A} \cdot \mathbf{L}_2^{\mathrm{T}} \cdot \mathbf{C}_2^{\mathrm{T}}. \tag{31}$$

Since the background state has been deconvolved out, we are now more confident that $\mathbf{A}$ is diagonal, and hence the generalised Klyshko method should be more reliable. We would expect also that the reconstruction of $\mathbf{A}$ would be more diagonal and more reliable than that carried out with the raw click statistics $\mathbf{P}$.

D TMD mode timings

TMD mode timings

Table 7 shows the timing of the pulses from the TMDs. The electronics must synchronise with these in order count correctly.

Mode times / ns	
TMD 1	TMD 2
0.0	0.0
131.7 ± 0.5	100.0 ± 0.5
248.9 ± 0.5	200.0 ± 0.5
381.2 ± 0.5	300.0 ± 0.5

Table 7: The time delay of the temporal modes in each multimode fibre TMD after the first temporal mode.

E Recent update

Since the main report was written, a further experiment was carried out using some of the recommendations in §6. Background measurements were taken in the manner suggested, and the resulting background click statistics were deconvolved from the data using the method described in Appendix C. This time, there was a beam splitter after the second harmonic generation. This was used to send blue power to a second KDP crystal, hence the power hitting our KDP this time was approximately half that of the previous experiment. Nevertheless, the power readings were still taken directly after the second harmonic generation.

Efficiency calculation

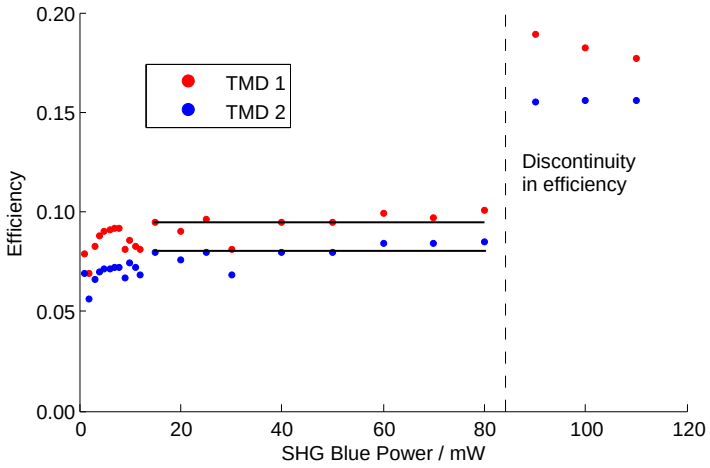


Figure 11: The TMD efficiencies.

Figure 11 shows the TMD efficiencies calculated using the method in §5.2. The measurement seems to have been improved by deconvolution of the background. The Klyshko calculation at lower power again shows unreliability, giving a large spread of data at power < 15mW. However, results at $15 - 80$mW revealed apparently consistent efficiencies, apart from one 'residual' at 30mW, which were used to form the mean efficiencies for the detectors.

However, at 90mW, we see the same thing that happened last time at a similar blue power from the SHG crystal - the Klyshko efficiency suddenly jumps up, approximately doubling this time. Given that this time 90mW SHG blue $\approx$ 45mW KDP blue, and last time 90mW SHG blue = 90mW KDP blue, whatever it is that is causing this discontinuity appears to be happening before the down conversion, most likely in the BBO.

Background measurement

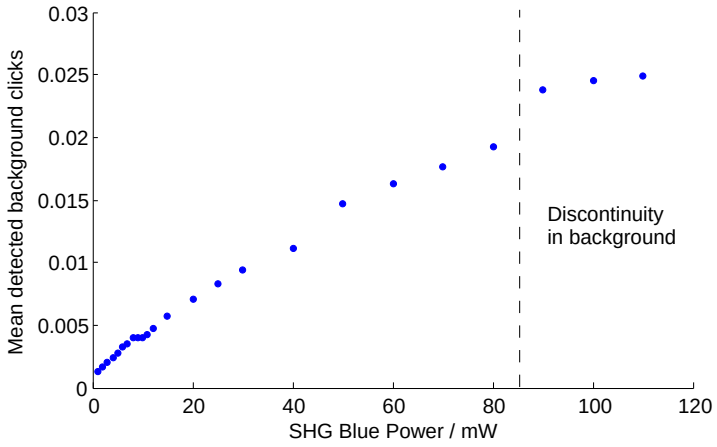


Figure 12: The mean detected background photons.

Figure 12 shows the mean background clicks detected. There is a discontinuity in the background at the same power as the discontinuity in the efficiencies. This suggests the cause of the difficulties in the Klyshko calculation, and in the state reconstruction at these powers is something independent of the down conversion process; it could be other processes in the KDP or the BBO.

Reconstructed mean photon number

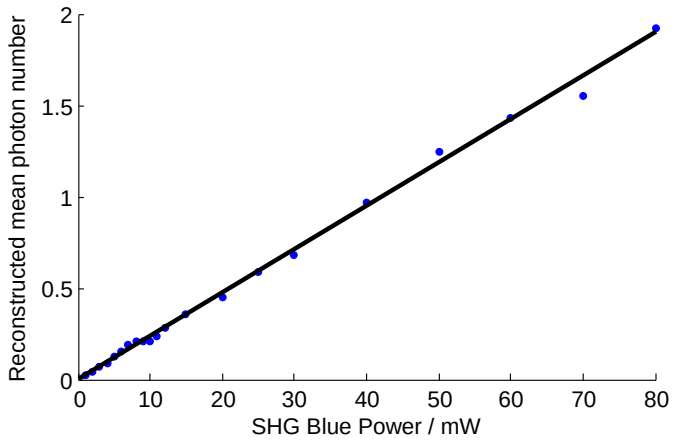


Figure 13: The reconstructed mean photon numbers.

Figure 13 shows the reconstructed mean photon number, using the mean efficiencies justified earlier. There is now an apparent linear relationship between SHG crystal blue power and mean number of pairs in the reconstructed state. The data at higher power was neglected because the mean efficiency is meaningless for them. This is an improvement on the previous experiment, since the fit must be at least linear to agree with the theoretical hyperbolic sine relationship described in §5.4.

Multithermal fit mode number

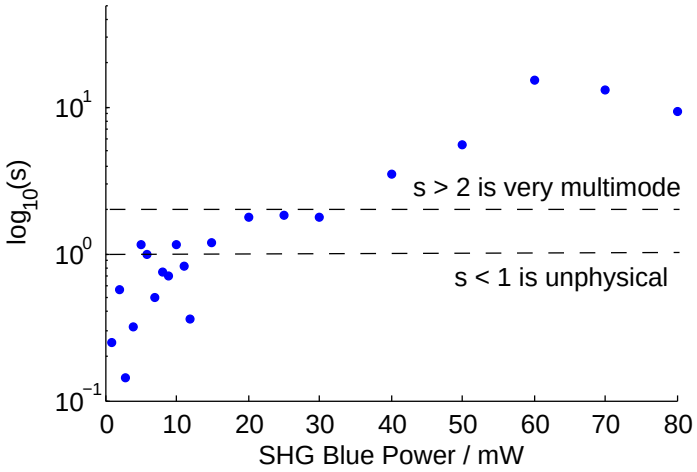


Figure 14: The number of modes produced by a multithermal fit.

Figure 14 shows the number of modes in a multithermal fit to the reconstructed states. In the previous experiment, the number of modes in the multithermal fit did not have an apparent pattern. This time, it seems that up to about 30mW, the fit produces less than two modes. At low power, it sometimes produces less than one mode, but is probably related to the uncertainty in the state at such low powers. At $40-80$mW, mode number starts to climb. Given that most of the background has now probably been taken into account, it seems increasingly likely that the difficulties in reconstructing a thermal state at higher powers are due to a problem with the pump beam incident on the KDP.

Outlook

These results indicate that the source appears to be close to a pure mode down converter a low powers, but quickly becomes multimode for many of the higher laser powers. This suggests that there is a problem with the pump beam from the second harmonic generation, which should be investigated.