Optimized Down-Conversion Source and State-Characterization Tools for Quantum Optics

Der Naturwissenschaftlichen Fakultät
der Universität Paderborn

zur

Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von

Georg Harder

Contents

1. Summary **1**

2. Zusammenfassung **3**

3. Introduction **5**

4. Basics **9**

4.1. Parametric Down-Conversion . . . 9
4.1.1. From Nonlinear Optics to the PDC-Hamiltonian . . . 9
4.1.2. Periodic Poling . . . 13
4.1.3. Spectral Decomposition . . . 14
4.1.4. The PDC State . . . 16
4.1.5. Effective Mode Number . . . 18
4.1.6. Interference-Based Characterization Techniques . . . 19
4.2. Phase Space Representation of Quantum States . . . 22
4.2.1. Quasiprobability Distributions . . . 22
4.2.2. Two-Mode Squeezed State . . . 26
4.3. Nonclassicality . . . 27
4.4. Correlation Functions . . . 29
4.5. Quantum Measurements . . . 31
4.5.1. Balanced Homodyne Detection . . . 32
4.5.2. Photon-Number Resolved Detection . . . 34
4.5.3. Losses . . . 34

5. A Source for Quantum Optics **37**

5.1. Source Engineering . . . 38
5.2. Experimental Setup . . . 41
5.2.1. Source . . . 41
5.2.2. Detectors . . . 42
5.3. Characterization in the Single-Photon Regime . . . 45
5.3.1. Measurement Settings . . . 46
5.3.2. Spectral Intensity . . . 46
5.3.3. Interference Measurements . . . 48
5.4. Characterization in the Multiphoton Regime . . . 50
5.4.1. Photon-Number Statistics . . . 51
5.4.2. Loss Inversion . . . 53
5.4.3. Potential in Continuous Variable Applications . . . 56

5.5. Conclusion . . . 57

6. Characterization of Quantum States **59**
6.1. Introduction to Tomography . . . 59
6.2. Time-Multiplexed Detection with InGaAs APDs . . . 61
6.3. Pattern Tomography . . . 63
6.3.1. Introduction . . . 63
6.3.2. Partial Tomography as Pattern Tomography . . . 65
6.3.3. Application to Two-Mode States . . . 65
6.4. Tomography by Noise . . . 76
6.4.1. Introduction . . . 76
6.4.2. Taking into Account the Mode Overlap . . . 78
6.4.3. Setup . . . 80
6.4.4. Results . . . 81
6.5. Squeezing Estimation without Phase-Reference . . . 84
6.5.1. Phase Averaged Two-Mode Squeezed State . . . 84
6.5.2. Direct Probing . . . 86
6.5.3. The Idea . . . 87
6.5.4. Experimental Results . . . 88
6.6. Correlation Functions and Nonclassicality . . . 91
6.6.1. Experimental Results for High-Order Correlation Functions . . . 93
6.7. Certifying Nonclassicality with Click Detectors . . . 95
6.7.1. Nonclassical Moments of the Click Statistics . . . 96
6.7.2. Experimental Results . . . 97

7. Conclusion and Outlook **101**

Appendix A. Operators **107**
Appendix A.1. Displacement Operator . . . 107
Appendix A.2. Quadrature Operators . . . 108
Appendix A.3. Beam Splitter . . . 108

Appendix B. PDC **111**
Appendix B.1. Gain from First Principles . . . 111
Appendix B.2. Chirp and Decorrelation . . . 112
Appendix B.3. Klyshko Efficiencies . . . 114

Appendix C. Miscellaneous **117**
Appendix C.1. Measurement Scheme for Second-Order Correlation . . . 117
Appendix C.2. Tomography using Convex Optimization . . . 118

Appendix D. Author Contributions **125**

Bibliography **127**

1. Summary

Parametric down-conversion (PDC) has become a common source for quantum states of light. Nevertheless, despite a wide adoption in proof of principle experiments, practical applications in large networks for communication or simulation remain unfeasible due to the high demands on the performance of sources and processes.

In this thesis, we continue to develop PDC sources with the goal of increasing the quality and size of nonclassical states. We realize a source which produced photon pairs with high indistinguishability into a single spatial and spectral mode [1], simplifying potential integration into a multisource quantum network. We further demonstrate that the single-mode operation is maintained in the multiphoton regime by measuring photon-number correlations between the two down-conversion ports with up to 80 photons in each port [2]. This demonstrates that the few-photon regime of current experiments based on PDC can in principle be extended into a many-photon regime, greatly increasing the dimensionality of accessible Hilbert spaces and the size of nonclassical states.

Complementing our work on sources, we investigate how quantum characterization schemes can be improved or simplified to meet experimental requirements like limited resources and imperfect detectors. We demonstrate that photon-number distributions can be inferred using only a single avalanche photodiode and a thermal noise source [3]. We further investigate a technique called pattern tomography [4] and apply it to measurements with imperfect time-multiplexing detectors, which cannot be described by a simple response model. This technique allows us to circumvent the challenging task of full detector tomography and reconstruct low-photon-number PDC states, heralded states and displaced states [5]. Moreover, we develop and demonstrate a technique to infer vacuum two-mode squeezing with a phase-randomized local oscillator [6]. This poses an alternative scheme to homodyne detection, which relies on a phase-locked local oscillator. Finally, we show that nonclassical correlations can be certified from raw click statistics, i.e. without loss inversion or tomography, even in a regime of high losses and relatively large photon-numbers [7].

2. Zusammenfassung

Parametrische Fluoreszenz (PDC) hat sich zu einem Standardprozess für die Erzeugung von Lichtquantenzuständen etabliert. Doch trotzt vieler erfolgreicher konzeptioneller Demonstrationen, bleiben echte Anwendungen wie etwa die Quantenkommunikation in großen Netzwerken oder die Simulation von Quantenprozessen unpraktikabel. Der Grund hierfür liegt unter anderem in den hohen Anforderungen an die Erzeugung, Manipulation und Messung von Quantenzuständen.

In dieser Arbeit führen wir die Entwicklung von PDC-Quellen fort mit dem Ziel, die Qualität und Größe von Quantenzuständen zu erhöhen. Wir implementieren eine Quelle, die nahezu ununterscheidbare Photonenpaare in nur eine spektrale und räumliche Mode emittiert [1]. Dies vereinfacht die potenzielle Integration in ein größeres Quantennetzwerk mit mehreren solchen Quellen. Darüber hinaus demonstrieren wir, dass der Einzelmodencharakter der Quelle für Mehrphotonenzustände erhalten bleibt, indem wir Photonenzahlmessungen auf den Paarmoden der PDC mit bis zu 80 Photon pro Mode durchführen [2]. Dies zeigt, dass aktuelle Experimente, die mit einzelnen oder wenigen Photon arbeiten, prinzipiell in ein Multiphotonenregime erweiterbar sind, sodass höherdimensionale Hilberträume und größere Nichtklassische Zustände realisierbar werden.

Ergänzend zur Implementierung der Quelle untersuchen wir, wie Techniken zur Charakterisierung von Quantenzuständen verbessert oder vereinfacht werden können, um experimentellen Anforderungen wie etwa begrenzten Ressourcen oder nichtidealen Detektoren entgegen zu kommen. Wir demonstrieren, dass Photonenzahlstatistiken mit nur einer Lawinenphotodiode als Klickdetektor und einem thermischen Referenzzustand ermittelt werden können [3]. Weiterhin untersuchen wir die so genannte Mustertomografie (pattern tomography) [4] und wenden sie auf Messungen mit Zeitmultiplexdetektoren an, welche nicht durch ein einfaches theoretisches Modell beschrieben werden können. Diese Methode erlaubt uns die vollständige Detektortomografie zu umgehen und diverse PDC-Zustände, konditionierte Zustände und im Phasenraum verschobene Zustände zu rekonstruieren [5]. Ferner entwickeln wir eine Methode zur Bestimmung der Stärke von zweimodengequetschten Zuständen mit einem phasenrandomisierten Lokaloszillator [6]. Dies stellt eine Alternative zur Homodyndetektion dar, welche auf phasenstabilen Lokaloszillatoren beruht.

Letztendlich zeigen wir, dass nichtklassische Korrelationen in direkt gemessenen Klickstatistiken, d.h. Statistiken ohne Tomografie oder Verlustinversion, nachgewiesen werden können. Dies Funktioniert sogar in einem Regime relativ hoher Photonenzahlen und hoher Verluste [7].

3. Introduction

Quantum physics show effects that are unfamiliar from our everyday world. The biggest difference being perhaps the fact that any measurement of a quantum system unavoidably changes its state. This interconnection between the measurement and the system under study lies at the heart of many paradoxes like the Schrödinger cat, which is dead and alive at the same time: Only when the observer opens the box and looks at the cat, its fate is decided. Another example is the Einstein-Podolsky-Rosen (EPR) paradox [8]. It describes how a quantum state can be used to allow two parties, A and B, separated by a great distance, to predict each other measurement outcomes instantly: If the measurement outcomes of A are described by a local probability distribution, then a measurement in B changes that probability distribution instantly, even though the parties are far apart. This paradox demonstrates that quantum correlations allow measurement results that cannot be described by classical probability theory. This statement can be proven by so called Bell tests [9], which have been verified without loophole assumptions in recent experiments [10–13].

If such nonclassicality could be exploited in large quantum systems, it would lead to numerous applications: Quantum computers can greatly outperform classical computers for some numeric problems or simulate physical systems that are intractable otherwise. Quantum communication networks can provide intrinsically secure communication. And quantum metrology allows measurements with greatly enhanced accuracy. These magnificent prospects have led to extensive research in very different physical systems like ion traps, superconducting qubits, ultracold atoms and of cause, quantum optics. In the race for quantum computing, superconducting qubits [14] seem to have taken the lead recently. For quantum communication, photons are ideal as they can be transmitted over large distances using fibers or free-space links. However, most applications are far from real-world realizations and the ideal implementations are unknown.

Apart from applications, there is also the fundamental question how macroscopic a quantum system can be. Where is the border between the quantum world and the classical world? How large can we make a Schrödinger cat and can we violate Bell inequalities for ever increasing system sizes?

Answering these questions as well as realizing useful applications both require the generation and measurement of large quantum states. In that process, one important step is to verify the nonclassical properties of the state, or more generally, characterize it.

In this thesis, we address the problems of **state generation** for large optical quantum states and **state characterization** with minimal resources.

State Generation: Quantum optics has the advantage over solid state or atomic systems, that photons do not decohere by coupling to the environment. In principle, they retain their quantum properties until the measurement and are therefore well suited to build up large quantum systems. In realistic systems, however, optical loss is unavoidable and constitutes another decoherence mechanism. Nevertheless, it is a well understood mechanism and continual progress is made to reduce loss experimentally. In the last few years, single-photon detectors have made a giant improvement, reaching almost 100 % detection efficiency whereas before, this number was limited to 30 %, in the technologically crucial telecom-wavelength regime. This advancement is very exciting since photon-number measurements are so called non-Gaussian measurements, which are necessary to overcome Gaussian no-go theorems [15]. Performing these measurements with high efficiency opens new possibilities to manipulate and measure quantum states of light that contain many photons.

Parametric down-conversion (PDC) is the most established process to generate optical nonclassical states of light. It can generate two beams of light, which are correlated in amplitude and phase. These correlations are equivalent to the EPR example above and therefore fundamentally interesting as well as useful for a broad range of applications. Remarkably, PDC beams can be relatively bright, containing billions of photons and still showing intensity correlations below the shot noise limit of coherent light [16–21]. In principle, many PDC sources can be combined to form up ever larger quantum states. In the continuous variable cluster state framework, 10,000 temporal modes from one PDC source have been entangled forming one of the largest quantum states to date [22]. If photon-number measurements are to be performed, building up large quantum states is more challenging due to the fact that photon-number measurements are sensitive to all spectral modes. Therefore, the states are required to be in a single-mode. This subtle point, on which we elaborate in more detail later in the thesis, is one of the reasons why the largest states measured with single photon detectors consisted of single photons rather than higher photon-number states or CV squeezed states. The largest states were eight entangled single photons [23, 24].

In this thesis, we present our implementation of a PDC source, that produces single-mode states and therefore matches well with single-photon detectors. We argue that it is ideally suited for large scale networks and the generation of large quantum states with high photon numbers. Combining it with state-of-the-art detectors, we demonstrate photon-number measurements up to 80 photons, one order of magnitude larger than previously shown, for our type of system.

State Characterization: Classical electromagnetic fields can be described in terms of their amplitude and phase. In a quantum description, amplitude and phase become represented by the so called quadratures of the electromagnetic field, which are mathematically analogous to position and momentum in quantum mechanics. This analogy gives an intuition for the quadrature-description of the electromagnetic fields. For example, the Heisenberg uncertainty principle precludes that both posi-

tion and momentum have an arbitrary narrow probability distribution. Moreover, if we have more than one field, the fields can be entangled in the same way quantum particles can be entangled. In particular, the PDC states we generate, produce two orthogonally polarized fields, which are entangled and show quadrature correlations below the Heisenberg limit. These are exactly the EPR-type correlations that formed the original paradox and are at the heart of continuous variable quantum optics. One important task is therefore to verify these correlations. This can be done in many different ways. The most generic approach is full quantum state tomography, which gives all information about the state.

In principle, any complete set of measurements can be used to fully reconstruct an unknown state, either in terms of quasi-probability distributions like the quadratures, or the density matrix formalism. However, such state tomography can easily double the complexity of the experiment. For example, phase-locking weak fields can be very challenging. If the detectors are unknown, fully characterizing them can be a complex task in itself. In this thesis, we address the question, whether it is possible to devise measurement settings that have less stringent requirements but still provide the interesting information about a state. Can we use phase-scrambled states, imperfect detectors, noisy reference fields and still reconstruct a quantum state or determine interesting properties? Well, of course, the answer is yes.

In this thesis, we show that it is possible to reconstruct photon-number distributions using only a single click-detector and thermal light, usually regarded simply as noise. We further develop a technique to reconstruct photon-number distributions with imperfect detectors, circumventing the additional step of detector tomography. Moreover, we demonstrate that nonclassical correlations can be verified from raw click-statistics of multiplexed click-detectors, avoiding any state reconstruction or loss-inversion. Finally, we demonstrate a technique to estimate the strength of EPR-correlations in a two-mode state without a phase reference. All these techniques add to the toolbox of quantum state-characterization to allow specialized measurements with lowest possible resources.

4. Basics

4.1. Parametric Down-Conversion

The process of parametric down-conversion (PDC) is a quantum-mechanical process where one photon decays into two daughter photons with quantum correlations between them. Before we go into detail about what this really means, let us consider, very generally, the propagation of a light wave through a dielectric medium.

4.1.1. From Nonlinear Optics to the PDC-Hamiltonian

If an electromagnetic wave travels through a medium, the charges inside the medium will respond in some way to the driving field. This can lead to a change of the material properties or influence the light field itself as all moving charges are themselves sources of electromagnetic radiation. Depending on this response of the medium, new frequency components can be generated or interactions between different frequency components mediated. In general, the polarization of the medium $\vec{P}(\omega)$ at frequency ω as a response to an electric field $\vec{E}$ can be expanded into a power series of the electric field:

$$\begin{aligned} P_i(\omega) = \epsilon_0[\ & \chi^{(1)}_{ij}(\omega)E_j(\omega) \\ & + \int \mathrm{d}\omega_1 \mathrm{d}\omega_2 \chi^{(2)}_{ijk}(\omega;\omega_1,\omega_2)E_j(\omega_1)E_k(\omega_2) \\ & + \int \mathrm{d}\omega_1 \mathrm{d}\omega_2 \mathrm{d}\omega_2 \chi^{(3)}_{ijkl}(\omega;\omega_1,\omega_2,\omega_3)E_j(\omega_1)E_k(\omega_2)E_l(\omega_3) + ...], \end{aligned} \tag{4.1}$$

where we use the convention to sum over pairs of indices. The expansion coefficients $\chi^{(i)}$ are referred to as electric susceptibilities of order i. Incorporating energy conservation, the notation with the semicolon in the arguments implies that $\omega = \omega_1 + \omega_2$ for the second term and $\omega = \omega_1 + \omega_2 + \omega_3$ for the third term. All quantities as a function of time can be obtained from the Fourier transform

$$P(t) = \int \mathrm{d}\omega P(\omega) e^{-i\omega t}. \tag{4.2}$$

For most processes in nonlinear optics, the susceptibilities χ can be assumed to be constant. This is justified, if the frequencies involved in a particular process are far from any resonances and the frequency widths are not very large. Then, applying the Fourier transform, the polarization simplifies to

$$P_i(t) = \epsilon_0[\ \chi^{(1)}_{ij}E_j(t) + \chi^{(2)}_{ijk}E_j(t)E_k(t) + \chi^{(3)}_{ijkl}E_j(t)E_k(t)E_l(t)\ ] + ... \tag{4.3}$$

It relates to the displacement field (external field) $\vec{D}$ by

$$D_i(\vec{r},t) = \epsilon_0 E_i(\vec{r},t) + P_i(\vec{r},t). \tag{4.4}$$

The first term in the expansion with coefficient $\chi^{(1)}$ is linear in the electric field and can be related to the refractive index of the material by $n^2 = 1 + \chi^{(1)}$. The second $\chi^{(2)}$-term is the first non-linear term and responsible for frequency-doubling and frequency-mixing between different frequency components. In a quantum description, it leads to the process of parametric down-conversion (PDC), which we are especially interested in. Moreover, this second term is responsible for the Pockels effect, i.e. a linear change of the refractive index with an applied electric field. The $\chi^{(3)}$-term can be associated with self-phase-modulation, four wave mixing and the Kerr effect, which is a quadratic change of the refractive index with an applied electric field. Since the most interesting term for us is the $\chi^{(2)}$-term and $\chi^{(3)}$ is significantly smaller than $\chi^{(2)}$, we drop the $\chi^{(3)}$-part in the following.

To derive a quantum Hamiltonian, one needs an expression for the energy of the system, which for electromagnetic fields is given by [25]

$$H(t) = \int \mathrm{d}^3 r \left(\int_0^{\vec{H}(t)} \vec{B}(t) \cdot \delta \vec{H}'(t) + \int_0^{\vec{D}(t)} \vec{E}(t) \cdot \delta \vec{D}'(t) \right), \tag{4.5}$$

whereas $\vec{H} = \vec{B}/\mu_0$ for a non-magnetic medium. Note that this integral is not equal to the well known expression $\frac{1}{2}\int \mathrm{d}^3 r(\vec{B}\vec{H} + \vec{E}\vec{D})$, which holds true only for a linear medium. In our case, we insert expression 4.4 into 4.5 and get

$$\begin{aligned} H &= \int \mathrm{d}^3 r \left(\frac{1}{\mu} \int_0^{\vec{B}} \vec{B}' \cdot \delta \vec{B}' + \epsilon_0 \int_0^{\vec{E}} E_i'(\delta E_i' + \chi_{ij}^{(1)} \delta E_j' + \chi_{ijk}^{(2)}[(\delta E_j')E_k' + E_j' \delta E_k'] \right) \\ &= \underbrace{\frac{1}{2} \int \mathrm{d}^3 r (\frac{1}{\mu_0} B_i B_i + \epsilon_0 (1 + \chi_{ii}^{(1)}) E_i E_i)}_{H_0} + \underbrace{\frac{2}{3} \epsilon_0 \int \mathrm{d}^3 r \chi_{ijk}^{(2)} E_i E_j E_k}_{H_\mathrm{I}}. \end{aligned} \tag{4.6}$$

Here we used the symmetry $\chi_{ijk}(\omega;\omega_1,\omega_2) = \chi_{ikj}(\omega;\omega_2,\omega_1)$. The summation is performed over all indices and $\chi_{ij}^{(1)}$ is assumed to be diagonal for simplicity. This means that our coordinate system is oriented along the optical axes of the crystal. Moving towards an interaction picture, we separate the Hamiltonian into a free part H_0, describing a free propagation with a refractive index $n_{ii}^2 = 1 + \chi_{ii}^{(1)}$ and the interaction part H_I, describing the mixing of three fields.

The quantization of the electromagnetic field in a nonlinear medium can be rather involved [26, 27]. Most people follow the approach for a free (vacuum) electromagnetic field, see e.g. [28, 29]. Essentially, the vector potential is expanded into plane waves and the following electric field operators are obtained:

$$\hat{E}_i(\vec{r},t) = i \sum_{\mathbf{k}} \sqrt{\frac{\hbar \omega_{\mathbf{k}}}{2\epsilon_0 n_{\mathbf{k}i} V}} \left(\hat{a}_{\mathbf{k}} \epsilon_{\mathbf{k}i} e^{i(\vec{k}\vec{r} - \omega_{\mathbf{k}} t)} - \hat{a}_{\mathbf{k}}^\dagger \epsilon_{\mathbf{k}i}^* e^{-i(\vec{k}\vec{r} - \omega_{\mathbf{k}} t)} \right), \tag{4.7}$$

where ω is the angular frequency, V a mode volume in which the field is contained (can tend to infinity) and $\vec{\epsilon}$ a unit polarization vector. The summation index $\mathbf{k} = (\vec{k}, s)$ labels the wavevector $\vec{k}$ and the polarization $s \in \{1, 2\}$. A similar expansion is found for the magnetic field. At the end of the day, operators $\hat{a}_\mathbf{k}, \hat{a}_\mathbf{k}^\dagger$ obey the standard commutation relations and

$$\hat{H}_0 = \sum_\mathbf{k} \hbar\omega_\mathbf{k}(\hat{a}_\mathbf{k}^\dagger \hat{a}_\mathbf{k} + \frac{1}{2}) \tag{4.8}$$

becomes the Hamiltonian of a harmonic oscillator. In the interaction picture, operators without an explicit time dependence evolve according to

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\hat{O}(t) = [\hat{O}(t), \hat{H}_0]. \tag{4.9}$$

For example, the annihilation operator becomes simply $\hat{a}_\mathbf{k}(t) = \hat{a}_\mathbf{k}(0)e^{-i\omega_\mathbf{k}t}$. The interesting dynamics of the system are contained in the interaction Hamiltonian, which is why $\hat{H}_0$ is often dropped entirely.

The state in the interaction picture evolves with H_I according to the Schrödinger equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \hat{H}_I(t)|\psi(t)\rangle \tag{4.10}$$

with the formal solution

$$|\psi_\mathrm{out}\rangle = \mathcal{T} \exp\left(\frac{-i}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}t \hat{H}_I(t)\right)|\psi_\mathrm{in}\rangle. \tag{4.11}$$

Note that $\hat{H}_I(t)$ is a function of time as the fields contain oscillating plane waves. The time ordering operator $\mathcal{T}$ is necessary because $\hat{H}(t)_I$ does not commute with itself for different times, leading to several non-trivial effects [30–32]. It turns out that time ordering does not play a role in the low pump-power regime. Even though we stay in this regime for the major part of this thesis, for the highest pump powers we use, time ordering could play a role [30]. Nevertheless, we neglect the effect in our theoretical model and just note that the model might become inaccurate at high pump powers. This allows us to write

$$|\psi_\mathrm{out}\rangle \approx \exp\left(\frac{-i}{\hbar}\int_{-\infty}^{\infty} \mathrm{d}t \hat{H}_I(t)\right)|\psi_\mathrm{in}\rangle. \tag{4.12}$$

Before we insert the fields $\hat{E}$ into $\hat{H}_I$ let us simplify the situation to a realistic scenario. We assume that the fields propagate collinear along z through a waveguide of length L with only one transverse mode, described by the field distribution $\xi(x, y)$, normalized as $\int \mathrm{d}x\mathrm{d}y|\xi(x, y)|^2 = 1$. That can be interpreted as reducing the mode volume V to the waveguide. The length of the waveguide still allows many longitudinal plane modes. However, in the $x-$ and $y-$ dimensions, only one mode is allowed. This reduces the problem to one dimension and allows us to simplify the summation

$\sum_{\mathbf{k}} = \sum_{k_z}$. Converting the sum to an integral, we can replace $\sum_{k_z} \to \frac{1}{\Delta\omega}\int \mathrm{d}\omega$ with the mode spacing $\Delta\omega = \frac{2\pi c}{L}$. This substitution also changes the delta function $\delta_{k,k'} \to \Delta\omega\delta(\omega-\omega')$ such that, to maintain the usual commutation relations, we have to introduce continuous mode field operators $\hat{a}_k \to \sqrt{\Delta\omega}\hat{a}(\omega)$. Then we obtain [29]

$$\hat{E}_{i,\omega_0}(\vec{r},t) = i\int \mathrm{d}\omega \sqrt{\frac{\hbar\omega_0}{4\pi c\epsilon_0 n(\omega_0)}}\left(\xi_i(x,y)\hat{a}_i(\omega)e^{i(k(\omega)z-\omega t)} - \xi_i^*(x,y)\hat{a}_i^\dagger(\omega)e^{-i(k(\omega)z-\omega t)}\right), \tag{4.13}$$

whereas $\omega \approx \omega_0$, known as the rotating wave approximation. Let us assume that we have three fields present in the process, which have all zero overlap with each other, either due to non-overlapping frequencies or orthogonal polarizations. We label their central frequencies by ω_1, ω_2 and ω_3. Each polarization component i is a sum of the three fields:

$$\hat{E}_i(\vec{r},t) = \hat{E}_{i,\omega_1}(\vec{r},t) + \hat{E}_{i,\omega_2}(\vec{r},t) + \hat{E}_{i,\omega_3}(\vec{r},t). \tag{4.14}$$

Inserting $\hat{E}_i(\vec{r},t)$ into $\hat{H}_I \sim \int \mathrm{d}z\chi^{(2)}_{ijk}\hat{E}_i\hat{E}_j\hat{E}_k$, we get 27 terms. Some of these terms are always zero. For example the term $\hat{E}^3_{\omega_1}$ can never fulfill energy conservation. Terms like $\hat{E}^2_{\omega_1}\hat{E}_{\omega_2}$ correspond to second harmonic generation or type I PDC, where type I means that both output ports have the same polarization. Again, many of these terms are zero because the nonlinearity is zero for a given crystal. For all other terms that are allowed in principle, we can engineer only those processes to be efficient, which are of interest to us. To discuss this further, let us look only at terms like $\hat{E}_{\omega_1}\hat{E}_{\omega_2}\hat{E}_{\omega_3}$, where all three fields are different. Due to the fact that the electric fields commute, we always get six terms with the same nonlinearity from the permutations of $(\omega_1,\omega_2,\omega_3)$ and we can write

$$\frac{-i}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\hat{H}_I(t) = \frac{-2i\epsilon_0}{3\hbar}\int \mathrm{d}t\mathrm{d}z\chi^{(2)}_{ijk}\hat{E}_i\hat{E}_j\hat{E}_k \tag{4.15}$$

$$= 4\chi^{(2)}\frac{-i\epsilon_0}{\hbar}\int \mathrm{d}t\mathrm{d}z\hat{E}_{\omega_1}\hat{E}_{\omega_2}\hat{E}_{\omega_3} \tag{4.16}$$

$$\begin{aligned} =B\int \mathrm{d}t\mathrm{d}z\mathrm{d}\omega_1\mathrm{d}\omega_2\mathrm{d}\omega_3\Big[& \\ O_1\exp\left\{i(k_1+k_2+k_3)z - i(\omega_1+\omega_2+\omega_3)t\right\}\hat{a}_1\hat{a}_2\hat{a}_3 & \\ +O_2\exp\left\{i(k_1+k_2-k_3)z - i(\omega_1+\omega_2-\omega_3)t\right\}\hat{a}_1\hat{a}_2\hat{a}_3^\dagger & \\ +O_3\exp\left\{i(k_1-k_2+k_3)z - i(\omega_1-\omega_2+\omega_3)t\right\}\hat{a}_1\hat{a}_2^\dagger\hat{a}_3 & \\ +O\exp\left\{i(k_1-k_2-k_3)z - i(\omega_1-\omega_2-\omega_3)t\right\}\hat{a}_1\hat{a}_2^\dagger\hat{a}_3^\dagger & \\ -\mathrm{h.c.}\Big] & \end{aligned} \tag{4.17}$$

Here, we have collected all constants into B and the overlaps O_i, which are only different with respect to the positions of the complex conjugate sign.

$$B = \frac{\chi^{(2)}}{2}\sqrt{\frac{\hbar\omega_{01}\omega_{02}\omega_{03}}{\epsilon_0\pi^3c^3n_1n_2n_3}} \qquad O = \int \mathrm{d}x\mathrm{d}y\,\xi_1(x,y)\xi_2^*(x,y)\xi_3^*(x,y) \tag{4.18}$$

Note that the dimension of the overlap term is m^{-1}. It can also be associated with an effective mode area $O = 1/\sqrt{A_{\mathrm{eff}}}$, which is roughly the area of the waveguide cross-section.

In Eq. 4.17 we can identify various processes. For example $\hat{a}_1\hat{a}_2^\dagger\hat{a}_3^\dagger$ means that one photon is annihilated at frequency ω_1 and two photons are created at frequencies ω_2 and ω_3. Similarly, $\hat{a}_1\hat{a}_2\hat{a}_3^\dagger$ means that two photons are annihilated at frequencies ω_1 and ω_2 and one photon is created at frequency ω_3. We can directly carry out the z integral using $\int_{-a}^{a} \mathrm{d}x e^{ixy} = \frac{2}{y}\sin(ay) = 2a\,\mathrm{sinc}(ay)$. The t integral results in delta functions with factors 2π and allows us to eliminate ω_1. These two integrals can be directly interpreted as the momentum and the energy conservation conditions of the process. These are slightly different for the four cases. In the first case, $\omega_1 + \omega_2 + \omega_3 = 0$ can never be fulfilled such that the strange looking term $a_1a_2a_3$ is always zero. The other terms correspond to sum-frequency generation, difference-frequency generation or parametric down conversion, depending on the labellings of the fields. Let us assume that energy and momentum conservation are only fulfilled for the last term, i.e. $\omega_1 = \omega_2 + \omega_3$ and $k_1 - k_2 - k_3 = \Delta k \approx 0$. This assumption also allows us to neglect all other terms like second harmonic generation. We further assume that the interaction happens in a crystal of length L. The output state can then be written as

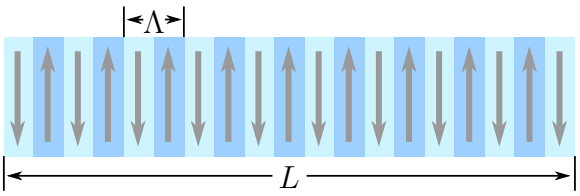
$$|\psi_{\mathrm{out}}\rangle = \exp\left\{2\pi BO \int \mathrm{d}\omega_2 \mathrm{d}\omega_3 L \underbrace{\mathrm{sinc}\left[\frac{\Delta k(\omega_2,\omega_3)L}{2}\right]}_{\phi(\omega_2,\omega_3)} \underbrace{\hat{a}_1(\omega_2+\omega_3)}_{A(\omega_2+\omega_3)} \hat{a}_2^\dagger(\omega_2)\hat{a}_3^\dagger(\omega_3) - \mathrm{h.c.}\right\}|\psi_{\mathrm{in}}\rangle. \tag{4.19}$$

The function $\phi(\omega_2, \omega_3)$ is called phasematching function. For a bright coherent pump field $|A\rangle$ in mode 1, which is the case for most PDC processes, we can replace $\hat{a}_1$ and $\hat{a}_1^\dagger$ by a complex amplitude A and A^* since $\hat{a}|A\rangle = A|A\rangle$ and $\hat{a}^\dagger|A\rangle \approx A^*|A\rangle$. Note that this approximation only holds for a large A, which is well justified for the pump field. This makes the Hamiltonian bilinear and easy solvable. For convenience, we can express the amplitude A in terms of the pulse energy I and a normalized frequency distribution $\alpha(\omega_1) = \alpha(\omega_2 + \omega_3)$ as

$$A(\omega_2+\omega_3) = \underbrace{\sqrt{\frac{I}{\hbar\omega_1}}}_{\alpha_0}\alpha(\omega_2+\omega_3) \qquad \int \mathrm{d}\omega_1 |\alpha(\omega_1)|^2 = 1, \tag{4.20}$$

4.1.2. Periodic Poling

Nonlinear crystals with a high $\chi^{(2)}$ nonlinearity usually have a crystal structure with a dipole moment of the unit cell. Some of these crystals are ferroelectric, meaning that the dipole moment can be inverted by applying a strong electric field. On an atomic level, this means that at least one atom per unit cell moves to a different position. If a certain symmetry is present, this mirrors the crystal structure, thus inverting the crystal axis and changing the sign of the nonlinear coefficient. [33]



Until now, we have assumed that $\chi^{(2)}$ is just a constant and collected it into the constant B. By applying a periodic electric field and inverting domains along the propagation direction, the nonlinear coefficient becomes a function of position $\chi^{(2)}(z) = (-1)^{\lfloor 2z/\Lambda \rfloor}\chi^{(2)}(0)$. Going back to Eq. 4.17, the z integral will be $\int \mathrm{d}z\chi^{(2)}(z)e^{i\Delta kz}$, which is essentially a Fourier transform of $\chi^{(2)}(z)$. The main Fourier component will be $\frac{2\pi}{\Lambda}$ and just shifts the phase matching as $\Delta k \to \Delta k \pm \frac{2\pi}{\Lambda}$. Additionally it reduces the efficiency of the process by a factor of $2/\pi$. For complicated poling patterns, more Fourier components play a role. One particular goal is to achieve a Gaussian phasematching function through special poling patterns [34, 35]. Interestingly, experimental imperfections, which change the waveguide properties along z, for example a temperature gradient in the diffusion process, have essentially the same effect. Experimentally measured phasematching functions thus often have a shape that is closer to a Gaussian function rather than to a sinc function.

4.1.3. Spectral Decomposition

The spectral properties of the PDC process are entirely described by the joint spectral amplitude (JSA)

$$f(\omega_s, \omega_i) = \phi(\omega_\mathrm{s}, \omega_\mathrm{i})\alpha(\omega_\mathrm{s} + \omega_\mathrm{i}), \tag{4.21}$$

where we have relabeled the fields 2 and 3 to signal (s) and idler (i). As can be seen above from Eq. 4.19, the phasematching function ϕ depends on the momentum mismatch and is therefore a function of the crystal properties and the periodic poling. The pump function α is the spectral amplitude of the pump beam.

Two examples of the JSA are shown in figures 4.1 and 4.2. The JSA visualizes the spectral correlations between the signal and idler fields. If the JSA is a circle, a frequency measurement of a signal photon does not yield any information about the frequency of the corresponding idler photon [36, 37]. If the JSA is a narrow line along the diagonal or anti-diagonal, signal and idler are perfectly correlated and will always be measured in frequency pairs. Mathematically, these correlations can be quantified by a Schmidt decomposition [38, 39]:

$$f(\omega_\mathrm{s}, \omega_\mathrm{i}) = F\sum_k c_k\phi_k(\omega_\mathrm{s})\psi_k(\omega_\mathrm{i}), \tag{4.22}$$

where c_k are the normalized eigenvalues of f with $\sum_k |c_k|^2 = 1$ and ϕ, ψ the orthonormal eigenvectors. The term

$$F = \sqrt{\int \mathrm{d}\omega_s \mathrm{d}\omega_i f(\omega_s, \omega_i) f^*(\omega_s, \omega_i)} \tag{4.23}$$

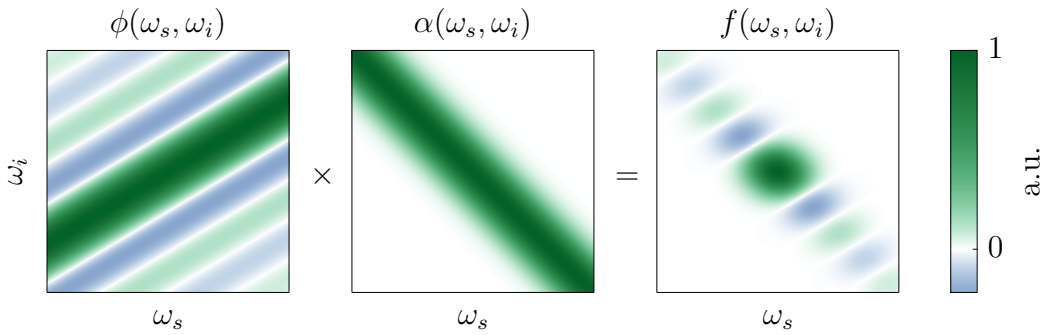


Figure 4.1.: Left: An ideal phasematching function for the material KTP at around 1550 nm. Center: A Gaussian pump envelope function. Right: Joint spectral amplitude, showing only small correlations between ω_s and ω_i due to the sinc side lobes in $\phi(\omega_s, \omega_i)$.

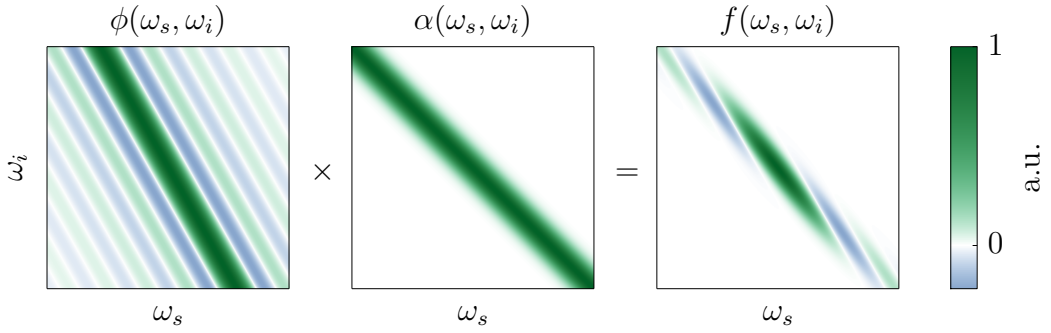


Figure 4.2.: Phasematching, pump-function and JSA for Lithium Niobate. Signal and Idler are correlated in frequency for any pump-width.

is a normalization constant. Plugging this into Eq. 4.19 the two ω integrals become independent and we can define new operators

$$\hat{A}_k^\dagger = \int \mathrm{d}\omega\ \phi_k(\omega)\hat{a}_\mathrm{s}^\dagger(\omega) \qquad \hat{B}_k^\dagger = \int \mathrm{d}\omega\ \psi_k(\omega)\hat{a}_\mathrm{i}^\dagger(\omega), \tag{4.24}$$

which all commute due to the orthonormality properties of ϕ_k and ψ_k. The functions ϕ_k and ψ_k define the spectral shapes of signal and idler for mode k. This transformation diagonalizes the frequency part of the Hamiltonian and the state reads

$$|\psi_\mathrm{out}\rangle = \exp\left\{B' \sum_k (c_k \hat{A}_k^\dagger \hat{B}_k^\dagger - c_k^* \hat{A}_k \hat{B}_k)\right\} |\psi_\mathrm{in}\rangle \tag{4.25}$$

$$= \bigotimes_k \hat{S}_k(\zeta_k)|\psi_\mathrm{in}\rangle. \tag{4.26}$$

where $\zeta_k = -B'c_k$ are often referred to as squeezing parameters and

$$\hat{S}_k(\zeta_k) = \exp(-\zeta_k \hat{A}_k^\dagger \hat{B}_k^\dagger + \zeta_k^* \hat{A}_k \hat{B}_k) \tag{4.27}$$

the squeezing operators [40]. Applied to vacuum, this operator produces PDC states that are discussed in the next section 4.1.4.

The constant B' is of interest if someone wants to predict the squeezing strengths in an experiment. It reads

$$B' = 4BO\alpha_0 FL \tag{4.28}$$

$$= 2\chi^{(2)}\sqrt{\frac{\hbar\omega_{01}\omega_{02}\omega_{03}}{\epsilon_0\pi^3c^3n_1n_2n_3}}O\alpha_0 FL, \tag{4.29}$$

where O is the field-overlap, defined in Eq. 4.18, α_0 the pump-field amplitude defined in Eq. 4.20, F the phasematching normalization term Eq. 4.23 and L the length of the crystal. We discuss the squeezing strength further in the next section 4.1.4. This result for B' is different from the coefficients obtained in [31] by a factor of 2 and by a factor of $2/\pi$ from periodic poling. While the latter one is expected, the factor of 2 clearly means that there is an error in one of the two derivations. Additionally, we tried to reproduced the classical equation for second harmonic generation and obtain a discrepancy of $\sqrt{\frac{2}{\pi}}$. Unfortunately, due to these discrepancies, we have to doubt the exact value of the efficiency-coefficient B'. Nevertheless, we give some realistic values for all the parameters in Appendix B.1 and compare them with measured PDC efficiencies.

Regarding Eq. 4.26 for a moment, we realize that we have a product of operators in orthonormal spaces, acting independently on the input state. This greatly simplifies the spectral degree of freedom in PDC. From the experimental point of view, it gives a natural basis to spectrally multiplex PDC states or operate on them [41].

4.1.4. The PDC State

The squeezing operator 4.27 applied to vacuum produces the PDC state [40]

$$|\zeta\rangle = \hat{S}(\zeta)|0\rangle \tag{4.30}$$

$$|\zeta\rangle = \frac{1}{\cosh r}\sum_{n=0}^{\infty}(-e^{i\phi}\tanh r)^n|n,n\rangle, \tag{4.31}$$

where $\zeta = re^{i\phi}$, or defining $\lambda = -e^{i\phi}\tanh r$

$$|\zeta\rangle = \sqrt{1-|\lambda|^2}\sum_{n=0}^{\infty}\lambda^n|n,n\rangle. \tag{4.32}$$

This is a state with perfect photon-number correlations. These correlations can be used to herald photon-number states (also called Fock states) by placing a photon-number detector in one mode and thus projecting the other mode on the measured photon number. In the low power regime, in which two- and higher-photon-number events can be neglected, this heralding technique is widely used to produce single photons.

Tracing out one mode, the photon-number probabilities in the remaining mode become $p_n = (1 - |\lambda|^2)|\lambda|^{2n}$. This is an exponential distribution just like a Bose-Einstein distribution of a thermal state:

$$P(n) = (1 - e^{-\frac{n\hbar\omega}{k_B T}})e^{-\frac{n\hbar\omega}{k_B T}} \quad \text{identifying} \quad |\lambda|^2 = e^{-\frac{\hbar\omega}{k_B T}}. \tag{4.33}$$

The mean photon number is given by

$$\langle n \rangle = \sum_n n p_n = \sinh^2 r \tag{4.34}$$

At this point we can elaborate on the scaling of the mean photon number with experimental parameters. Staying in the low-photon-number-per-mode regime, we approximate $\langle n_k \rangle \approx r_k^2$. From Eq. 4.29 we see that the squeezing parameter is proportional to the nonlinear coefficient $\chi^{(2)}$, the square root of the pump energy, $\sqrt{I}$ and the mode overlap O. That means:

$$\langle n_k \rangle \sim (\chi^{(2)})^2 \qquad \langle n_k \rangle \sim I \qquad \langle n_k \rangle \sim O^2 = \frac{1}{A_{\text{eff}}} \tag{4.35}$$

The scaling with length is a bit more complicated as it depends on the phasematching function. Let us assume that the phasematching function shows no curvature on the relevant scale (excluding degenerate type I for now) and is not oriented exactly at $-45\,^\circ$. Two examples would be the phasematchings depicted in figures 4.1 and 4.2. We can discuss two cases:

(a) The pump width σ_α is much narrower than the phasematching width $\sigma_\phi \sim 1/L$, i.e. we have a large number of modes $N \sim \sigma_\phi/\sigma_\alpha \sim 1/L$ and $c_k \sim 1/\sqrt{N} \sim \sqrt{L}$. Furthermore, $F \sim \sigma_\phi \sim 1/L$. That means that $\zeta_k \sim FLc_k \sim \sqrt{L}$ and

$$\langle n_k \rangle \sim L \qquad \langle n_{\text{total}} \rangle \approx N \langle n_k \rangle \sim \text{const} \tag{4.36}$$

(b) The pump width σ_α is adjusted to keep the number of modes constant as the length increases. In that case $\zeta_k \sim FL \sim$ const is independent of the crystal length. These considerations also apply to the special case of Fig. 4.1 where mainly one mode is present.

Remarkably, neither the crystal length nor the pump width can change the total number of photons produced in a PDC process (in this low gain regime). Note that in the literature, in particular for continuous wave systems, the number of photons per bandwidth is more relevant. According to Eq. 4.36, the number of photons per bandwidth is proportional to L.

Now let us briefly comment on degenerate type I. Here the phasematching function is oriented exactly at $-45\,^\circ$ and we assume that it has a quadratic curvature. Repeating the considerations of case (a), we find that $N \sim 1/\sqrt{L}$ and $F \sim 1/\sqrt{L}$. That leads to $\langle n_k \rangle \sim L^{3/2}$ and $\langle n_{\text{total}} \rangle \approx N \langle n_k \rangle \sim L$.

4.1.5. Effective Mode Number

The modal properties of the PDC state are entirely described by the set $\{B', c_k, \phi_k, \psi_k\}$ of squeezing parameters and their corresponding modes. However, these parameters are experimentally hard to measure. Besides, we are trying to generate PDC states with only one mode in this thesis. Therefore it is often sufficient to specify just an effective mode number

$$K = \frac{1}{\sum_k |c_k|^4}. \tag{4.37}$$

To motivate this definition, assume that we pump the PDC very weakly to generate single photons and neglect all two-photon components. The state then becomes

$$|\psi\rangle = \bigotimes_k (|0,0\rangle_k + B' c_k |11\rangle_k) \tag{4.38}$$

$$\propto \sum_k c_k |11\rangle_k. \tag{4.39}$$

The probability for each mode is $|c_k|^2$ and normalization ensures $\sum_k |c_k|^2 = 1$. This is a state of two, frequency-entangled, single photons. If all coefficients have the same value $|c_k|^2 = 1/N$, then $K = N$ is exactly the number of modes.

If we trace out one of the two spatial modes and relabel $|1\rangle_k \equiv |k\rangle$, the density matrix becomes

$$\rho_{|1\rangle} = \sum_k |c_k|^2 |k\rangle\langle k|, \tag{4.40}$$

or in matrix representation:

$$\rho_{|1\rangle} = \begin{pmatrix} |c_1|^2 & 0 & \cdots \\ 0 & |c_2|^2 & \\ \vdots & & \ddots \end{pmatrix}. \tag{4.41}$$

The purity of this state is

$$P = \mathrm{tr}\rho^2 = \sum_k |c_k|^4 = \frac{1}{K}. \tag{4.42}$$

Thus, the purity of heralded single photons is inversely proportional to the effective mode number.

A similar connection exists between K and the photon-number statistics of the marginal beam. Here, we also take higher photon-number components into account. Each mode k in 4.32 has exponentially decaying photon-number probabilities $p_k(n) = (1 - |\lambda_k|^2)|\lambda_k|^{2n}$. A photon-number measurement just adds up the photon numbers of all modes. The probability to measure N photons is then

$$P(N) = \sum_{n_1,\ldots,n_{k_{\max}}} p_1(n_1) p_2(n_2 - n_1) p_3(n_3 - n_2) \cdots p(N - n_{k_{\max}}) \tag{4.43}$$

$$= p_1 * p_2 * \ldots * p_{k_{\max}}, \tag{4.44}$$

a convolution of all photon-number probability distributions. Therefore the photon-number statistics between a single-mode state and a multimode state are very different. In the limit of many modes, the distribution becomes Poissonian, as can be seen by evaluating the sum directly in Eq. 4.43.

In a more quantitative analysis, the photon-number distribution can be conveniently used to measure the number of modes. As shown in e.g. [42], the function $g^{(2)} = \frac{\langle N^2\rangle - \langle N\rangle}{\langle N\rangle^2}$, only requiring the mean and the variance of the distribution, is directly related to the number of modes by

$$K = \frac{1}{g^{(2)} - 1}, \tag{4.45}$$

provided that the number of photons per mode is not much bigger than one. Such correlation functions as $g^{(2)}$ are further discussed in Sec. 4.4. To sum up, the purity of heralded single photons, the effective mode number and the marginal second order correlation function are all directly related to each other:

$$K = \frac{1}{P} = \frac{1}{g^{(2)} - 1}. \tag{4.46}$$

4.1.6. Interference-Based Characterization Techniques

Hong-Ou-Mandel Interference between Signal and Idler

The interference of two photons at a beamsplitter followed by a coincidence measurement with single-photon detectors, sketched in Fig. 4.3, is a well established method to deduce information about a biphoton quantum state. First demonstrated by Hong, Ou and Mandel [43] to measure the duration of photon wavepackets, it can be used as a measure of indistinguishability [44–46] and even for reconstruction of the spectral properties of a state [47–49]. The Hong-Ou-Mandel (HOM) interference is unaffected by loss and has been widely applied for the characterization of single-photon states from PDC.

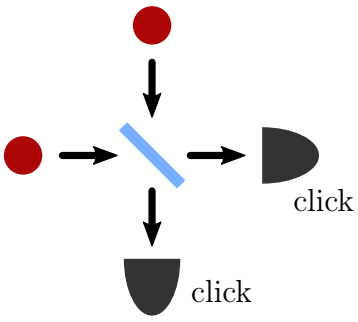


Figure 4.3.: Two photons interfering at a beam splitter. The photon-bunching effect of indistinguishable photons can be observed in the coincidence-click probabilities.

Assuming that signal and idler originate from a perfect, weakly pumped PDC source (negligible two-photon components) with a JSA $f(\omega_s, \omega_i)$, the first non-vacuum

term can be described as

$$|\psi\rangle = \int \mathrm{d}\omega_s \mathrm{d}\omega_i f(\omega_s, \omega_i)|\omega_s, \omega_i\rangle. \tag{4.47}$$

The probability to obtain one photon in each output port of a 50/50 beam splitter after a time delay τ between signal and idler is given by

$$p(\tau) = \frac{1}{2} - \frac{1}{2}\mathrm{Re}\int \mathrm{d}\omega_s \mathrm{d}\omega_i f^*(\omega_s, \omega_i) f(\omega_i, \omega_s) e^{-i\tau(\omega_s - \omega_i)}. \tag{4.48}$$

To derive this formula, we have described the time delay as $e^{i\omega\tau}$ and applied the beam splitter transformation, see Appendix A.3. We can further define the visibility of a HOM dip as

$$V = 1 - \frac{p(\infty) - p(0)}{p(\infty)} = 1 - 2p(0). \tag{4.49}$$

For $\tau = 0$, the integral term is the overlap between the JSA and its mirrored counterpart along the 45° line in the (ω_s, ω_i) plot. A HOM measurement hence probes the symmetry of the state under the exchange of signal and idler, as illustrated in Fig. 4.4. Even for strongly correlated signal and idler, the visibility of a HOM interference can be high, if they show this symmetry. For a decorrelated state as depicted in the central plot, the visibility represents the indistinguishability between signal and idler in terms of their spectral overlap, as one might intuitively imagine. Perfect indistinguishability in this case requires the angle of the phasematching function to be exactly 45 °. For the plots here we assumed an angle of 59 °, according to our experimental situation in Chapter 5.

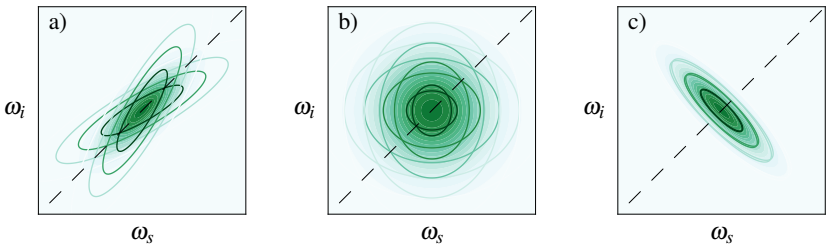


Figure 4.4.: The visibility of the HOM interference is proportional to the overlap (shaded area) of the JSA with itself under the exchange of signal and idler (contour plots). The phasematching function is similar to Fig. 4.1 but has a Gaussian shape. Three cases are shown: a) Positive frequency correlation with $K = 2$ resulting in a visibility of $V = 70\%$. b) Decorrelation with $K = 1$ resulting in $V = 96.7\%$. c) Negative frequency correlation with $K = 2$ resulting in $V = 99.8\%$.

Hong Ou Mandel Interference with a Reference Field

High visibilities in a HOM interference between signal and idler can be obtained in the decorrelated as well as in the negatively correlated case. If our goal is to distinguish these two scenarios, we need a second measure. One possibility is to interfere the state with an independent reference beam. This can be understood as a purity measurement of the marginal beams.

If signal and idler are correlated, tracing out the idler mode results in a mixed signal state, described in frequency space by a density matrix $\rho_s(\omega_s, \omega_s')$. Being Hermitian, this matrix must be symmetric with respect to the principal diagonal. In the following we assume that it is a 2-dimensional Gaussian function, which is well justified for our source as we will see in Sec. 5.3.2. It follows from straight forward calculations that the purity of the state is given by $\mathrm{Tr}(\rho_s^2) = \sigma_2/\sigma_1$, where σ_1 and σ_2 are the major and minor axes of the Gaussian function $\rho_s(\omega_s, \omega_s')$. The major axis is the spectral width of the signal beam and can be measured directly with a spectrometer. The minor axis cannot be accessed directly with a spectral measurement but defines the coherence length of the state and hence the interference pattern, as depicted in Fig. 4.5. By a HOM interference measurement with a known reference field, the minor axis can be deduced from the temporal width of the HOM dip, which is given by

$$\delta^2 = \frac{1}{2\sigma_2^2} + \frac{1}{2\sigma_\beta^2} \tag{4.50}$$

where σ_β is the width of the reference field. A detailed analysis of this technique can be found in [50]. Note that a measurement of the widths is much more robust compared to a measurement of the visibility because it is not affected by higher photon number components and imperfect overlaps of the state in the spectral and polarization degree of freedom.

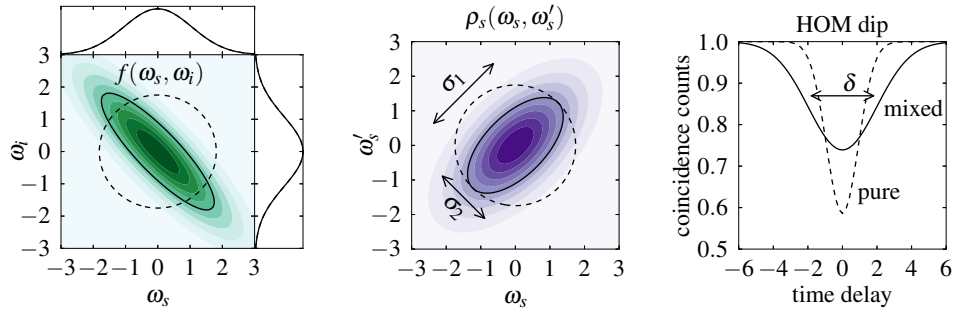


Figure 4.5.: Comparison of an uncorrelated PDC state (dotted black lines) to a correlated PDC state (green, purple or solid black lines) . Left: JSA. Center: reduced density matrix of signal. Right: HOM dip between signal and reference. Both cases, correlated and uncorrelated, have the same marginal spectral widths. Nevertheless, the width of the HOM dip reveals the purity of the marginal state and hence the amount of correlations or number of modes.

4.2. Phase Space Representation of Quantum States

4.2.1. Quasiprobability Distributions

In quantum mechanics, measurements are statistical processes. Even though the time evolution of a quantum state ρ is fully deterministic, measurement outcomes can in general only be predicted with some uncertainty. For example, the position and momentum of a particle can never be determined both with high precision at the same time, prohibited by the Heisenberg uncertainty $\sigma_x \sigma_y \geq \frac{\hbar}{2}$. More generally, such an uncertainty exists for all non-commuting observables:

$$\sigma_A \sigma_B \geq \frac{1}{2} \langle |[A, B]| \rangle \tag{4.51}$$

For the light field, two such non-commuting observables are

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger) \qquad \hat{p} = \frac{-i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger), \tag{4.52}$$

also referred to as position and momentum, since they obey the commutation relation $[\hat{x}, \hat{p}] = i\hbar$. It results from the definition of the fields in Eq. 4.7 and explains why the unperturbed Hamiltonian Eq. 4.8 is just a harmonic oscillator.

In the density matrix formalism of quantum mechanics, the states are described by a density matrix $\hat{\rho}$ and the expectation values of operators are given by

$$\langle \hat{O} \rangle = \mathrm{tr}(\hat{\rho}\hat{O}). \tag{4.53}$$

One can ask the question if there is an alternative description in terms of probability distributions. In the case of the position and momentum operators, it would be a probability distribution $W(x, p)$ with $x, p \in \mathbb{R}$, that fully describes any measurement outcomes and allows to calculate expectation values for $\hat{x}$, $\hat{p}$ or higher moments of them. In general, the answer to this question is no. Essentially, this is the reason why it is not possible to explain quantum mechanics in terms of everyday life phenomena, which can all be described by classical probability theory. However, it turns out that such a $W(x, p)$ exists, if it is allowed to have negative values. For the lack of an interpretation for negative probabilities, it is termed a *quasi*-probability distribution. The negativities are quantum features that distinguish it from classical probability theory.

One way to formally introduce quasiprobability distributions [40, 51] is as follows. We can regard any operator $\hat{O}$ as a function of creation and annihilation operators $\hat{O} = \hat{f}(\hat{a}, \hat{a}^\dagger)$. Now, we need to identify

$$\hat{\rho} \leftrightarrow P(\alpha) \quad \text{and} \quad \hat{f}(\hat{a}, \hat{a}^\dagger) \leftrightarrow f(\alpha, \alpha^*) \tag{4.54}$$

such that the expectation value of $\hat{O}$ can be calculated similar to classical probability theory:

$$\langle \hat{O} \rangle = \int \mathrm{d}^2\alpha P(\alpha) f(\alpha, \alpha^*) \tag{4.55}$$

To find $f(\alpha, \alpha^*)$ we would like to just replace the operators $\hat{a}^\dagger, \hat{a}$ by complex variables α^*, α. According to the inverse of 4.52, they parametrize the quadratures as $\alpha = \frac{1}{\sqrt{2}}(x + ip)$. The problem arising at this point is that $\hat{f}$ depends on the order of the operators $\hat{a}$ and $\hat{a}^\dagger$ but $f(\alpha)$ does not. Therefore, $\hat{f}$ must be rearranged into a specific order, denoted by $:\hat{f}:$ before replacing $\hat{a}$ by α. Three types of ordering are common:

- normal ordering: all creation operators are placed to the left of all annihilation operators, e.g. $:\hat{a}\hat{a}^\dagger\hat{a}:_{\mathrm{N}} = \hat{a}^\dagger\hat{a}^2$.
- symmetric ordering: an average over all possible arrangements, e.g. $:\hat{a}\hat{a}^\dagger\hat{a}:_{\mathrm{S}} = \frac{1}{3}(\hat{a}^2\hat{a}^\dagger + \hat{a}\hat{a}^\dagger\hat{a} + a^\dagger\hat{a}^2)$.
- antinormal ordering: all creation operators are placed to the right of all annihilation operators, e.g. $:\hat{a}\hat{a}^\dagger\hat{a}:_{\mathrm{A}} = \hat{a}^2\hat{a}^\dagger$.

If the ordering is fixed, we can use the identity

$$:f(\hat{a}, \hat{a}^\dagger): = \int \mathrm{d}^2\alpha \, :\delta(\hat{a} - \alpha): f(\alpha, \alpha^*) \tag{4.56}$$

to rewrite

$$\langle :f(\hat{a}, \hat{a}^\dagger):\rangle = \int \mathrm{d}^2\alpha \, \underbrace{\langle :\delta(\hat{a} - \alpha):\rangle}_{P(\alpha)} f(\alpha, \alpha^*), \tag{4.57}$$

which is exactly the expression 4.55 we are looking for. For different orderings, we get different distributions $P(\alpha)$. These are the Glauber-Sudarshan P-function $P(\alpha)$ for normal ordering, the Wigner function $W(\alpha)$ for symmetric ordering and the Husimi Q-function $Q(\alpha)$ for antinormal ordering.

The Wigner Function

Since x and p are already symmetrically ordered in their definition 4.52, their corresponding quasiprobability distribution is the Wigner function which can be derived the following way:

We rewrite the delta-function as

$$:\hat{\delta}(\hat{a} - \alpha): = :\frac{1}{\pi^2}\int \mathrm{d}^2\beta \, \exp\{(\hat{a}^\dagger - \alpha^*)\beta - (\hat{a} - \alpha)\beta^*\}: \tag{4.58}$$

$$= :\frac{1}{\pi^2}\int \mathrm{d}^2\beta \, \hat{D}(\beta) \exp\{-\alpha^*\beta + \alpha\beta^*\}:. \tag{4.59}$$

The operator $\hat{D}(\beta) = \exp\{\beta\hat{a}^\dagger - \beta^*\hat{a}\}$ is known as the displacement operator. It has the property of displacing the annihilation operator: $\hat{D}(\alpha)\hat{a}\hat{D}(-\alpha) = \hat{a} + \alpha$. Since the displacement operator is symmetrically ordered already, we can drop the ordering symbol and continue rewriting it as

$$:\delta(\hat{a} - \alpha):_{\mathrm{S}} = \frac{1}{\pi^2}\int \mathrm{d}^2\beta \, \hat{D}(\alpha)\hat{D}(\beta)\hat{D}(-\alpha). \tag{4.60}$$

Then the Wigner function becomes

$$W(\alpha) = \langle :\delta(\hat{a} - \alpha):_{\mathrm{S}} \rangle = \frac{1}{\pi^2} \int \mathrm{d}^2\beta \, \mathrm{Tr}\{\hat{D}(-\alpha)\hat{\rho}\hat{D}(\alpha)\hat{D}(\beta)\} \tag{4.61}$$

$$= \frac{2}{\pi} \sum_n (-1)^n \langle n|\hat{D}(-\alpha)\hat{\rho}\hat{D}(\alpha)|n\rangle \tag{4.62}$$

where we have used $\langle m| \int \mathrm{d}^2\beta \, \hat{D}(\beta)|n\rangle = 2\pi(-1)^n\delta_{mn}$, derived in Appendix A.1. This is one way to relate the Wigner function to the density matrix. In this representation, the Wigner function at the origin ($\alpha = 0$) is the expectation value of the parity operator. At any coordinate α in phase space, the state just needs to be displaced by $-\alpha$ and its parity measured. In this sense, photon-number measurements are ideally suited to measure single points of the Wigner function because they directly sample the parity.

In Fig. 4.6 we show a few examples of Wigner functions. The Wigner function of the coherent state is a Gaussian function whose width corresponds to the uncertainties in the quadratures. Alternatively, one can think of the distance to the origin as the amplitude of the state (since the photon-number operator corresponds to $\frac{1}{2}[x^2 + p^2 - 1]$) and the angle in the x-p-plane as the phase of the field. The squeezed state has a Gaussian distribution as well but the widths of the quadratures are different. One quadrature has a lower uncertainty at the expense of the other, constrained by the Heisenberg uncertainty principle $\Delta x \Delta p \geq \frac{1}{2}$. The Wigner function of the three-photon Fock state is quite dissimilar to the Gaussian examples. It shows oscillations and has negative values. Since it is rotationally symmetric, any marginal distribution for x, p or a linear combination $\cos(\phi)x + \sin(\phi)p$ shows the same probability distribution with four peaks. It is obvious, that this property cannot be achieved with a positive two-dimensional distribution and demonstrates the nonclassicality of this Fock state. The last example is the cat state which is a superposition of two coherent states with opposite phase showing interference fringes between them and also clear nonclassicality features.

The Glauber-Sudarshan P-function

The P-function is associated with normally ordered operators. It can be used to calculate

$$\langle \hat{a}^{\dagger m} \hat{a}^n \rangle = \int \mathrm{d}^2\alpha P(\alpha) \alpha^{*m} \alpha^n. \tag{4.63}$$

Interestingly, the P-function has a rather intuitive interpretation: It corresponds to the expansion of the density matrix in terms of coherent states:

$$\hat{\rho} = \int \mathrm{d}^2\alpha P(\alpha) |\alpha\rangle\langle\alpha|. \tag{4.64}$$

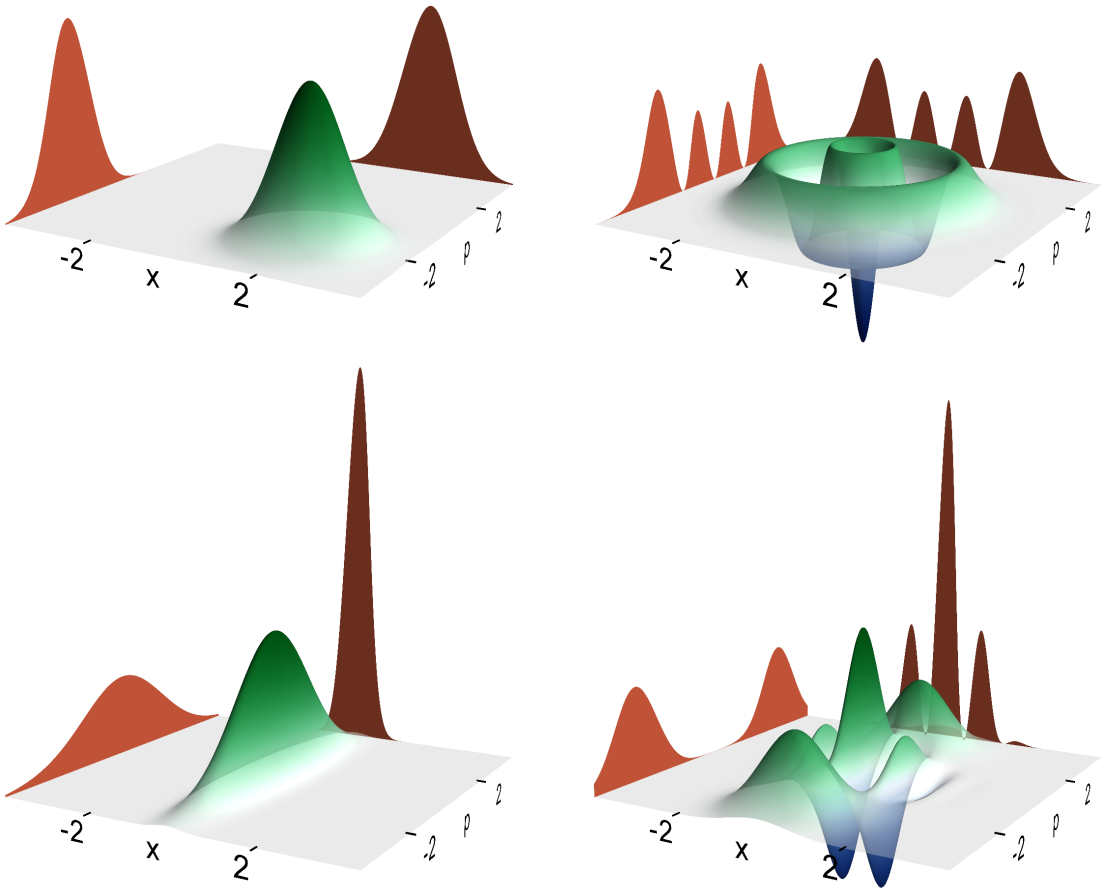


Figure 4.6.: Wigner functions $W(x,p)$. Top-Left: Coherent state $|\alpha\rangle$ with $\alpha = (\sqrt{2}+\sqrt{2}i)$, centered at $x = p = 2$. Top-Right: Fock state $|3\rangle$. Bottom-Left: squeezed state with $r = 0.8$. Bottom-Right: Cat state $|\alpha\rangle + |-\alpha\rangle$ with $\alpha = 2i$. Green indicates positive values and blue indicates negative values. Projections are marginal probability distributions for x and p.

This can be easily verified by

$$\langle \hat{a}^{\dagger m}\hat{a}^n\rangle = \mathrm{Tr}(\hat{a}^n\hat{\rho}\hat{a}^{\dagger m}) \tag{4.65}$$

$$= \mathrm{Tr}\left(\int \mathrm{d}^2\alpha P(\alpha)|\alpha\rangle\langle\alpha|\alpha^{*m}\alpha^n\right) \tag{4.66}$$

$$= \int \mathrm{d}^2\alpha P(\alpha)\alpha^{*m}\alpha^n. \tag{4.67}$$

If the state is a coherent state, the P-function is a delta function, i.e. a dot in phase-space. If the state is 'narrower' than a coherent state in some sense, as the squeezed state, the Fock state or the cat state, the P-function becomes not well behaved, including derivatives of delta functions. It is thus not very well suited to illustrate quantum states in the same way as can be done using the Wigner function. However,

it plays a central role in the definition of nonclassicality as we will discuss in Sec. 4.3.

Husimi Q-function

The Husimi Q-function corresponds to antinormal ordering. It can be written as

$$Q(\alpha) = \langle\alpha|\hat{\rho}|\alpha\rangle. \tag{4.68}$$

Again, this can be verified by

$$\begin{aligned}\langle\hat{a}^n\hat{a}^{\dagger m}\rangle &= \mathrm{Tr}(\hat{a}^{\dagger m}\hat{\rho}\hat{a}^n) &(4.69)\\ &= \int \mathrm{d}^2\alpha\langle\alpha|\alpha^{*m}\hat{\rho}\alpha^n|\alpha\rangle &(4.70)\\ &= \int \mathrm{d}^2 Q(\alpha)\alpha^{*m}\alpha^n. &(4.71)\end{aligned}$$

This form also means that the Q-function is always positive, i.e. a classical probability distribution.

The three quasiprobability distributions introduced here can be transformed into one another by (de-)convolutions with a Gaussian function [51]. The relation is $P(\alpha) \xrightarrow{\text{convolution}} W(\alpha) \xrightarrow{\text{convolution}} Q(\alpha)$. This directly shows that the P-function is the 'narrowest' quasiprobability distribution and hence destined to define nonclassicality for quantum states.

4.2.2. Two-Mode Squeezed State

The PDC state is also referred to as the two-mode squeezed state. Its Wigner function reads [40].

$$\begin{aligned}W(\alpha,\beta) = \frac{4}{\pi^2}\exp\{-2[\,&\cosh(2r)(|\alpha|^2+|\beta|^2)\\ &+\sinh(2r)(\alpha\beta e^{i\phi}+\alpha^*\beta^* e^{-i\phi})]\}\end{aligned} \tag{4.72}$$

or

$$W(\alpha,\beta) = \frac{4}{\pi^2}\exp\{-e^{2r}|\alpha+\beta^* e^{-i\phi}|^2 - e^{-2r}|\alpha-\beta^* e^{-i\phi}|^2\}. \tag{4.73}$$

In the limit of infinite squeezing $r\to\infty$, the quadratures are correlated as $|\alpha+\beta^* e^{-i\phi}| = 0$. For $\phi = 0$, the positions $\sqrt{2}\alpha_r$ and $\sqrt{2}\beta_r$ are perfectly anti-correlated and the momenta $\sqrt{2}\alpha_i$ and $\sqrt{2}\beta_i$ are perfectly correlated. For finite r, the variance of the correlations is $V(|\alpha+\beta^* e^{-i\phi}|) = e^{-2r}$. By measuring any quadrature in one mode, the corresponding correlated quadrature in the other mode can be predicted with an uncertainty smaller than the Heisenberg limit by a factor of e^{-2r}, schematically illustrated in Fig. 4.7. Of course, this does not contradict the Heisenberg uncertainty principle but is a manifestation of the entanglement between the two modes A and B.

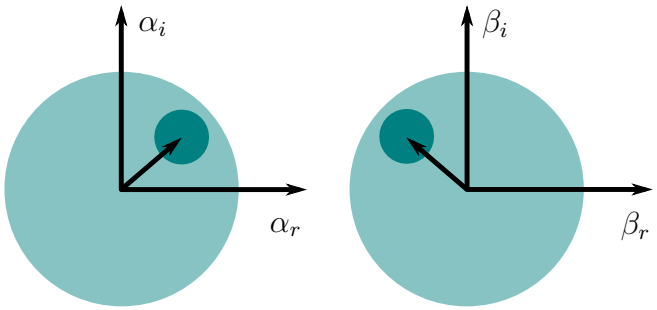


Figure 4.7.: Schematic depiction of a two-mode PDC state. The marginal probability distributions for each mode alone are indicated by the large circles. They correspond to a thermal state, i.e. a 'noisy' light field. The small circles indicate the correlations between the two modes. If both systems are measured in the quadrature pairs (α_r, β_r) or (α_i, β_i), the values will be strongly correlated, indicated by the small circles. However, a non-matching combination like (α_r, β_i) will not show strong correlations. Note that the frame of reference can be rotated by simply introducing a relative phase delay between the two modes.

The famous paper by Einstein, Podolski and Rosen [8] started a long lasting debate about how it might be, that by entangling two particles, it is possible to measure the position of one particle and the momentum of the other particle, and due to their entanglement, conclude the values for both position and momentum to an arbitrary precision. In a modern interpretation of quantum mechanics, this paradox only contradicts local variable theories of quantum mechanics which attempt to describe each system A and B independently of each other. In such a theory, the statistics of each system are predetermined and cannot be influenced by a measurement in the other, spatially separated system. It turns out that it is not possible to formulate such a theory in a convenient way, meaning that any quasiprobability function, in particular the Wigner function, cannot be written as a product of two separate functions for the systems A and B. The different sized circles in Fig. 4.7 attempt to illustrate this in the sense that one gets big or small variations in one mode depending on the choice of measurement in the other mode.

4.3. Nonclassicality

There are several definitions of nonclassicality, since it is such a general term. We would like to mention two popular notions here: nonclassical correlations [52] motivated by information theory and nonclassicality motivated by quasiprobability distributions [51].

The first case is closely related to entanglement. If two systems A and B share some information that cannot be explained by classical information theory, then they share non-classical correlations. A number of specific definitions exist from quantum

discord [53] to EPR entanglement [9, 54], with a hierarchy as sketched in Fig. 4.8. A typical example for an EPR entangled state is the Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, which features perfect correlations in the zz basis as well as in the xx and yy bases. This allows to construct Bell inequalities that witness the EPR entanglement.

However, all these definitions focus on the correlations between two or more systems. The type of states used, is irrelevant for the correlations. For example, classical correlations can be shared by single photons, if they are only used to transmit classical bits. One such example is a mixture of photon-number correlated states:

$$\hat{\rho} = \sum_n p_n |n, n\rangle\langle n, n| \tag{4.74}$$

This state has no off-diagonal terms meaning that it has no phase information. Since it is a mixture of separable states, it features no entanglement. Furthermore it has no quantum discord, as is shown for example in [55]. One could also say that this is due to the fact that we cannot find two non-commuting observables, like amplitude and phase, which both show strong correlations. Only the amplitude-, i.e. photon-number correlations are present, which can be well described classically, by using balls instead of photons, for example. This seems strange because the state can still be used to herald single photons or Fock states in general. And single photons can be send onto a beam splitter to generate entanglement [56]. The output modes of the beam splitter just need to lead to A and B and we have an entangled state again. Hence, with a very common element from an optics lab, a beam splitter, a separable state can be converted into an entangled state. Note that this does not violate any local-operation-and-classical-communication (LOCC) [57] rules, since distributing the state is a non-local operation. But using a beam splitter and sending photons through space does not require quantum operations and is not particularly hard to do, which motivates the following definition of nonclassicality.

The second nonclassicality criterion focuses on the type of state, rather than on correlations between two parties. We have already introduced quasiprobability distributions, which are functions that describe the distribution of observables. Depending on the operator ordering of the observables, there is the Q function, the Wigner function and the P function. It turns out that P is the 'sharpest' of the quasiprobability functions, i.e. if P is positive, the other functions are positive too. Nonclassicality is now defined as P having negative values anywhere in phase space. If P is positive everywhere, then all possible observables of a state can be described by a positive, i.e. classical probability distribution and the state is defined as classical. This rather formal definition has a simple, intuitive interpretation, since the P-function is the expansion of the density matrix in terms of coherent states:

$$\hat{\rho} = \int \mathrm{d}^2\alpha P(\alpha) |\alpha\rangle\langle\alpha|. \tag{4.75}$$

It means that coherent states, represented by a delta-P-function, are the 'narrowest' classical states. All states that can be expressed as a classical mixture of coherent

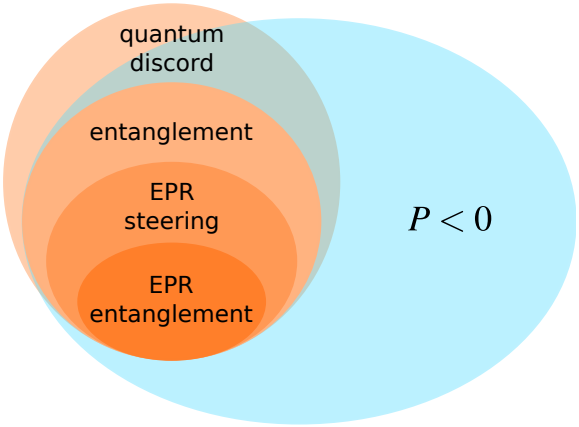


Figure 4.8.: Nonclassicality criteria. Orange definitions are based on information theory, going from the weakest quantum discord to the strongest EPR entanglement (also called Bell-nonlocality). The definition in terms of the P-function stands apart as it is a more fundamental definition, not based on information-theoretical arguments, but on the phase-space representation of the state. In particular, quantum discord and P-nonclassicality are entirely different as there exist states, which are part of one but not of the other. However, entanglement is a subclass of P-nonclassicality.

states are also classical states. If a state cannot be expressed in that way, i.e. $P < 0$ for some α, the state must be non-classical. Single photons for example are very non-classical with a P-function that is not very well behaved, let alone positive. Our example from above, where we used a non-entangled state to produce an entangled state by simple operations, can now be explained by the fact that we used a non-classical resource, namely the single photon. From this perspective, we can also define nonclassical correlations in the quasiprobability framework. If a bipartite state is non-classical, but the local states (i.e. traced out over the other system) are classical, the nonclassicality must lie in the correlations between the two systems. We can call this nonclassical correlations, which is not to be confused with nonclassical correlations in the information-theoretical framework.

In general, the definitions of nonclassicality based on quasiprobability distributions and based on information-theory are inequivalent [55]. We illustrate this point in Fig. 4.8.

4.4. Correlation Functions

Correlation functions, first introduced by Glauber [58], are one way to characterize quantum states. Most generally, the n-th order correlation function for a single mode can be defined by

$$g^{(n)}(t_1, ..., t_n) = \frac{\langle \hat{a}^\dagger(t_1) \cdots \hat{a}^\dagger(t_n) \hat{a}(t_1) \cdots \hat{a}(t_n) \rangle}{\langle \hat{a}^\dagger(t_1) a(t_1) \rangle \cdots \langle \hat{a}^\dagger(t_n) \hat{a}(t_n) \rangle}. \tag{4.76}$$

Intuitively, it might be understood in the following way: We destroy n particles at times $t_1 ... t_n$, create them again at the same times, and calculate the overlap with the original state. The denominator is just a normalization with respect to the mean photon numbers. For example, we consider the second order correlation function (integrating over the second time variable)

$$g^{(2)}(\tau) = \frac{\int \mathrm{d}t \ \langle \hat{a}^\dagger(\tau + t)\hat{a}^\dagger(t)\hat{a}(\tau + t)\hat{a}(t)\rangle}{[\int \mathrm{d}t \ \langle \hat{a}^\dagger(t)\hat{a}(t)\rangle]^2}. \tag{4.77}$$

If the state is a train of single photons, then for $\tau = 0$ the value will be zero, but for $\tau \to \infty$, the value will be one. This is the anti-bunching effect of single photons in a continuous field as emitted by true single photon sources. For coherent continuous wave states, $g^{(2)}(\tau) = 1$ is completely independent of τ, since coherent states are eigenstates of a. And for a train of pulses, i.e. light with periodic amplitude modulations, $g^{(2)}(\tau)$ will contain peaks with the distance of the pulses, as expected from a classical intensity-correlated signal.

In the scope of this thesis, we work with ultra-short pulses of $\sim 1\,\mathrm{ps}$ length and single photon detectors with integration times of about $\sim 1\,\mathrm{ns}$. In that case, we have to integrate the correlation function 4.76 over the whole pulse. It turns out [42], that the resulting correlation function can be written as

$$g^{(n)} = \frac{\langle : (\sum_k \hat{A}_k^\dagger \hat{A}_k)^n : \rangle}{\langle \sum_k \hat{A}_k^\dagger \hat{A}_k \rangle^n}, \tag{4.78}$$

where A_k are broadband operators, usually associated with the Schmidt decomposition of our PDC state, and $: \ldots :$ donates normal ordering. We can now simply plug in the PDC mode operators from the Heisenberg picture,

$$\hat{A}_k^{\mathrm{out}} = \cosh(r_k)\hat{A}_k^{\mathrm{in}} + \sinh(r_k)\hat{B}_k^{\mathrm{in}\dagger} \tag{4.79}$$

$$\hat{B}_k^{\mathrm{out}} = \cosh(r_k)\hat{B}_k^{\mathrm{in}} + \sinh(r_k)\hat{A}_k^{\mathrm{in}\dagger}, \tag{4.80}$$

and calculate all correlation functions for our PDC state. This leads, for example, to the well known result that $g^{(2)} = 1 + \frac{1}{K}$ in the low pump regime, where K is the Schmidt mode number.

From the experimental point of view, we can just write

$$g^{(n)} = \frac{\langle : \hat{N}_{\mathrm{tot}}^n : \rangle}{\langle \hat{N}_{\mathrm{tot}} \rangle^n}, \tag{4.81}$$

where $\hat{N}_{\mathrm{tot}} = \hat{a}^\dagger\hat{a}$ is the total number of photons in all the modes, and $\hat{a} = \sum_k \hat{A}_k = \int \mathrm{d}\omega \ \hat{a}(\omega) = \int \mathrm{d}t \ \hat{a}(t)$. Such correlation functions can be measured directly in an experiment using a multiplexing network with at least n ports and a click-detector in each mode [59]. In Appendix C.1 we show this for $g^{(2)}$.

Correlation functions are closely related to photon-number measurements. In fact, we can use (omitting the hats above the operators)

$$\begin{aligned} a^{\dagger n}a^n &= a^{\dagger(n-1)}(aa^\dagger - 1)a^{n-1} && (4.82)\\ &= a^{\dagger(n-1)}a(a^\dagger a - 1)a^{n-2} && (4.83)\\ &= a^{\dagger(n-1)}a^2(a^\dagger a - 2)a^{n-3} && (4.84)\\ &= a^{\dagger(n-1)}a^{n-1}(a^\dagger a - n + 1). && (4.85) \end{aligned}$$

and hence

$$a^{\dagger n}a^n = \prod_{l=0}^{n-1}(a^\dagger a - l), \tag{4.86}$$

to obtain

$$\langle \hat{a}^{\dagger n}\hat{a}^n \rangle = \sum_k \prod_{l=0}^{n-1}(k-l)p_k, \tag{4.87}$$

where p_k are the photon-number probabilities. That means that if we measure photon-number probabilities, we can directly calculate correlation functions. It also points out that photon probabilities and correlation functions are equivalent descriptions of a part of a quantum state. Nevertheless, there is a fundamental difference between correlation functions and photon-number probabilities: The correlation functions are loss tolerant, due to their normalization. When a state undergoes losses, $\hat{N}_{\mathrm{tot}} \to \eta\hat{N}_{\mathrm{tot}}$, the transmission η cancels out in Eq. 4.81. This is one of the reasons for their popularity in quantum optics experiments. However, the problem of losses is by no means solved by switching to correlation functions. It is merely hidden in the normalization. Losses still impact the precision of a measured value. Especially for higher order correlations, the precision will still scale roughly as η^n, because the probability to register an n-photon event scales roughly as η^n compared to a lossless case. Therefore, statistical and systematic errors will have a stronger impact. For the simplest model of Poissonian statistical errors, the uncertainty for f events is $\sqrt{f}$ and thus for n-photon events, the errors increase by $\eta^{-n/2}$. In this sense, the loss tolerance of correlation functions is deceitful.

4.5. Quantum Measurements

Most measurements in quantum mechanics can be thought of as projective measurements onto a sub-Hilbert space. However, a general description of measurement requires the so called positive operator valued measure (POVM) [54]. It is realized by operators that are not projections, i.e. performing a measurement a second time changes the state again. The probability to measure the outcome associated with a POVM $\hat{\pi}_n$ and a state $\hat{\rho}$ is

$$p_{n\rho} = \mathrm{tr}(\hat{\rho}\hat{\pi}_n). \tag{4.88}$$

The $\hat{\pi}_n$ are Hermitian and sum to unity but are not orthonormal and not projective. Therefore there can be more operators than dimensions of the Hilbert space. The

density matrix after the measurement becomes

$$\hat{\rho}' = \frac{\hat{M}_n \hat{\rho} \hat{M}_n^\dagger}{\mathrm{tr}\ \hat{M}_n \hat{\rho} \hat{M}_n^\dagger}, \tag{4.89}$$

where $\hat{\pi}_n = \hat{M}_n^\dagger \hat{M}_n$. The $\hat{M}$'s are not unique and not Hermitian.

Non-projective measurements can be interpreted as projective measurements on a larger Hilbert space, stated by Neumark's dilation theorem [60]. Since the measurement generally takes place in a larger Hilbert space given by the actual experimental implementation, the number of measurement outcomes can be much larger than the dimensions of $\hat{\rho}$. Restricting the description to the smaller space of $\hat{\rho}$ requires the introduction of the POVMs $\hat{\pi}_n$.

Eq. 4.88 is the starting point of all quantum tomography schemes. In state tomography, the $\hat{\pi}_n$ are assumed to be known and the task is to find the state $\hat{\rho}$ best explaining the measured probabilities $f_{n\rho}$ sampled from $p_{n\rho}$. In detector tomography, a set of states $\hat{\rho}_i$ is assumed to be known and the best fitting $\hat{\pi}_n$ are sought. In general, one can think of quantum tomography as a multidimensional fitting problem. We introduce quantum tomography more thoroughly in Chapter 6.

In the following we would like to elaborate on two basic measurement schemes available in quantum optics, balanced homodyne detection and photon-number resolved detection.

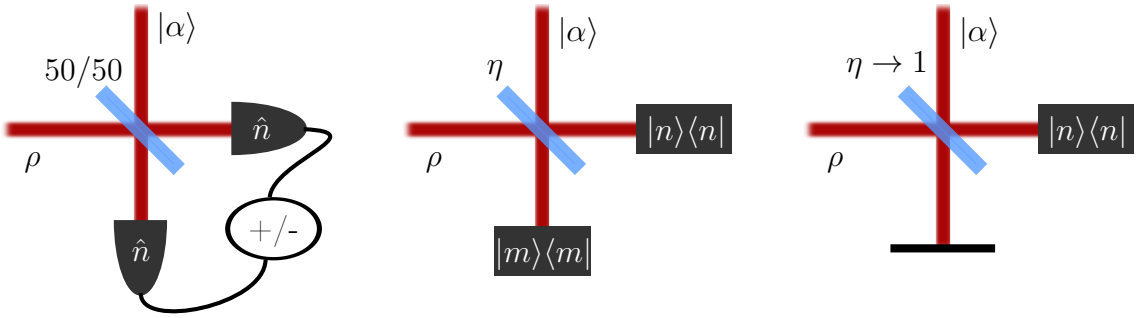


Figure 4.9.: Left: Balanced homodyne detection. A bright coherent field $|\alpha\rangle$ is overlapped at a 50/50 splitter and impinges onto linear photodiodes. Center: Weak-field homodyne detection. A weak coherent field is used with photon-number resolving detectors. Right: Direct Probing. A nearly transmissive beam splitter and only one photon-number resolving detector is used.

4.5.1. Balanced Homodyne Detection

Balanced homodyne detection [61] is the method of choice if one wants to measure the phase-space quadratures $\hat{x}$ and $\hat{p}$ as introduced in Sec. 4.2. A schematic is shown in Fig. 4.9. The advantage of homodyne detection is that a bright local oscillator can be used and only the intensities $\langle n_1 \rangle_\tau$ and $\langle n_1 \rangle_\tau$ have to be measured. The brackets

indicate that all detectors have some integration time τ. In a pulsed system, we could assume that the integration time is much smaller than the pulse distance such that the photon number for each pulse is measured. Even in a continuous wave system, the integration time does not play a role as long as one is not interested in high frequency components beyond the bandwidth $1/\tau$. Hence, we can drop the time filtering and calculate the difference of the two output currents:

$$\begin{aligned}
\hat{S} &= \hat{n}_1 - \hat{n}_2 && (4.90)\\
&= \frac{1}{2}(\hat{a}^\dagger + \hat{b}^\dagger)(\hat{a} + \hat{b}) - \frac{1}{2}(\hat{a}^\dagger - \hat{b}^\dagger)(\hat{a} - \hat{b}) && (4.91)\\
&= \hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a} && (4.92)\\
&= \hat{a}^\dagger\alpha + \hat{a}\alpha^* && (4.93)\\
&= |\alpha|(\hat{a}^\dagger e^{i\phi} + \hat{a}e^{-i\phi}) && (4.94)\\
&= |\alpha|\hat{x}_\phi, && (4.95)
\end{aligned}$$

where we have used the beam splitter transformation, see Appendix A.3. Depending on the phase of the local oscillator, all possible quadratures can be measured, in particular $\hat{x}$ for $\phi = 0$ and $\hat{p}$ for $\phi = i$.

If we have multiple modes in our signal, or signal and local oscillator are not perfectly matched, we can expand the local oscillator in modes of our signal (confer sec. 4.1.3) and obtain

$$\hat{S} = \sum_k \hat{n}_1^{(k)} - \hat{n}_2^{(k)} = \sum_k |\alpha^{(k)}|\hat{x}_\phi^{(k)}. \tag{4.96}$$

The quadratures of each mode are scaled by the corresponding local oscillator. If the local oscillator is matched to the one single mode we are interested in, all other modes do not contribute to the measurement. Similarly, if there is a mismatch, part of the local oscillator will measure vacuum. This is equivalent to introducing losses to the system, as losses are modelled by a beam splitter with a transformation $\hat{x}_\phi \to \tau\hat{x}_\phi + \sigma\hat{x}_\phi^{\text{vac}}$. This corresponds exactly to Eq. 4.96 with $|\alpha^{(0)}| = \tau$ and $|\alpha^{(1)}| = \sigma$. Hence, imperfect mode overlap is indistinguishable from losses in homodyne detection.

The experimental realization of such a homodyne setup [62] is challenging because both output ports have to be balanced perfectly despite bright light on the detectors and sensitive electrical circuitry. For many measurements, however, the aim is not to sample the full quadrature distribution, but only measure its variance [63]. In that case, it is common to apply a frequency filter to the output signal to remove the mean value $|\alpha|\langle x_\phi\rangle$ together with slow varying drifts in the electronics. For continuous wave experiments, those frequencies are usually chosen to lie above the bandwidth of the laser, filtering out most instabilities in the setup. At any time step $\Delta t > \tau_c$, where τ_c is the coherence time of the laser, the quadrature value is completely random, sampled from its probability distribution $p(\hat{x}_\phi) = \int \mathrm{d}x_{\phi+\frac{\pi}{2}} W(x,p)$. Therefore, information like its variance can be deduced at any filtering frequency below the detector bandwidth.

4.5.2. Photon-Number Resolved Detection

Photon-number measurements on the state itself or after an interference with a reference state as depicted in Fig. 4.9 are an alternative to homodyne detection. Let us regard the configuration depicted in Fig. 4.9 on the right. The highly transmissive beam splitter with a coherent state at the second input port acts as a displacement operator $\hat{D}(\alpha')$, where $\alpha' = \alpha\sqrt{1-\eta}$, see Appendix A.3. Therefore, such a detector measures photon-number statistics of the displaced state. The outcome probabilities are given by

$$p_n = \langle n|\hat{D}(\alpha')\hat{\rho}\hat{D}(-\alpha')|n\rangle. \tag{4.97}$$

You might recall, that this is exactly what we need to measure the Wigner function at specific points in phase space, see Eq. 4.62. For this fact, this scheme is named 'direct probing' [64–66] and has been used to reconstruct the Wigner function of heralded single photons [67].

The direct-probing scheme as well as the more general scheme sketched in Fig. 4.9 (center), are also referred to as weak-field homodyne detection [68] in the literature. Such photon-number resolved detection with a reference field, usually using only one port of the beam splitter, has been applied to show the violation of a Bell-type inequality [69], characterize superposition states [70] or probe coherences across Fock layers [71]. The scheme in the central sketch of Fig. 4.9 offers the most flexibility in the choice of reference states and beam-splitter transmittivities. In the limit of very high photon numbers, at which the discrete character of photon numbers can be neglected, this setup becomes a standard homodyne setup.

There are two fundamental differences between homodyne detection and the photon-number based approaches. First, photon-number measurements, in contrast to quadrature measurements, are so-called non-Gaussian measurements. They allow to overcome Gaussian no-go theorems [15], enabling continuous-variable entanglement distillation [72, 73] or the preparation of exotic states like cat states [74, 75]. Second, photon-number measurements are sensitive to all modes. In Eq. 4.96, unmatched modes for which $\alpha^k = 0$ do not contribute to the measurement. This can be seen as an advantage if those modes are considered to be a noise source. With photon-number detectors, all modes are measured because they are not amplified by the local onscillator. The unmatched modes just don't interfere with the reference field. This can also be seen as an advantage, if characterization of the source is the primary goal of the setup. If non-overlapping modes show distinct features from overlapping modes [76], more complete information about the source can be obtained.

4.5.3. Losses

Linear optical loss can be modeled by a beam splitter. If n photons impinge on a beam splitter with transmittivity η, the probability for m photons to pass the beam splitter is given by the binomial distribution

$$L_{mn}(\eta) = \binom{n}{m} \eta^m (1-\eta)^{n-m}. \tag{4.98}$$

The photon-number probabilities before and after losses are related by

$$p'_m = \sum_n L_{mn} p_n. \tag{4.99}$$

For a perfect photon-number detector with an efficiency η, the POVMs in the photon-number basis are given by

$$\hat{\pi}_m = \sum_n L_{mn}(\eta) |n\rangle\langle n|. \tag{4.100}$$

Another common type of detector is the so called on/off or click detector, which is a binary detector sensitive to a single photon but incapable of distinguishing between different numbers of photons. It has two POVM elements which are

$$\hat{\pi}_0 = \sum_n (1-\eta)^n |n\rangle\langle n| \quad \text{and} \quad \hat{\pi}_1 = 1 - \hat{\pi}_0 \tag{4.101}$$

where $(1-\eta)^n$ is the probability to reflect all n photons at the beam splitter.

The matrix L_{mn} can be inverted analytically and reads [76]

$$(L^{-1})_{mn} = \binom{n}{m} \frac{(-1+\eta)^{n-m}}{\eta^n}. \tag{4.102}$$

In principle, no information is lost and losses can be fully inverted by

$$\sum_n (L^{-1})_{mn} p'_n = p_m. \tag{4.103}$$

However, the alternating sign and the η^{-n} scaling already shows that it becomes experimentally unfeasible for very low efficiencies and high photon numbers. This turns out to be a very fundamental problem. Even with more sophisticated tomographic techniques, the photon-number space that can be reconstructed ultimately remains limited by the losses in the experiment.

5. A Source for Quantum Optics

Quantum optics is a broad field with a wide range of applications in communication [77], metrology [78] and computation [79–83]. Without going into detail about specific designs and protocols, we can ask, what the main ingredients are, to achieve a quantum enhancement in some application over classical alternatives.

Most importantly, nonclassicality has to come in at some point. This can be a nonclassical state to start with or a nonclassical measurement that effectively projects onto a nonclassical state.

The second ingredient is high-dimensionality. Many sources have to be entangled with each other to cover a large distance in communication or outscale a classical system in metrology or computation. Since photons do not interact easily with each other, the basic element to achieve high dimensionality is a beam splitter. Even with one beam splitter and two single photons, the Hong-Ou-Mandel (HOM) effect can be demonstrated, where both photons always end up in the same detector due to a genuine quantum effect, namely photon bunching. Most applications in photon based systems rely on such a quantum interference at beam splitters [79, 81, 83, 84].

However, in real systems, imperfections degrade the quantum interference. In particular, distinguishability or mixedness of quantum states are one of the main limitations in quantum applications today. In the simplest case of the HOM interference, distinguishability leads towards the classical 50% chance for both photons to end up in different detectors.

Another main limitation in most applications are losses. In a communication link, photon loss limits the data rate and requires complicated structures like repeaters to compensate for it. In computation, loss is even more critical. The least demanding error-correction codes require losses below 1% [85, 86]. All experimental implementations with optical states fall short of this demand by a large amount. Nevertheless, progress is being made with current system efficiencies close to 80% [12, 13].

Optical states have one significant advantage over all other physical implementations. They do not decohere in the same way as other systems because they do not interact with their environment. While this is certainly a problem when building up a large system, it is an advantage when maintaining a large system and moving it around in the lab or between distant locations. This makes photonic systems excellent candidates to study quantum effects of large or even macroscopic states, provided they can be generated. In principle, optical quantum states with thousands of photons are generated in laboratories today. The challenge is to measure their quantum character and more importantly, access and characterize that huge space. Utilizing higher photon numbers with photon-number detectors is one possible route in that direction.

All in all, the ideal source we envision should generate high dimensional pure quantum states that can be measured with high efficiency. In this chapter we present the implementation and characterization of a PDC source optimized towards that goal.

5.1. Source Engineering

In section 4.1.3 we have introduced the Schmidt decomposition as a means to quantify the spectral properties of a PDC source. Generic PDC sources produce a multitude of spectral Schmidt modes. Each mode contains a two mode squeezed state with orthogonal polarizations, which can be used to split the two modes into channel A and B. Unfortunately, all single photon detectors available today are insensitive to the spectral degree of freedom directly[1]. Suppose we use such a detector on a spectrally multimode PDC source to herald single photons by placing it in channel A. Then after the measurement, confer Sec. 4.1.5, the state in channel B is left as a mixture $\rho_B \sim \sum_k c_k \ |1\rangle_k\langle 1|_k$. Interfering two such states at a beam splitter will reduce the HOM visibility depending on the number of modes and render the source useless for multiplexing in a larger network.

The most elegant solution to this problem is to develop a mode-selective measurement like a quantum pulse gate [41] which selectively operates on single spectral modes by utilizing frequency conversion. Unfortunately, this rather new method still introduces significant losses. Homodyne detection is a more established measurement-based approach, which only measures one spectral mode and ignores all others. However, this limits possible measurements to quadrature measurements and excludes photon-number measurements.

The most simple and widely used approach to achieve single mode operation is spectral filtering [36, 44, 87–92]. The drawback here is that perfect decorrelation can only be achieved in the limit of a small bandwidth, leading to longer pulses and reducing the brightness of the source significantly [93]. Even more critically, spectral filtering unavoidably introduces losses [94].

Our approach is to engineer the PDC source such that it produces only one spectral mode. Then, ideally, no spectral filtering is required [37, 95]. To achieve that, the JSA has to be separable. In the (ω_s, ω_i)-plane, the phasematching function needs to have a positive slope, which in turn depends on the group velocities as follows [96]. To find an expression for the orientation of the phasematching function, we expand the phasematching condition $\Delta_k = 0$ around ω_s^0, ω_i^0 and $\omega_p^0 = \omega_s^0 + \omega_i^0$ and get

$$0 = \Delta k = \frac{\partial k_s}{\partial \omega_s}\Delta\omega_s + \frac{\partial k_i}{\partial \omega_i}\Delta\omega_i - \frac{\partial k_p}{\partial \omega_p}(\Delta\omega_s + \Delta\omega_i), \tag{5.1}$$

[1] This is not quite true for transition edge sensors, since they resolve the energy of the photon. However, the resolution is very low.

where $\Delta\omega = \omega - \omega^0$ and $\frac{\partial k}{\partial \omega} = \frac{1}{v}$ is the inverse of the group velocity. The slope is

$$\frac{\Delta\omega_i}{\Delta\omega_s} = \frac{\frac{1}{v_p} - \frac{1}{v_i}}{\frac{1}{v_s} - \frac{1}{v_p}} = \frac{1 - \frac{v_p}{v_i}}{\frac{v_p}{v_s} - 1}. \tag{5.2}$$

It is positive, if the pump group velocity lies between the signal and idler group velocities:

$$v_s < v_p < v_i \qquad \text{or} \qquad v_i < v_p < v_s \tag{5.3}$$

One suitable material fulfilling this condition is potassium titanyl phosphate ($KTiOPO_4$ or KTP) for wavelengths above 1300 nm, in particular at telecommunication wavelengths around 1550 nm for the process $y \to y + z$, meaning that the pump is y-polarized and signal and idler are y- and z-polarized, respectively [33]. This process is associated with the d_{24} nonlinear coefficient which is approximately $d_{24} = 2.6\,\frac{\mathrm{pm}}{\mathrm{V}} = \chi^{(2)}/2$ [97, 98][2]. If the slope of the phasematching is positive, the width of the pump beam can be chosen to produce a decorrelated JSA.

This principle has first been demonstrated in a different material, potassium dihydrogen phosphate (KDP), at 830 nm using a bulk crystal [99]. Already in this first experiment, very high single-photon purities of 95 % could be achieved. However, the wavelengths were not ideal and signal and idler had very different spectral widths. The process described above was then realized in a KTP waveguide at telecom wavelengths around 1540 nm [100]. Again, good single-mode operation with a Schmidt-mode number of $K = 1.25$, obtained from a marginal $g^{(2)}$ measurement, could be achieved. Here, the increased conversion efficiency due to a waveguide structure was remarkable: A pulse energy of only 75 pJ[3] was sufficient to generate PDC states with a mean photon number of 2.5. In this thesis, we use essentially the same source.

Additionally to decorrelation, indistinguishability between signal and idler is desirable, as it allows both modes to be used in a network. Ideal indistinguishability could be realized if the angle of the phasematching function in the (ω_s, ω_i)-plane was exactly 45 °, corresponding to $v_s + v_i = 2v_p$. In that case, signal and idler would have the same spectral widths. Using the bulk group-velocities of KTP, the corresponding wavelength is at 1585 nm, very close to the ideal telecom-wavelength of 1550 nm. Such decorrelated and degenerate sources in bulk KTP have been realized in different experiments [34, 101–105] and showed remarkable properties in terms of single-photon purities and signal-idler indistinguishabilities.

For a waveguide source such as ours, the wavelengths for perfect indistinguishability are difficult to predict, because the group velocities depend on the waveguide structure. Using a realistic waveguide model[4], it shifts to around 1700 nm, see Fig.

[2] Note that there exist two conventions for the crystal axes of KTP swapping $x \leftrightarrow y$ and hence $d_{24} \leftrightarrow d_{15}$. This leads to some confusion in the literature.

[3] This was measured behind the waveguide, whereas in this thesis we measure pump-pulse energies in front of the waveguide.

[4] The refractive indices were provided by Helge Rütz. They are based on a waveguide model presented in [106] and the commercial eigenmode-solver RSoft.

5.1. The predicted angles are in good agreement with a measured value of 59 ° [96] at 1572 nm. Unfortunately, wavelengths around 1700 nm are incompatible with standard telecommunication technology and the common InGaAs avalanche photodiodes are insensitive in that range. We therefore choose to operate at around 1550 nm and accept a slight distinguishability between signal and idler.

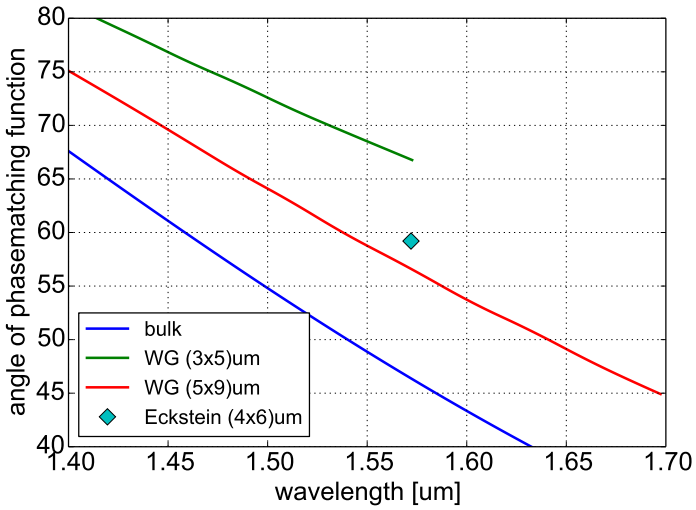


Figure 5.1.: Angle of the phasematching function for degenerate type II PDC in KTP. An angle of 45 ° is of particular interest because signal and idler become perfectly indistinguishable. For our wavelengths around 1550 nm, the angle is around 60 °, strongly dependent on the exact waveguide dimensions. Note that the small waveguide of 3 μm × 5 μm only guides light below $\sim$ 1600 nm.

Concerning the crystal length, we have already seen in sec. 4.1.3 that it does not affect the efficiency of the process. Therefore, we choose experimentally feasible parameters of a crystal length of $L = 8$ mm requiring a pump FWHM of around 2 nm, corresponding to 440 fs long pulses. Shorter crystal lengths are generally less susceptible to imperfections in the waveguide production process. However, too short pulses are more susceptible to dispersion effects if we want to interfere different beams like the signal and local oscillator, even though dispersion does not affect the separability of the JSA in this regime very much [107, 108]. We briefly analyze the effect of dispersion in Appendix B.2. Another important limitation of short pulses is self phase modulation in the crystal, which is a third-order effect in the electric field caused by a non-zero $\chi^{(3)}$ coefficient or cascaded second order processes [109]. Self phase modulation changes the pump spectrum significantly, limiting the usable pump peak-intensities.

5.2. Experimental Setup

5.2.1. Source

A schematic of our experimental setup is shown in Fig. 5.2. The nonlinear medium consists of a periodically poled KTP waveguide engineered to produce decorrelated and degenerate signal and idler modes at 1536 nm. Our chip was commercially purchased from ADVR and has the following parameters: a length of 8 mm, waveguide dimensions of about 4 μm × 6 μm and a poling period of 117 μm. To pump the crystal, we use pulsed light from a Ti:sapphire laser (Chameleon from Coherent) or a frequency doubled beam from an OPO (Chameleon Compact OPO from APE).

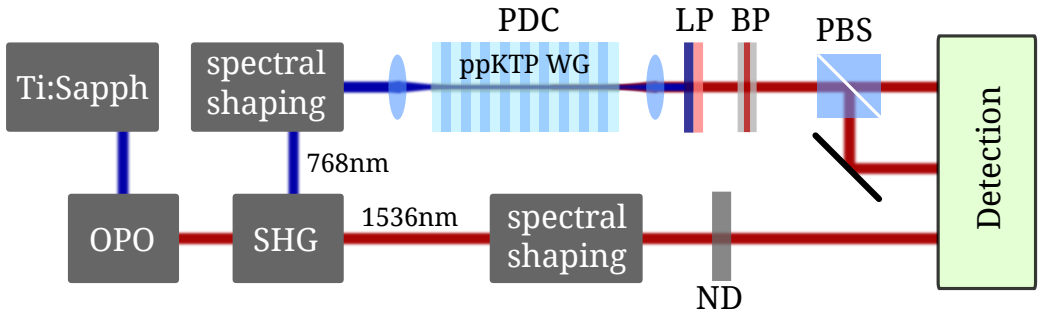


Figure 5.2.: Setup. A Ti:sapphire laser pumps an optical parametric oscillator (OPO), producing pulsed light at 1536 nm. Part of that light is frequency doubled by second harmonic generation (SHG). Both beams are spectrally shaped using a 4-f line to a length of around 0.5 ps. The pump beam is coupled into the periodically poled KTP waveguide. A long pass (LP) filter removes the pump after the down-conversion process and a band pass (BP) filter suppresses the sinc-sidelobes of the phasematching function. Signal and idler are split at a polarizing beam splitter (PBS) and analyzed with different detection configurations.

A crucial part of the experiment is the spectral shaping of the pump. We employ a $4f$-spectrometer, consisting of two gratings, two lenses and one slit in the center, all separated by the focal length of the lenses. The width of the spectrum can be narrowed by reducing the width of the slit. Additionally, the slit is tilted, effectively reducing the resolution of the spectrometer, to produce Gaussian spectral shapes rather than square-like shapes. We adjust the slit to filter the $\sim$ 5 nm pump down to 2 nm (corresponding to 0.45 ps) and the $\sim$ 15 nm reference down to $\sim$ 4 nm (corresponding to 0.9 ps). Note that the pump- and down-conversion wavelengths differ by a factor of two. The spectrum of both, pump and reference beams, is measured and monitored by an optical spectrum analyzer (OSA).

If we use the Ti:sapph directly to pump the source, we can reach pulse energies of up to 2.5 nJ in front of the waveguide.

We use aspheric lenses for in- and outcoupling of the waveguide and into fibers in the detection part of the setup. Behind the waveguide, we employ a longpass

filter (Semrock LP02-808RU-25) to block out the pump beam and a bandpass filter (Semrock NIR01-1535/3-25) with an FWHM of 8 nm. This bandpass filter is the second crucial part of the experiment because the PDC generates background signal over a wide spectral range [110]. The bandpass filter is chosen such that it blocks as much background signal as possible without affecting the PDC signal itself.

5.2.2. Detectors

Before continuing with the characterization of the source, we would like to take a detailed look at the three different types of detectors used in this thesis. These are avalanche photo diodes, superconducting nanowire detectors and transition edge sensors.

Avalanche Photodiodes

An avalanche photodiode (APD) is a diode which is biased high enough for a single photo electron to be accelerated above the energy needed to excite other electrons and cause an avalanche. As soon as the avalanche becomes electrically detectable, a quenching circuit brings the diode back into its original state. APDs are the cheapest single photon detectors available. However, their quantum efficiency is limited especially in the infrared wavelength range. The best commercially available detectors feature up to 30 % efficiency at 1550 nm. To reach these efficiencies they need to be gated, meaning that they are active only in a predefined time-window. There are two types of noise in APDs: Dark counts and afterpulses. Dark counts are probabilistic counts that have not been triggered by an optical signal. Afterpulses are dark counts conditioned on previous count events. In figure 5.3 we plot the click probabilities as a function of time after any click-event. The empirical fit

$$p_{\text{ap}} = \alpha \exp\left\{\frac{n^\beta}{\gamma}\right\} + \delta(1-\delta)^n \tag{5.4}$$

seems to describe these afterpulse-probabilities very well, where n counts the gates since a click-event and $(\alpha, \beta, \gamma, \delta)$ are fit parameters. The last term $\delta(1-\delta)^n$ is the probability to get a normal click after exactly n events, where δ is the click-probability per event. We used two different APD models: The id201 from ID Quantique and the CPDS1000 from Nucrypt. These two devices differ primarily in their allowed repetition rate. The parameters for the fits shown in the plots are

$$\text{id201}: \quad (\alpha, \beta, \gamma, \delta) = (0.064, 0.59, 1.21, 0.010) \tag{5.5}$$

$$\text{CPDS1000}: \quad (\alpha, \beta, \gamma, \delta) = (0.031, 0.40, 1.08, 0.003) \tag{5.6}$$

Dark-count probabilities are usually of the order of 10^{-4} and do not affect most measurements. If they do, they can be modelled quite accurately, since they are independent of the signal. Afterpulses on the other hand occur at a significant fraction of the signal rate and are very hard to model theoretically. They can even dominate over the actual signal in time-multiplexing experiments.

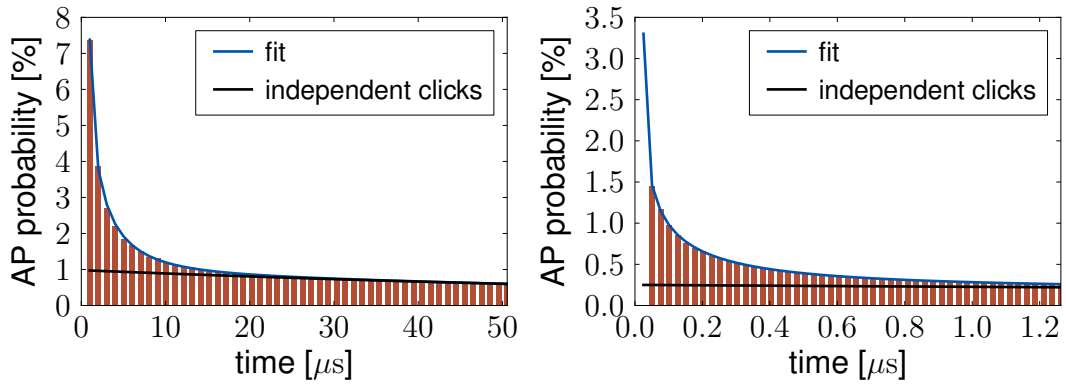


Figure 5.3.: Afterpulse probabilities. Left: idQuantique id201 gated at 1 MHz operating at around 22 % detection efficiency; Right: NuCrypt CPDS-1000-4 gated at 40 MHz with unknown detection efficiency. The total afterpulse-probability for both plots is around 12 %. This value is highly dependent on the biasing and hence the detection efficiency of the detectors. Note that the first bar in the right plot is missing due to details in the electronics of that device.

Superconducting Nanowire Detectors

Superconducting nanowire single-photon detectors (SNSPD) [111, 112] are narrow wires cooled down to about 1 K and biased close to the critical current. If a photon gets absorbed, it locally breaks the superconductivity and creates a normally-conducting region, which expands over the whole wire due to the inductance of the wire. The resulting jump in the resistance can be directly detected electronically. The current drops and stays low for a certain time given by the inductance of the circuit. During that time, the wire cools down and becomes superconducting again.

SNSPDs can have detection efficiencies $> 90\,\%$[112], low jitters < 40 ps and low dead-times < 20 ns. They have no afterpulses and almost no dark counts, behaving like perfect click-detectors. Commercial SNSPDs (Quantum Opus, Photon Spot, SingleQuantum, Scontel, idQuantique) have emerged in the past two years advancing many experiments that crucially depend on detection efficiencies.

Transition Edge Sensors

A transition edge sensor (TES) [113] is a calorimetric detector. It operates at the transition between the superconducting and normally-conducting state, making it very sensitive to tiny temperature variations. By coupling it to a superconducting quantum interference device (SQUID), the energy absorbed in the TES can be directly measured with a resolution below the energy of a single photon. This is usually the case up to about 20 photons. Beyond that point sub-shot-noise resolution can be maintained to about 1000 photons [114]. Typical voltage-traces after amplification are shown in Fig. 5.4. These are actual PDC traces we took at the National Institute of Standards and Technology (NIST) in Boulder, Colorado, USA, in the scope of a

collaboration. The height and shape of the pulses contains information about the

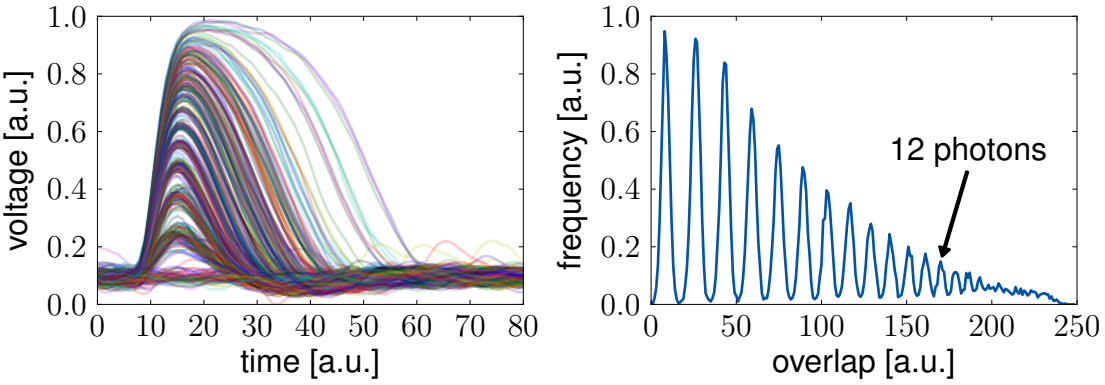


Figure 5.4.: Left: 500 individual traces from a PDC beam at 1535 nm. The discretization of the energy absorbed can be directly seen in the raw pulses. Right: Histogram of overlap-integrals (see main text for details). The calibration trace for this specific histogram was obtained from coherent light with a mean photon number of 12.

energy and hence the number of photons absorbed. Since the traces are relatively noisy, some further analysis is required to assign a photon-number to every voltage pulse with low uncertainty. We choose a method based on an overlap integral with a trace-template $\bar{V}(t)$ [115]

$$O = \int \mathrm{d}t V(t)\bar{V}(t), \tag{5.7}$$

where $V(t)$ is the trace of a particular event and

$$\bar{V}(t) = \frac{1}{N}\sum_{i=1}^{N} V_i(t) \tag{5.8}$$

is the average trace over a large number of events. This trace-template can in principle be taken from the data itself. This works well if the traces are very similar, for example for coherent states with a limited spread of photon numbers. In our case, we expect a very broad distribution due to the photon-number statistics of PDC states with a low Schmidt number. Averaging over all traces has the disadvantage that traces with high photon number, which occur less likely and have a different shape than traces with low photon number, will have less distinct overlap values.

Therefore, we take a multi-template approach where we calculate several templates from the responses to different coherent states. To get a good resolution for all photon numbers, we use 20 different input-power settings for the coherent light with mean photon numbers up to 100, calculate the average traces $\bar{V}_n(t)$ for each of them and assign photon-number bins to overlap values assuming a Poissonian photon-number distribution. The overlap histogram with one such template is shown in Fig. 5.4,

clearly showing the discretization of the overlaps and hence the photon number. Since Poissonian distributions are relatively narrow, this calibration is only reliable around the respective mean photon numbers of its template. For each unknown detection event, we calculate the overlaps with all 20 templates giving 20 photon-number estimations, ideally all the same, and take that particular estimation that is closest to the mean photon number of its template. This method extends the range over which we can reliably resolve photon numbers as compared to the one-template approach and even gives reliable estimations of photon numbers beyond the single-photon resolution regime. The clustering of the overlaps can still be seen up to 20 photons in a histogram, allowing for cross checking the calibration simply by counting peaks. This also allows us to estimate systematic errors coming from changes in the detector response between the calibration measurement and the real measurement. In that case, the peaks of the histogram will not lie in the center of the photon-number-bins. By rescaling the templates slightly, we can deliberately over or underestimate photon numbers, giving an estimate of systematic uncertainties.

5.3. Characterization in the Single-Photon Regime

Most PDC sources are used in the single photon regime, meaning that they are weakly pumped to produce much less than one photon on average per pulse. For this regime, a number of characterization techniques have been established to benchmark the quality of the states, in particular of the heralded single photons. We have already introduced Hong-Ou-Mandel (HOM) interference experiments in Sec. 4.1.6. They can be used to deduce the indistinguishability between signal and idler as well as the purity of heralded single photons, which in turn is defined by the effective Schmidt number.

Alternatively to HOM interference measurements, information about the separability of the state can be inferred from the joint spectral intensity (JSI) $|f(\omega_s, \omega_i)|^2$. This function can be measured by a two-mode spectrometer capable of measuring the wavelength shot-to-shot in a coincidence configuration. We implement this measurement with dispersive fibers and APDs, which we will shortly describe in more detail.

A third measure is the second order correlation function (see Sec. 4.4)

$$g^{(2)} = \frac{\langle a^{\dagger 2} a^2 \rangle}{\langle a^\dagger a \rangle^2} = \frac{\langle n^2 \rangle - \langle n \rangle}{\langle n \rangle^2}. \tag{5.9}$$

It gives some information about the photon-number statistics of the marginal beam. For a decorrelated PDC state, the marginal statistics are thermal with $g^{(2)} = 2$ and for a strongly correlated state, the marginal statistics are Poissonian with $g^{(2)} = 1$ [92]. The relation between the Schmidt number and the marginal Glauber correlation function is given by $g^{(2)}(0) = 1 + 1/K$ [42].

All three methods for source characterization have already been applied in the literature. However, there is no ideal measurement: A JSI measurement is affected by

a low spectral resolution and cannot resolve correlations hidden in the phase of the JSA. Furthermore, it is blind to any non spectral correlations. A $g^{(2)}$ measurement is sensitive to correlations in all degrees of freedom, but is, unfortunately, strongly affected by detector dark counts and non-PDC background signal. The HOM interference measurement requires the exact knowledge of spectral widths of signal and reference fields. Therefore it makes sense to apply all three techniques and carefully compare their results as we do in the following.

5.3.1. Measurement Settings

All the measurements in this section are done with APDs (either Id Quantique id201 or NuCrypt CPDS-1000-4). The repetition rate of the id201 is 1 MHz and of the CPDS-1000-4 40 MHz. We measure raw coincidence vs single ratios, i.e. Klyshko efficiencies [116–119] , defined in Appendix B.3, of up to 20.5% for the signal beam and 15.5% for the idler beam. Corrected for the detector (id201) efficiencies of approximately 25% and 22%, these values correspond to coupling efficiencies into the SMFs of roughly 80% in one arm and 70% in the other arm. All the different measurement configurations are shown in Fig. 5.5.

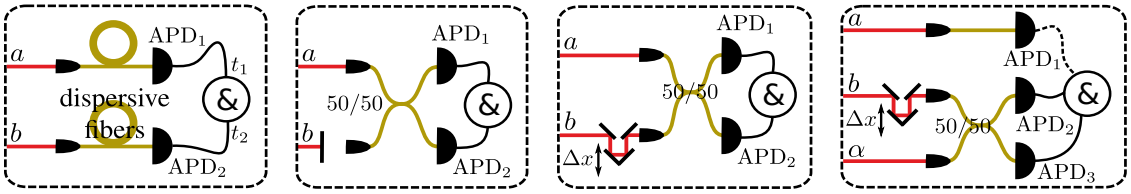


Figure 5.5.: Scheme of different measurement settings. a) JSI measurement; b) $g^{(2)}(0)$ measurement; c) HOM interference between signal and idler; d) 2-fold or 3-fold HOM interference with a reference field. The Id Quantique detectors(1 MHz) are used for the JSI measurement and the NuCrypt detectors (40 MHz) for the rest.

5.3.2. Spectral Intensity

We characterize the spectral properties of our source with a fiber spectrometer [120] (see first frame of Fig. 5.5). Signal and idler travel through dispersive fibers before impinging onto two APDs. Different wavelengths arrive at different times. By scanning the gating times of the APDs, a joint spectral intensity (JSI) distribution is obtained from coincidence click rates. Similarly, the marginal spectral distribution of signal and idler are obtained from the single click rates. The resolution of the fiber spectrometer is limited by the gate width of the APDs and the length of the fiber. We use the id201 with a gate width of approximately 1.5 ns and about 8 km of dispersion-shifted fiber per mode, resulting in a spectral resolution of 1.8 nm for the JSI and 0.9 nm (both fibers combined) for the marginal spectra.

The JSI and the marginal spectral distributions with and without the bandpass filter are shown in Fig. 5.6. The bandpass filter has a width of 8 nm and one can see from the comparison of the unfiltered with the filtered spectrum in Fig. 5.6 that background signal is suppressed directly outside of the PDC range while the PDC spectrum itself is mostly unaffected. The spectra of signal and idler have Gaussian shapes and their widths at FWHM, obtained from Gaussian fits, are 5.2 nm and 4.0 nm, respectively. Assuming that the finite resolution is effectively a convolution with an 0.9 nm wide Gaussian, the true widths can be estimated to be 5.1 nm and 3.9 nm.

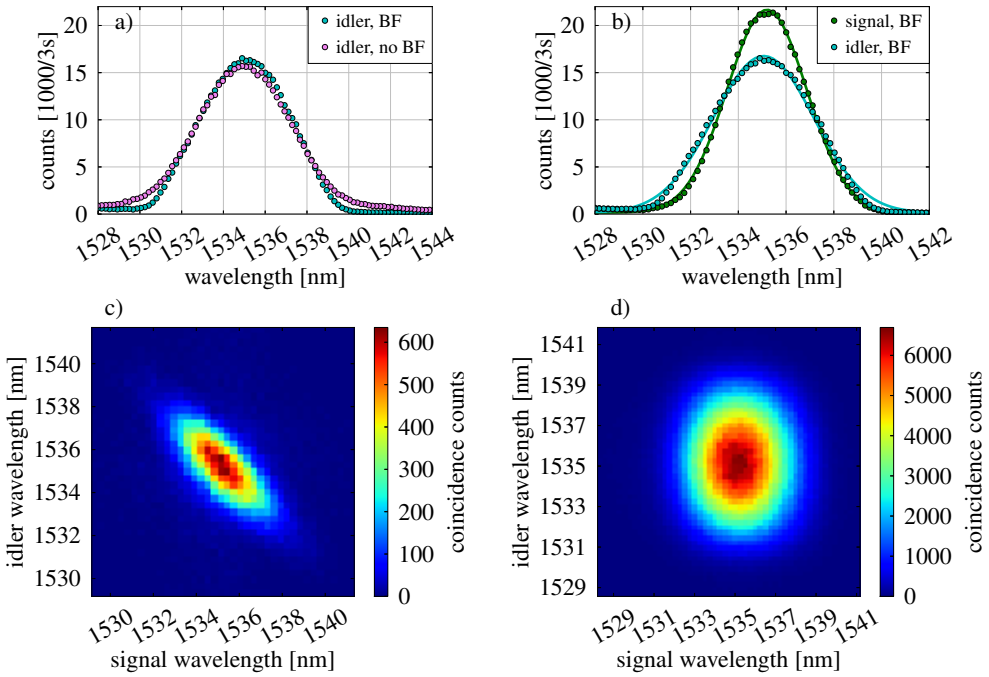


Figure 5.6.: a) Marginal idler spectra for the decorrelated case with and without the bandpass filter (BF). One can see that the spectral filter fits nicely the idler spectrum, touching it only slightly at the very edges. b) Signal and idler spectra with the bandpass filter. Bottom: Measured joint spectral intensity for anticorrelation (c) and decorrelation (d). The pump widths are 0.5 nm and 2.1 nm, respectively.

The JSI measurement in the decorrelated case shows a smooth, slightly elliptical shape as expected from the marginal measurements. The Schmidt decomposition of the JSI yields a Schmidt number of only 1.001. This suggests that the state is indeed decorrelated, even though we have to keep in mind that this measurement gives only a lower bound on the Schmidt number due to the fact that it is only an intensity measurement and the resolution is rather low. However, the smooth shape of the JSA as well as of the marginal spectra supports the assumption of Gaussian phasematching- and pump spectral profiles.

Despite the nice shape of the JSI, there might still be correlations hidden in the phase of the JSA or correlations in other degrees of freedom. To compare this result with a non-spectral measurement, we measure the second order Glauber correlation function $g^{(2)}$ with a 50/50 fiber coupler [92], as depicted in the second frame of Fig. 5.5 and explained in more detail in Appendix C.1. We find raw values of 1.83 ± 0.02[5] and 1.86 ± 0.02 for signal and idler, respectively. These values pose an upper bound on the correlations present in the PDC source. The corresponding Schmidt numbers are 1.20 and 1.16. This seems to be in slight disagreement with the JSI. One possible explanation are background events. From comparison of the filtered with the unfiltered spectra, we conclude that the background still remaining under the PDC spectrum makes up for approximately 1.9 % of the total count rates. Since background signal has a Poissonian photon-number distribution, it strongly degrades the $g^{(2)}$ value. Correcting for these background events [110], we get values of 1.90 and 1.94, which is in better agreement with the JSI measurement. We would like to note that the raw value of $g^{(2)} = 1.86$ is among the highest compared to other PDC sources [92, 100, 121, 122] indicating that the amount of background events is relatively low.

5.3.3. Interference Measurements

To demonstrate the indistinguishability between signal and idler we interfere them at a 50/50 fiber coupler, as sketched in the third frame of Fig. 5.5. The coincidence rate versus a delay of one of the beams is shown in Fig. 5.7. The measurement is done with a pump energy as low as 0.6 pJ per pulse leading to a mean photon number of 0.002. We obtain a visibility of $(94.8 \pm 0.6)\%$, which is among the highest values reported in the literature for PDC sources without narrow-band filtering. It is close to the theoretical value of 96.7% for the decorrelated case. Part of that small deviation from the theoretical value is caused by our fiber coupler which has a slightly uneven coupling ratio of 49.1/50.9. The high agreement between measurement and theory shows that the states overlap perfectly in all degrees of freedom and the indistinguishability between signal and idler is indeed very high.

As discussed in the theory section, a robust method for verifying the decorrelation of the PDC state is to measure the purity of the marginal beams by HOM interference with a reference field. As the reference field, we use part of the original laser beam and attenuate it to the single photon level. All three beams, signal, idler and reference are coupled into single mode fibers and the count rates are recorded with the APDs, as sketched in Fig. 5.5. We record two-fold and three-fold coincidences, where by two-fold we mean coincidence events behind the beam splitter disregarding the third APD and by three-fold we mean triple coincidences between all three output ports. The heralding with the second PDC beam in the three-fold case increases the visibility of the interference [88]. The results are shown in Fig. 5.8. As expected, the visibility in the three-fold case is higher than in the two-fold case. It depends on the spectral

[5] All uncertainties given in this whole Chapter correspond to the 1σ standard deviation.

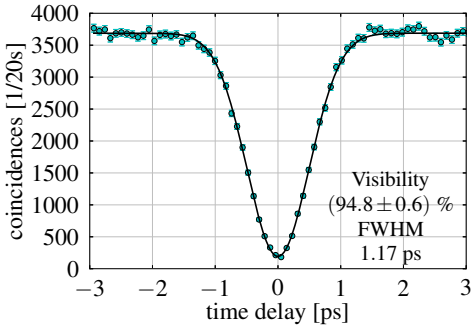


Figure 5.7.: HOM interference between signal and idler.

overlap between the two beams as well as the mean photon numbers which for this measurement are 0.006 for the marginal beam and 0.08 for the reference beam. For calculating the purity of the state, we use the width of the dip rather than the visibility, as discussed in Sec. 4.1.6. Both two-fold and three-fold curves have similar widths of 1.33 ± 0.02 ps and 1.28 ± 0.04 ps. Taking into account the spectral width of the Gaussian reference field of 4.5 nm and the signal spectral width of 3.9 nm, we calculate a purity of 82.1 ± 1.7 % and 86.7 ± 4.3 %. These values are in good agreement with the raw $g^{(2)}(0)$ values. Compared to other PDC sources without narrow-band filtering, Mosley et. al [99] have shown higher purities around 95 % by measuring the HOM interference between two independent PDC sources. Our measured purities are slightly below this value but still in the same range, demonstrating excellent source performance. Similar values have been obtained in [123], who also utilized HOM interference with a coherent field.

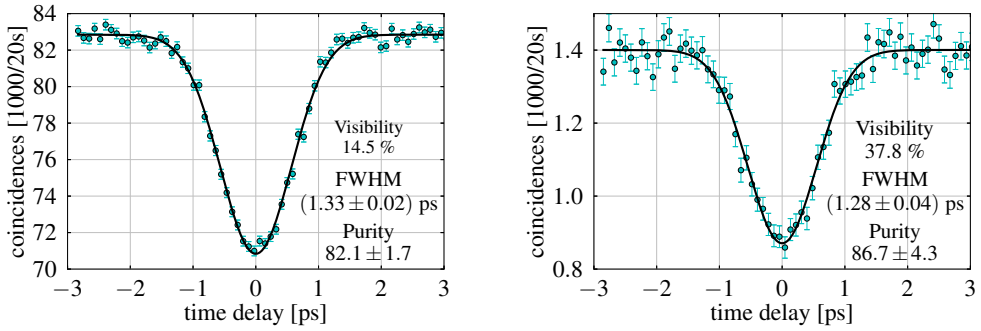


Figure 5.8.: two-fold (left) and three-fold (right) HOM interference between signal and reference.

In conclusion, we have implemented a source with remarkable properties in terms of pump-efficiency, purity and symmetry. The HOM interference between signal and idler showed a high visibility of 94.8 %. The purity values for signal and idler obtained from $g(2)$, JSI and interference with a reference beam reveal a purity of above 80 %.

5.4. Characterization in the Multiphoton Regime

Here we pump the PDC in a regime with mean photon numbers of the order of 20 and measure these numbers shot-to-shot with TES detectors. This experiment was done in collaboration with the National Institute of Standards and Technology (NIST) in Boulder, Colorado, USA. The high photon-number components at these pump powers dramatically increase the size of the Hilbert space from $2^2 = 4$ (no photon or one photon per mode) to $80^2 = 6400$ (up to 80 photons per mode). An interesting question from the experimental point of view is whether the state remains single mode despite the high pump powers. We show that this is indeed the case and characterize the photon-number distribution of the state in various ways in this section.

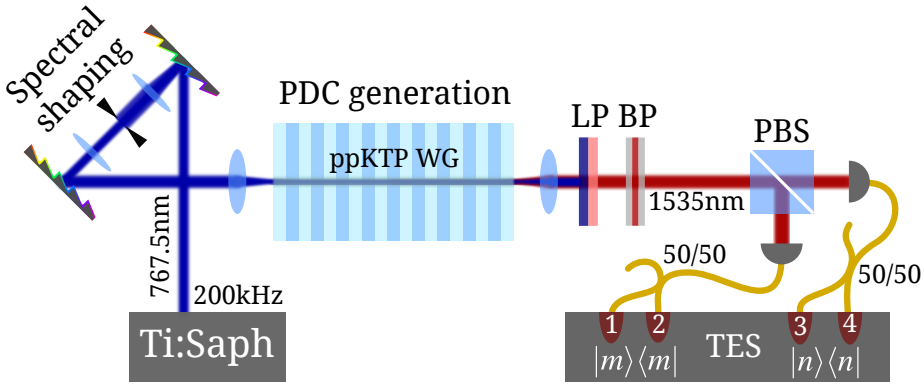


Figure 5.9.: Setup with TES. We pump the process with a Ti:Sapph directly to get high pump powers up to 1.5 nJ per pulse. We use two TES-configurations, where we either send each mode onto one TES or split it at a 50/50 coupler and send onto two TES. In the first case, we get detection efficiencies up to 64 % and 68 %, including all losses in the setup.

Measurement Settings

We employ TES detectors as depicted in Fig. 5.9 to measure the photon numbers of signal and idler. We either use two TES per mode with a 50/50 fiber beam splitter to increase the photon-number resolution at higher photon numbers or one TES per mode for lowest possible losses. The repetition rate of the experiment is around 200 kHz and the system detection efficiencies (as we will determine in Sec. 5.4.2) are up to 64 % and 68 % for signal and idler respectively, including all losses. To analyze the traces of the TES, we apply the scheme described in the detectors-section 5.2.2. We get single-photon resolution up to 10 to 20 photons depending on the detector. Of course, we can also measure higher photon numbers than 20, but with an uncertainty bigger than one photon.

Additionally, we use an APD with known Klyshko-efficiency and calibrated attenuators to measure mean photon numbers up to 80. This different measurement

configuration is due to the fact that it took place in Paderborn, where we had slightly more power available (2.5 nJ rather than 1.5 nJ).

5.4.1. Photon-Number Statistics

The measured photon-number probabilities for a state with a mean photon number of 20 are shown in Fig. 5.10. The statistics are very close to what we expect from a single-mode PDC state. The distribution decays exponentially along the diagonal as can be seen from the logarithmic plot. The vacuum component is still the highest element despite measured mean photon numbers of 11 and 9 in each mode. This directly reveals the single-mode character of the state; for a multimode state, the mixture of different thermal distributions would tend towards a Poissonian distribution as the number of modes increases. For a coherent state, the vacuum component wold be almost zero at these mean photon numbers. To quantify the singlemodeness, we calculate the second order autocorrelation function (confer Sec. 4.4) $g^{(2)} = \frac{\langle n^2\rangle - \langle n\rangle}{\langle n\rangle^2}$, where n is the photon number, on the marginal distribution of each mode. For thermal statistics, $g^{(2)} = 2$ and for Poissonian statistics $g^{(2)} = 1$. For the state shown in Fig. 5.10 we obtain 1.89(3) and 1.87(3) for signal and idler, respectively. These values are slightly higher but in good agreement with the values measured in the single-photon regime. They correspond to effective mode numbers $K = 1/(g^{(2)}(0) - 1)$ of 1.12(4) and 1.15(4), where 1 would be the ideal case.

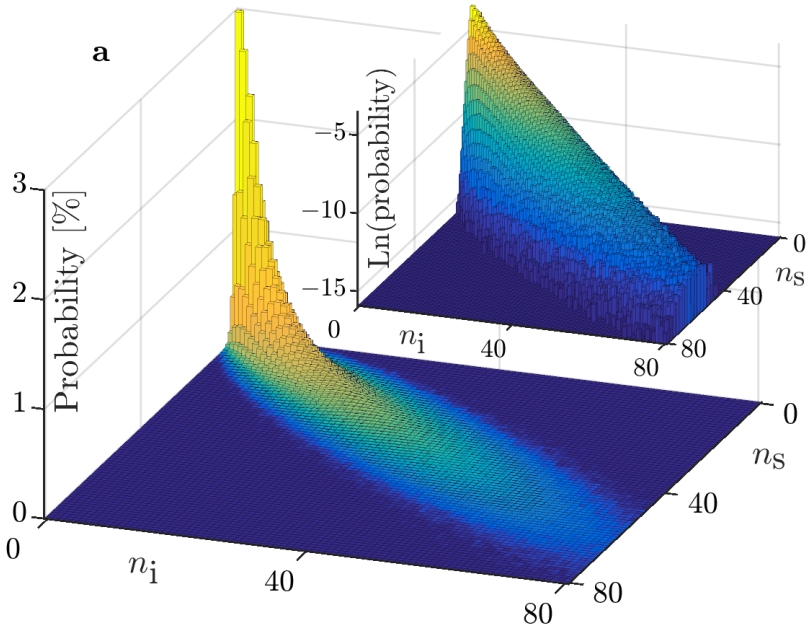


Figure 5.10.: Raw photon-number correlation matrix of the state $\langle n\rangle = 20$ with exponentially decaying diagonal elements. Inset shows the probabilities on a logarithmic scale.

We see no dependence of the effective mode number on pump power. In our

setup with the highest possible pump powers and an APD for mean photon-number measurements, the source generates states with a mean photon number of 80, corresponding to a gain of $r = 2.9$. The mean photon number as a function of pump power is shown in Fig. 5.11 and follows the expected curve of a single mode state. On the one hand, this confirms that the process remains single mode beyond the regime that we can access with our TES setup. On the other hand, mean photon numbers of 80 are already in a regime, where time ordering can be expected to play a role [30, 32, 124] which should cause a deviation from the shown theory curve. Whether an ignored effect like self-phase modulation or broadening of the PDC spectra beyond the filter bandwidth is compensating the mean photon number accidentally, or time ordering only sets on beyond mean photon numbers of 80, remains to be investigated in further measurements.

Coming back to the photon-number distribution, the correlations between signal and idler should be apparent in the width along the diagonal $n_s - n_i$. To encapsulate this criterion, one figure of merit is the noise reduction factor [17] NRF $= \frac{\mathrm{Var}(n_s - n_i)}{\langle n_s + n_i \rangle}$. For ideal PDC with a detection efficiency of η, the NRF is equal to $1-\eta$. We measure values below 0.4, see Fig. 5.11, in those cases where we use one TES on each mode, in agreement with the measured efficiencies of around 66%. This corresponds to 4.2 dB of correlated photon-number squeezing not corrected for losses. In the case where we use two TES on each mode, the NRF is higher due to slightly lower and more asymmetric efficiencies in that configuration.

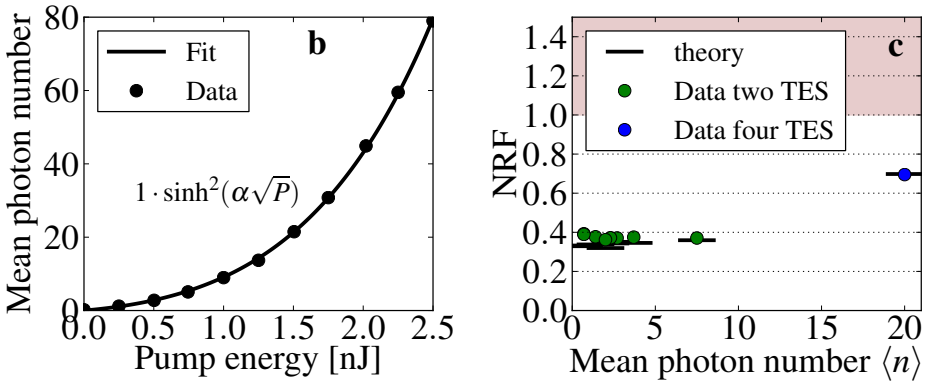


Figure 5.11.: Left: Mean photon number in one mode versus pump power measured with a low efficiency APD. The excellent fit with only one fit parameter α indicates that the state stays single-mode up to at least 80 photons. Right: Noise reduction factor (NRF) for different mean photon numbers, showing the nonclassical correlations of the state. Statistical error bars in both plots are smaller than the data points.

Photon-number correlations can be used to herald states with sub-Poissonian photon-number distributions, by conditioning on a certain outcome in one mode. One figure of merit for sub-Poissonian statistics is the heralded $g^{(2)}$ value, i.e. the $g^{(2)}$ in one mode conditioned on a certain outcome in the other mode. For ideal n-photon Fock states $g^{(2)} = 1-1/n$. Values below 1 indicate nonclassical sub-Poissonian

statistics. Even heralding on a 50-photon event, the measured states fulfill this nonclassicality criterion as shown in Fig. 5.12. With increasing photon number, the transition from strongly nonclassical states to classical states becomes apparent as they become harder to distinguish. Producing larger nonclassical states would require reducing the losses in the heralding mode. At the current efficiencies, the 50 photon event happens about twice per second with a PDC mean photon number of 7.

As a benchmark for the experimental quality of the measurements and the source, we calculate the parity $\langle(-1)^n\rangle$ of heralded Fock states, see Fig. 5.12. Negative parity can only be observed with detection efficiencies above 50 %, in principle. For one and three-photon heralded states, we see negative parities of $-0.131(1)$ and $-0.013(2)$ in the raw heralded data, which again is a sufficient condition for nonclassicality. For higher heralded states, the parity tends to zero and is obscured by statistical errors.

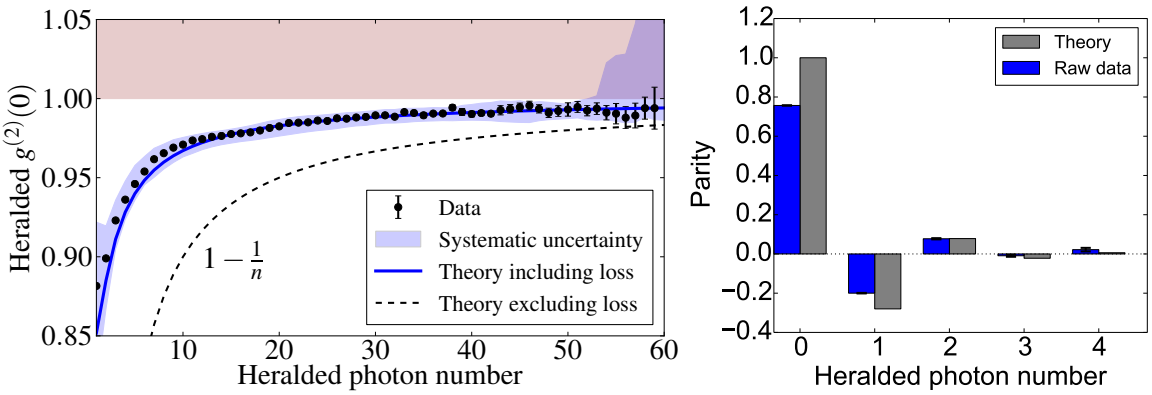


Figure 5.12.: Left: Heralded $g^{(2)}$ as a nonclassicality measure for a state with $\langle n\rangle = 7$. The shaded blue area accounts for worst-case systematic errors stemming from the analysis of the TES response. Error bars are statistical errors. The heralded states stay nonclassical up to around 50 photons. Right: Parity of heralded states. The one- and three-photon states show negative parity, which can only be observed for system efficiencies above 50 %, thus demonstrating the high quality of our data.

5.4.2. Loss Inversion

The excellent agreement with theory indicates that the limiting factor is the loss in our setup. To get a glimpse of how our states would look like without losses, we fit a model to the data. The model consists of a state that can be described as a mixture of a (spectrally) multimode PDC state, a coherent state and a thermal state:

$$\rho = \rho_{\mathrm{PDC}}(n^{\mathrm{PDC}}, K) \otimes \rho_{\alpha}(n_{\mathrm{s}}^{\alpha}, n_{\mathrm{i}}^{\alpha}) \otimes \rho_{\mathrm{th}}(n_{\mathrm{s}}^{\mathrm{th}}, n_{\mathrm{i}}^{\mathrm{th}}), \tag{5.10}$$

where n are the respective mean photon numbers and K the effective mode number of the PDC state. We expect K to be low since the marginal $g^{(2)}$ measurements

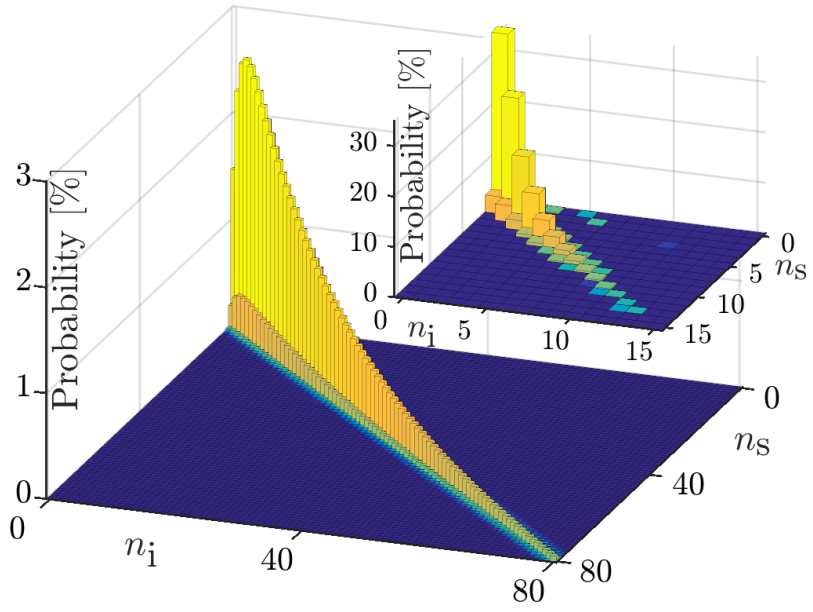


Figure 5.13.: Inferred states before losses. Main graph: High power state ($\langle n \rangle = 20$) using a parameter fit to the data. Inset: Low power state ($\langle n \rangle = 1.4$) using full loss inversion without assumptions about the state.

suggest $K \approx 1.13$. We hence choose exponentially decaying coefficients c_k^2 for each (spectral) mode, whereas $\sum_k c_k^2 = 1$ and $K = 1/\sum_k c_k^4$. Such exponentially decaying coefficients are a reasonable approximation for low effective mode numbers [42].

The losses are described by a standard beam-splitter model with transmissions η_s and η_i in the two beam paths. The photon-number probabilities are given by $p_{kl}^{\mathrm{out}} = \sum_{mn} L_{km}^{\mathrm{s}}(\eta_\mathrm{s}) L_{ln}^{\mathrm{i}}(\eta_\mathrm{i}) p_{mn}^{\mathrm{in}}$, and $L_{kn}(\eta) = \binom{n}{k} \eta^k (1-\eta)^{n-k}$. In Eq. 5.10, the photon-number distributions of the three contributions are independent. That means that the total photon-number distribution p^{in} is a convolution of the three individual distribution. This can be implemented numerically in a straight forward way.

Finally, we minimize the weighted sum of the least square differences

$$\sum_{mn} ((p_{mn}^{\mathrm{meas}} - p_{mn}^{\mathrm{out}})/\sigma_{mn})^2, \tag{5.11}$$

where p_{mn}^{meas} are the measured photon-number probabilities, p_{mn}^{out} the probabilities of the expected state after losses and $\sigma_{mn} = 1/N + \sqrt{p_{mn}^{\mathrm{meas}}/N}$ estimates for the statistical error due to N total events. Effectively, this is a fit with eight parameters $(\eta_s, \eta_i, n_{\mathrm{PDC}}, K, n_\mathrm{s}^\alpha, n_\mathrm{i}^\alpha, n_\mathrm{s}^{\mathrm{th}}, n_\mathrm{i}^{\mathrm{th}})$. Allowing Poissonian and thermal background statistics covers most optical and electrical background signals while keeping the number of free parameters very low.

The fit result for the state with $\langle n \rangle = 20$ is shown in Fig. 5.13 and has the fit parameters

η_s	η_i	n^{PDC}	K	n_s^α	n_i^α	n_s^{th}	n_i^{th}
43.13(3)%	52.12(4)%	20.30(2)	1.097(1)	0.14(12)	0.38(5)	0.00(12)	0.00(5)

The fidelity with the data is 99.98%. The largest contribution in photon number by almost two orders of magnitude is the PDC. Possible sources for the background contributions might be fluorescence or so called nonlinear Cherenkov radiation in waveguides [125, 126], which is a non-collinear process where only one of the down-converted modes is guided spatially and the other mode is radiated into the substrate. Additionally, there might be numerous non-optical sources for Poissonian background statistics.

Compared to our previous measurements, this fit allows us to evaluate the effects of losses and background independently of the PDC state. The standard Klyshko-approach to measure quantum efficiencies assumes perfect photon-number correlations. If background is present, this method underestimates the true detection efficiencies of the setup. In the case of our more efficient two-TES setup configuration with the $\langle n \rangle = 7$ state, the fit suggests efficiencies of 64% and 68%. In comparison, a Klyshko based calculation [127] yields 60% and 64%, which is 4% lower. From the fit-result, we can claim that our maximum efficiencies are up to 68%. The systematic uncertainties on this value, estimated by varying parameters in the TES-trace analysis, are around 3%. The efficiencies for the large state of 43% and 52% are lower, due to an extra pair of fiber beam splitters and probably a worse alignment of the setup.

The mode number $K = 1.10$ given by the fit is also not affected by the background. We can compare this number to the other measurements of K in this chapter, based on the marginal $g^{(2)}$ in the high-power regime with $\langle n \rangle = 20$ as well as in the single photon regime with $\langle n \rangle < 1$, and based on the interference measurement with a local oscillator in the single photon regime:

	$\langle n \rangle < 1$	$\langle n \rangle = 20$
$g^{(2)} \rightarrow K$	1.16(3) and 1.20(3)	1.126(40) and 1.143(40)
purity $\rightarrow K$ (interf. with ref.)	1.15(6) and 1.22(3)	-
fit $\rightarrow K$	-	1.097(1)
fit $\rightarrow g^{(2)} \rightarrow K$	-	1.116 and 1.142

These numbers are surprisingly consistent with each other and the differences can be explained by the presence of background. For the bright state, we can calculate the $g^{(2)}$ of the loss-inverted state and conclude that the background increases the inferred K slightly, by 0.019 and 0.045, consistent with the values inferred from the $g^{(2)}$ of the raw statistics. In the single photon regime the effect of background can be expected to be larger, simply due to the exponential scaling of the PDC photon numbers with pump power opposed to a presumably linear scaling of the background.

As a side remark, we also perform a general loss inversion [128] with physical constraints but no assumptions about the photon-number distributions. More details about the method will be given in Sec. 6.1 and Appendix C.2. This method is limited to low-power states due to the large parameter space. For a state with $\langle n \rangle = 1.4$,

shown in Fig. 5.13(inset), we restrict the photon number space to < 15 photons, resulting in $15^2 - 1$ free parameters. Again, the inverted state resembles the expected PDC state very well.

5.4.3. Potential in Continuous Variable Applications

In continuous variable experiments, squeezing [29] is the most important resource. The main limitations for squeezing are usually pump power, noise, overlap with a reference field and optical loss. From our photon-number distribution we know that pump power and noise are not the limiting factors. If we assume that our spectra are mainly Gaussian, a high overlap with a reference beam should also be achievable in a realistic setup. The main limitation is thus the optical loss of 32% in the current setup. A direct measurement of continuous variable (CV) squeezing with these values, assuming a good mode overlap, should show above 4 dB of two mode squeezing. For comparison, the highest squeezing directly measured in a single pass, pulsed system is 5 dB [129] and in a continuous-wave cavity system 12.7 dB [63]. Measuring high squeezing would be the first step if we really want to combine several such sources to form very large quantum states.

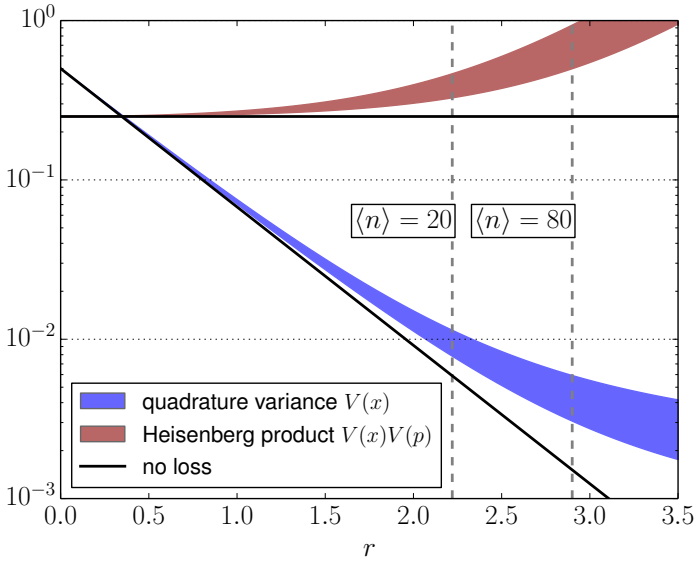


Figure 5.14.: Squeezing of a type I source with uniform losses along the waveguide. We assume here losses between 3.6 % (0.2 dB/cm) and 10 % (0.66 dB/cm). At a mean photon number of 20, squeezing of 17 dB is realistic. With increasing pump power, the Heisenberg limit $V(x)v(y) \geq 1/4$ becomes increasingly exceeded. However, the squeezing is not limited in principle. Even 25 dB of squeezing might be experimentally feasible.

To improve the performance of the source at this stage would require to reduce the losses further. The main loss contributions in our setup come from the coupling

to single-mode fibers of around 80% and 85% for the two ports and the linear optical elements with a total transmission of about 90%. With on chip integration of polarizing beam splitters and detectors, of which both have been demonstrated [130, 131], the total efficiencies could go up to above 90%. This would push the size of possible nonclassical states to hundreds of photons and allow world record squeezing measurements. The ultimate goal would be an efficiency around 99%, allowing 20 dB of squeezing, at which fault tolerant quantum computation with CV cluster states becomes possible [86].

Ultimately, the squeezing producible with the source is limited by the losses inside the crystal. One might expect that this sets a bound on the achievable squeezing similar to cavity-based systems where the limit is given by the round-trip loss. For a single-pass source however, such a hard limit is not present [132]. In Fig. 5.14 we plot the expected squeezing for type I PDC for realistic loss parameters, using the equation given in [132]. Even though it is a type I process, we expect the corresponding plot for a type II process to be very similar. We conclude that squeezing above 20 dB is experimentally feasible, provided that the detection itself allows to measure such values.

5.5. Conclusion

We have shown that our ppKTP source is capable of producing near single mode states with about 1.1 spectral modes and heralded single photon purities around 85 %. Remarkably, the single mode character remains unchanged in the high-pump regime. With state-of-the-art photon-number detectors, we demonstrated that the higher photon numbers of the state are accessible, drastically increasing the dimensions per optical mode. We found that the performance of the setup, e.g. the size of nonclassical states or potential squeezing, is limited mainly by optical loss. Even though we achieved detection efficiencies up to 68 %, further improvement is possible and constitutes the next step in the realization of very large optical quantum states in such a pulsed, photon-number based implementation.

6. Characterization of Quantum States

6.1. Introduction to Tomography

In imaging, tomography is the process of reconstructing an object from projective measurements. Say a two dimensional object, like a slice through the human body, is described by a density $f(x, y)$. Then a transmission measurement under some angle θ in the (x, y) plane, e.g. an X-ray scan, gives information about the integrated density

$$g(\theta, y) = \int \mathrm{d}x f_\theta(x, y), \tag{6.1}$$

where $f_\theta(x, y)$ is rotated by θ with respect to $f(x, y)$. The map $f(x, y) \to g(\theta, y)$ is known as the Radon transform. By taking measurements from all possible angles, no information is lost and the image can be reconstructed by the inverse Radon transform.

Homodyne detection is exactly analogous as it projects the Wigner function $W(x, y)$ onto a quadrature x_θ. Explicitly, distributions

$$g(q, \theta) = \int_{-\infty}^{\infty} W(q\cos\theta - p\sin\theta, q\sin\theta + p\cos\theta)\mathrm{d}p \tag{6.2}$$

are measured and the inverse Radon transform can be written as [61]

$$W(q, p) = -\frac{\mathcal{P}}{2\pi^2}\int_0^\pi \int_{-\infty}^{\infty} \frac{g(q, \theta)}{(q\cos\theta + p\sin\theta - x)^2}\ \mathrm{d}x\ \mathrm{d}\theta, \tag{6.3}$$

where $\mathcal{P}$ is the Cauchy's principal value. The numerical implementation of this formula is challenging due to the singularity such that some filtering of the data is always required. However, algorithms from classical imaging can be readily applied. See e.g. [62] for the first tomography of squeezed states. The inverse Radon transform is one example from the toolbox of continuous-variable optical quantum state tomography [133].

For the description of a general measurement, including homodyne tomography, we have to start from

$$p_n = \mathrm{tr}(\rho\pi_n), \tag{6.4}$$

where ρ is the density matrix of the quantum state, π_n the POVM operators and p_n the probability to get outcome n. In an experiment, the measurement is performed

N times, resulting in a measured distribution f_n, sampled from the real probability distribution p_n. The distribution f_n has statical uncertainties due to a final number of measurements and systematic errors resulting from a non-ideal setup, or in other words, an inaccurate description of the measurement apparatus. Now, what is a good estimation for the state ρ that produced the measured distribution f_n?

There are various different approaches to answer this question and estimate the errors of the estimation. A nice comparison of different methods for the reconstruction of a single qubit, the simplest possible quantum state, is given in [134]. Maximum likelihood approaches [135, 136] are conceptually the best, since they search for the most probable state, given the measured data. It has been applied for time multiplexed photon-number measurements in e.g. [137]. We however prefer the least square approach, mainly for its simplicity and straight forward numerical implementation. We follow the formulation in [128].

For state tomography, the least square approach is the minimization of the distance

$$\underset{\rho}{\text{minimize}} \{||\mathbf{f} - \mathbf{p}(\rho)||\}, \tag{6.5}$$

with the standard L^2 norm $||\mathbf{a}|| = \sqrt{\sum_i (|a_i|^2)}$. $\mathbf{f}$ is the vector of the measured probabilities f_n and $\mathbf{p}$ is the vector of the theoretical predictions $p_n = \text{tr}(\rho\pi_n)$. They can be expressed in matrix form by choosing any specific basis, e.g. the photon-number basis

$$\rho = \sum_{ij} r_{ij}|i\rangle\langle j| \qquad \pi_n = \sum_{kl} \theta^n_{kl}|k\rangle\langle l|. \tag{6.6}$$

Then

$$p_n = \text{tr}(\rho\pi_n) = \sum_{ij} r_{ij}\theta^n_{ji}. \tag{6.7}$$

For detectors, which are insensitive to the optical phase, the POVMs are diagonal in the photon-number basis, which reduces the state to a vector r_{ii}. In any case, this least square approach is a convex optimization problem [138], since the L^2 norm is a convex function. The physical constraints are that $\mathbf{r}$ is positive semi-definite and has a trace of one. Such convex optimization problems with constraints are well known and can be solved very efficiently.

Detector tomography can be formulated in a very similar way:

$$\underset{\Theta}{\text{minimize}} \{||\mathbf{f} - \mathbf{p}(\Theta)||\}, \tag{6.8}$$

where Θ is a tensor with elements θ^n_{ij}. Here, we are looking for POVMs π_n, assuming a set of known states ρ_m. The physical constraints are that $\sum_n \pi_n = 1$ and $\pi_n > 0$. Note that each π_n has the dimension of ρ. Compared to state tomography, the size of the problem (number of elements in $\mathbf{f}$) becomes roughly squared.

Since the number of free parameters can become very large, especially in detector tomography, it is beneficial to incorporate some of our knowledge about the physical implementation into the problem. When we did the loss inversion in section 5.4.2, we reduced the number of parameters from 6400 to eight, by restricting ourselves to

particular states that we expect to be present. This is an extreme example that might not deserve the term tomography any more but is better described as a least square fit of a model to the data. A more general example of using additional information, is detector tomography with lossy photon-number detectors. We know that their probabilities should be distributed smoothly, meaning that the probability to detect e.g. 4 photons conditioned on an input of 6 photons should not be too different from the same event conditioned on an input of 7 photons. In [128] the idea was put forward to use a regularization term to penalize such unrealistic results. The authors added a cost term $S = \sum_k (\theta_k^n - \theta_{k+1}^n)^2$, which penalizes large differences between neighboring POVM elements. Then

$$||\mathbf{f} - \mathbf{p}(\Theta)|| + \gamma S \tag{6.9}$$

needs to be minimized. The factor γ should be chosen as small as possible such that it does not affect the result too much. The regularization term helps to select the most realistic solution out of many, almost equally minimizing, solutions. Our numerical implementation of this least square approach is further discussed in Appendix C.2. In the following, we use this method to characterize a time multiplexing detector.

6.2. Time-Multiplexed Detection with InGaAs APDs

To obtain photon-number resolution with 'click' detectors, a common approach is to split the signal pulse into several spatial or temporal modes and count the clicks. Multiplexing in time [137, 139, 140] is particularly favorable, because only one or two detectors are sufficient. Such a scheme is shown in Fig. 6.1 with a total of eight bins.

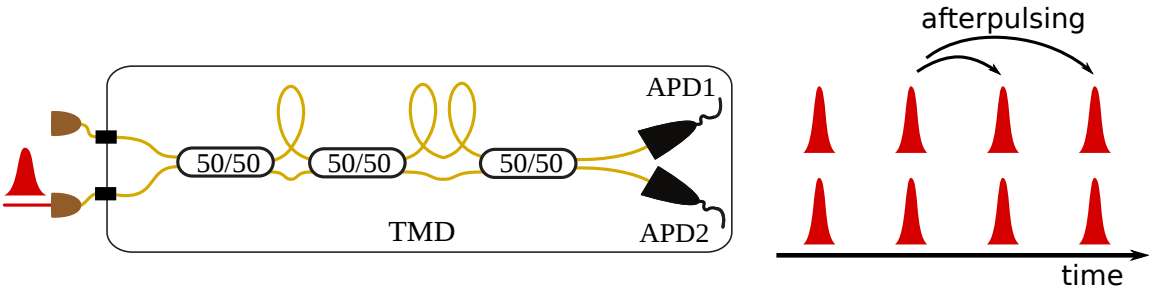


Figure 6.1.: Schematic of a fiber-integrated time-multiplexing detector. One input pulse is distributed into 8 bins by beam splitters and delay lines. With ideal detectors, photons are distributed statistically into the eight bins. However, in particular for InGaAs APDs, dark counts and afterpulses are detrimental effects that cannot be neglected.

Since there is no interference happening, the photons can be treated as independent classical particles. If P_m is the probability that m photons enter the TMD, we can

relate the output probability by a linear model

$$p_\alpha = \sum_{m=0}^{D-1} M_{\alpha m} P_m, \tag{6.10}$$

where D is a cutoff dimension to keep the state-space finite. The outcome distribution p_α can be any detector response we choose. For example, we can choose to sum up the number of clicks. In that case, for a TMD with n bins, there are $n+1$ indices α. We can also take the full detector response, where each bin takes a value of zero or one. In that case there are 2^n responses α.

The matrix M contains all information about the TMD, including afterpulses and dark-counts. However, non-linear effects with respect to the photon-number distribution are not included. One such example would be a change of detection efficiencies with increasing click-rates. This can happen if the APDs do not have enough time to recover between clicks. The afterpulse probabilities shown in Fig. 5.3 indicate that even though the APDs can be gated with a gate distance of $1\,\mu$s, some relaxation process takes up to $20\,\mu$s. If the average distance between two clicks is significantly lower than that, such effects might build up and change the parameters of the APD. Therefore, the bin distances of the TMD as well as the overall repetition rate of the experiment must be chosen carefully to avoid saturation effects and stay in a regime where the linear model is accurate.

For a theoretical model of M, let us assume that we do not have any dark-counts or afterpulses. Further, we take the number of clicks as our detector response. Then the TMD model, introduced in [139, 141], has to take into account only two effects: losses and the statistical distribution into bins. Surprisingly, losses occurring at any stage in the TMD can be modeled by a single beam splitter in front of the TMD. This can be described by the matrix equation $P_i(\eta) = \sum_j L_{ij} P_j$, with the binomial distribution

$$L_{ij} = \binom{j}{i} \eta^i (1-\eta)^{(j-i)}. \tag{6.11}$$

The statistical distribution of m photons into n bins is a more involved combinatorial problem. For equal bin-probabilities, an expression can be found in [140]. For unequal bin-probabilities please see [141] and [142]. This results in an upper diagonal matrix C, such that the measured click-probabilities are given by

$$p_n^{\text{click}} = \sum_{km} C_{nk} L_{km} P_m. \tag{6.12}$$

The model 6.12 works well for detectors like SNSPDs or Si-APDs which show no or very low afterpulsing. In principle, afterpulsing can be incorporated into the model as well, because it is again a combinatorial problem, however, a very complicated one. Instead, we perform detector tomography as described above. We use a set of well calibrated coherent input states and reconstruct the POVMs. The result is shown in Fig. 6.2. The similarity between the reconstruction and the CL-model is strikingly

high. However, when we look at the difference, there is a systematic shift to higher click numbers. This directly shows the effect of afterpulsing on the click statistics. Whether or not this renders the standard model inaccurate depends on the goal of a specific measurement. After all, the absolute deviations are only of the order of 1 %. Nevertheless, for reconstructions of full photon statistics, which are very sensitive to small errors due to the low efficiency of the detector, such errors might already be too large. In that case, detector tomography becomes necessary.

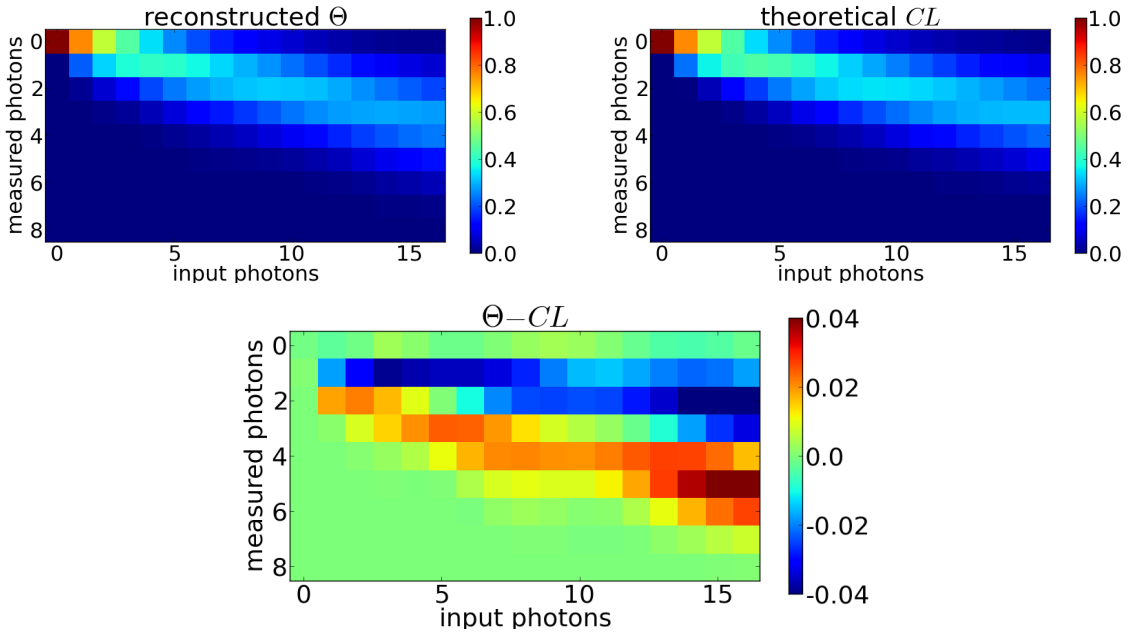


Figure 6.2.: Detector tomography of an eight-bin TMD. Color-coded are the probabilities to count a certain number of clicks (measured photons) conditioned on a certain number of input photons. Top-left: Reconstructed POVM elements. Top-right: Model-POVM elements for an indeal TMD. Bottom: The difference of the two. The shift to higher click counts can be explained by the afterpulsing effect. For example, there is a chance to get two clicks, if just one single photon was present.

6.3. Pattern Tomography

6.3.1. Introduction

There is a large number of publications on state tomography and detector tomography. Interestingly, these two types are often performed independently of each other, probably because the detectors can be described sufficiently well by a model and the detector tomography only functions to verify that. However, restricting the class

of detectors to such that can be described well by a model unnecessarily reduces available resources.

If we want to include all detectors to our toolbox, in principle, we have to perform detector tomography first before we use them for state tomography. Not only is this one additional step, but a much more complex one due to the huge parameter space of detector tomography. Pattern tomography [143–145], is an answer to exactly this question: How to do state tomography with unknown detectors? It works as follows. We start with a limited number of probe states $\sigma^{(\xi)}$, similar to detector tomography, to measure the detector responses to these states with outcome probabilities

$$p_i^{(\xi)} = \mathrm{tr}(\Pi_i \sigma^{(\xi)}). \tag{6.13}$$

These responses are called patterns. The trick is now to expand the unknown state in the same basis of probe states,

$$\rho \approx \sum_\xi c_\xi \sigma^{(\xi)}. \tag{6.14}$$

In the experimentally favorable scenario, where $\sigma^{(\xi)}$ are coherent states, the obtained coefficients c_ξ correspond to a discrete version of the Glauber-Sudarshan P-function. The more probe states are used, the closer one gets to the actual P-function of the original state ρ. Now we can express the response of the detector to this state by

$$p_i^{(\rho)} = \sum_\xi c_\xi \mathrm{tr}(\Pi_i \sigma^{(\xi)}) = \sum_\xi c_\xi p_i^{(\xi)}. \tag{6.15}$$

Due to a finite number of measurements, we measure the distributions $f_i^{(\rho)}$ and $f_i^{(\xi)}$, drawn from $p_i^{(\rho)}$ and $p_i^{(\xi)}$. Therefore we get

$$f_i^{(\rho)} \simeq \sum_\xi c_\xi f_i^{(\xi)}. \tag{6.16}$$

From this equation, we can find coefficients c_ξ, and thus the state, under the physicality constraints that ρ is positive semidefinite and $\mathrm{tr}(\rho) = 1$. One way to do this is to minimize the sum of least squares

$$\sum_i |f_i^{(\rho)} - \sum_\xi c_\xi f_i^{(\xi)}|^2 \tag{6.17}$$

with appropriate constraints. In [4], our theory collaboraturs and we present in detail an efficient algorithm and illustrate it with an example of homodyne detection.

Pattern tomography has two main advantages over a two-step approach. Firstly, the detectors don't have to be fully characterized, significantly reducing the required number of probe states. Secondly, the probe states give a natural truncation of the Hilbert space. It is not necessary to artificially cut the Hilbert space at a certain photon number, thus avoiding systematic errors or an overly large Hilbert space.

6.3.2. Partial Tomography as Pattern Tomography

Partial tomography is the estimation of a subset of parameters of the full state. Ideally, such an estimation requires a much lower number of measurements compared to full state tomography. For example, one could be interested in only a few photon-number components of a two-mode state. To check the strength of the correlation between signal and idler, only one anti-diagonal line of probabilities ρ_{mn} with $m+n =$ const might be sufficient. Another example is the parity of a single-mode state $S = \sum_n (-1)^n \rho_n$, where $\rho_n = \langle n|\rho|n\rangle$. The parity is an extreme case as it is only a single parameter to estimate.

Pattern tomography allows to perform partial tomography in the same way as full tomography. To illustrate this, let us assume that we are only interested in the first four photon-number components ρ_n with $n < 4$ of a single-mode state. These components can be expressed in the photon-number components of the probe states:

$$\rho_n = \sum_\xi c_\xi \sigma_n^{(\xi)} \tag{6.18}$$

Choosing four linearly independent probe states $\sigma^{(\xi)}$ with photon numbers below four is sufficient to find ρ_n in this subspace. Finding the coefficients c_ξ is done in the same way as in full tomography, for example by minimizing the least squares in Eq. 6.17. The fact that we have much more measurement outcomes than free parameters is not a problem and actually the more common case for least-square fits. Interestingly, the state itself can live in a much larger photon-number space. This only means that we cannot use the constraint that all probabilities sum to one. However, the probe states should be confined to the subspace of interest as good as possible. This is crucial to obtain an accurate estimation. Otherwise, we will be fitting the higher photon-number components as well, smearing out our information about the subset of interest.

Applying this to parity estimation, we get

$$S = \sum_\xi c_\xi \sum_n (-1)^n \sigma_n^{(\xi)}. \tag{6.19}$$

In this case only two probe states are sufficient, one with parity $+1$ and one with parity -1. These probe states should have equal photon-number probabilities in the full photon-number space of the state. Unfortunately, such probe states are very hard to come by. If we use coherent states instead, we have to use much more than two to sample the whole space sufficiently.

6.3.3. Application to Two-Mode States

Experimental Implementation

In the following, we apply pattern tomography to two-mode states, using a two-mode TMD, sketched in Fig. 6.3.

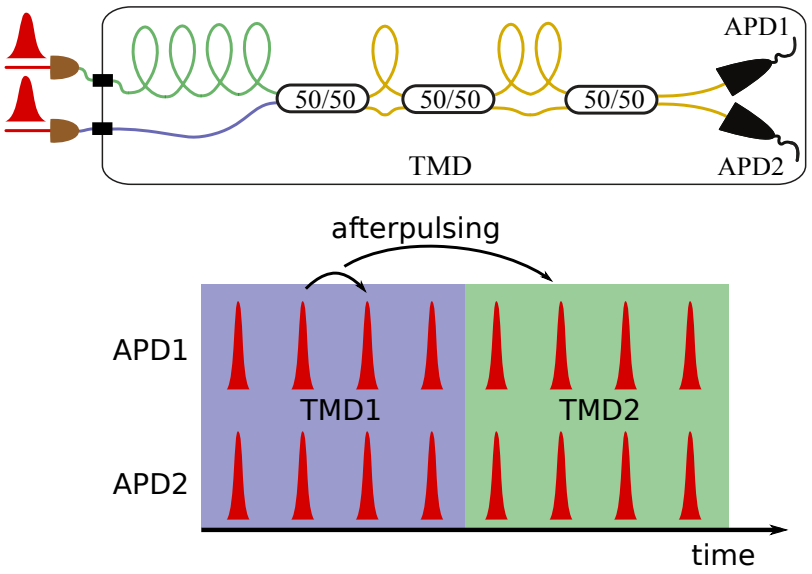


Figure 6.3.: Schematic of a two-mode TMD. Both inputs of the TMD are used, whereas one is delayed by a large amount to arrive at a later time. The purple shaded pulses correspond to the first pulse and the green shaded to the second pulse.

The model for a two-mode TMD is a straight forward extension of Eq. 6.10:

$$p_{\alpha\beta} = \sum_{m,n} M_{\alpha m \beta n} P_{mn} \tag{6.20}$$

Due to the specific implementation, where the signals of the two modes arrive consecutively at the same APDs, and the afterpulse effect, the response cannot be factorized into two independent TMDs. This complicates the analysis of the data but minimizes the experimental resources: We still need only two APDs for two-mode photon-number measurements. The details of the experimental setup are shown in Fig. 6.4.

The experiment is performed in two steps. First, a set of coherent states is measured. Second, the quantum states are measured with the same detector settings. Generating coherent states at the single photon level with known mean-photon numbers is a non-trivial task. We use the simple approach of measuring a reference power with a standard power meter and attenuating the beam with calibrated neutral density (ND) filters by roughly nine orders of magnitude.

For the attenuation, we tested two types of ND filters: free space and fiber-coupled. We observed that some of the free space filters introduce a significant beam displacement and thus change the fiber-coupling behind the filter. We thus had to perform the calibration by measuring the power behind the fiber and take care to place the attenuators in exactly the same way for the actual experiment. The fiber-coupled attenuators did not suffer from this difficulties and are therefore preferable. An-

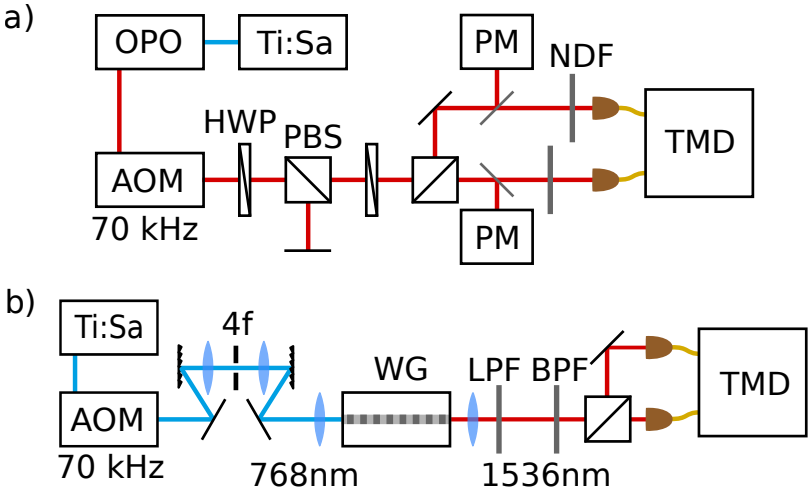


Figure 6.4.: The two setups for pattern tomography. (a) Probe-state generation and measurement. Coherent states are generated using an optical parametric oscillator (OPO) ass before. The repetition rate is lowered by an acousto-optical modulator (AOM). Motorized half-wave plates (HWP) followed by polarizing beam splitters (PBS) allow us to vary the attenuation and generate a set of probe states. The intensity of the probe states is measured with standard power meters (PM), then further attenuated by neutral density filters (NDF) and coupled into the single-mode fibers of the time multiplexing detector (TMD). (b) State generation, identical to the previous chapter.

other source of error is a drift in the offset of the power meter. The power meter should therefore not be used close to its sensitivity limits. According to the manufacturer (Thorlabs), our power meter has an uncertainty of 5 %. Additionally, we get statistical uncertainties from the calibration of the attenuators. By repeating the calibration measurement several times, we estimate that the total uncertainty there is also around 5 %. The absolute uncertainty should therefore not exceed 10 %. This is a conservative estimation for the absolute power calibration. Additional to that, there are drifts in the power meter and the fiber incoupling. From comparing the incoupling efficiency before and after a long measurement, we measure a difference of 5 %. All these errors are given in relative units, because they scale linearly with power. Other errors, such as a constant background signal, might also be present, but are much smaller than these relative errors. To give an example of the raw reference signal of the power, we show the raw monitoring signal in Fig. 6.5.

The absolute calibration error does not produce significant systematic errors. Since coherent states remain coherent states after losses, a miscalibration of their power will result in a misestimation of the total detection efficiency. To understand what effect that has on state tomography, it is useful to divide the setup into two parts. One part can be regarded as state preparation and the other part as detection. A meaningful position for the border is inside the fiber that leads to the TMD, because that is the

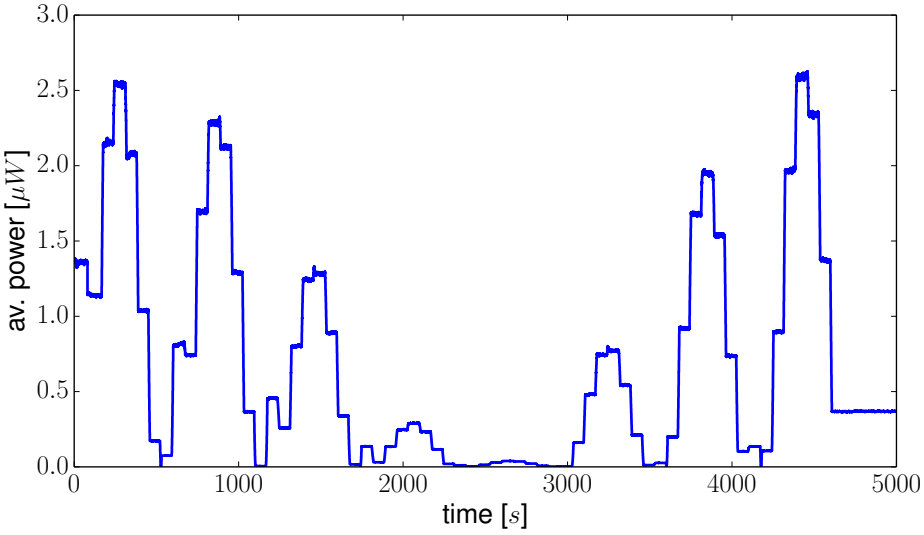


Figure 6.5.: Reference power for one of the two input ports. Each plateau is about 60 s long and is linked to the corresponding measurement. The power is changed with half-wave plates (HWPs), rotated by a constant amount at each step. The little spikes at high and low power levels correspond to the HWP going through a maximum or minimum. No measurement is taken, while the HWPs move. Several such runs produced 235 two-mode probe states to be used for pattern tomography.

part at which the state could be actually used in an imaginable larger network to combine it with other states. The calibration of the coherent states makes the state tomography tolerant to losses that happen after the reference point of the calibration, which is also inside the fiber that leads to the TMD. That means that all TMD losses and APD efficiencies are accounted for by this tomographic reconstruction. Losses that happen on the state preparation side, however, are not accounted for and will be present in the reconstructed state. If we now misestimate the powers of the reference states, we effectively shift the border between state preparation and measurement to either direction. If it is on the state preparation side, we automatically perform a slight loss-inversion on the state. If it is on the measurement side, the states will look more lossy than they actually are. Both effects are not dramatic, as the system efficiency up into the fibers is only around 70 %. Making an error of 10 % in the calibration would be equivalent to spending a few more hours on optimizing the incoupling to 80 %. Only for significant calibration errors, above the total losses in the state preparation part, the reconstruction would produce unrealistic states that never existed in the setup.

Reconstruction of Coherent states and PDC states

Here, we present our results for the reconstruction of different states with pattern tomography, performed in collaboration with our theory partners from the group of Zdeněk Hradil in Olomouc, Czech Republic, and with Luis Sánchez-Soto in Madrid, Spain. We take into consideration a fixed number of patterns with amplitudes below a given threshold $\alpha_{\max} \approx 2$. This threshold is important, because afterpulsing seems to be more pronounced for stronger states, indicating that the total power on the APDs is getting too large and the linear model becomes invalid.

Each probe-measurement consists of the two amplitudes of the probe states α_1 and α_2 and the patterns, or responses, $f_{\alpha,\beta}$. (Note that the indices here do not refer to the amplitudes of coherent states but just count the 2^{16} different detector responses, corresponding to all possible click combinations.) For each probe-pattern, we register about $N_\xi = 4.2 \times 10^6$ events. This corresponds to a measurement time of 60 s per setting. In total, we use 235 different settings.

The states that we want to characterize, are produced with the same PDC source, that was characterized in chapter 5. For most of the states, the number of events per measurement is $N_{\mathrm{PDC}} = 21 \times 10^6$. With these numbers, the statistical noise is insignificant (except, perhaps, for heralded detections) and the reconstruction accuracy is governed by systematic errors and afterpulsing effects.

The reconstruction of a state is repeated 100 times with randomly chosen probe subsets of size $M < 235$. These subsets are sampled with replacement from the initial set of probe states, i.e. the same probe can be drawn multiple times. Such resampling is known as nonparametric bootstrapping [146]. It allows to estimate errors without error propagation or assumptions about the underlying distributions. In this way, the redundancy in the data is propagated into the final estimate by averaging over the samples and the variation between the samples is used to estimate the associated errors.

To check the performance for different parameter sets, we first perform a cross-validation [147], to verify whether the estimated state is consistent with the observed data sample. To this end, we check the quality of the reconstruction with random sets of coherent states discarded from the probes, but with the same amplitude threshold (in this case, $\alpha < 1$). We quantify the accuracy in terms of the well-motivated infidelity $1 - F(\rho_{\mathrm{est}}, \rho_{\mathrm{true}})$, where F is the fidelity between the estimate and the true state. We compare these results with the standard approach of performing detector tomography first and doing state tomography with the obtained POVMs. In Fig. 6.6, we plot the infidelity for both methods as a function of $n_{\max}$, which is the maximum dimension of the matrix P_{mn} (i.e., $0 \leq m, n \leq n_{\max}$). Except for very small reconstruction dimensions, patterns largely outperform detector tomography. Indeed, this latter technique fails for $n_{\max}$ greater than 5, whereas data-pattern tomography improves remarkably as $n_{\max}$ increases. In both cases, the number of probes slightly enhances the performance. To get the data in the Fig. 6.6, we have used linear estimation and we have ignored the positivity constraint. Once one does that, we get infidelities of the order of 1 %, which amounts to errors of a few percent for

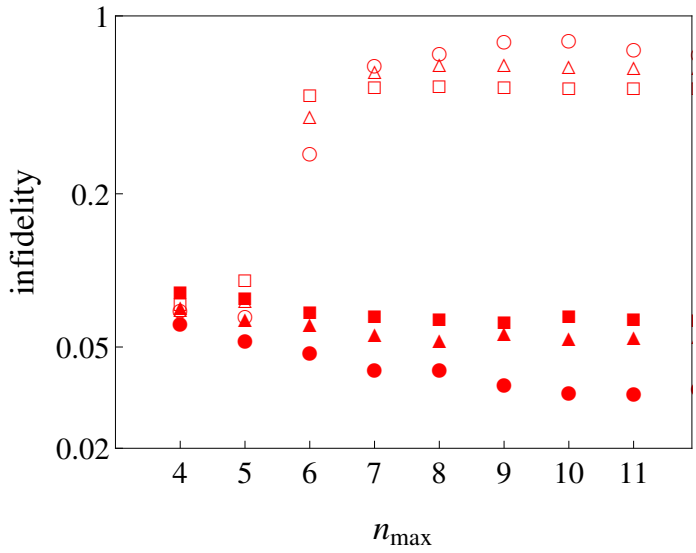


Figure 6.6.: (Color online) Average infidelity versus the dimension of the reconstruction space $n_{\max}$. The number of patterns employed is indicated by symbols: squares (50 patterns), triangles (80 patterns) and circles (150 patterns). Open symbols are used for standard detector tomography and filled ones for data-pattern tomography.

the reconstructed elements of P_{mn}. More probes give somewhat better results, but small sets of probes can be surprisingly good. This is due to a small variation of patterns with respect to the probe states, characterized by a small number of principal components in the singular value decomposition of the joint probability distribution of probes and patterns. Formulated in another way, the TMD has a certain symmetry between different outcomes. If it was a perfect TMD with equal splitting ratios and detection efficiencies, there would be only 9 different outcomes (photon numbers 0-8). This is one of the rare cases where asymmetry in detection efficiencies and splitting ratios is useful, increasing the number of principal components and hence the capability to reconstruct large states.

To compare reconstructed PDC states with theory, we assume the PDC distribution

$$P_{mn}^{\text{lossless}} = \frac{\langle n\rangle^n}{(1+\langle n\rangle)^{1+n}}\delta_{mn}\,, \tag{6.21}$$

together with a finite detection efficiency η which is taken into account by a Bernoulli distribution:

$$P_{mn} = \sum_{k=m}^{\infty}\sum_{\ell=n}^{\infty}\binom{k}{m}\binom{\ell}{n}\eta^{n+m}(1-\eta)^{k+\ell-n-m}P_{k\ell}^{\text{lossless}}\,. \tag{6.22}$$

From the zero-detection probabilities of coherent probes with known amplitudes, the quantum efficiency of detectors was estimated to be 0.22 ± 0.01. Further, the

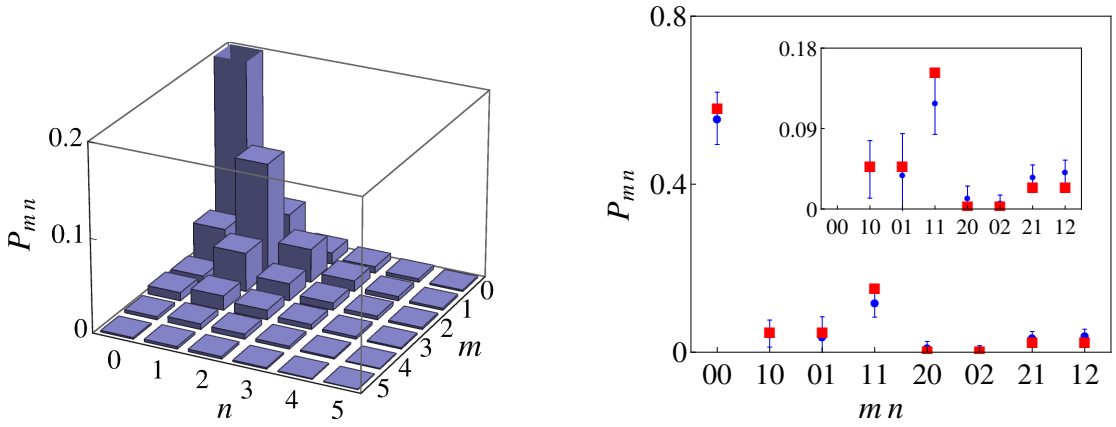


Figure 6.7.: (Color online) Two-mode reconstruction of the PDC_2 state. Left panel: Complete photon-number distribution, with $M = 50$ probes. Right: Comparison with the theoretical model. Circles with error bars represent the experimental results, whereas theory is indicated by squares.

coupling efficiency was estimated to 75%. This, in turn, enables to calculate the mean photon numbers of the PDC states: we generated three different ones, denoted PDC_1, PDC_2 and PDC_3, with $\langle n_1\rangle = 0.11$, $\langle n_2\rangle = 0.76$, and $\langle n_3\rangle = 1.34$, respectively. These numbers were used to predict the two-mode statistics through Eqs. (6.21) and (6.22).

In Fig. 6.7 we plot typical results of two-mode TMD measurements for PDC_2. Strong signal-idler correlations are observed and the agreement with the theory is pretty good. In Fig. 6.8 we show the reconstructions of the marginal signal states, i.e. the state with traced out idler mode, for two different pump intensities. Best fits to Bose-Einstein distributions are almost indistinguishable from the experimental results.

Next, we analyze heralded states, conditioned on single or double detection in the signal mode. By double detection we mean here a click at detector A accompanied by a simultaneous click at detector B. Double detections at any single detector are discarded to avoid doubles caused by afterpulsing. Heralded single- and especially two-photon states are difficult to reconstruct, since we are picking out quite a small subset of all the detection events. Besides, afterpulsing creates artificial signal-idler correlations, whose strength depends on the distance of the signal detection from the first idler time bin. This leads to larger reconstruction errors.

Reconstructed single-and two-photon heralded idler states from two different PDC states are shown in Fig. 6.9. To get theoretical predictions, we again assume an inefficient coupling (0.75) of the PDC state and calculate the post-measurement idler state P_i from the pre-measurement P as

$$P_i = \frac{\mathrm{Tr}_s(E\,P\,E^\dagger)}{\mathrm{Tr}_{s,i}(E\,P\,E^\dagger)}, \tag{6.23}$$

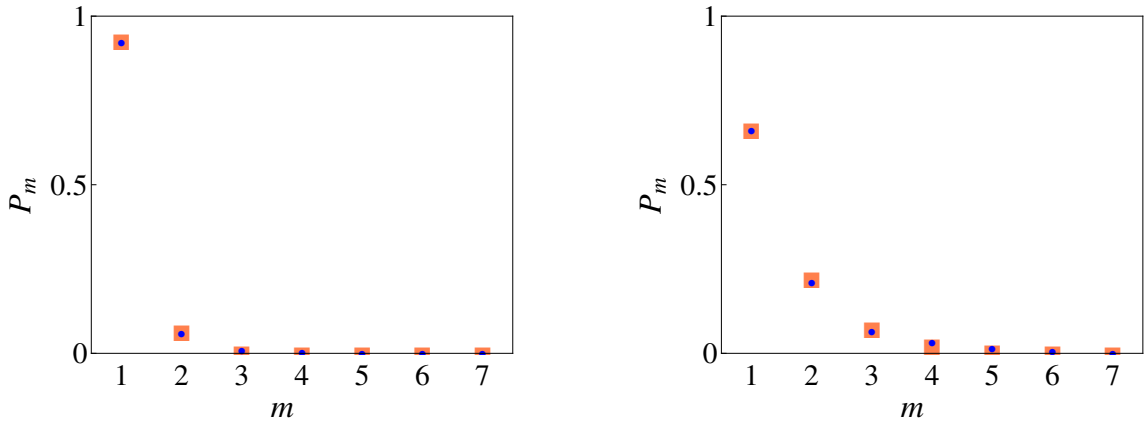


Figure 6.8.: (Color online) Single-mode signal reconstructions (dots) of PDC_1 (left) and PDC_2 (right), both with $M = 30$ probes. Best fits to Bose-Einstein distributions (squares) are also shown. The reconstruction errors are almost negligible and cannot be appreciated.

where $\hat{E}^\dagger \hat{E}$ is the POVM element describing the single/double detection in the signal mode and $\mathrm{Tr}_{s,i}$ indicates trace over the signal/idler. All states and POVM elements are diagonal here. From the measured photon-number probabilities, we calculate the parity of the states and hence estimate the value of the Wigner function at the origin for the single-photon heralded states. The results are $W(0) = -0.72 \pm 0.06$ (PDC_1) and $W(0) = -0.30 \pm 0.09$ (PDC_2). This agrees with the calculated values $W(0) = -0.77$ (PDC_1) and $W(0) = -0.29$ (PDC_2), respectively and confirms the nonclassicality of these states. With more intense PDC inputs, single detection in the signal tends to leave a mixture of Fock states in the idler. This explains why the nonclassicality of heralded states decreases with increasing pump intensity.

Finally, we also simulate heralded states as post-measurement states based on the results of full two-mode tomography. To this end, we perform 100 two-mode reconstructions for each measured PDC state. The idler post-measurement state is calculated based on a thought single or double signal detection. The statistics of the resulting ensemble of heralded states is shown in Fig. 6.10, along with the theoretical predictions. These predictions based on the full two-mode reconstructions are less accurate than the single-mode heralded ones. The latter is more direct. In heralded detections, what helps is that the dimension of the search space is reduced and the dominating vacuum or even single-photon terms are eliminated, which improves the accuracy. Nevertheless, it is nice to see that the agreement between single- and two-mode measurements is actually pretty good. The two-mode predictions improve with increasing intensity, as one could expect. More intense PDC states have larger higher-order P_{mn} components, which are easier to extract.

In conclusion, this data analysis shows that pattern tomography works, when standard detector tomography fails. It is thus a very robust method. With only two

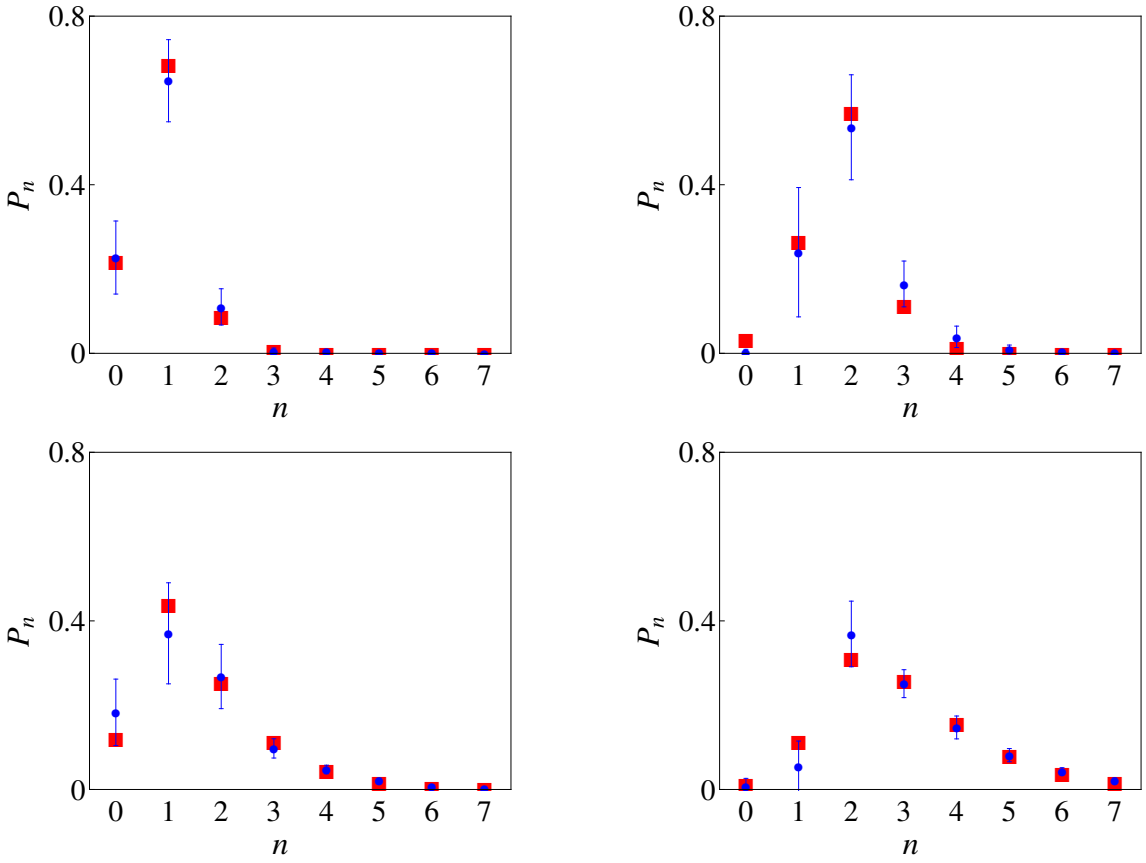


Figure 6.9.: (Color online) Reconstructed single-photon (left) and two-photon (right) heralded idler states generated from PDC_1 (top) and PDC_2 (bottom), with $M = 80$ probes. Squares denote again the corresponding theoretical predictions.

low-cost single photon detectors, that are plagued by afterpulses and low detection efficiencies, pattern tomography can reconstruct low photon-number states and help characterizing quantum states.

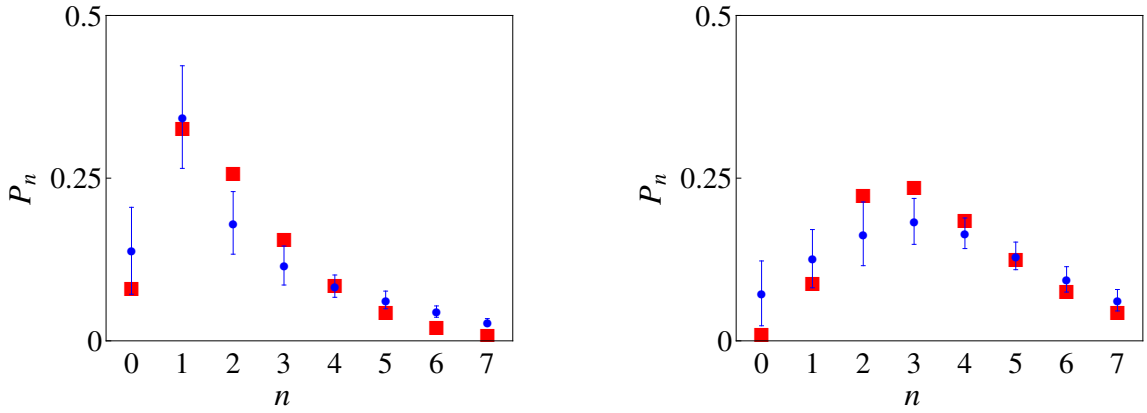


Figure 6.10.: (Color online) Heralded single-photon (left) and double-photon (right) idler states as predicted from the reconstructed two-mode photon-number distributions of PDC_3, with $M = 80$ probes.

Partial Tomography on Displaced PDC States

In anticipation of the next chapter 6.5, in which we estimate squeezing in a mode selective way, we discuss here the reconstruction of displaced states. The setup we use is an extension of the previously mentioned setup with the addition of a reference field, that overlaps with the signal, as sketched in Fig. 6.11. Our goal here is to use this data to test partial tomography. Following our two examples in the introduction, we try to estimate the antidiagonal photon-number probabilities (elements with $m+n = \text{const}$) of a two-mode state and the parity of heralded states.

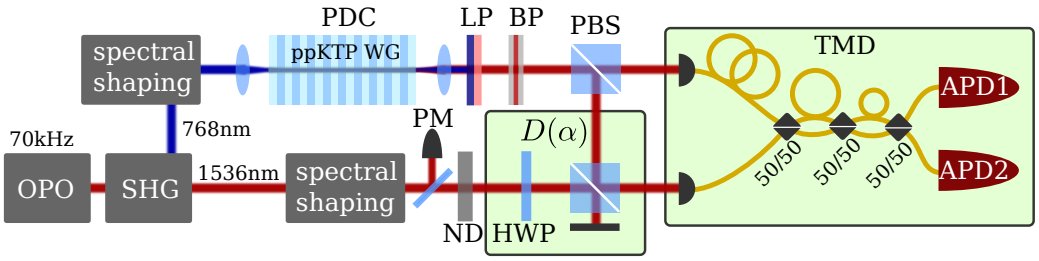


Figure 6.11.: The setup is essentially equivalent to the setup used for HOM interferences in Capter 5. The new element here is the displacement operator, consisting of a half wave plate (HWP) and a PBS, set to reflect 90 % of the PDC mode. The power of the reference, and hence the amplitude of the displacement, is monitored with a power meter (PM).

For the anti-diagonal terms, we restrict P_{mn} up to four photons in each mode and we address exclusively the anti-diagonal elements $P_{40}, P_{31}, P_{22}, P_{13}, P_{04}$. Without displacement, the anti-diagonal should be symmetric around P_{22}; with displacement, it is expected to become biased towards P_{40}. This is because the displacement adds

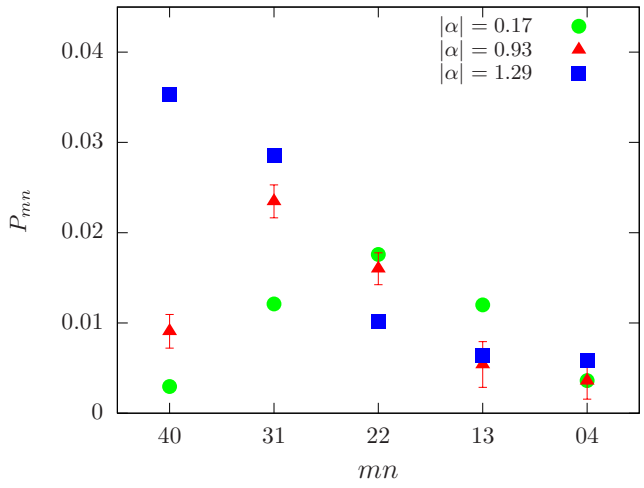


Figure 6.12.: Partial tomography of antidiagonal photon-number probabilities for two-mode states with varying displacements in one mode. Error bars are only shown for one set of points to avoid symbol clustering.

intensity and the photon-number correlations become weaker. In full tomography, with the same photon-number limit ($m, n \leq 4$), the space has a dimension $5 \times 5 = 25$. Thus we need at least 25 linearly independent probe states of size 25 each. For partial tomography, we only have a space dimension of 5. And hence our states need only have a size of 5 (the five anti-diagonal terms). In Fig. 6.12 we use 10 probes with their five relevant terms each to reconstruct the anti-diagonal elements of interest. The data follows the expected behavior and the data points have moderate error bars.

The second example is the parity. Here, we compare the results obtained from partial tomography with full tomography. In the first case, we directly obtain a parity in terms of the parities of the probe states. In the second case we reconstruct photon numbers and calculate the parity from them. The result is shown in Fig. 6.13. The two methods agree up to the two-click heralded states. For the three-click heralded state, partial tomography fails entirely, while the full tomography still works. This seems surprising at the first glance, because partial tomography requires less information. However, the three-click events are rather unlikely (in a PDC state with a mean photon number of 1.1 in this case) such that statistical noise becomes present and systematic errors like afterpulsing have a stronger effect. The essential difference between full tomography and partial tomography in this case is that we implement physical constraints for full tomography, whereas this cannot be done for partial tomography. The former rejects unphysical estimates and effectively performs better with noisy data. In Fig. 6.13 we compare full tomography with and without physical constraints. It is apparent that the latter case produces unphysical results that lead to an unphysical parity estimation with huge error bars.

In conclusion, partial tomography is very powerful in reducing the experimental

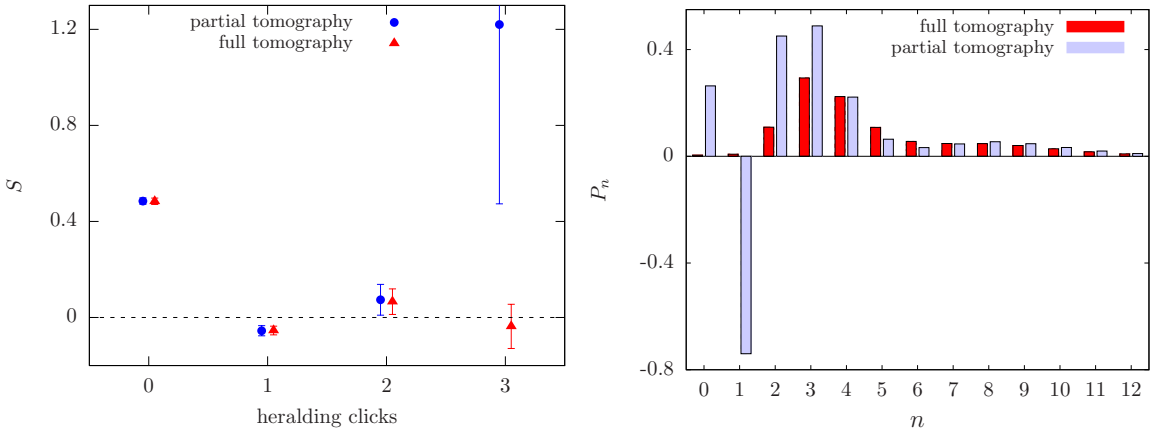


Figure 6.13.: Partial tomography vs. full tomography of the parity S of heralded, non-displaced states. While the agreement is very good up to two-click states, partial tomography fails for the three-click herald due to the fact that the reconstruction is not constrained to physical results. On the right, we compare the reconstructed photon-number probabilities between full tomography with constraints and full tomography without constraints, labeled partial tomography. This leads to negative photon-number probabilities and hence meaningless parities with huge error bars.

and numerical costs. The number and dimensionality of probe states can be drastically lowered. However, care must be taken when noisy data is analyzed. In contrast to partial tomography, full tomography allows to implement physical constraints and thus reduce the search space to the physical world. It is therefore more robust than partial tomography when noise and systematic errors are present.

6.4. Tomography by Noise

6.4.1. Introduction

Noise is a quite general term that is used for many unwanted processes in an experiment. Quite generally, unwanted processes can be described by a coupling to an environment, that introduces excitations of the system. If that environment is large and at thermal equilibrium, it will introduce thermal excitations. That means that the probability for an excitation decays exponentially with its energy. In optics, such thermal excitations do not exist directly, because black body radiation at room temperature contains photons in the far-infrared range and has close to no photons in the visible or near-infrared range. Imagine though that we opened our light-shielded windows in the lab and let some sunlight enter the experiment. Then it could be described as a thermal noise source, provided that we filter the sunlight to the mode of interest to us. Specifically, the light would need to be filtered in spectrum, time and space to match with the quantum states. Such filtering is quite demanding and

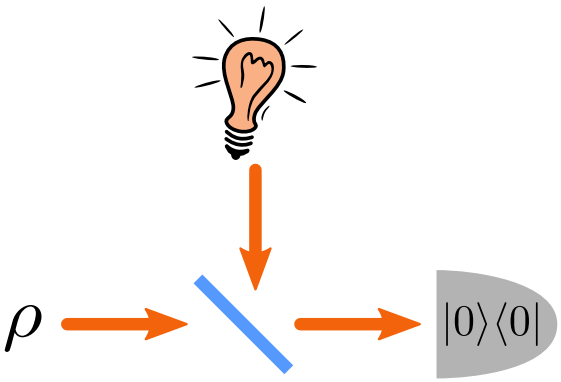


Figure 6.14.: Scheme for tomography by noise. A thermal light source is overlapped with the state at a beam splitter and impinges onto a click detector.

would probably leave close to no photons in the mode of interest. Fortunately, we can produce thermal light by other means.

The general idea of tomography by noise is to take a well known noise source, in our case thermal light, and use it for state characterization. The essence of the scheme is simple: using a beam splitter, we mix a known thermal source with the signal and measure it with a click detector, as sketched in Fig. 6.14. Varying the intensity of the noise, we can build a set of measurements, sufficient for the inference of the photon-number distribution of the state. This detection scheme is even more minimalistic than the multiplexing scheme as it requires only one mode and one single-click detector. The idea for such a tomography by noise comes from our collaborator Dmitri Mogilevtsev from Minsk, Belarus. The analysis of the data presented in this section was done in collaboration with him and with Natalia Korolkova from St Andrews, UK.

The probability of registering a signal for a phase insensitive detector is generally given by

$$p_j = \sum_{m=0}^{N} \Pi_{jm} \rho_{mm}, \tag{6.24}$$

where the elements $\Pi_{jm} = \langle m|\Pi_j|m\rangle$ are related to the jth element of positive valued operator measure (POVM), Π_j, which describes a measurement performed on the signal. $N+1$ is the dimension of the subspace of all possible signal states. We have only one on/off detector (with two POVM elements) and hence we have to find a way to generate at least $N+1$ different POVM elements. This can be achieved with our noise-scheme by varying the mean photon number of the probe, i.e. varying the temperature of the thermal noise. Let us suppose that the probe completely overlaps with the signal at the detector, which has a detector efficiency η. If we now register 'no click' events, that is the detector does not fire when a signal impinges on it, we obtain POVM operator matrix elements for such a measurement [148, 149]:

$$\Pi_{jm} = y_j(1 - y_j\eta)^m, , \tag{6.25}$$

where $y_j = 1/(1+\eta n_j)$ and n_j is the mean photon number of the thermal probe. Thus, we get $N+1$ different POVM elements for $N+1$ different probe mean photon numbers. Since we can represent the system (Eq. 6.24) as $p_j/y_j = \sum_{m=0}^{N}(1-y_j\eta)^m\rho_{mm}$, it means that using thermal state probes provides us with measurements, which should provide enough information to reconstruct elements ρ_{mm}. This result is very similar to another simple scheme [150], also requiring only one on/off detector. That scheme relies on a set of precalibrated attenuators to change the efficiency of the detector, which results in a set $\Pi_{jm} = (1-\eta_j)^m$, where η_j are the detector efficiencies corresponding to different attenuator settings. Comparing these two methods, our approach does not need additional calibration beyond the calibration of the on/off detector. To collect the necessary information about the mean photon number of the probe, it is sufficient to block the signal and use the detector itself to measure n_j.

6.4.2. Taking into Account the Mode Overlap

In practice, when mixing the signal and probe field at a beam splitter, we have to take into account an imperfect mode overlap of the fields. Before deriving the expressions for a mode overlap $0 < \mu < 1$, let us regard the case $\mu = 0$, i.e. no overlap at all between the signal and the probe field. In that case, the probability to detect a no-click event, is just the product of probabilities

$$\Pi_{jm} = y_j(1-\eta)^m. \tag{6.26}$$

Note that y_j is the probability to have no photons in a thermal state and $(1-\eta)^m$ the probability to loose m photons. Comparing this equation with Eq. 6.25 for perfect mode overlap $\mu = 1$, all elements are linearly dependent with respect to the photon-number distribution. In other words, changing y_j only multiplies a factor to the vacuum probability and we cannot estimate the photon-number distributions.

Now we move to the general case with two imperfectly overlapping fields. If the signal a and the probe b interfere on a beam-splitter and afterwards impinge on the detector, the probability to register 'no click' is given by [76]:

$$\begin{aligned} p_j = \mathrm{Tr}\{:\exp\{-\eta(Ta^\dagger a + (1-T)b^\dagger b + \\ x(a^\dagger b + b^\dagger a))\} : \rho\sigma_j\}, \end{aligned} \tag{6.27}$$

where $a^\dagger$, a and $b^\dagger$, b are the creation and annihilation operators of signal and probe modes; σ_j is the density matrix of the jth probe field; T is the transmissivity of the beam-splitter; $x = \sqrt{\mu T(1-T)}$, μ is the overlap parameter; :: denotes the normal ordering operator. For the perfect overlap, Eq. 6.27 results in a straightforward relation:

$$\Pi_{jm} = \sum_{n,k,l=0}^{N}(1-\eta)^k\sigma_{nj}|U^{kl}_{mn}|^2. \tag{6.28}$$

Quantities $\sigma_{nj} = (n_j)^n/(1+n_j)^{n+1}$ are diagonal matrix elements of the jth thermal probe state. The operator U describes the rotation performed by the beam splitter. It has the following matrix elements in the Fock-state basis:

$$U^{kl}_{mn} = \sqrt{k!l!m!n!}\sum_{g=0}^{k}\sum_{h=0}^{l}\frac{t^{g+h}r^{k+l-g-h}(-1)^{k-g}}{g!h!(k-g)!(h-l)!}\times \\ \delta_{m,l+g-h}\delta_{n,k+h-g}. \tag{6.29}$$

Here $t=\sqrt{T}$ is the amplitude transmission coefficient and $r=\sqrt{1-T}$ is the amplitude reflection coefficient, respectively. For the zero-temperature noise, $n_j = 0$, Eq. 6.28 gives a well-known expression for a 'no click' event probability for a damped signal:

$$p_{\text{signal}} = \sum_{k=0}^{N}(1-T\eta)^k\rho_{kk}. \tag{6.30}$$

Now let us represent the probe thermal state as a mixture of coherent states, $|\alpha\rangle$, [148]:

$$\sigma_j = \frac{1}{\pi n_j}\int d^2\alpha\exp\{-|\alpha|^2/n_j\}|\alpha\rangle\langle\alpha|. \tag{6.31}$$

Then, in the case of the perfect overlap, the probability of 'no clicks' for the thermal probe can be expressed through the probability of "no clicks" for the coherent probe (given in Ref.[76]):

$$p_j = \frac{1}{\pi n_j}\int d^2\alpha\exp\{-|\alpha|^2/n_j\}\times \\ \langle:\exp\{-\eta T(a^\dagger+\nu\alpha^*)(a+\nu\alpha)\}:\rangle_a, \tag{6.32}$$

where $\nu=\sqrt{(1-T)/T}$. Notice, that this formula is equivalent to the expression for p_j given by POVM elements in Eq. 6.28 for $N\to\infty$. The POVM elements for an imperfect overlap can be derived representing Eq. 6.27 for the probability of 'no click' for the thermal probe in the form similar to Eq. 6.32:

$$p_j = \frac{\bar{n}_j}{\mu n_j}\frac{1}{\pi\bar{n}_j}\int d^2\alpha\exp\{-|\alpha|^2/\bar{n}_j\}\times \\ \langle:\exp\{-\eta T(a^\dagger+\nu\alpha^*)(a+\nu\alpha)\}:\rangle_a. \tag{6.33}$$

The quantity $\bar{n}_j$ is the 'modified' average number of thermal photons of that part of the thermal field which is actually overlapping with the signal:

$$\bar{n}_j = \mu n_j/(1+(1-\mu)(1-T)\eta n_j). \tag{6.34}$$

Comparing the expression 6.33 with Eqs. 6.32, 6.28 for the perfect overlap, we obtain the relation for the POVM elements in case of the imperfect overlap:

$$\Pi^{\text{overlap}}_{jm} = \frac{\bar{n}_j}{\mu n_j}\sum_{n,k,l=0}^{N}(1-\eta)^k\bar{\sigma}_{nj}|U^{kl}_{mn}|^2, \tag{6.35}$$

where $\bar{\sigma}_{nj}$ are the diagonal matrix elements of the thermal state with the average number of photons $\bar{n}_j$ in the Fock state basis. Eqs. 6.34 and 6.35 point to a number of important conclusions. First of all, for the zero overlap, the 'modified' average number of photons is also zero, $\bar{n}_j = 0$. As follows from Eqs. 6.30 and 6.35, the resulting 'no click' probability factorizes,

$$p_j(\mu \to 0) \to p_{\text{signal}} p_{\text{term}}, \tag{6.36}$$

where $p_{\text{term}} = 1/(1+(1-T)\eta n_j)$ is the 'no click' probability for the thermal probe with the vacuum instead of the signal. For a weak probe, when $(1-\mu)(1-T)\eta n_j << 1$, the actual situation can be modelled by having two probe modes, the one completely overlapping with the signal with average number of photons equal to μn_j, and the non-overlapping one with average number of photons equal to $(1-\mu)n_j$. This behavior has also been found in [76] for coherent probe states. When the probe is strong, $(1-\mu)(1-T)\eta n_j >> 1$, it follows from Eq.(6.34), that the part of the probe actually interfering with the signal remains constant, $\bar{n}_j \approx \mu/(1-\mu)(1-T)\eta$. In other words, too strong thermal probe will wash out effects of interference and destroy a possibility to reconstruct the signal. The optimal regime is the moderate levels of thermal probe.

6.4.3. Setup

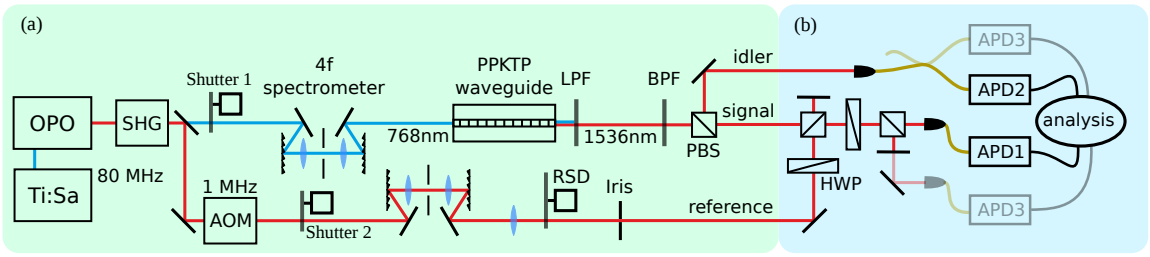


Figure 6.15.: Again the setup is very similar to all our setups with a reference beam. The difference here is that the reference has thermal statistics, realized with a rotating speckle disk (RSD) followed by irises. Instead of using the TMD, we use a single APD in the mode that overlaps with the reference. Two more APDs are used to herald single- and two-photon states as well as to estimate the mode overlap between signal and reference by HOM interference.

Here we use again the same source as in Chapter 5. The setup is depicted in Fig. 6.15. The new element now is the rotating speckle disk. It generates pseudo-thermal light in the following way. For each position of the speckle disk, a random interference pattern is created. After spatial filtering by irises and the final fiber incoupling, the intensity shows an exponential, hence thermal, probability distribution. We verify this by slowly rotating the disk and using an APD as a power meter. The distribution of the mean photon number is shown in Fig. 6.16. For the actual measurement, it

is not required to have independent mean photon numbers from pulse to pulse, as long as enough settings are sampled. We verify this by a measurement of $g^{(2)} > 1.9$, confirming close-to-perfect thermal statics.

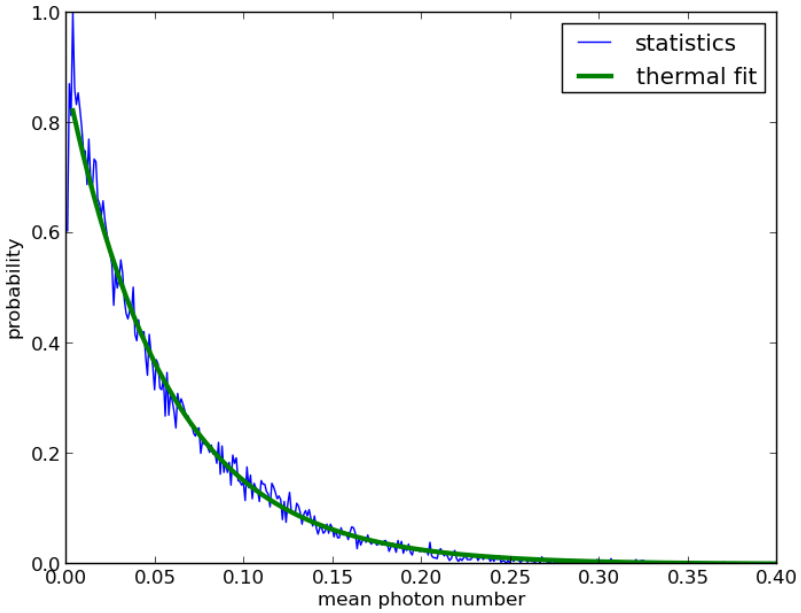


Figure 6.16.: Pseudo-thermal statistics. For each position of the rotating speckle disk, the photon-number statistics are Poissonian with a certain mean photon number. The probability distribution (in arbitrary units) of this mean photon number for different spot positions on the speckle disk is shown here. The exponential fit is almost perfect, indicating that only a small coherent background remains in the beam.

The calibration parameters of our scheme are the mode overlap between signal and probe μ and the overall efficiency η. To determine the mode overlap, we adjust our variable beam splitter to 50/50 and measure a Hong-Ou-Mandel dip. The overlap is calculated from the visibility of the dip as described in [88] to be $\mu = 0.45$. The decrease from unity comes possibly from a spectral mismatch in the 4-f setup or a spatial mismatch while coupling into the fiber. The detection efficiency is measured using the Klyshko scheme (confer Appendix B.3) from which we obtain $\eta = 0.15$. To generate a set of probe states, we rotate a HWP and measure the mean photon numbers n_j for each setting from counts in APD1 (see setup Fig. 6.15) with a physically blocked PDC beam.

6.4.4. Results

Fig. 6.17 shows the results of reconstruction for the heralded single-photon state. A total of 150 measurement points were used for the inference. For each thermal-

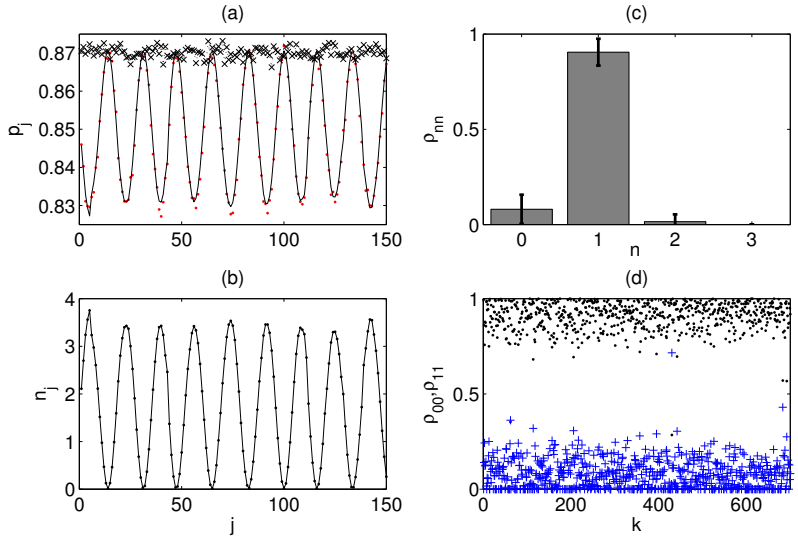


Figure 6.17.: Reconstruction of a heralded single-photon state. (a) 'No click' probability (red dots) versus measurement setting, produced by stepwise rotation of the HWP (hence the oscillation). Crosses depict the same probability of 'no clicks' on the APD for the signal alone; solid line shows probabilities estimated by Eq. 6.35 for the result shown (c). (b) Average number of thermal photons, n_j, of the reference field. (c) Inferred photon-number probabilities of the heralded single-photon state. (d) Estimated values of vacuum (blue crosses) and single-photon (black dots) components of the signal obtained via bootstrapping.

probe setting 10^7 events were recorded. The reconstruction was done using least-square estimation with non-negativity constraints [144]. The detection efficiency $\eta = 0.15$ and the overlap $\mu = 0.45$ were assumed. Fig. 6.17 (d) visualizes the estimated values of vacuum and single-photon components of the signal obtained via bootstrapping the data, similar to our analysis in Sec. 6.3. Our reconstruction procedure for the single-photon state gives the following value of the single-photon component $\rho_{11} \approx 0.905 \pm 0.07$. This estimate conforms well with the result from pattern tomography shown in the previous chapter.

Fig. 6.18 shows experimentally obtained data for the heralded two-photon state and the thermal state. In Fig. 6.18 (a) only a part of the measured data is shown. Here, when varying the thermal probe intensity, 600 different values of the reference-field intensity were taken. The average photon-number distribution shown in Fig. 6.18 (b) is close to a thermal distribution with an average number of photons equal to 0.17. Relatively large values of variances might be explained by a drift of detection efficiency, suggested by the count rates of heralded states as shown in Fig. 6.18 (c). The drift corresponds to a relative efficiency change of about 15%. This was not present in the measurement used for the heralded single photon, presented in Fig. 6.17. The exact experimental reason for the drift is unknown. It could be explained by a drift

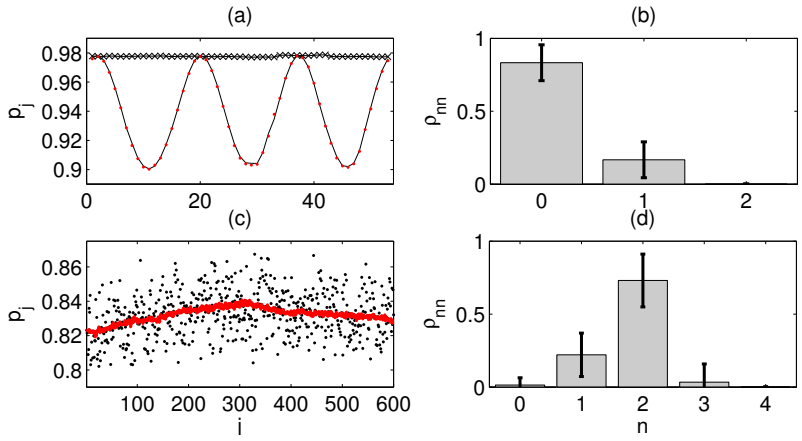


Figure 6.18.: (a, b) Reconstruction of a thermal state. (a) Red dots show experimentally collected probability of 'no clicks' for the unheralded signal overlapped with the thermal probe. Crosses depict the relative number of 'no clicks' on the APD for the signal alone; solid line shows probabilities estimated by Eq. 6.35 for the average reconstruction result shown in (b). (c) Efficiency drift over the time of the measurement. Shown are the 'no-click' probabilities without the reference for heralded single photons (red line) and heralded two photons (black dots). (d) Experimentally inferred photon number probabilities of the heralded two-photon state, obtained accounting for the efficiency drift.

of the single-mode-fiber incoupling over the total measurement time of two hours.

One of the powerful features of our method is the possibility to account for these variations in efficiency. For the estimation of the detection efficiency, we need to use the data for the signal state not mixed with the reference. For example, if we take the single-photon state, use Eq. 6.30 and the experimentally measured probability p_{signal} as shown in Fig. 6.17, we can compute the actual values of η. The drift in the detection efficiency η is reflected in the varying value of p_j for the heralded single-photon signal without the reference, as depicted by the solid red line in Fig. 6.18 (c). Ideally, this should be a straight line, as it is approximately for the data set with a low field intensity, used for the single-photon state reconstruction (p_j in Fig. 6.17 (a), solid line). For the data set with the higher field intensity, as used for the two-photon reconstruction, this is not the case anymore (Fig. 6.18 (c)). To account for this, we incorporated the calculated actual efficiency values in the expression for the POVM elements, Eq. 6.35, when inferring ρ_{nn} for the generated two-photon state, shown in Fig. 6.18 (d). The obtained results for the two-photon signal are again quite similar to the results obtained with the data-pattern technique and the TMD, shown in Fig. 6.9. It should be emphasized that imperfections of the setup do not lead to reconstruction artifacts in our scheme. For example, the vacuum component of the reconstructed signal remains very low despite a rather noisy character of the data. Also, the result of reconstruction does unambiguously show that despite low efficiencies of the detection,

the scheme produces states with large two-photon component. All these features are preserved even if no correction for varying detection efficiency is performed for the two-photon state, although the relative errors are much higher then.

In conclusion, we have demonstrated both theoretically and experimentally that reconstruction by noise is indeed feasible and provides a lucid, robust tomographic tool. By merely mixing the signal with a thermal noise and measuring statistics of the resulting field on the on/off detector, we can collect data sufficient for inferring photon-number distributions of different signal fields. Moreover, the reference field for each setting can be calibrated using the same on/off detector. The scheme has proven to be quite robust with respect to noise and drifts affecting the measurement set-up. We believe that such a scheme can become a simple, inexpensive and efficient working tool of quantum diagnostics. Potentially, even spectrally filtered light from such incoherent sources as an incandescent lamp can be used as the probe. Perhaps, one could even do a quantum state reconstruction with daylight.

6.5. Squeezing Estimation without Phase-Reference

In Sec. 4.2.2, we have introduced the two-mode squeezed state and discussed the correlations between its quadratures x and p. Here, our goal is to characterize these correlations experimentally. This is a challenging task, because any known scheme, most prominently balanced homodyne detection, requires two reference fields with known phases relative to each other and to the two squeezed modes. While conceptually simple, phase-locking pulses at the single photon level is experimentally difficult. Nevertheless, it is possible and has been done multiple times in the literature. Here, however, we ask the question whether it is possible to estimate the amount of squeezing without a phase-locked local oscillator.

6.5.1. Phase Averaged Two-Mode Squeezed State

First let us reformulate the Wigner function of the two-mode squeezed state from Eq. 4.72:[1]

$$\begin{aligned} W(\alpha,\beta) &= \frac{4}{\pi^2}\exp\{-2[\cosh(2r)(|\alpha|^2+|\beta|^2) \\ &\qquad + \sinh(2r)(\alpha\beta e^{i\phi}+\alpha^*\beta^* e^{-i\phi})]\} && (6.37)\\ &= f(|\alpha|,|\beta|)\exp\{-2\sinh(2r)(\alpha\beta e^{i\phi}+\alpha^*\beta^* e^{-i\phi}\} && (6.38)\\ &= f(|\alpha|,|\beta|)\exp\{-2\sinh(2r)(|\alpha||\beta|e^{i\phi_{\mathrm{sum}}}+|\alpha||\beta|e^{-i\phi_{\mathrm{sum}}}\} && (6.39)\\ &= f(|\alpha|,|\beta|)\exp\{-4\sinh(2r)|\alpha||\beta|\cos(\phi_{\mathrm{sum}})\}, && (6.40) \end{aligned}$$

[1] We regard a pure two-mode-squeezed state here for simplicity. The Wigner function can be extended to mixed Gaussian states simply by allowing independent strengths for the squeezed and anti-squeezed quadratures.

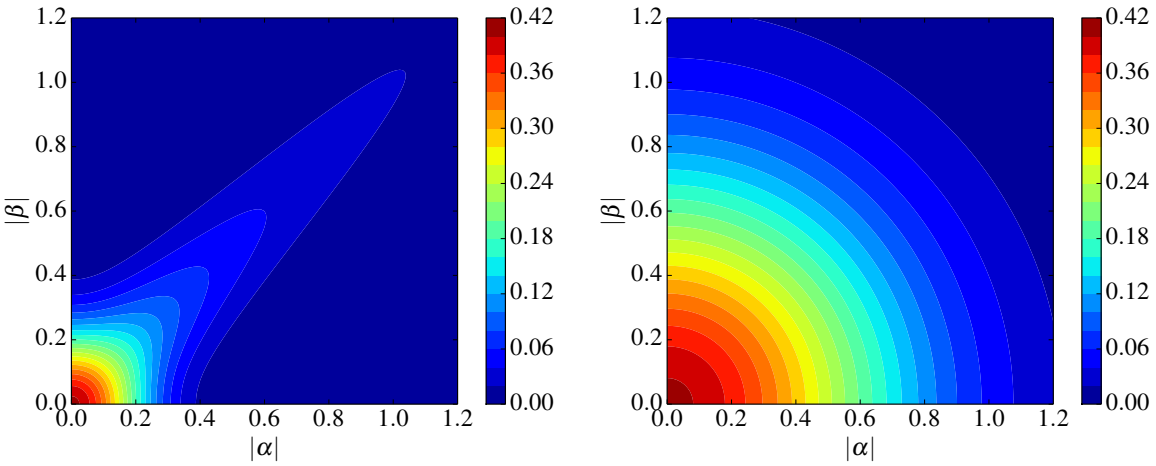


Figure 6.19.: Left: Amplitude correlations in the phase-averaged Wigner function of the two-mode squeezed state with squeezing parameter $r = 1.5$. Right: Vacuum state as a special coherent state, which has the strongest possible classical correlations.

where ϕ_{sum} is the sum of the signal-, idler- and pump-phase. In this form it becomes immediately clear that averaging over one of the fields completely destroys the phase-correlations of the state. Doing so, we get

$$W_{\text{avg}}(\alpha,\beta) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi_{\text{sum}} W(\alpha,\beta) \tag{6.41}$$

$$= \frac{4}{\pi^2}\exp[-2\cosh(2r)(|\alpha|^2+|\beta|^2)]\; I_0\left(-4\sinh(2r)|\alpha||\beta|\right), \tag{6.42}$$

where $I_0 = \int_0^{2\pi} \mathrm{d}\ \phi e^{x\cos(\phi)}$ is the modified Bessel function of the first kind. For completeness, we can also write the state in density matrix formalism

$$\rho_{\text{av}} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi\ |re^{i\phi}\rangle\langle re^{i\phi}| \tag{6.43}$$

$$= \frac{1}{2\pi}(1-|\lambda|^2)\int_0^{2\pi}\sum_{m,n}|\lambda|^n|\lambda|^m e^{i\phi(n-m)}|n,n\rangle\langle m,m| \tag{6.44}$$

$$= (1-|\lambda|^2)\sum_n |\lambda|^{2n}|n,n\rangle\langle n,n|. \tag{6.45}$$

This is a classical mixture of photon-number correlated states without any entanglement or quantum discord. The amplitude correlations in phase space, however, remain strong as can be seen in Fig. 6.19. Compared to vacuum or two coherent states in A and B, these correlations are stronger and show the nonclassicality of the state.

6.5.2. Direct Probing

The amplitude correlations that remain after phase-averaging the two-mode squeezed state cannot be seen in correlations of the real or imaginary parts of the quadratures any more, as would be measured by homodyne detection. This is obvious, since we averaged over the squeezed- and anti-squeezed variables. In fact, the quadratures of the two fields are completely uncorrelated. Only with a full state tomography, the Wigner function can be reconstructed and the amplitude correlations revealed.

Using a photon counting scheme, however, we can sample the Wigner function directly. We have shown in Eq. 4.62, that the Wigner function at a point α in phase space can be expressed as

$$W(\alpha) = \frac{2}{\pi}\sum_n (-1)^n \langle n|D(-\alpha)\rho D(\alpha)|n\rangle \tag{6.46}$$

This can be straightforwardly extended to the two-mode case:

$$W(\alpha,\beta) = \frac{4}{\pi^2}\sum_{m,n}(-1)^{m+n}\langle m,n|D_A(-\alpha)D_B(-\beta)\rho D_A(\alpha)D_B(\beta)|m,n\rangle. \tag{6.47}$$

The displacement operation can be implemented by an asymmetric beam splitter and a coherent state, as already noted in Sec. 4.5.2.

A crucial point of this direct-probing scheme is the mode overlap between the reference field and the signal. While in homodyne detection a non-perfect mode overlap is equivalent to losses, in direct probing, it has a different signature. Let us assume that the state consists of different modes, for example spectral or spatial modes. The state can then be written as a product state in an orthogonal basis

$$\rho = \bigotimes_m \rho_m. \tag{6.48}$$

From Eq. 6.46 we get

$$\begin{aligned} W(\alpha) &= \textstyle\sum_n (-1)^n \sum_{\{n_m|\sum_m n_m=1\}} \prod_m \langle n_m|\rho_m(\alpha_m)|n_m\rangle && (6.49)\\ &= \textstyle\sum_{\{n_m\}} (-1)^{\sum_m n_m} \prod_m \langle n_m|\rho_m(\alpha_m)|n_m\rangle && (6.50)\\ &= \textstyle\prod_m W_m(\alpha_m). && (6.51) \end{aligned}$$

The measured Wigner function is the product of all Wigner functions. One important example is a single-mode state and a non-unity overlap M with the local oscillator. In that case, we decompose the local oscillator into two parts, one part with perfect overlap and one orthogonal part with no overlap. Then the measured Wigner function is a product of the state- and vacuum Wigner functions with modified displacements:

$$W(\alpha) = W_{\text{state}}(\sqrt{M}\alpha)W_{\text{vac}}(\sqrt{1-M}\alpha). \tag{6.52}$$

This behavior is fundamentally different from loss and can be used to estimate both the loss of the state and the overlap with the local oscillator. For the purpose of state characterization, this is certainly an advantage as it gives more information about the state.

6.5.3. The Idea

The phase averaged Wigner function in Fig. 6.19 features correlations between α and β for all values, including $\beta = 0$. This suggests a very simple experiment to measure these correlations: We restrict the measurement to $\beta = 0$, meaning that no displacement in B is needed, and sample the Wigner function for different values of α. This idea is illustrated in Fig. 6.20. The resulting function $W(\alpha, 0)$ should show a distribution, which width depends on the squeezing. We plot $W(\alpha, 0)$ in Fig. 6.21 for realistic parameters including non-perfect overlap with the reference. In principle we can extract two quantities from a curve like that. The value at the origin is related to the total losses in the setup and the width of the distribution is related to the squeezing. Unfortunately, the overlap is also related to the width of the distribution as it corresponds to a rescaling of the overlapping local-oscillator amplitude and hence the width of the Wigner functions. However, the effect of a reduced mode overlap is rather small, since even for a low value of 70%, the curves widen only slightly. Therefore, we just interpret it as a reduction on the measurable squeezing.

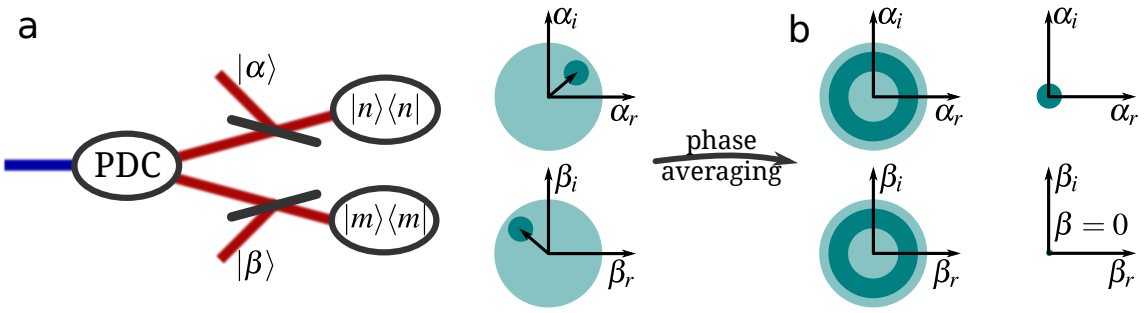


Figure 6.20.: (a) The scheme to sample the full Wigner function. If the phases of all fields are known, we see correlations (small circles) between systems A and B within broad probability distributions for each system alone (large circles). (b) Phase averaging the state destroys the phase correlations, but keeps the amplitude correlations intact. In particular, choosing the point $\beta = 0$ in B reduces the spread of α in A below the Heisenberg limit of a vacuum state.

In principle, the overlap could be estimated independently, if we used a Wigner function with a known shape, preferably with oscillations for 'sharp' features. Fock states are examples which fulfill this requirement; confer the Wigner function of the three photon Fock state in Fig. 4.6. If the overlap rescales the oscillations, while losses affect the value at the origin, we can determine both values independently.

Let us now briefly compare our method with other schemes that are based on photon-number measurements alone, without a reference field. In principle, such schemes can also be used to infer the squeezing of a state [151], because a phase-averaged state like in Eq. 6.45 only has diagonal elements in its density matrix. However, a photon-number measurement ignores the modal structure of the state. Therefore, we have to make assumptions about this structure, if we want to estimate

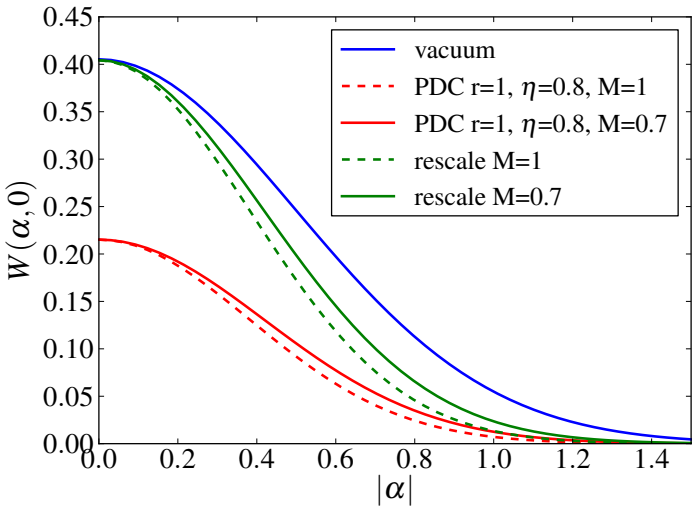


Figure 6.21.: Wigner function of a PDC state for $\beta = 0$ as it would be measured by direct probing for realistic parameters (red solid line) with gain r, efficiency η and mode overlap M. The losses affect the value at $\alpha = 0$ and the overlap broadens the distribution. The width of the curve is smaller than the vacuum reference due to squeezing.

the squeezing. Such an assumption can be that the state is a pure state or, more realistically, that it can be described by a certain modal composition as we did in our eight-parameter fit of the bright state in Sec. 5.4.2. The direct-probing scheme we present here does not require any assumptions about the modal properties due to the fact that it is mode-selectiv. The squeezing we estimate is the squeezing in the mode of the local oscillator, similar to homodyne detection. In principle, we could also estimate the anti-squeezing by calculating the marginal $W(|\alpha|)$, disregarding the second mode. This could then be used to estimate the purity of the state.

6.5.4. Experimental Results

Here we use again the setup shown in Fig. 6.11, with our PDC source, a reference field in one of the two arms, and the TMD with InGaAs APDs. We use pattern tomography to obtain photon-number distributions for different displacements α. From these photon-number distributions we calculate the parity and hence the Wigner function. The error estimations are done with bootstrapping as in the pattern-tomography section of this thesis.

The main result is the measured Wigner function, shown in Fig. 6.22 with the reduction of its width due to squeezing. To compare the experimentally obtained results with theoretical predictions, one has to know the reference displacements, as well as the squeezing r and the coupling efficiency η. Displacements are determined by measuring the reference beam alone and comparing its probabilities for vacuum-detection events with those of the coherent probes. A least-square fit is sufficient for this purpose. Additionally, the original undisplaced PDC state is measured to

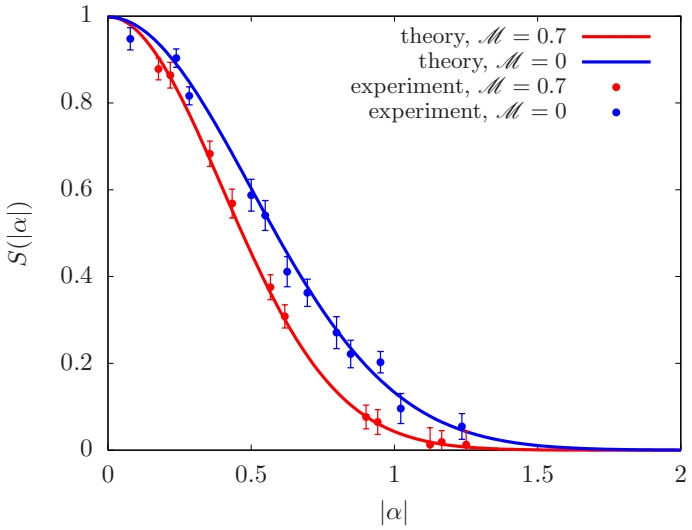


Figure 6.22.: Measured (normalized) Wigner functions $S = W(\alpha, 0)/W(0, 0)$ for the two-mode squeezed state with high overlap and with zero overlap. The zero overlap case is equivalent to measuring vacuum, since any non-displaced state is a constant factor, which is removed by the normalization. The reduction in width is strongly within the error bars and agrees well with theory.

estimate the squeezing parameter $r \approx 0.6$ and coupling efficiency $\eta \approx 0.75$: these are the values for which the theory provides the best fit.

With the same set of data, we can also look at displaced heralded states. The simplest, nontrivial case is the heralded single photon. Unfortunately, heralding with our TMD becomes very nontrivial due to the afterpulse effect. Since the heralding event is taken from the 'raw' data, pattern tomography does not account for it. Afterpulsing can be ignored for weak undisplaced PDC states, but, for large displacements, afterpulses make a significant fraction of the small number of genuine idler detections (more than 10% in our case). This is illustrated in Fig. 6.23 (top), where the idler single-click detection rate is plotted as a function of the reference-beam intensity $|\alpha|^2$. Without afterpulsing, the idler rates should stay constant, while we observe a linear increase of the idler detection rate with $|\alpha|^2$. After a single idler detection, the post-measurement state of the signal mode is $P_s = \mathrm{Tr}_i(E\,P\,E^\dagger)/\,\mathrm{Tr}_{s,i}(E\,P\,E^\dagger)$, where $E^\dagger E = \Pi_{1s}$ is the operator describing the idler detection, and $\mathrm{Tr}_{s,i}$ indicates tracing over both signal and idler modes. To deal with afterpulses, we construct the measurement operator as an incoherent superposition of signal and idler single detections, $E^\dagger E = [1 - x(\alpha)]\Pi_{1i} + x(\alpha)\Pi_{1s}$. The contribution of afterpulses for a given displacement is estimated from Fig. 6.23 (top).

The measurement for the displaced heralded single-click states is summarized in Fig. 6.23 (bottom). For high overlap, we expect the parity to trace the Wigner function of the heralded signal state, which can be approximated by a single-photon state. The negativity around the origin is apparent in the red curves. For no overlap,

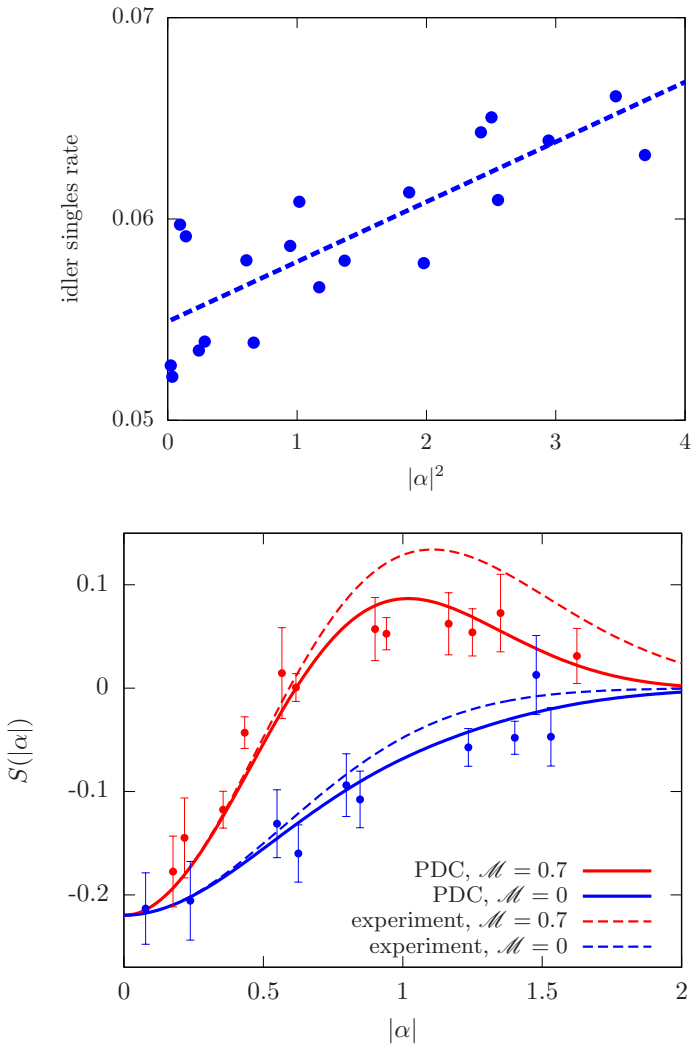


Figure 6.23.: Top: Observed idler singles rate (dots) for different signal displacements of a PDC state with $r \approx 0.6$. Best linear fit is also shown (line). Bottom: Parities of single-click heralded states measured with high overlap (red symbols) and no overlap (blue symbols) and $r \approx 0.6$. Solid (broken) lines show theory with afterpulses included (ignored), respectively. Notice that the afterpulses tend to decrease the measured parity of the displaced signal mode.

the measured parities remain negative for all displacements. The two cases can be discriminated, showing the distinct effect of the overlap.

In conclusion, we have demonstrated that squeezing can be estimated without a phase reference in a mode-sensitive way. The mode sensitivity makes this method superior to a photon-number measurement without a reference field. However, the afterpulsing effect proved obstructive in the measurement of heralded states and the error bars in the squeezing measurement do not yet allow to use this method as a precise measurement of the actual squeezing. Further improvements, especially on the detector side, could overcome these issues to make these measurements more reliable.

6.6. Correlation Functions and Nonclassicality

We have introduced nonclassicality based on the phase-space description of quantum states in Sec. 4.3. By that definition, a state is classical if and only if the P-function is positive everywhere. However, measuring the P-function or even the Wigner function can be challenging as we have seen in the previous chapter. If we just want to verify the nonclassicality of a state, the knowledge about the full quantum state is not required. Can we formulate simpler conditions to verify nonclassicality? It turns out that correlation functions, which we introduced in Sec. 4.4, are well suited for this task. From the experimental point of view, they can be measured in a loss-tolerant way using photon-number detectors or multiplexed click-detectors. From the theoretical point of view, they can be used to formulate nonclassicality criteria in quite a general way. We will describe this connection in this section and then present experimental results using our data from the TES measurement as well as data from a TMD measurement.

Single Mode States

The connection between correlation functions and the P-function is simple: Since the correlation functions are normally ordered, their expectation value can be directly calculated using the P-function:

$$\langle a^{\dagger n} a^n \rangle = \int \mathrm{d}^2\alpha P(\alpha) \alpha^{*n} \alpha^n = \int \mathrm{d}^2\alpha P(\alpha) |\alpha|^{2n}. \tag{6.53}$$

Imagine now the two-dimensional phase space with a classical distribution $P(\alpha)$. Correlation functions are just expectation values for moments of $|\alpha|^2$, which is the distance to the origin in phase space. In the case of a dot, i.e. ρ is a coherent state, $\langle |\alpha|^{2n} \rangle = \langle |\alpha|^2 \rangle^n$ and hence all $g^{(n)} = 1$. In the case of a broader distribution, $\langle |\alpha|^{2n} \rangle > \langle |\alpha|^2 \rangle^n$ and hence $g^{(n)} > 1$. This means that $g^{(n)} \geq 1$ is a necessary condition for classicality and therefore $g^{(n)} < 1$ is a sufficient condition for nonclassicality.

Two Mode States

Correlation functions can be straightforwardly extended to the two-mode scenario, by defining

$$g^{(m,n)} = \frac{\langle a^{\dagger m} a^m b^{\dagger n} b^n \rangle}{\langle a^\dagger a \rangle^m \langle b^\dagger b \rangle^n}. \tag{6.54}$$

In this case, nonclassicality conditions are a bit more difficult to define, since correlations between the two fields can be large for quantum as well as for classical states. We follow here the approach presented in Refs. [152, 153]. The idea is that we can formulate a nonclassicality condition for any operator $\hat{f}(\hat{a}^\dagger, \hat{a}, \hat{b}^\dagger, \hat{b})$ by using

$$\langle : \hat{f}^\dagger \hat{f} : \rangle = \int \mathrm{d}^2\alpha \; \mathrm{d}^2\beta \; P(\alpha, \beta) f(\alpha^*, \alpha, \beta^*, \beta). \tag{6.55}$$

If the P-function is a classical probability distribution, $P \geq 0$, then $\langle : \hat{f}^\dagger \hat{f} : \rangle > 0$ for all $\hat{f}$. Consequently, if

$$\langle : \hat{f}^\dagger \hat{f} : \rangle < 0 \tag{6.56}$$

for some $\hat{f}$, then P is not a classical distribution. Fortunately, there is a systematic way of testing different functions $\hat{f}$, by defining the vector

$$\vec{\hat{f}} = (\;\; 1, \quad \hat{n}_a, \hat{n}_b, \quad \hat{n}_a^2, \hat{n}_a\hat{n}_b, \hat{n}_b^2, \quad \hat{n}_a^3, \hat{n}_a^2 n_b, \hat{n}_a\hat{n}_b^2, \hat{n}_b^3, \quad ..., \hat{n}_b^N \;\;). \tag{6.57}$$

This vector contains increasing powers of $\hat{n}_a$ and $\hat{n}_b$ in all possible arrangements. The precise form of $\vec{\hat{f}}$ is not fixed. We could also choose any other function. In this way, however, the powers of the operators are ordered and give us a hierarchy of correlation functions, depending on the value of N.

The condition in Eq. 6.56 can be checked by constructing the matrix

$$M = \langle : \vec{\hat{f}}^{\,\dagger} \vec{\hat{f}} : \rangle = \begin{pmatrix} \langle : \hat{f}_1^\dagger \hat{f}_1 : \rangle & \langle : \hat{f}_1^\dagger \hat{f}_2 : \rangle & \cdots & \langle : \hat{f}_1^\dagger \hat{f}_N : \rangle \\ \langle : \hat{f}_2^\dagger \hat{f}_1 : \rangle & \langle : \hat{f}_2^\dagger \hat{f}_2 : \rangle & \cdots & \langle : \hat{f}_2^\dagger \hat{f}_N : \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle : \hat{f}_N^\dagger \hat{f}_1 : \rangle & \langle : \hat{f}_N^\dagger \hat{f}_2 : \rangle & \cdots & \langle : \hat{f}_N^\dagger \hat{f}_N : \rangle \end{pmatrix}. \tag{6.58}$$

Each element of M is essentially a correlation function with different powers. It is shown in [152] that if M has at least one negative principal minor, the state is nonclassical. We can also formulate that if M has at least one negative eigenvalue, the state is nonclassical. This condition can be easily checked numerically, if we have access to correlation functions of the form of Eq. 6.54 (where the normalization is not important).

In conclusion, this construction for two-mode fields poses a nonclassicality condition based on correlation functions. It contains a hierarchy taking into account higher and higher orders of correlations. Some states might reveal their nonclassicality only, if higher orders of correlations are considered.

6.6.1. Experimental Results for High-Order Correlation Functions

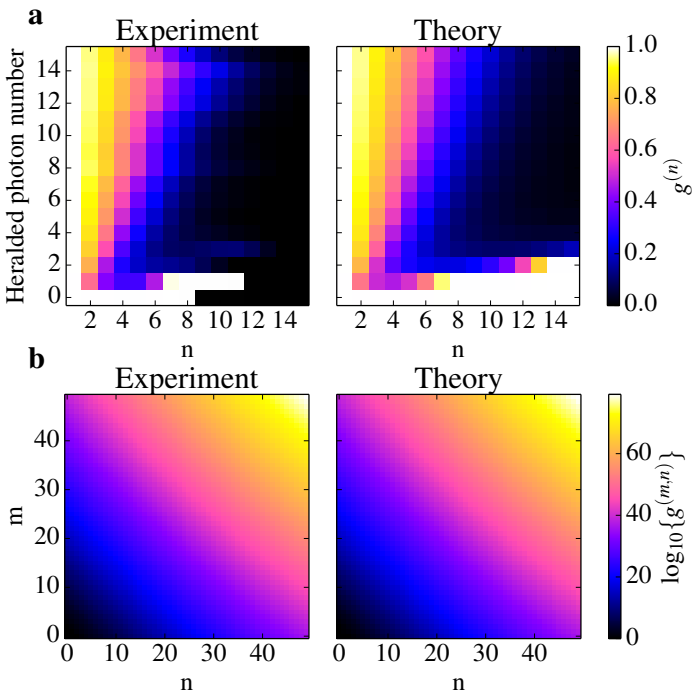


Figure 6.24.: Correlation functions. (a) $g^{(n)}$ for heralded states from a PDC state with $\langle n \rangle = 1.4$. Experimental results are shown on the left and theoretical predictions on the right. Values smaller than one imply nonclassicality. Note that correlation functions of higher order than the heralded photon number (lower triangular part of the matrix) should be treated with care. (b) $g^{(m,n)}$ for a PDC state with $\langle n \rangle = 20$.

From the photon-number measurements presented in Sec 5.4, we can calculate any correlation function according to Eq. 4.87. In Fig. 6.24, we show single mode $g^{(n)}$ of heralded states and two mode $g^{(m,n)}$. For the single-mode case, remember that $g^{(n)} < 1$ is a sufficient condition for nonclassicality. This is fulfilled for all heralded states except vacuum, which has a $g^{(n)} = 1$ as expected for Poissonian background noise. Remarkably, the values become lower for higher orders. For example, the $g^{(6)}$ for the heralded 10-photon state is below 0.5. On the one hand, this increases the confidence of the nonclassicality criterion. On the other hand, the statistical and possibly systematic errors for higher-order correlations increase significantly.

We estimate the uncertainties of the values by performing a simple Monte-Carlo simulation: In the two-mode case, based on the measured photon-number probabilities, we draw measurement frequencies from $8.2 \cdot 10^6$ events and calculate the corresponding correlation functions. From 10000 such trials, we calculate the standard deviations. In the single-mode case, we perform the same Monte-Carlo simulation, though this time based on theoretical probability distributions, due to relatively low

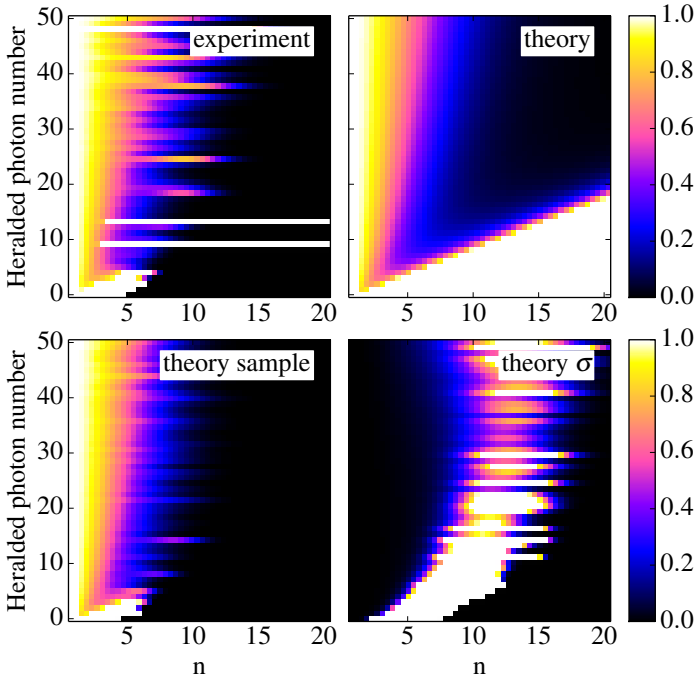


Figure 6.25.: Correlation functions $g^{(n)}$ for heralded states from a PDC state with $\langle n \rangle = 7$. Top: Comparison of experiment and theory. Bottom left: Simulation with $8.2 \cdot 10^6$ events. Bottom right: Expected standard deviation from a Monte-Carlo simulation. Only the region left of the bright area is reliable with the given statistics.

numbers of heralded events for higher photon numbers. For example, the 35-photon herald happens only about 1000 times. The results of this error analysis are shown in Fig. 6.26 for the two-mode $g^{(m,n)}$ and in Fig. 6.25 for the heralded $g^{(n)}$. The statistical uncertainties dominate already above $g^{(10)}$. Also, errors for correlation functions of order n for heralded states on photon numbers $< n$ are large, as can be expected, since the dimensionality of the heralded state is not large enough. However, the general agreement with theory is very good. The two-mode correlations allow us to calculate values up to $g^{(40,40)}$ with relatively low uncertainties. Since the $g^{(m,n)}$ has no upper bound, the values become very large.

To show nonclassicality, we can apply Eq. 6.58. Unsurprisingly, the nonclassicality condition is fulfilled for all orders with very high significances. Instead of going into detail here, we rather move on to the next section, in which we demonstrate this technique on data from less perfect detectors.

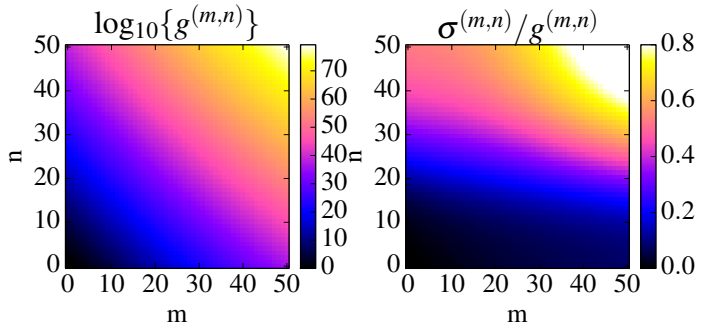


Figure 6.26.: Left: $g^{(m,n)}$ for the state with $\langle n \rangle = 20$. Right: Relative error obtained from a Monte-Carlo simulation based on the measured probability distribution. Values up to $g^{(40,40)}$ seem reliable. The asymmetry in the two modes arises from asymmetric detection efficiencies.

6.7. Certifying Nonclassicality with Click Detectors

In the previous section we used TES detectors to measure photon numbers. We assumed that the TES are perfect photon-number detectors, which is a very good approximation in that case. However, TES are very expensive as they need to operate in the mK-range and require complicated electronics. Typically, detectors which directly measure photon numbers like the TES are not available. Using a multiplexing scheme with click detectors, like our TMD with APDs, is much more common. To measure the photon-number distribution with such a scheme requires advanced data post-processing, for example using pattern tomography as we did in sec. 6.3. In this section, our goal is not to obtain the photon-number statistics of the state, but simply verify its nonclassicality. Remarkably, this can be done using the raw click statistics, without loss inversion or any other type of tomography. To this end, we collaborate with Jan Sperling, Martin Bohmann and Werner Vogel from Rostock, who are experts on nonclassicality and click statistics.

The model of a TMD detector we introduced in Sec. 6.2, Eq. 6.12 contains a loss term and a convolution term. The convolution term in particular is crucial for the analysis of nonclassicality. Imagine for example a classical Poissonian photon-number distribution as the input into the TMD. Due to the chance of photons ending up in the same bin, the measured click statistics might show a sub-Poissonian (narrower than Poissonian) distribution, which is a signature of nonclassicality. This effect is also present if the number of photons is one order of magnitude below the number of bins and might lead to fake nonclassical signatures. The proper theoretical detection model for such devices is a quantum version of the binomial statistics [154]. It can be used to derive nonclassicality criteria for the raw statistics as described in the following.

6.7.1. Nonclassical Moments of the Click Statistics

Perfect photon-number measurements (for one mode) can be described by [28, 51]

$$p_k = \langle : \frac{(\eta\hat{n})^k}{k!} e^{-\eta\hat{n}} : \rangle, \tag{6.59}$$

where p_k is the probability to measure k photons and η is the efficiency of the detector. From this starting point, it is shown in [154, 155] that the probability to measure k_A clicks within the $N = 8$ time-bins assigned to the signal A together with k_B clicks from the signal B is described through the joint click-counting:

$$\begin{aligned} c_{k_A,k_B} = \langle : & \binom{N}{k_A} \hat{m}_A^{N-k_A} (\hat{1}_A - \hat{m}_A)^{k_A} \\ & \times \binom{N}{k_B} \hat{m}_B^{N-k_B} (\hat{1}_B - \hat{m}_B)^{k_B} : \rangle, \end{aligned} \tag{6.60}$$

with $: \cdot :$ denoting the normally ordering prescription and $\hat{m}_i = e^{-\Gamma(\hat{n}_i/N)}$ for the modes $i = A, B$. In general, Γ can be an unknown detector response being a function of the photon-number operators $\hat{n}_i$. For example, a linear form of the response function is $\Gamma(\hat{n}_i/N) = \eta\hat{n}_i/N + \nu$, with η and ν being the quantum efficiency and the dark-count rate, respectively.

In Ref. [155] it has been demonstrated that the matrix of click moments $M^{(K_A,K_B)}$ is non-negative for any classical light field,

$$0 \leq M^{(K_A,K_B)}, \tag{6.61}$$

where M is defined as in Eq. 6.58. The superscript (K_A, K_B) defines the highest moment of each subsystems within the matrix M. For instance, the single-mode and bipartite, second-order matrices of click moments are

$$\begin{aligned} M^{(2,0)} &= \begin{pmatrix} 1 & \langle : \hat{m}_A : \rangle \\ \langle : \hat{m}_A : \rangle & \langle : \hat{m}_A^2 : \rangle \end{pmatrix}, \quad M^{(0,2)} = \begin{pmatrix} 1 & \langle : \hat{m}_B : \rangle \\ \langle : \hat{m}_B : \rangle & \langle : \hat{m}_B^2 : \rangle \end{pmatrix}, \\ \text{and } M^{(2,2)} &= \begin{pmatrix} 1 & \langle : \hat{m}_A : \rangle & \langle : \hat{m}_B : \rangle \\ \langle : \hat{m}_A : \rangle & \langle : \hat{m}_A^2 : \rangle & \langle : \hat{m}_A \hat{m}_B : \rangle \\ \langle : \hat{m}_B : \rangle & \langle : \hat{m}_A \hat{m}_B : \rangle & \langle : \hat{m}_B^2 : \rangle \end{pmatrix}, \end{aligned} \tag{6.62}$$

respectively. The needed moments can be directly retrieved from the measured click counting statistics [155],

$$\langle : \hat{m}_A^{l_A} \hat{m}_B^{l_B} : \rangle = \sum_{k_A=0}^{N-l_A} \sum_{k_B=0}^{N-l_B} \frac{\binom{N-k_A}{l_A}\binom{N-k_B}{l_B}}{\binom{N}{l_A}\binom{N}{l_B}} c_{k_A,k_B}. \tag{6.63}$$

This directly allows us to formulate nonclassicality conditions for the click-statistics. We might further define the term 'nonclassical correlations' in the following way:

The marginal states at A and B are classical, while the total state is nonclassical. Then we have nonclassical Kth-order click correlations if

$$M^{(K,0)} \geq 0, \; M^{(0,K)} \geq 0, \text{ and } M^{(K,K)} \ngeq 0. \tag{6.64}$$

This means that both Kth-order single-mode marginals are classical and the bimodal, Kth-order correlation matrix is nonclassical. Intuitively, applied to our photon-number correlated state, this condition means that the marginal distributions are broad, while the photon-number correlations are narrow. In order to genuinely certify such nonclassical correlations, it is sufficient to consider the minimal eigenvalues of the click-moment matrices e_A, e_B, and e_{AB} of $M^{(K,0)}$, $M^{(0,K)}$ and $M^{(K,K)}$, respectively. Now, the definition in Eq. 6.64 is rewritten as

$$e_A \geq 0, \quad e_B \geq 0 \quad \text{and} \quad e_{AB} < 0. \tag{6.65}$$

This method will serve as our approach to determine Kth order quantum correlations between the subsystems A and B.

6.7.2. Experimental Results

Second-order correlations

First, let us focus on second-order click correlations, cf. Eq. 6.62, which include the information about the mean values, the variances, and the covariance of the joint click counting statistics [155]. In Fig. 6.27, we plot the measurement results. Using the approach in Eq. 6.65, the minimal eigenvalues e_{AB} (top), e_A (bottom, left), and e_B (bottom, right) are shown in dependence of the energy per pulse. The single-mode matrices are non-negative, $e_A \geq 0$ and $e_B \geq 0$, whereas the cross correlations are nonclassical, $e_{AB} < 0$. Thus, we verified the quantum nature of the second-order click correlations between the spatial modes A and B.

In order to compare our measured results, a simple theoretical model is used. We assume that a pure two-mode squeezed state is generated, and the detectors are described via a plain linear response function: $\Gamma(\hat{n}/N) = \eta\hat{n}/N + \nu$. For a two-mode squeezed state, the gain parameter r depends on the pump power P as $r = \xi_0\sqrt{P}$. The proportionality constant ξ_0, the quantum efficiency η, and the dark count rate have been fitted using the standard method of least squares. They are $\eta = 9.6\%$, $\nu = 0.51$, and $\xi_0 = 0.087\,(\mu\text{W})^{-1/2}$. With these values, we estimate total mean photon numbers to be $0.9 \ldots 15$ photons for the pump powers $50 \ldots 403\,\mu$W or energies $0.7 \ldots 5.8\,$nJ. Already this simplified model yields a good agreement with the measured results, which highlights the excellent performance of the engineered PDC source; cf. dashed lines in Fig. 6.27. The inset in the upper plot shows the extrapolation of the quantum correlations, $e_{AB} < 0$, for higher pump energies. At some point, the mean photon number is so high, that these correlations saturate and eventually vanish. A high squeezing level and a classical laser light with a large coherent amplitude result in the same signal – all time-bins are occupied with a large number of

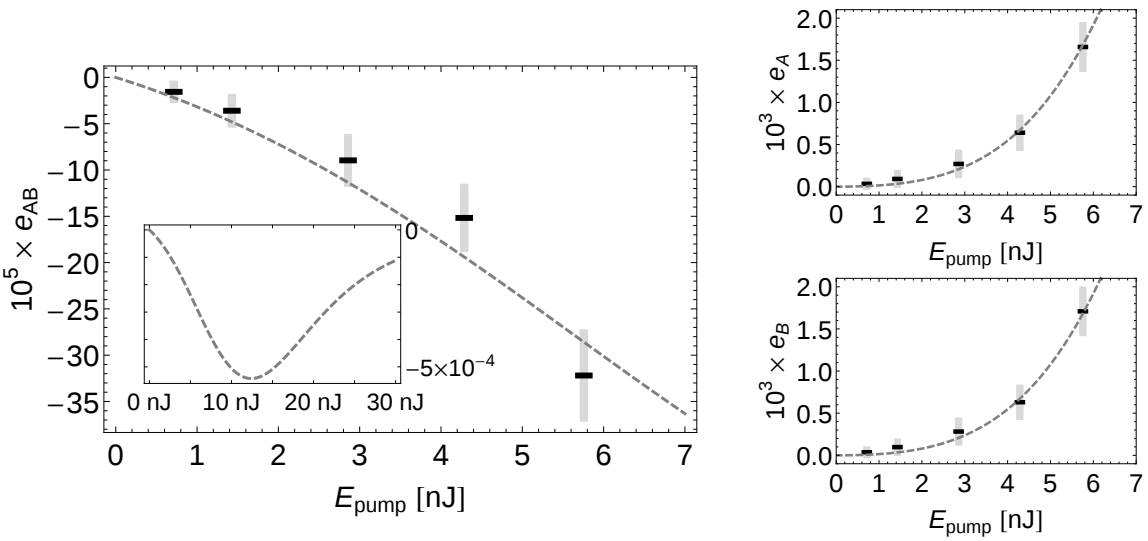


Figure 6.27.: The left plot shows the minimal eigenvalue e_{AB} (black bars) of the experimentally obtained matrix of moments $M^{(2,2)}$ depending on the pump energy E_{pump}. Error bars are given as gray areas. The dashed curve is the theoretical prediction. The inset depicts the continuation of the theoretical curve for higher energies including saturation effects. The right plots show the minimal eigenvalues e_A and e_B of $M^{(2,0)}$ (left) and $M^{(0,2)}$ (right), respectively, a ten standard deviations error bar, and the theoretical prediction. Since the bimodal correlations are negative and the single-mode reductions are non-negative, we successfully determined nonclassical correlations between the modes A and B.

photons. Therefore, at high photon numbers, the signals of nonclassical and classical states cannot be discriminated. Thus, the recorded quantum correlations of the former state must decrease due to the saturation of click counting devices, which is automatically included in the click counting theory. Let us emphasize again that the states have been generated for pump powers ranging over almost one order of magnitude. Verifying nonclassical photon-photon correlations in this comparably large domain is typically considered a challenging task, but can easily be accomplished with our TMD click counters. To illustrate the broad range of amplitudes for which this method could be used, we plot in Fig. 6.28 the minimal eigenvalues of the theoretically obtained second order matrix of click moments. For very small and very large mean photon numbers $\bar{n}$, the minimal eigenvalues converge to 0. For large $\bar{n}$, the saturation effects of the APDs diminish correlations, which is correctly displayed by the theory via the decay of correlations. However, still for a mean photon number of 150, we could reliably certify nonclassicality with an eight-bin TMD.

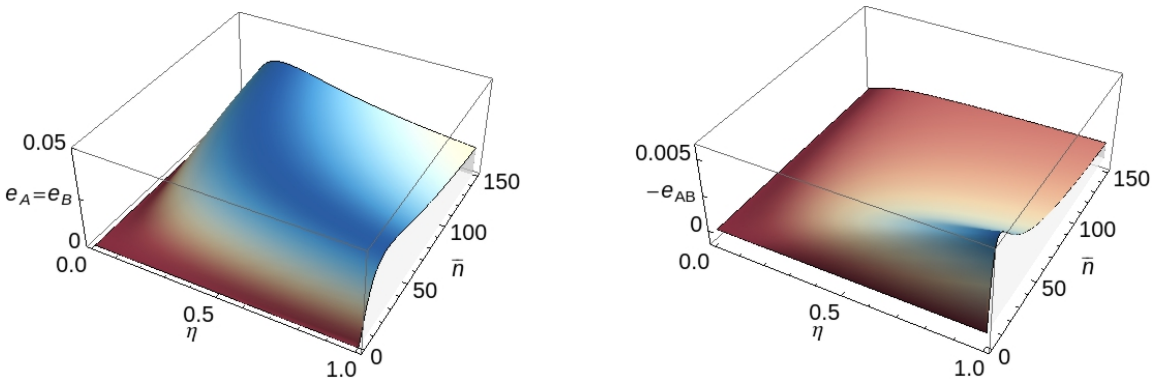


Figure 6.28.: The minimal eigenvalues $e_A = e_B$ of $M^{(0,2)} = M^{(2,0)}$ (left) and e_{AB} of $M^{(2,2)}$ (right) are plotted depending on mean photon number $\bar{n}$ and the quantum efficiency η (fixed dark count rate $\nu = 0.51$; for a better visualization, we plot $-e_{AB}$). The single-mode eigenvalues $e_A = e_B$ are positive for any value of $\bar{n} > 0$ and $\eta > 0$, while the bipartite case shows negativities for all these parameters, $-e_{AB} > 0$.

Higher-order correlations

Let us study higher-order quantum correlations. They become particularly meaningful when second-order criteria fail to properly characterize the state. For example, the third and fourth order moment relate to the so-called skewness and the flatness (or kurtosis), respectively. The highest possible order of moments one can infer from $N = 8$ time bins for each mode is given by $K = 8$ in Eq. 6.65 yielding a full characterization of the measured click counting statistics.

The bound for a classical signal, $M^{(K,K)} \geq 0$, is given by the eigenvalue $e_{\text{cl.}} = 0$. Thus, the signed distance, in units of standard deviations, of the experimentally obtained minimal eigenvalue, $e = \overline{e} \pm \Delta e$, to this classical bound leads to a signed

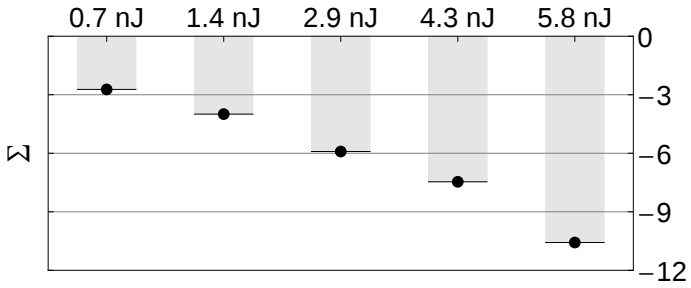


Figure 6.29.: The signed significances Σ of the largest negativities of the full matrix of click moments $M^{(8,8)}$ are shown. The highest significance levels of nonclassicality between the subsystems is verified for our highest energy. In this region the impact of the click-counting theory is most pronounced.

significance,

$$\Sigma = \frac{\overline{e} - e_{\mathrm{cl.}}}{\Delta e} = \frac{\overline{e}}{\Delta e}, \tag{6.66}$$

representing a signed relative error. A negative significance, $\Sigma < 0$, verifies the nonclassicality with a significance of $|\Sigma|$. Typically, $\Sigma \lesssim -3$ is a significant verification of the negativity, whereas $\Sigma \approx 0$ cannot be distinguished from the classical bound 0.

In Fig. 6.29, the significance levels of the full eighth order quantum correlation are given. The single-mode, signed significances for $M^{(8,0)}$ and $M^{(0,8)}$ can be found in [7]. There are no significant eighth order, single-mode correlations, $M^{(8,0)}, M^{(0,8)} \gtrsim 0$ – the largest single-mode negativities are in the order of $\Sigma \sim -10^{-5}$. Additionally, the significance of negativities in $M^{(8,8)}$, cf. Fig. 6.29, are in the range of $2.7 \ldots 10.6$ for the energies $0.7 \ldots 5.8\,\mathrm{nJ}$, respectively. Hence, for most of the energies, a significant eighth-order nonclassical correlations between the modes A and B is certified. Remarkably, the significance even increases with the energy which is due to an improved signal-to-noise ratio with increasing mean photon numbers, because the no-click event has a much lower probability in this regime.

On the one hand, one would typically not use such comparably high intensities in our measurement setup, because the difference between photo-counting theory and click-counting is quite pronounced in this regime. This can also be seen when analyzing the data on purpose with the inappropriate photoelectric detection model. Then, the single-mode signed significances evaluate to $-2.4 \ldots -13$, cf. [7], suggesting fake nonclassicality which worsens with increasing pump power. On the other hand, our consistent treatment in terms of the click statistics correctly identifies this higher-order bimodal correlations while showing, as expected for our source, no nonclassicality in the single-mode marginals.

In summary, we have shown in this section that we can certify nonclassicality using time-multiplexed detectors over one order of magnitude of intensities, and mean photon numbers up to 15. This is possible despite low detection efficiencies of 10% and without any corrections for imperfections or related post-processing techniques.

7. Conclusion and Outlook

Conclusion

In this thesis, we generated quantum states using an engineered type-II parametric down-conversion (PDC) source and characterized them with a variety of methods.

In the low pump-power regime, we heralded single photons and showed that they have a high purity of 85 %. Additionally, the signal and idler photons where nearly indistinguishable with a Hong-Ou-Mandel visibility of 95 %. These results are comparable with achieved benchmarks in the literature. Additionally, the source showed a high down-conversion efficiency, which can be attributed to the single-mode operation in the spectral degree of freedom and the use of a waveguide. The exceptional performance lies in the combination of all these characteristics in one single source. The singlemodeness allows potentially to combine several such sources without introducing further losses and to perform photon-number measurements with very high efficiency.

In our collaboration with NIST, where we used transition edge sensors (TES), we achieved efficiencies of up to 68 %. Here, we investigated the properties of the quantum states in a high-pump-power regime with mean photon numbers up to 20. Remarkably, the single-mode character persisted in this regime, which is very fortunate, since a number of nontrivial effects could be expected to degrade that property. With the TES, we were able to show photon-number correlations up to 80 photons in a single-mode state. Such large states can already be considered macroscopic and constitute a resource for the generation of a variety of nonclassical states. We showed that high-order correlations are present and accessible by calculating two-mode correlation functions up to order 40, again demonstrating the huge dimensionality of the state and its potential as a resource for quantum optics.

The second part of this thesis was concerned with the characterization of quantum states. One of our main results here was the inference of squeezing without a phase reference. The fact that this is possible is surprising at first, since the phase-averaged state does not contain entanglement, but the photon-number correlations allowed us to circumvent this problem. Using photon-number detectors and a reference field with an unknown phase, we probed the two-mode Wigner function at several points around the origin, where one mode was fixed exactly at the origin. This allowed us to estimate the squeezing in a mode-selective way and without assumptions about the purity of the state. The experimental simplicity makes this method an appealing alternative to the established homodyne detection scheme.

We further investigated how InGaAs APDs, which have low detection efficiencies

and relatively high afterpulsing probabilities can be used to reconstruct the photon-number distribution of a state or prove its nonclassicality. To reconstruct photon-number distributions, we applied two schemes: Pattern tomography and inference by noise.

Pattern tomography reconstructs the state in a basis of probe states, rather than a pre-chosen basis like photon number. It avoids truncation errors and the need for full detector tomography. We showed that with pattern tomography, a time multiplexing scheme, which is particularly susceptible to afterpulsing effects, can be used to reconstruct the photon-number distribution of quantum states in a few-photon regime. This included the distributions of heralded single- and two-photon states and the reconstruction of the Wigner function from statistics of displaced single photons and of displaced two-mode squeezed states. Even though pattern tomography proved to be able to deal with the afterpulsing effect, it is reliable only in the linear regime of the detector response. If the count rates increase, the detectors show accumulation effects, meaning that their response depends on the mean photon number of the state and not just on the detected photon number of the current pulse. This generally limits such detectors to a low-click-probability regime.

Tomography by noise is a minimalistic scheme to measure the photon-number distribution of a state, requiring only a single click detector and a reference field with a thermal photon-number distribution. With this method, we successfully reconstructed single-photon and two-photon heralded states, as well as thermal states. It is surprising that this method worked despite the low efficiencies of 13 %. It shows the power of tomographic techniques to extract information from strongly convolved data.

Finally, we applied the theory of click statistics to our two-mode time-multiplexed detection. Without performing any loss inversion or tomography, we proved nonclassical correlations between the two modes. Since the raw data was used directly, this method worked even for mean photon numbers of 15 on an 8-bin detector with low efficiencies of 10 %. It is a simple and reliable tool to test the classicality of quantum states.

Outlook

The source we built up during this thesis showed remarkable properties in the single- as well as in the multiphoton regime. The Fock states and squeezed states it generates are very basic resources in quantum optics and therefore have numerous potential applications.

At this time, there are already experiments in preparation that use very similar sources. The group of Ian Walmsley in Oxford is currently implementing multi-source experiments where the singlemodeness of the states is exploited. Within our own group, such a source is used to generate indistinguishable photons for a time-multiplexed quantum walk [156, 157]. Furthermore, our group uses the source to generate states for the quantum pulse gate, a device that selectively operates on

spectral modes [158]. Here, the spectral tunability of the source is exploited, to generate single-mode as well as few-mode states. Interestingly, by shaping the pump spectrum to a first Hermite-Gaussian mode, the source should produce exactly two modes with equal weights. In the single-photon approximation, this state becomes a Bell state in the spectral degree of freedom. Such Bell states in combination with the pulse gate allow to formulate a complete framework for quantum information science [159].

A continuation of the source characterization would be to implement a phase sensitive measurement to fully access the phase-space properties of the state. We estimated in this thesis that the source should show high squeezing values, surpassing current world records for single pass sources and possibly even cavity based sources. The measurement would require a homodyne setup or a phase-sensitive direct probing scheme. The key challenge in both cases is to implement the measurements with a very high quantum efficiency. Besides measuring high squeezing, an interesting question is where the limitations of the source are when going to ever higher photon numbers. Can the single-mode character be maintained for stronger pump powers? How large can single-mode quantum states become, generated from this source?

In the course of this thesis, we had fruitful collaborations with theory groups specialized in tomography and other sophisticated characterization techniques. There seem to be plenty of ideas to continue this work. Currently, we are investigating with Dmitri Mogilevtsev, the group of Zdenek Hradil, and Luis Sanchez-Soto a scheme to find the smallest possible subspace that fully describes an unknown state. This is interesting as it allows one to choose an appropriately small state-space for the reconstruction as well as for refining measurements. We also plan to continue our collaboration with Dmitri Mogilevtsev and Natalia Korolkova to further investigate what information about a quantum state can be inferred with a single click-detector and appropriate reference fields. Moreover, the certification of nonclassicality we demonstrated with Jan Sperling, Martin Bohmann and Werner Vogel could be extended to a phase-sensitive measurement [160, 161] for inferring squeezing or entanglement.

Finally, a large new field of research is the investigation of the macroscopic character of a quantum state. We know that PDC can produce billions of photons per mode [21]. How macroscopic is such a state? Can the single mode character of our source be exploited here? Do losses ultimately limit the macroscopicity despite huge photon numbers or can we play some tricks with photon-number detectors? Besides such fundamental questions, there is the technological dream to integrate state-manipulation and detection on chip. This would drastically decrease losses and improve scalability of multiplexing schemes. All in all, there are enough fundamental as well as technological questions waiting to be answered. Indeed, we see a bright future ahead for our source.

Acknowledgment

First of all I would like to thank Christine Silberhorn for taking me on into her group and giving me the opportunity to do some research. It was a dream come true for me to work as a scientist, and I actually mean it. During that time I have seen a lot and learned a lot. My project involved a mix of experiment and theory, which allowed me to do exciting work in the lab as well as take time to understand all underlying theory.

I very much appreciated the relaxed atmosphere in the group and Christine's amazing motivation skills. Somehow she never fails to spark excitement about a project. I further thank Christine for letting me travel around the world to numerous conferences and the collaboration with NIST, which was a highlight of my PhD.

I would also like to thank all my colleagues for a great time at work and outside of work. First of all Kaisa Laiho, who introduced me to the magic of a quantum optical setup. Thanks also to Andreas Christ, who always had an answer to a PDC question, or when not, reassured you that the question is hard and a current research topic. Thanks to Benjamin Brecht for helping out on many occasions and the rest of the Erlangen-crew, Malte Avenhaus and Andreas Eckstein, for support with the experiment.

Great thanks to Helge Rütz, my valued office mate, who I had so many discussions with that some of them, I am sure, we repeated several times and it was no less fun. Further I would like to thank Vahid Ansari for a great time in the lab. I think we were a pretty good team and I am looking forward to continue that work on the quantum pulse gate.

Many thanks to Thomas Gerrits, who made the experiment at NIST possible. He miraculously made sure that quite a complex setup was just running. Thanks also to Tim Bartley, together with whom we built up the experiment at NIST in ten times the normal speed.

Further, I would like to thank all our theory collaborators, in particular Jaroslav Rehacek, Bohumil Stoklasa, Luis Sanchez-Soto, Dmitri Mogilevtsev and Jan Sperling for amazing theory work and paper writing.

And of course the rest of the IQO team: Fabian Elster and Stephan Krapick for a great time in California and at other conferences. Thomas Nitsche for keeping the mood up with incredibly smart jokes. Michael Stefszky for organizing all sorts of non-physics activities. Kai Hong Luo, Regina Kruse, Johannes Tiedau, Sonja Barkhofen, Christoph Eigner, Markus Allgaier, Nicola Montaut, Marcello Massaro, Viktor Quiring, Matteo Santandrea, Linda Sansori, Harald Herrmann, Raimund Ricken, Rita Prevor, Kerstin Falke and Irmgard Zimmermann for lots of fun and a great time.

Finally, and most importantly, my biggest thanks goes to my parents for their

unlimited support over all these years. Without their love and their appreciation for my work, I would never have come this far. Thank you so so much.

Appendix A.

Operators

A.1. Displacement Operator

The displacement operator is defined as

$$\hat{D}(\alpha) = \exp\{\alpha\hat{a}^\dagger - \alpha^*\hat{a}\}. \tag{A.1}$$

It has the properties:

$$\hat{D}(\alpha)|0\rangle = |\alpha\rangle \tag{A.2}$$

$$\hat{D}^\dagger(\alpha) = \hat{D}(-\alpha) \tag{A.3}$$

$$\hat{D}(\alpha)D(\beta) = \hat{D}(\alpha+\beta)\exp\{\frac{1}{2}(\alpha\beta^* - \alpha^*\beta)\} \tag{A.4}$$

$$\hat{D}(\alpha)\hat{a}\hat{D}(-\alpha) = \hat{a} + \alpha \tag{A.5}$$

In the main text we use:

$$\int \mathrm{d}^2\alpha\, \hat{D}(\alpha) = \int \mathrm{d}^2\alpha\, \exp\{\alpha\hat{a}^\dagger - \alpha^*\hat{a}\} \tag{A.6}$$

$$= \int \mathrm{d}^2\alpha\, e^{\alpha\hat{a}} e^{-\alpha^*\hat{a}} e^{-|\alpha|^2/2} \tag{A.7}$$

$$= \sum_{mn} \int \mathrm{d}\phi \mathrm{d}r\, r e^{-r^2/2} r^{m+n} (-1)^n e^{i\phi(m-n)} \frac{\hat{a}^{\dagger m}\hat{a}^n}{m!n!} \tag{A.8}$$

$$= 2\pi \sum_n \int_0^\infty \mathrm{d}r\, r e^{-r^2/2} r^{2n} (-1)^n \frac{\hat{a}^{\dagger n}\hat{a}^n}{(n!)^2} \tag{A.9}$$

$$= 2\pi \sum_n (-1)^n 2^n \frac{\hat{a}^{\dagger n}\hat{a}^n}{n!} \tag{A.10}$$

and hence

$$\langle m| \int \mathrm{d}^2\alpha\, \hat{D}(\alpha)|n\rangle = 2\pi(-1)^n \delta_{mn} \tag{A.11}$$

A.2. Quadrature Operators

The standard quadrature operators are defined as:

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger) \qquad \hat{p} = \frac{-i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger) \tag{A.12}$$

The Wigner function is usually parametrized as $W(x, p) = W(\alpha)$ with

$$\alpha = \frac{1}{\sqrt{2}}(x + ip), \tag{A.13}$$

whereas the $\frac{1}{\sqrt{2}}$-factor is sometimes omitted in the literature. For a rotation of the quadratures by an angle ϕ, we can write them as

$$x_\phi = \sqrt{2}(\alpha_r \cos\phi + \alpha_i \sin\phi) \tag{A.14}$$

$$x_{\phi+\frac{\pi}{2}} = \sqrt{2}(\alpha_i \cos\phi - \alpha_r \sin\phi) \tag{A.15}$$

A.3. Beam Splitter

A beam splitter is a linear unitary transformation from two input ports to two output ports that can be described as (omitting the hats on the operators)

$$\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \begin{pmatrix} \tau & -\rho \\ \rho & \tau \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{A.16}$$

where a_1 and a_2 are the annihilation operators in the two ports and $|\tau|^2 + |\rho^2| = 1$. This can be used to calculate states after a beamsplitter by replacing the creation operators that generate the state. For example:

$$|0, n\rangle' = \sum_{l=0}^{n} \sqrt{\binom{n}{l}} \tau^l \rho^{n-l} |l, n-l\rangle \tag{A.17}$$

$$|1, 1\rangle' = \frac{1}{\sqrt{2}}(|2, 0\rangle - |0, 2\rangle) \tag{A.18}$$

The latter state (where we assumed a 50/50 beam splitter) is a perfectly correlated state, which shows no coincidence counts between the two output ports. This effect is known as the Hong-Ou-Mandel interference [43].

The transformation in Eq. A.16 corresponds to the Hamiltonian

$$H = i\theta(a_1^\dagger a_2 - a_2^\dagger a_1), \tag{A.19}$$

where $\tau = \cos(\theta)$ and $\rho = \sin(\theta)$. Because of this form, the beam splitter can be used to realize the displacement operator

$$D(\alpha) = e^{\alpha a^\dagger - \alpha^* a} \tag{A.20}$$

by using an almost transparent beam splitter ($\theta \to 0$) and approximating a strong coherent field by its complex field amplitude $\theta a_2 \to \alpha$.

From equation A.16 we can also write down the transformation for the quadratures:

$$\begin{pmatrix} x_1' \\ x_2' \\ p_1' \\ p_2' \end{pmatrix} = \begin{pmatrix} \tau & -\rho & 0 & 0 \\ \rho & \tau & 0 & 0 \\ 0 & 0 & \tau & -\rho \\ 0 & 0 & \rho & \tau \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ p_1 \\ p_2 \end{pmatrix} \tag{A.21}$$

Appendix B.

PDC

B.1. Gain from First Principles

According to Eq. 4.29, the gain is given by

$$B' = 4BO\alpha_0 FL \tag{B.1}$$

To calculate the phasematching normalization term F, we assume a decorrelated process and approximate the phasematching as

$$\phi(\omega_s, \omega_i) = e^{-\tau^2(\omega_s - \omega_i)^2}. \tag{B.2}$$

The pump spectrum is given by

$$\alpha(\omega_s, \omega_i) = \sqrt[4]{\frac{2\tau^2}{\pi}} e^{-\tau^2(\omega_s + \omega_i)^2}. \tag{B.3}$$

Then we obtain

$$F = \sqrt{\int \mathrm{d}\omega_s \mathrm{d}\omega_i |\phi(\omega_s, \omega_i)\alpha(\omega_s, \omega_i)|^2} = \frac{\sqrt[4]{2\pi}}{2\sqrt{\tau}}. \tag{B.4}$$

We use the following parameters:

$\chi^{(2)}$ [m/V]	$5.2 \cdot 10^{-9}$
(n_1, n_2, n_3)	$(1.76, 1.74, 1.82)$
$(\omega_1, \omega_2, \omega_3)$ [1/s]	$(2.42, 1.71, 1.71) \cdot 10^{15}$
τ [s]	$1.9 \cdot 10^{-13}$
L [m]	$8 \cdot 10^{-3}$
O [1/m]	$1.5 \cdot 10^5$
α_0	$0 - 10^5$

The pulse duration corresponds to a spectral-intensity FWHM of about 2 nm at a wavelength of 775 nm. (The more common definition of pulse duration in terms of its intensity FWHM would be $\tau_{\mathrm{FWHM}} = 2\sqrt{2\ln 2}\,\tau = 447\,\mathrm{fs}$.)

The overlap of $1.5 \cdot 10^5 \frac{1}{\mathrm{m}}$ is based on the eigenmodes of a realistic waveguide. This value was calculated by Helge Rütz.

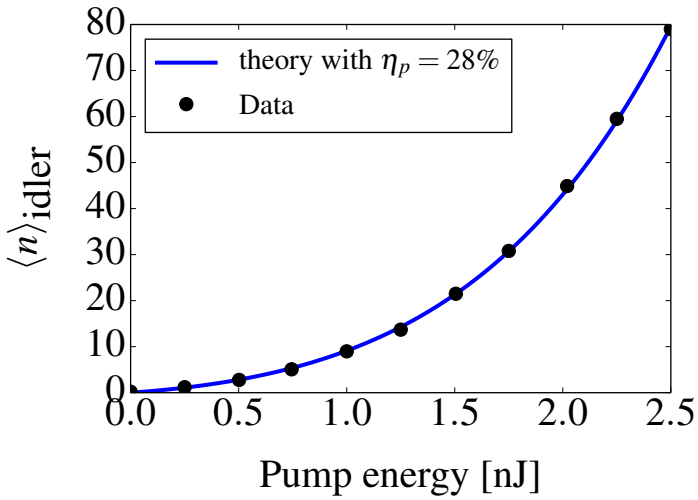


Figure B.1.: Mean photon number versus pump power. The fit has only the coupling efficiency into the waveguide as a free parameter. All other parameters are derived from first principles. The resulting coupling efficiency is lower than expected but still reasonable.

The value of the nonlinear coefficient for our process is quite difficult to extract from the literature. There seem to be only two papers which measure the relevant coefficients [97, 98]. Unfortunately, they use different nomenclatures for KTP, which means that the tensor elements d_{24} and d_{15} are swapped. It seems that this has not been noticed when referencing this work in [33]. This means that the nonlinearity $\chi^{(2)}$ of our process, which we believe to be $y \rightarrow y+z$ corresponding to the element d_{24}, can be either 2.8 $\frac{\text{pm}}{\text{V}}$ or 5.2 $\frac{\text{pm}}{\text{V}}$, taking the values from [97]. We believe the latter one is correct. Additionally, this value was measured at 1313 nm, which means that the value for our process at 1535 nm should be slightly lower. As if those uncertainties were not enough, we cannot reproduce the SHG efficiency from classical optics up to a factor of $\frac{2}{\pi}$ with our theory. This is unfortunate and does not allow us to reliably compare the measured mean photon numbers with theory. Nevertheless, to get a rough comparison between theory and experiment, we use $\chi^{(2)} = 5.2\,\frac{\text{pm}}{\text{V}}$ and multiply it by our correction term $\sqrt{\frac{2}{\pi}}$. The only free parameter remaining is the coupling efficiency of the pump beam into the waveguide. In Fig. B.1 we show the result of this fit to our measured data. The data itself and the experimental setup are discussed in Sec. 5.4.

B.2. Chirp and Decorrelation

For simplicity, we assume a Gaussian spectral envelope function. Chirp in the spectral domain can be described as a frequency dependent phase term in the spectrum of

the pulse:

$$\alpha(\omega_\mathrm{p}) = \alpha_0 \exp\left\{-\frac{\Delta\omega_\mathrm{p}^2}{2\sigma_\mathrm{p}^2} - i(a\Delta\omega_\mathrm{p}^2 + b\Delta\omega_\mathrm{p})\right\}, \tag{B.5}$$

where $\Delta\omega_\mathrm{p} = \Delta\omega_\mathrm{s} + \Delta\omega_\mathrm{i}$ and a and b are chirp parameters. $2a$ is usually referred to as the group velocity dispersion (GVD). The linear term only shifts the pulse by the group velocity. It can also be separated into independent signal and idler terms and can be omitted from here on. To see the effect of the GVD, we also assume a Gaussian phasematching function oriented at 45° in phase space:

$$\phi = \exp\left\{-\frac{(\Delta\omega_\mathrm{s} - \Delta\omega_\mathrm{i})^2}{2\sigma_\phi^2}\right\}, \tag{B.6}$$

leading to the phasematching function

$$f(\omega_\mathrm{s}, \omega_\mathrm{i}) = \alpha_0 \exp\left\{-\left(\frac{1}{2\sigma_\mathrm{p}^2} + ia\right)(\Delta\omega_\mathrm{s} + \Delta\omega_\mathrm{i})^2 - \frac{1}{2\sigma_\phi^2}(\Delta\omega_\mathrm{s} - \Delta\omega_\mathrm{i})^2\right\} \tag{B.7}$$

$$= \alpha_0 \exp\left\{-\left(\frac{1}{2\sigma_\mathrm{p}^2} + ia + \frac{1}{2\sigma_\phi^2}\right)(\Delta\omega_\mathrm{s}^2 + \Delta\omega_\mathrm{i}^2) - \left(\frac{1}{2\sigma_\mathrm{p}^2} + ia - \frac{1}{2\sigma_\phi^2}\right)2\Delta\omega_\mathrm{s}\Delta\omega_\mathrm{i}\right\}. \tag{B.8}$$

If we set $\sigma_\mathrm{p} = \sigma_\phi$, the correlation term almost vanishes and we get

$$f(\omega_\mathrm{s}, \omega_\mathrm{i}) = g(\omega_\mathrm{s})h(\omega_\mathrm{i})e^{-2ia\Delta\omega_\mathrm{s}\Delta\omega_\mathrm{i}}. \tag{B.9}$$

The phase term becomes negligible if

$$\sigma_\mathrm{p}^2 a << 1. \tag{B.10}$$

Assuming a pump spectral width of $\sigma_{\lambda_\mathrm{p}} = 2\,\mathrm{nm}$ at $\lambda_\mathrm{p} = 775\,\mathrm{nm}$ (meaning $\sigma_\mathrm{p} = 1\,\mathrm{THz}$) and $2a = 2 \cdot 10^3\,\mathrm{fs}^2$, which corresponds to 5 cm of glass or 1 cm of KTP, we get

$$\sigma_\mathrm{p}^2 a \approx 10^{-3} << 1. \tag{B.11}$$

That means that dispersion does not play a role with the experimental parameters chosen for this thesis. Even one meter of fiber would just introduce a dispersion of $a = 2 \cdot 10^4\,\mathrm{fs}^2$. Only above these values, dispersion leads to correlations in the JSA which significantly affect its separability.

Quite interestingly, regarding Eq. B.8, we could eliminate the correlation term by a corresponding chirp in the phasematching function. This is possible in principle by designing the poling function $\chi(z)$ accordingly. Since the phasematching function is essentially a Fourier transform of $\chi(z)$, the inverse Fourier transform of $\exp\{-(\frac{1}{2\sigma_\mathrm{p}^2} +$

$ia)\Delta k^2\}$ should do the trick, which is

$$\chi(z) = \exp\left\{\frac{i2\pi z}{\Lambda}\right\}\exp\left\{-\frac{z^2}{2\left(\frac{1}{2\sigma_\mathrm{p}^2}+ia\right)}\right\} \tag{B.12}$$

$$= \exp\left\{\frac{i2\pi z}{\Lambda}\right\}\exp\left\{-\frac{z^2\left(\frac{1}{2\sigma_\mathrm{p}^2}-ia\right)}{2\left(\frac{1}{4\sigma_\mathrm{p}^4}+a^2\right)}\right\} \tag{B.13}$$

$$= \exp\left\{i\left(\frac{2\pi z}{\Lambda}+\frac{z^2}{2\left(\frac{1}{4a\sigma_\mathrm{p}^4}+a\right)}\right)\right\}\exp\left\{-\frac{z^2}{2\left(\frac{1}{2\sigma_\mathrm{p}^2}+2\sigma_\mathrm{p}^2a^2\right)}\right\} \tag{B.14}$$

or something as close as possible with the constraint that χ can only take the values $+1$ or -1. We have added the $2\pi/\Lambda$ term here, even though it was included in the Δk previously. The Gaussian envelope accounts for the Gaussian envelope of the phasematching. In case of a sinc-phasematching this would be a box-function as used in the theory section of PDC with a length L. The new part here is the quadratic phase, very similar to the chirp in Eq. B.5. In conclusion, a quadratic chirp of the pump can be accounted for by a quadratic chirp of the poling momentum vector. This does not mean that signal and idler have no chirp, but that they are separable despite a chirp.

Reversing that argument, imperfections in the waveguide could reduce the separability of the JSA. For example, the effective refractive index can vary over the length of the waveguide, making the effective domain width a function of position $\Lambda(z)$. A temperature variation over the length of the waveguide would have the same effect. Here, one might try to chirp the pump on purpose to achieve a better separability.

B.3. Klyshko Efficiencies

Klyshko efficiencies [116, 118, 119] are the easiest way to characterize loss in a PDC setup. Both signal and idler are sent onto single-photon detectors and the coincidence- and single-click probabilities are recorded. In the low-pump-power regime ($\langle n\rangle << 1$) the probability for a photon-pair is given by

$$p_\mathrm{cc} = p_\mathrm{PDC}\eta_\mathrm{s}\eta_\mathrm{i}. \tag{B.15}$$

Here η_s and η_i are the probabilities for a single photon to make it to the detector and produce a click. The probabilities for single clicks are

$$p_\mathrm{s} = p_\mathrm{PDC}\eta_\mathrm{s} \tag{B.16}$$

$$p_\mathrm{i} = p_\mathrm{PDC}\eta_\mathrm{i}. \tag{B.17}$$

Now we see that

$$\eta_\mathrm{s} = \frac{p_\mathrm{cc}}{p_\mathrm{i}} \qquad \eta_\mathrm{i} = \frac{p_\mathrm{cc}}{p_\mathrm{s}}. \tag{B.18}$$

Two important aspects of a real setup are not taken into account in this expressions. Firstly, higher-order photons inevitably affect the measured values. This can be accounted for by analyzing the scaling of the measured η_s with pump power. η_s will depend roughly linear on the pump power. Therefore, an extrapolation to zero power yields the correct value. Secondly, background photons, which do not come in pairs increase the single counts and therefore reduce the calculated efficiencies. This effect cannot be easily avoided. One option is to use photon-number resolving detectors and fit a reasonable model to the data as we do in Sec. 5.4.2.

Appendix C.

Miscellaneous

C.1. Measurement Scheme for Second-Order Correlation

To measure the second-order correlation function $g^{(2)}$, it is sufficient to split the state at a 50/50 beam splitter and measure it with low-efficiency click detectors, as sketched in Fig. C.1. This works, because in the limit of a low click probability, a click detector acts as a linear detector and can be described by the operator $\eta\hat{n}$, where η is the efficiency and $\hat{n}$ the number operator. Then, the coincidence probability (a joint click in modes a_1 and a_2) divided by the single-click probability gives

$$\frac{p_{\mathrm{cc}}}{p_{\mathrm{s}_1}p_{\mathrm{s}_2}} = \frac{\langle \eta_1\eta_2 a_1^\dagger a_1 a_2^\dagger a_2\rangle}{\langle \eta_1 a_1^\dagger a_1\rangle\langle \eta_2 a_2^\dagger a_2\rangle} = \frac{\langle a_1^\dagger a_1 a_2^\dagger a_2\rangle}{\langle a_1^\dagger a_1\rangle\langle a_2^\dagger a_2\rangle}. \tag{C.1}$$

Applying the beam splitter transformations for $a \leftrightarrow b$ and using the fact that one of the ports is a vacuum port, we get

$$a_1^\dagger a_1 a_2^\dagger a_2 \propto (b_1^\dagger + b_2^\dagger)(b_1 + b_2)(b_1^\dagger - b_2^\dagger)(b_1 - b_2) = b_1^\dagger(b_1 + b_2)(b_1^\dagger - b_2^\dagger)b_1 \tag{C.2}$$

$$= b_1^\dagger b_1 b_1^\dagger b_1 - b_1^\dagger b_2 b_2^\dagger b_1 = b_1^{\dagger 2} b_1^2 \tag{C.3}$$

and hence

$$\frac{p_{\mathrm{cc}}}{p_{\mathrm{s}_1}p_{\mathrm{s}_2}} = \frac{\langle b_1^{\dagger 2} b_1^2\rangle}{\langle b_1^\dagger b_1\rangle^2} = g^{(2)}. \tag{C.4}$$

This is a very simple yet powerful measurement because it gives information about the number of modes in a PDC state or identifies sub-Poissonian, nonclassical states.

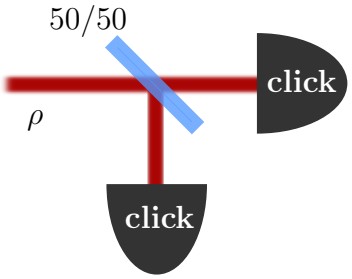


Figure C.1.: Measurement scheme for $g^{(2)}$.

C.2. Tomography using Convex Optimization

Here, we discuss the numerical implementation of a convex quadratic minimization problem for detector tomography and state tomography. The method was introduced for detector tomography in [128] and used in [162, 163]. However, the details of the numerical implementation are not given in the paper. Here, we describe our implementation in python.

The Problem

We start with detector tomography of a single phase insensitive photon-number detector (though the method can be extended to multidimensional and phase sensitive detectors). The POVM is given by

$$\Pi_k = \sum_n \theta_{kn} |n\rangle\langle n| \tag{C.5}$$

and the outcome probabilities by

$$p_k = \mathrm{Tr}(\Pi_k \rho). \tag{C.6}$$

In detector tomography, the states are known and the probabilities are measured. We use coherent states and append another index to the equation:

$$p_{k\alpha} = \mathrm{Tr}(\Pi_k \rho_\alpha). \tag{C.7}$$

In the photon-number basis this equation becomes

$$p_{k\alpha} = \sum_n \theta_{kn} r_{n\alpha}, \tag{C.8}$$

where $r_{nk} = |\langle n|\alpha\rangle|^2$. In matrix notation it can be written as

$$P = \Theta R. \tag{C.9}$$

Since the number of measurements is always finite, the measured probabilities will not be exactly P but drawn from a multinomial distribution according to probability distribution P. We call the matrix with measured probabilities F. The minimization problem is then formulated as follows:

$$\text{minimize} \quad \{||F - \Theta R||_2 + \gamma S\} \tag{C.10}$$

$$\theta_{kn} \geq 0 \qquad \sum_k \theta_{kn} = 1. \tag{C.11}$$

The two constraints ensure that all probabilities are positive and sum to one. The 2-norm is defined as $||A||_2 = \sqrt{\sum_{ij} |a_{ij}|^2}$. S is a regularization factor defined as

$$S = \sum_{kn} (\theta_{kn} - \theta_{k,n+1})^2. \tag{C.12}$$

The influence of S can be tuned with the parameter γ. Since S penalizes high differences in the value of adjacent matrix elements, it ensures a smooth looking matrix. This assumption is justified by the effect of losses on the detector.

The minimization problem is convex, quadratic (norm-2) and semidefinite ($\pi \geq 0$) and is therefore numerically feasible.

Implementation for Detector Tomography

Our implementation of choice is done in python using the module cvxopt. This is a basic, powerful module for convex optimization problems. We use the solver cvxopt.solvers.cp(). It understands the following problem:

$$\begin{aligned} \text{minimize} \quad & y(x) && \text{(C.13)} \\ \text{subject to} \quad & Ax = b && \text{(C.14)} \\ & Gx \preceq h && \text{(C.15)} \end{aligned}$$

where y is a scalar function, x the vector to be optimized, A and G matrices and b and h vectors. The inequality $\preceq$ is understood element-wise. Since our problem is formulated for a matrix Θ, we have to identify the vector elements of x with our matrix elements. If Θ is a $K \times N$ matrix, a convenient way is to 'flatten' the matrix by identifying

$$\theta_{kn} = x_{Nk+n}. \tag{C.16}$$

The matrices A, G and vectors g, h consist of ones and zeros:

$$A_{ij} = \sum_k \delta_{i,j+Nk} \qquad b_i = 1 \tag{C.17}$$

$$G_{ij} = -\delta_{i,j} \qquad h = 0 \tag{C.18}$$

To incorporate the cost-term, we minimize $y(x) = g(x) + \gamma S(x)$, where $g(x) = ||F - \Theta R||_2^2$ and S given by Eq. C.12. Note that we minimize the square of the distance rather than the distance itself because we fount that it is numerically faster. This should not make a difference in the final solution. The solver requires the specific definitions of the first and second order derivatives of $y(x)$. After careful calculations we find that

$$\frac{\mathrm{d}g}{\mathrm{d}\theta_{k'n'}} = -2\sum_\alpha (f_{k'\alpha} - \sum_n \theta_{k'n} r_{n\alpha}) r_{n'\alpha} \tag{C.19}$$

and

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\theta_{k''n''}\theta_{k'n'}} = 2\sum_\alpha \delta_{k'k''} r_{n''\alpha} r_{n'\alpha}. \tag{C.20}$$

In this notation, the first derivative is a $K \times N$ matrix and the second derivative is a $K \times N \times K \times N$ tensor. This has to be translated to a KN vector and $KN \times$

KN matrix, respectively, which is just a redefinition of indices. We also need the derivatives of S:

$$\frac{\mathrm{d}S}{\mathrm{d}\theta_{k'n'}} = \begin{cases} 2(\theta_{k'0} - \theta_{k'1}) & n' = 0 \\ 2(2\theta_{k'n'} - \theta_{k'n'+1} - \theta_{k'n'-1}) & 0 < n' < N \\ 2(\theta_{k'N} - \theta_{k'N-1}) & n' = N \end{cases} \tag{C.21}$$

and

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\theta_{k''n''}\theta_{k'n'}} = \begin{cases} 2\delta_{k'k''}(\delta_{n''0} - \delta_{n''1}) & n' = 0 \\ 2\delta_{k'k''}(2\delta_{n''n'} - \delta_{n''n'+1} - \delta_{n''n'-1}) & 0 < n' < N \\ 2\delta_{k'k''}(\delta_{n''N} - \delta_{n''N-1}) & n' = N \end{cases} \tag{C.22}$$

Eq. C.20 and all of Eqs. C.22 are not dependent on Θ and have to be defined only ones, using for-loops if yo want. The solver is amazingly fast which is probably due to the knowledge of the derivatives. Please look into the python code (not provided with the thesis) and the users guide of cvxopt under the topic nonlinear convex optimization for further details on the implementation.

In Fig. C.2 we show simulation result for an 8−bin TMD with a detection efficiency of 20%. This was calculated using 100 different coherent states with $0 \leq \alpha < 3.5$, 10^6 events, setting photon-number contributions above 20 to zero and $\gamma = 0.01$.

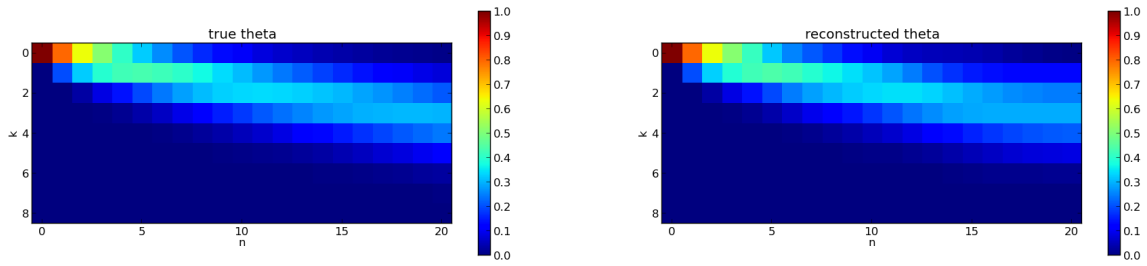


Figure C.2.: The true and the reconstructed POVM elements θ_{kn}. The reconstruction works quite well. Without the regularization function S, the elements for higher n would fluctuate extremely between zero and one. A direct inversion method for these parameters would not work at all, yielding unphysical results.

Implementation for One-Mode State Tomography

Since state tomography and detector tomography are very similar, the same approach can be used. State tomography is even simpler, because it needs one index less. We still assume phase insensitive detectors, such that only the photon-number distribution of the state can be measured. Therefore, we can assume that the state is diagonal in the photon-number basis and hence described by a vector ρ. The matrix equation is

$$p = \Theta\rho, \tag{C.23}$$

where p are the click probabilities, ρ the state to be measured and Θ the detector matrix, possibly known from preceding detector tomography. Again, the measured frequencies are denoted by f and we can formulate the problem:

$$\text{minimize} \quad \{||f - \Theta\rho||_2^2 + \gamma S\} \tag{C.24}$$

$$\rho_n \geq 0 \qquad \sum_n \rho_n = 1. \tag{C.25}$$

The 2-norm is just the length of the vector. The regularization factor S can be defined as

$$S = \sum_n (\rho_n - \rho_{n+1})^2. \tag{C.26}$$

Whereas in detector tomography, we could justify the regularization by the presence of losses, in this case there is no such justification. The regularization will artificially broaden the photon-number distribution. For Fock states, the effect should be the strongest. Therefore, we can use Fock states to estimate the influence on photon-number statistics.

The derivatives can be calculated to be

$$\frac{\mathrm{d}g}{\mathrm{d}\rho_\alpha} = -2\sum_i (f_i - \sum_j \theta_{ij}\rho_j)\theta_{i\alpha} \tag{C.27}$$

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\rho_\beta \mathrm{d}\rho_\alpha} = 2\sum_i \theta_{i\alpha}\theta_{i\beta} \tag{C.28}$$

Furthermore

$$\frac{\mathrm{d}S}{\mathrm{d}\rho_\alpha} = \begin{cases} 2(\rho_0 - \rho_1) & \alpha = 0 \\ 2(2\rho_\alpha - \rho_{\alpha+1} - \rho_{\alpha-1}) & 0 < \alpha < N \\ 2(\rho_N - \rho_{N-1}) & \alpha = N \end{cases} \tag{C.29}$$

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\rho_\beta \mathrm{d}\rho_\alpha} = \begin{cases} 2(\delta_{\beta,0} - \delta_{\beta,1}) & \alpha = 0 \\ 2(2\delta_{\alpha,\beta} - \delta_{\alpha+1,\beta} - \delta_{\alpha-1,\beta}) & 0 < \alpha < N \\ 2(\delta_{N,\beta} - \delta_{N-1,\beta}) & \alpha = N. \end{cases} \tag{C.30}$$

We find that $\gamma \approx 10^{-7}$ is sufficient to smoothen the reconstructed photon-number probabilities whereas still maintaining fidelities above 99% for Fock states and coherent states. Similar to detector tomography, the smoothing improves the results.

Implementation for Two-Mode State Tomography

Only small changes have to be made to apply the previous code to two-mode tomography. We assume that the two TMDs are independent. The detector is then characterized by

$$\Theta = \Theta_1 \otimes \Theta_2, \tag{C.31}$$

where Θ_i are the POVMs of TMD i and $\otimes$ refers to the tensor product. The measured frequencies can be characterized by a matrix $\tilde{f}_{m,n}$, which gives the measured click probabilities for m photons in TMD 1 and n photons in TMD 2. This matrix has to be flattened to $f_i = f_{km+n} = \tilde{f}_{m,n}$, where $k = N + 1$ and N is the number of TMD bins. Now we can use the formulation of Eq. C.24. We just have to be careful about the regularization S. It should be changed to

$$S = \sum_{mn} [(\rho_{mk+n} - \rho_{mk+n+1})^2 + (\rho_{mk+n} - \rho_{(m+1)k+n})^2] \tag{C.32}$$

with the appropriate derivatives $\frac{\mathrm{d}S}{\mathrm{d}\rho_{\alpha k+\beta}} =$

$$\begin{cases} 2(\rho_0 - \rho_1) & \\ +2(\rho_0 - \rho_k) & \alpha = 0, \beta = 0 \\ 2(2\rho_\beta - \rho_{\beta+1} - \rho_{\beta-1}) & \\ +2(\rho_\beta - \rho_{k+\beta}) & \alpha = 0, 0 < \beta < N \\ 2(\rho_{\alpha k} - \rho_{\alpha k+1}) & \\ +2(2\rho_{\alpha k} - \rho_{(\alpha+1)k} - \rho_{(\alpha-1)k}) & 0 < \alpha < N, \beta = 0 \\ 2(2\rho_{\alpha k+\beta} - \rho_{\alpha k+\beta+1} - \rho_{\alpha k+\beta-1}) & \\ +2(2\rho_{\alpha k+\beta} - \rho_{(\alpha+1)k+\beta} - \rho_{(\alpha-1)k+\beta}) & 0 < \alpha < N, 0 < \beta < N \\ 2(2\rho_{Nk+\beta} - \rho_{Nk+\beta+1} - \rho_{Nk+\beta-1}) & \\ +2(\rho_{Nk+\beta} - \rho_{(N-1)k+\beta}) & \alpha = N, 0 < \beta < N \\ 2(\rho_{\alpha k+N} - \rho_{\alpha k+N-1}) & \\ +2(2\rho_{\alpha k+N} - \rho_{(\alpha+1)k+N} - \rho_{(\alpha-1)k+N}) & 0 < \alpha < N, \beta = N \\ 2(\rho_{Nk+N} - \rho_{Nk+N-1}) & \\ +2(\rho_{Nk+N} - \rho_{(N-1)k+N}) & \alpha = N, \beta = N. \end{cases} \tag{C.33}$$

All these cases mean that for each entry in the $k \times k$ matrix you have to subtract its direct neighbors (four in the middle, three at the borders and two at the corners).

And now the second derivatives $\frac{\mathrm{d}S}{\mathrm{d}\rho_{\alpha k+\beta}\mathrm{d}\rho_{\gamma k+\delta}} =$

$$\begin{cases} 2\delta_{0,\gamma}(\delta_{0,\delta}-\delta_{1,\delta}) & \\ +2\delta_{0,\delta}(\delta_{0,\gamma}-\delta_{1,\gamma}) & \alpha=0,\beta=0 \\ 2\delta_{0,\gamma}(2\delta_{\beta,\delta}-\delta_{\beta+1,\delta}-\rho_{\beta-1,\delta}) & \\ +2\delta_{\beta,\delta}(\delta_{0,\gamma}-\delta_{1,\gamma}) & \alpha=0,0<\beta<N \\ 2\delta_{\alpha,\gamma}(\delta_{0,\delta}-\delta_{1,\delta}) & \\ +2\delta_{0,\delta}(2\delta_{\alpha,\gamma}-\delta_{(\alpha+1),\gamma}-\delta_{(\alpha-1),\gamma}) & 0<\alpha<N,\beta=0 \\ 2\delta_{\alpha,\gamma}(2\delta_{\beta,\delta}-\delta_{\beta+1,\delta}-\delta_{\beta-1,\delta}) & \\ +2\delta_{\beta,\delta}(2\delta_{\alpha,\gamma}-\delta_{\alpha+1,\gamma}-\delta_{\alpha-1,\gamma}) & 0<\alpha<N,0<\beta<N \\ 2\delta_{N,\gamma}(2\delta_{\beta,\delta}-\delta_{\beta+1,\delta}-\delta_{\beta-1,\delta}) & \\ +2\delta_{\beta,\delta}(\delta_{N,\gamma}-\delta_{N-1,\gamma}) & \alpha=N,0<\beta<N \\ 2\delta_{\alpha,\gamma}(\delta_{N,\delta}-\delta_{N-1,\delta}) & \\ +2\delta_{N,\delta}(2\delta_{\alpha,\gamma}-\delta_{\alpha+1,\gamma}-\delta_{\alpha-1,\gamma}) & 0<\alpha<N,\beta=N \\ 2\delta_{N,\gamma}(\delta_{N,\delta}-\delta_{N-1,\delta}) & \\ +2\delta_{N,\delta}(\delta_{N,\gamma}-\delta_{N-1,\gamma}) & \alpha=N,\beta=N. \end{cases} \tag{C.34}$$

Appendix D.

Author Contributions

[1] G. Harder, V. Ansari, B. Brecht, T. Dirmeier, C. Marquardt, and C. Silberhorn: "An optimized photon pair source for quantum circuits". *Opt. Express* 21, 13975–13985 (June 2013)

Georg Harder and Vahid Ansari performed the experiment and analyzed the data. Benjamin Brecht, Thomas Dirmeier, Christoph Marquardt and Christine Silberhorn took part in scientific discussions. Georg Harder wrote the manuscript with contributions from Vahid Ansari and feedback from all authors. Christine Silberhorn supervised the work.

[2] G. Harder, T. J. Bartley, A. E. Lita, S. W. Nam, T. Gerrits, and C. Silberhorn: "Single-Mode Parametric-Down-Conversion States with 50 Photons as a Source for Mesoscopic Quantum Optics". *Phys. Rev. Lett.* 116, 143601 (Apr. 2016)

Georg Harder and Tim Bartley equally contributed to this work in taking the data, analyzing it and writing the manuscript. Thomas Gerrits supervised and took part in the experiment. Adriana Lita and Sae Woo Nam developed the TES detectors used in this experiment. Christine Silberhorn, Sae Woo Nam and Thomas Gerrits supervised the work.

[3] G. Harder, D. Mogilevtsev, N. Korolkova, and C. Silberhorn: "Tomography by Noise". *Phys. Rev. Lett.* 113, 070403 (Aug. 2014)

Dmitri Mogilevtsev and Natalia Korolkova conceived the idea for the scheme. Georg Harder carried out the experiment. Dmitri Mogilevtsev analyzed the data and wrote the manuscript with contributions from Georg Harder. Natalia Korolkova and Christine Silberhorn supervised the work.

[4] L. Motka, B. Stoklasa, J. Rehacek, Z. Hradil, V. Karasek, D. Mogilevtsev, G. Harder, C. Silberhorn, and L. L. Sánchez-Soto: "Efficient algorithm for optimizing data-pattern tomography". *Phys. Rev. A* 89, 054102 (May 2014)

Libor Motka, Bohumil Stoklasa, Jaroslav Rehacek, Zdenek Hradil, V. Karasek, Dmitri Mogilevtsev and Luis Sanchez-Soto developed the algorithm and wrote the manuscript. Georg Harder contributed in discussions with experimental considerations. Georg Harder and Christine Silberhorn gave feedback on the manuscript.

[5] G. Harder, C. Silberhorn, J. Rehacek, Z. Hradil, L. Motka, B. Stoklasa, and L. L. Sánchez-Soto: "Time-multiplexed measurements of nonclassical light at telecom wavelengths". *Phys. Rev. A* 90, 042105 (Oct. 2014)

Georg Harder carried out the experiment. Jaroslav Rehacek, Libor Motka and Bohumil Stoklasa analyzed the data. Luis Sanchez-Soto wrote the manuscript with contributions from Georg Harder, Jaroslav Rehacek, Libor Motka and Bohumil Stoklasa. Christine Silberhorn and Zdenek Hradil supervised the work.

[6] G. Harder, C. Silberhorn, J. Rehacek, Z. Hradil, L. Motka, B. Stoklasa, and L. L. Sánchez-Soto: "Local Sampling of the Wigner Function at Telecom Wavelength with Loss-Tolerant Detection of Photon Statistics". *Phys. Rev. Lett.* 116, 133601 (Mar. 2016)

Georg Harder and Christine Silberhorn developed the idea. Georg Harder carried out the experiment. Jaroslav Rehacek, Libor Motka and Bohumil Stoklasa analyzed the data. Luis Sanchez-Soto wrote the manuscript with contributions from Georg Harder, Jaroslav Rehacek, Libor Motka and Bohumil Stoklasa. Christine Silberhorn and Zdenek Hradil supervised the work.

[7] J. Sperling, M. Bohmann, W. Vogel, G. Harder, B. Brecht, V. Ansari, and C. Silberhorn: "Uncovering Quantum Correlations with Time-Multiplexed Click Detection". *Phys. Rev. Lett.* 115, 023601 (July 2015)

Jan Sperling, Martin Bohmann and Werner Vogel developed the theory and analyzed the data. Georg Harder carried out the experiment. Benjamin Brecht and Vahid Ansari initiated this project and contributed in discussions. Christine Silberhorn and Werner Vogel supervised the work.

Bibliography

[1] G. Harder, V. Ansari, B. Brecht, T. Dirmeier, C. Marquardt, and C. Silberhorn: "An optimized photon pair source for quantum circuits". *Opt. Express* 21, 13975–13985 (2013) (cit. on pp. 1, 3, 125)

[2] G. Harder, T. J. Bartley, A. E. Lita, S. W. Nam, T. Gerrits, and C. Silberhorn: "Single-Mode Parametric-Down-Conversion States with 50 Photons as a Source for Mesoscopic Quantum Optics". *Phys. Rev. Lett.* 116, 143601 (2016) (cit. on pp. 1, 3, 125)

[3] G. Harder, D. Mogilevtsev, N. Korolkova, and C. Silberhorn: "Tomography by Noise". *Phys. Rev. Lett.* 113, 070403 (2014) (cit. on pp. 1, 3, 125)

[4] L. Motka, B. Stoklasa, J. Rehacek, Z. Hradil, V. Karasek, D. Mogilevtsev, G. Harder, C. Silberhorn, and L. L. Sánchez-Soto: "Efficient algorithm for optimizing data-pattern tomography". *Phys. Rev. A* 89, 054102 (2014) (cit. on pp. 1, 3, 64, 125)

[5] G. Harder, C. Silberhorn, J. Rehacek, Z. Hradil, L. Motka, B. Stoklasa, and L. L. Sánchez-Soto: "Time-multiplexed measurements of nonclassical light at telecom wavelengths". *Phys. Rev. A* 90, 042105 (2014) (cit. on pp. 1, 3, 126)

[6] G. Harder, C. Silberhorn, J. Rehacek, Z. Hradil, L. Motka, B. Stoklasa, and L. L. Sánchez-Soto: "Local Sampling of the Wigner Function at Telecom Wavelength with Loss-Tolerant Detection of Photon Statistics". *Phys. Rev. Lett.* 116, 133601 (2016) (cit. on pp. 1, 3, 126)

[7] J. Sperling, M. Bohmann, W. Vogel, G. Harder, B. Brecht, V. Ansari, and C. Silberhorn: "Uncovering Quantum Correlations with Time-Multiplexed Click Detection". *Phys. Rev. Lett.* 115, 023601 (2015) (cit. on pp. 1, 3, 100, 126)

[8] A. Einstein, B. Podolsky, and N. Rosen: "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Phys. Rev.* 47, 777–780 (1935) (cit. on pp. 5, 27)

[9] J. S. Bell: "On the Einstein Podolsky Rosen paradox". *Physics* 1, 195 (1964) (cit. on pp. 5, 28)

[10] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat: "Detection-Loophole-Free Test of Quantum Nonlocality, and Applications". *Phys. Rev. Lett.* 111, 130406 (2013) (cit. on p. 5)

[11] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson: "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres". *Nature* 526, 682–686 (2015) (cit. on p. 5)

[12] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam: "Strong Loophole-Free Test of Local Realism*". *Phys. Rev. Lett.* 115, 250402 (2015) (cit. on pp. 5, 37)

[13] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger: "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons". *Phys. Rev. Lett.* 115, 250401 (2015) (cit. on pp. 5, 37)

[14] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis: "State preservation by repetitive error detection in a superconducting quantum circuit". *Nature* 519, 66–69 (2015) (cit. on p. 5)

[15] J. Eisert, S. Scheel, and M. B. Plenio: "Distilling Gaussian States with Gaussian Operations is Impossible". *Phys. Rev. Lett.* 89, 137903 (2002) (cit. on pp. 6, 34)

[16] A. Heidmann, R. J. Horowicz, S. Reynaud, E. Giacobino, C. Fabre, and G. Camy: "Observation of Quantum Noise Reduction on Twin Laser Beams". *Phys. Rev. Lett.* 59, 2555–2557 (1987) (cit. on p. 6)

[17] O. Aytür and P. Kumar: "Pulsed twin beams of light". *Phys. Rev. Lett.* 65, 1551–1554 (1990) (cit. on pp. 6, 52)

[18] D. T. Smithey, M. Beck, M. Belsley, and M. G. Raymer: "Sub-shot-noise correlation of total photon number using macroscopic twin pulses of light". *Phys. Rev. Lett.* 69, 2650–2653 (1992) (cit. on p. 6)

[19] M. Bondani, A. Allevi, G. Zambra, M. G. A. Paris, and A. Andreoni: "Sub-shot-noise photon-number correlation in a mesoscopic twin beam of light". *Phys. Rev. A* 76, 013833 (2007) (cit. on p. 6)

[20] I. N. Agafonov, M. V. Chekhova, and G. Leuchs: "Two-color bright squeezed vacuum". *Phys. Rev. A* 82, 011801 (2010) (cit. on p. 6)

[21] P. Sharapova, A. M. Pérez, O. V. Tikhonova, and M. V. Chekhova: "Schmidt modes in the angular spectrum of bright squeezed vacuum". *Phys. Rev. A* 91, 043816 (2015) (cit. on pp. 6, 103)

[22] S. Yokoyama, R. Ukai, S. C. Armstrong, C. Sornphiphatphong, T. Kaji, S. Suzuki, J.-i. Yoshikawa, H. Yonezawa, N. C. Menicucci, and A. Furusawa: "Ultra-large-scale continuous-variable cluster states multiplexed in the time domain". *Nat Photon* 7, 982–986 (2013) (cit. on p. 6)

[23] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan: "Observation of eight-photon entanglement". *Nat Photon* 6, 225–228 (2012) (cit. on p. 6)

[24] Y.-F. Huang, B.-H. Liu, L. Peng, Y.-H. Li, L. Li, C.-F. Li, and G.-C. Guo: "Experimental generation of an eight-photon Greenberger–Horne–Zeilinger state". *Nature Communications* 2, 546 (2011) (cit. on p. 6)

[25] J. D. Jackson: *Classical electrodynamics*. Wiley, 1999 (cit. on p. 10)

[26] F. Dell'Anno, S. De Siena, and F. Illuminati: "Multiphoton quantum optics and quantum state engineering". *Physics Reports* 428, 53–168 (2006) (cit. on p. 10)

[27] N. A. R. Bhat and J. E. Sipe: "Hamiltonian treatment of the electromagnetic field in dispersive and absorptive structured media". *Phys. Rev. A* 73, 063808 (2006) (cit. on p. 10)

[28] L. Mandel and E. Wolf: *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995 (cit. on pp. 10, 96)

[29] R. Loudon: *The Quantum Theory of Light*. OUP Oxford, 2000 (cit. on pp. 10, 12, 56)

[30] A. Christ, B. Brecht, W. Mauerer, and C. Silberhorn: "Theory of quantum frequency conversion and type-II parametric down-conversion in the high-gain regime". *New J. Phys.* 15, 053038 (2013) (cit. on pp. 11, 52)

[31] N. Quesada and J. E. Sipe: "Effects of time ordering in quantum nonlinear optics". *Phys. Rev. A* 90, 063840 (2014) (cit. on pp. 11, 16)

[32] N. Quesada and J. E. Sipe: "Time-Ordering Effects in the Generation of Entangled Photons Using Nonlinear Optical Processes". *Phys. Rev. Lett.* 114, 093903 (2015) (cit. on pp. 11, 52)

[33] T. Suhara and M. Fujimura: *Waveguide Nonlinear-Optic Devices*. Springer Science & Business Media, 2013 (cit. on pp. 13, 39, 112)

[34] P. Ben Dixon, J. H. Shapiro, and F. N. C. Wong: "Spectral engineering by Gaussian phase-matching for quantum photonics". *Optics Express* 21, 5879 (2013) (cit. on pp. 14, 39)

[35] A. Dosseva, Ł. Cincio, and A. M. Brańczyk: "Shaping the joint spectrum of down-converted photons through optimized custom poling". *Phys. Rev. A* 93, 013801 (2016) (cit. on p. 14)

[36] A. B. U'Ren, C. Silberhorn, K. Banaszek, and I. A. Walmsley: "Efficient Conditional Preparation of High-Fidelity Single Photon States for Fiber-Optic Quantum Networks". *Phys. Rev. Lett.* 93, 093601 (2004) (cit. on pp. 14, 38)

[37] A. B. U'Ren, C. Silberhorn, K. Banaszek, I. A. Walmsley, R. Erdmann, W. P. Grice, and M. G. Raymer: "Generation of Pure-State Single-Photon Wavepackets by Conditional Preparation Based on Spontaneous Parametric Downconversion". *Laser Phys.* 15, 146 (2005) (cit. on pp. 14, 38)

[38] C. K. Law, I. A. Walmsley, and J. H. Eberly: "Continuous Frequency Entanglement: Effective Finite Hilbert Space and Entropy Control". *Phys. Rev. Lett.* 84, 5304–5307 (2000) (cit. on p. 14)

[39] W. P. Grice and I. A. Walmsley: "Spectral information and distinguishability in type-II down-conversion with a broadband pump". *Phys. Rev. A* 56, 1627–1634 (1997) (cit. on p. 14)

[40] S. Barnett and P. M. Radmore: *Methods in Theoretical Quantum Optics*. Clarendon Press, 2002 (cit. on pp. 16, 22, 26)

[41] B. Brecht, A. Eckstein, A. Christ, H. Suche, and C. Silberhorn: "From quantum pulse gate to quantum pulse shaper—engineered frequency conversion in nonlinear optical waveguides". *New Journal of Physics* 13, 065029 (2011) (cit. on pp. 16, 38)

[42] A. Christ, K. Laiho, A. Eckstein, K. N. Cassemiro, and C. Silberhorn: "Probing multimode squeezing with correlation functions". *New J. Phys.* 13, 033027 (2011) (cit. on pp. 19, 30, 45, 54)

[43] C. K. Hong, Z. Y. Ou, and L. Mandel: "Measurement of subpicosecond time intervals between two photons by interference". *Phys. Rev. Lett.* 59, 2044–2046 (1987) (cit. on pp. 19, 108)

[44] T. Pittman, B. Jacobs, and J. Franson: "Heralding single photons from pulsed parametric down-conversion". *Optics Communications* 246, 545–550 (2005) (cit. on pp. 19, 38)

[45] P. J. Mosley, J. S. Lundeen, B. J. Smith, and I. A. Walmsley: "Conditional preparation of single photons using parametric downconversion: a recipe for purity". *New J. Phys.* 10, 093011 (2008) (cit. on p. 19)

[46] F. W. Sun and C. W. Wong: "Indistinguishability of independent single photons". *Phys. Rev. A* 79, 013824 (2009) (cit. on p. 19)

[47] M. Avenhaus, M. V. Chekhova, L. A. Krivitsky, G. Leuchs, and C. Silberhorn: "Experimental verification of high spectral entanglement for pulsed waveguided spontaneous parametric down-conversion". *Phys. Rev. A* 79, 043836 (2009) (cit. on p. 19)

[48] W. Wasilewski, P. Kolenderski, and R. Frankowski: "Spectral Density Matrix of a Single Photon Measured". *Phys. Rev. Lett.* 99, 123601 (2007) (cit. on p. 19)

[49] P. Kolenderski and W. Wasilewski: "Derivation of the density matrix of a single photon produced in parametric down-conversion". *Phys. Rev. A* 80, 015801 (2009) (cit. on p. 19)

[50] K. N. Cassemiro, K. Laiho, and C. Silberhorn: "Accessing the purity of a single photon by the width of the Hong–Ou–Mandel interference". *New J. Phys.* 12, 113052 (2010) (cit. on p. 21)

[51] W. Vogel, D.-G. Welsch, and S. Wallentowitz: *Quantum Optics: An Introduction*. Wiley, 2001 (cit. on pp. 22, 26, 27, 96)

[52] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, and A. Acín: "Almost all quantum states have nonclassical correlations". *Phys. Rev. A* 81, 052318 (2010) (cit. on p. 27)

[53] H. Ollivier and W. H. Zurek: "Quantum Discord: A Measure of the Quantumness of Correlations". *Phys. Rev. Lett.* 88, 017901 (2001) (cit. on p. 28)

[54] M. A. Nielsen and I. L. Chuang: *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010 (cit. on pp. 28, 31)

[55] A. Ferraro and M. G. A. Paris: "Nonclassicality Criteria from Phase-Space Representations and Information-Theoretical Constraints Are Maximally Inequivalent". *Phys. Rev. Lett.* 108, 260403 (2012) (cit. on pp. 28, 29)

[56] M. S. Kim, W. Son, V. Bužek, and P. L. Knight: "Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement". *Phys. Rev. A* 65, 032323 (2002) (cit. on p. 28)

[57] M. A. Nielsen: "Conditions for a Class of Entanglement Transformations". *Phys. Rev. Lett.* 83, 436–439 (1999) (cit. on p. 28)

[58] R. J. Glauber: "The Quantum Theory of Optical Coherence". *Phys. Rev.* 130, 2529–2539 (1963) (cit. on p. 29)

[59] M. Avenhaus, K. Laiho, M. V. Chekhova, and C. Silberhorn: "Accessing Higher Order Correlations in Quantum Optical States by Time Multiplexing". *Phys. Rev. Lett.* 104, 063602 (2010) (cit. on p. 30)

[60] A. Peres: "Neumark's theorem and quantum inseparability". *Found Phys* 20, 1441–1453 (1990) (cit. on p. 32)

[61] U. Leonhardt: *Essential Quantum Optics: From Quantum Measurements to Black Holes*. Cambridge University Press, 2010 (cit. on pp. 32, 59)

[62] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani: "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum". *Phys. Rev. Lett.* 70, 1244–1247 (1993) (cit. on pp. 33, 59)

[63] T. Eberle, S. Steinlechner, J. Bauchrowitz, V. Händchen, H. Vahlbruch, M. Mehmet, H. Müller-Ebhardt, and R. Schnabel: "Quantum Enhancement of the Zero-Area Sagnac Interferometer Topology for Gravitational Wave Detection". *Phys. Rev. Lett.* 104, 251102 (2010) (cit. on pp. 33, 56)

[64] K. Banaszek and K. Wódkiewicz: "Direct Probing of Quantum Phase Space by Photon Counting". *Phys. Rev. Lett.* 76, 4344 (1996) (cit. on p. 34)

[65] S. Wallentowitz and W. Vogel: "Unbalanced homodyning for quantum state measurements". *Phys. Rev. A* 53, 4528–4533 (1996) (cit. on p. 34)

[66] K. Laiho: "Characterization of Quantum Optical States in Discrete Variables". PhD thesis. 2011 (cit. on p. 34)

[67] K. Laiho, K. N. Cassemiro, D. Gross, and C. Silberhorn: "Probing the Negative Wigner Function of a Pulsed Single Photon Point by Point". *Phys. Rev. Lett.* 105, 253603 (2010) (cit. on p. 34)

[68] G. Puentes, J. S. Lundeen, M. P. A. Branderhorst, H. B. Coldenstrodt-Ronge, B. J. Smith, and I. A. Walmsley: "Bridging Particle and Wave Sensitivity in a Configurable Detector of Positive Operator-Valued Measures". *Phys. Rev. Lett.* 102, 080404 (2009) (cit. on p. 34)

[69] A. Kuzmich, I. A. Walmsley, and L. Mandel: "Violation of Bell's Inequality by a Generalized Einstein-Podolsky-Rosen State Using Homodyne Detection". *Phys. Rev. Lett.* 85, 1349 (2000) (cit. on p. 34)

[70] K. J. Resch, J. S. Lundeen, and A. M. Steinberg: "Quantum State Preparation and Conditional Coherence". *Phys. Rev. Lett.* 88, 113601 (2002) (cit. on p. 34)

[71] G. Donati, T. J. Bartley, X.-M. Jin, M.-D. Vidrighin, A. Datta, M. Barbieri, and I. A. Walmsley: "Observing optical coherence across Fock layers with weak-field homodyne detectors". *Nat Commun* 5 (2014) (cit. on p. 34)

[72] H. Takahashi, J. S. Neergaard-Nielsen, M. Takeuchi, M. Takeoka, K. Hayasaka, A. Furusawa, and M. Sasaki: "Entanglement distillation from Gaussian input states". *Nat Photon* 4, 178–181 (2010) (cit. on p. 34)

[73] Y. Kurochkin, A. S. Prasad, and A. I. Lvovsky: "Distillation of The Two-Mode Squeezed State". *Phys. Rev. Lett.* 112, 070402 (2014) (cit. on p. 34)

[74] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier: "Generation of optical 'Schrödinger cats' from photon number states". *Nature* 448, 784–786 (2007) (cit. on p. 34)

[75] T. Gerrits, S. Glancy, T. S. Clement, B. Calkins, A. E. Lita, A. J. Miller, A. L. Migdall, S. W. Nam, R. P. Mirin, and E. Knill: "Generation of optical coherent-state superpositions by number-resolved photon subtraction from the squeezed vacuum". *Phys. Rev. A* 82, 031802 (2010) (cit. on p. 34)

[76] K. Laiho, M. Avenhaus, K. N. Cassemiro, and C. Silberhorn: "Direct probing of the Wigner function by time-multiplexed detection of photon statistics". *New Journal of Physics* 11, 043012 (2009) (cit. on pp. 34, 35, 78–80)

[77] N. Gisin and R. Thew: "Quantum communication". *Nat Photon* 1, 165–171 (2007) (cit. on p. 37)

[78] V. Giovannetti, S. Lloyd, and L. Maccone: "Quantum-Enhanced Measurements: Beating the Standard Quantum Limit". *Science* 306, 1330–1336 (2004) (cit. on p. 37)

[79] E. Knill, R. Laflamme, and G. J. Milburn: "A scheme for efficient quantum computation with linear optics". *Nature* 409, 46–52 (2001) (cit. on p. 37)

[80] S. Aaronson and A. Arkhipov: "The Computational Complexity of Linear Optics". *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*. STOC '11. New York, NY, USA: ACM, 2011, 333–342 (cit. on p. 37)

[81] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley: "Boson Sampling on a Photonic Chip". *Science* 339, 798–801 (2012) (cit. on p. 37)

[82] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther: "Experimental boson sampling". *Nat Photon* 7, 540–544 (2013) (cit. on p. 37)

[83] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White: "Photonic Boson Sampling in a Tunable Circuit". *Science* 339, 794–798 (2013) (cit. on p. 37)

[84] T. C. Ralph: "Quantum optical systems for the implementation of quantum information processing". *Rep. Prog. Phys.* 69, 853–898 (2006) (cit. on p. 37)

[85] R. Raussendorf, J. Harrington, and K. Goyal: "A fault-tolerant one-way quantum computer". *Annals of Physics* 321, 2242–2270 (2006) (cit. on p. 37)

[86] N. C. Menicucci: "Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States". *Phys. Rev. Lett.* 112, 120504 (2014) (cit. on pp. 37, 57)

[87] A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller: "Quantum State Reconstruction of the Single-Photon Fock State". *Phys. Rev. Lett.* 87, 050402 (2001) (cit. on p. 38)

[88] K. Laiho, K. N. Cassemiro, and C. Silberhorn: "Producing high fidelity single photons with optimal brightness via waveguided parametric down-conversion". *Opt. Express* 17, 22823–22837 (2009) (cit. on pp. 38, 48, 81)

[89] A. Zavatta, S. Viciani, and M. Bellini: "Quantum-to-Classical Transition with Single-Photon-Added Coherent States of Light". *Science* 306, 660–662 (2004) (cit. on p. 38)

[90] E. Bimbard, N. Jain, A. MacRae, and A. I. Lvovsky: "Quantum-optical state engineering up to the two-photon level". *Nature Photon.* 4, 243–247 (2010) (cit. on p. 38)

[91] T. B. Pittman and J. D. Franson: "Violation of Bell's Inequality with Photons from Independent Sources". *Phys. Rev. Lett.* 90, 240401 (2003) (cit. on p. 38)

[92] P. R. Tapster and J. G. Rarity: "Photon statistics of pulsed parametric light". *J. Mod. Opt.* 45, 595–604 (1998) (cit. on pp. 38, 45, 48)

[93] A. M. Brańczyk, T. C. Ralph, W. Helwig, and C. Silberhorn: "Optimized generation of heralded Fock states using parametric down-conversion". *New J. Phys.* 12, 063001 (2010) (cit. on p. 38)

[94] A. Christ, C. Lupo, M. Reichelt, T. Meier, and C. Silberhorn: "Theory of filtered type-II parametric down-conversion in the continuous-variable domain: Quantifying the impacts of filtering". *Phys. Rev. A* 90, 023823 (2014) (cit. on p. 38)

[95] W. P. Grice, A. B. U'Ren, and I. A. Walmsley: "Eliminating frequency and space-time correlations in multiphoton states". *Phys. Rev. A* 64, 063815 (2001) (cit. on p. 38)

[96] A. Eckstein: "Mastering quantum light pulses with nonlinear waveguide interactions". PhD thesis. Naturwissenschaftliche Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg, 2012 (cit. on pp. 38, 40)

[97] I. Shoji, T. Kondo, A. Kitamoto, M. Shirane, and R. Ito: "Absolute scale of second-order nonlinear-optical coefficients". *Journal of the Optical Society of America B* 14, 2268 (1997) (cit. on pp. 39, 112)

[98] H. Vanherzeele and J. D. Bierlein: "Magnitude of the nonlinear-optical coefficients of KTiOPO_4". *Optics Letters* 17, 982 (1992) (cit. on pp. 39, 112)

[99] 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". *Phys. Rev. Lett.* 100, 133601 (2008) (cit. on pp. 39, 49)

[100] A. Eckstein, A. Christ, P. J. Mosley, and C. Silberhorn: "Highly Efficient Single-Pass Source of Pulsed Single-Mode Twin Beams of Light". *Phys. Rev. Lett.* 106, 013603 (2011) (cit. on pp. 39, 48)

[101] T. Gerrits, M. J. Stevens, B. Baek, B. Calkins, A. Lita, S. Glancy, E. Knill, S. W. Nam, R. P. Mirin, R. H. Hadfield, R. S. Bennink, W. P. Grice, S. Dorenbos, T. Zijlstra, T. Klapwijk, and V. Zwiller: "Generation of degenerate, factorizable, pulsed squeezed light at telecom wavelengths". *Opt. Express* 19, 24434–24447 (2011) (cit. on p. 39)

[102] M. Yabuno, R. Shimizu, Y. Mitsumori, H. Kosaka, and K. Edamatsu: "Four-photon quantum interferometry at a telecom wavelength". *Phys. Rev. A* 86, 010302 (2012) (cit. on p. 39)

[103] W. P. Grice, R. S. Bennink, P. G. Evans, T. S. Humble, and J. C. Schaake: "Auxiliary entanglement in photon pairs for multi-photon entanglement". *Journal of Modern Optics* 59, 1538–1545 (2012) (cit. on p. 39)

[104] R.-B. Jin, K. Wakui, R. Shimizu, H. Benichi, S. Miki, T. Yamashita, H. Terai, Z. Wang, M. Fujiwara, and M. Sasaki: "Nonclassical interference between independent intrinsically pure single photons at telecommunication wavelength". *Phys. Rev. A* 87, 063801 (2013) (cit. on p. 39)

[105] R.-B. Jin, R. Shimizu, K. Wakui, H. Benichi, and M. Sasaki: "Widely tunable single photon source with high purity at telecom wavelength". *Optics Express* 21, 10659 (2013) (cit. on p. 39)

[106] P. T. Callahan, K. Safak, P. Battle, T. D. Roberts, and F. X. Kärtner: "Fiber-coupled balanced optical cross-correlator using PPKTP waveguides". *Optics Express* 22, 9749 (2014) (cit. on p. 39)

[107] B. Brecht and C. Silberhorn: "Characterizing entanglement in pulsed parametric down-conversion using chronocyclic Wigner functions". *Phys. Rev. A* 87, 053810 (2013) (cit. on p. 40)

[108] B. Brecht: "Engineering ultrafast quantum frequency conversion". PhD thesis. University of Paderborn, 2014 (cit. on p. 40)

[109] M. L. Sundheimer, C. Bosshard, E. W. Van Stryland, G. I. Stegeman, and J. D. Bierlein: "Large nonlinear phase modulation in quasi-phase-matched KTP waveguides as a result of cascaded second-order processes". *Optics Letters* 18, 1397 (1993) (cit. on p. 40)

[110] A. Eckstein, A. Christ, P. J. Mosley, and C. Silberhorn: "Realistic g(2) measurement of a PDC source with single photon detectors in the presence of background". *Phys. Status Solidi* 8, 1216–1219 (2011) (cit. on pp. 42, 48)

[111] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield: "Superconducting nanowire single-photon detectors: physics and applications". *arXiv:1204.5560* (2012) (cit. on p. 43)

[112] F. Marsili, V. B. Verma, J. A. Stern, S. Harrington, A. E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. D. Shaw, R. P. Mirin, and S. W. Nam: "Detecting single infrared photons with 93% system efficiency". *Nat Photon* 7, 210–214 (2013) (cit. on p. 43)

[113] A. E. Lita, A. J. Miller, and S. W. Nam: "Counting near-infrared single-photons with 95% efficiency". *Opt. Express* 16, 3032–3040 (2008) (cit. on p. 43)

[114] T. Gerrits, B. Calkins, N. Tomlin, A. E. Lita, A. Migdall, R. Mirin, and S. W. Nam: "Extending single-photon optimized superconducting transition edge sensors beyond the single-photon counting regime". *Opt. Express* 20, 23798–23810 (2012) (cit. on p. 43)

[115] Z. H. Levine, T. Gerrits, A. L. Migdall, D. V. Samarov, B. Calkins, A. E. Lita, and S. W. Nam: "Algorithm for finding clusters with a known distribution and its application to photon-number resolution using a superconducting transition-edge sensor". *J. Opt. Soc. Am. B* 29, 2066–2073 (2012) (cit. on p. 44)

[116] D. N. Klyshko: "Utilization of vacuum fluctuations as an optical brightness standard". *Sov. J. Quantum Electron.* 7, 591–595 (1977) (cit. on pp. 46, 114)

[117] D. N. Klyshko: "Use of two-photon light for absolute calibration of photoelectric detectors". *Sov. J. Quantum Electron.* 10, 1112 (1980) (cit. on p. 46)

[118] J. G. Rarity, K. D. Ridley, and P. R. Tapster: "Absolute measurement of detector quantum efficiency using parametric downconversion". *Appl. Opt.* 26, 4616–4619 (1987) (cit. on pp. 46, 114)

[119] D. Achilles, C. Silberhorn, and I. A. Walmsley: "Direct, Loss-Tolerant Characterization of Nonclassical Photon Statistics". *Phys. Rev. Lett.* 97, 043602 (2006) (cit. on pp. 46, 114)

[120] M. Avenhaus, A. Eckstein, P. J. Mosley, and C. Silberhorn: "Fiber-assisted single-photon spectrograph". *Opt. Lett.* 34, 2873 (2009) (cit. on p. 46)

[121] X. Li, L. Yang, L. Cui, Z. Y. Ou, and D. Yu: "Fiber-based source of photon pairs at telecom band with high temporal coherence and brightness for quantum information processing". *Opt. Lett.* 33, 593 (2008) (cit. on p. 48)

[122] H. Takesue and B. Miquel: "Entanglement swapping using telecom-band photons generated in fibers". *Opt. Express* 17, 10748–10756 (2009) (cit. on p. 48)

[123] R.-B. Jin, J. Zhang, R. Shimizu, N. Matsuda, Y. Mitsumori, H. Kosaka, and K. Edamatsu: "High-visibility nonclassical interference between intrinsically pure heralded single photons and photons from a weak coherent field". *Phys. Rev. A* 83, 031805 (2011) (cit. on p. 49)

[124] N. Quesada and J. E. Sipe: "Limits in high efficiency quantum frequency conversion". *arXiv:1508.03361 [quant-ph]* (2015) (cit. on p. 52)

[125] P. K. Tien, R. Ulrich, and R. J. Martin: "OPTICAL SECOND HARMONIC GENERATION IN FORM OF COHERENT CERENKOV RADIATION FROM A THIN-FILM WAVEGUIDE". *Applied Physics Letters* 17, 447–450 (1970) (cit. on p. 55)

[126] Y. Zhang, Z. Qi, W. Wang, and S. N. Zhu: "Quasi-phase-matched Čerenkov second-harmonic generation in a hexagonally poled LiTaO3 waveguide". *Applied Physics Letters* 89, 171113 (2006) (cit. on p. 55)

[127] A. P. Worsley, H. B. Coldenstrodt-Ronge, J. S. Lundeen, P. J. Mosley, B. J. Smith, G. Puentes, N. Thomas-Peter, and I. A. Walmsley: "Absolute efficiency estimation of photon-number-resolving detectors using twin beams". *Opt. Express* 17, 4397–4411 (2009) (cit. on p. 55)

[128] A. Feito, J. S. Lundeen, H. Coldenstrodt-Ronge, J. Eisert, M. B. Plenio, and I. A. Walmsley: "Measuring measurement: theory and practice". *New J. Phys.* 11, 093038 (2009) (cit. on pp. 55, 60, 61, 118)

[129] Y. Eto, A. Koshio, A. Ohshiro, J. Sakurai, K. Horie, T. Hirano, and M. Sasaki: "Efficient homodyne measurement of picosecond squeezed pulses with pulse shaping technique". *Optics Letters* 36, 4653 (2011) (cit. on p. 56)

[130] R. C. Alferness and L. L. Buhl: "Low-cross-talk waveguide polarization multiplexer/demultiplexer for lam = 1.32 um". *Opt. Lett.* 9, 140–142 (1984) (cit. on p. 57)

[131] B. Calkins, P. L. Mennea, A. E. Lita, B. J. Metcalf, W. S. Kolthammer, A. Lamas-Linares, J. B. Spring, P. C. Humphreys, R. P. Mirin, J. C. Gates, P. G. R. Smith, I. A. Walmsley, T. Gerrits, and S. W. Nam: "High quantum-efficiency photon-number-resolving detector for photonic on-chip information processing". *Optics Express* 21, 22657 (2013) (cit. on p. 57)

[132] R. Chrapkiewicz and W. Wasilewski: "Multimode spontaneous parametric down-conversion in a lossy medium". *Journal of Modern Optics* 57, 345–355 (2010) (cit. on p. 57)

[133] A. I. Lvovsky and M. G. Raymer: "Continuous-variable optical quantum-state tomography". *Rev. Mod. Phys.* 81, 299–332 (2009) (cit. on p. 59)

[134] R. Schmied: "Quantum State Tomography of a Single Qubit: Comparison of Methods". *arXiv:1407.4759 [quant-ph]* (2014) (cit. on p. 60)

[135] Z. Hradil, J. Řeháček, J. Fiurášek, and M. Ježek: "3 Maximum-Likelihood Methodsin Quantum Mechanics". *Quantum State Estimation*. Ed. by M. Paris and J. Řeháček. Lecture Notes in Physics 649. Springer Berlin Heidelberg, 2004, 59–112 (cit. on p. 60)

[136] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky: "Diluted maximum-likelihood algorithm for quantum tomography". *Phys. Rev. A* 75, 042108 (2007) (cit. on p. 60)

[137] J. Řeháček, Z. Hradil, O. Haderka, J. Peřina, and M. Hamar: "Multiple-photon resolving fiber-loop detector". *Phys. Rev. A* 67, 061801 (2003) (cit. on pp. 60, 61)

[138] S. P. Boyd and L. Vandenberghe: *Convex Optimization*. Cambridge University Press, 2004 (cit. on p. 60)

[139] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, and I. A. Walmsley: "Fiber-assisted detection with photon number resolution". *Opt. Lett.* 28, 2387–2389 (2003) (cit. on pp. 61, 62)

[140] M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson: "Photon-number resolution using time-multiplexed single-photon detectors". *Phys. Rev. A* 68, 043814 (2003) (cit. on pp. 61, 62)

[141] 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, 1499–1515 (2004) (cit. on p. 62)

[142] M. Avenhaus: "Advances in the Spectral and Statistical Characterization of Quantum States of Light". PhD thesis. 2014 (cit. on p. 62)

[143] J. Řeháček, D. Mogilevtsev, and Z. Hradil: "Operational Tomography: Fitting of Data Patterns". *Phys. Rev. Lett.* 105, 010402 (2010) (cit. on p. 64)

[144] D. Mogilevtsev, A. Ignatenko, A. Maloshtan, B. Stoklasa, J. Rehacek, and Z. Hradil: "Data pattern tomography: reconstruction with an unknown apparatus". *New J. Phys.* 15, 025038 (2013) (cit. on pp. 64, 82)

[145] M. Cooper, M. Karpiński, and B. J. Smith: "Local mapping of detector response for reliable quantum state estimation". *Nat Commun* 5 (2014) (cit. on p. 64)

[146] A. C. Davison and D. V. Hinkley: *Bootstrap Methods and Their Application*. Cambridge University Press, 1997 (cit. on p. 69)

[147] D. Mogilevtsev, Z. Hradil, J. Rehacek, and V. S. Shchesnovich: "Cross-Validated Tomography". *Phys. Rev. Lett.* 111, 120403 (2013) (cit. on p. 69)

[148] J. Peřina: *Quantum Statistics of Linear and Nonlinear Optical Phenomena*. Dordrecht: Springer Netherlands, 1991 (cit. on pp. 77, 79)

[149] E. B. Rockower: "Calculating the quantum characteristic function and the photon-number generating function in quantum optics". *Phys. Rev. A* 37, 4309–4318 (1988) (cit. on p. 77)

[150] D. Mogilevtsev: "Diagonal element inference by direct detection". *Optics Communications* 156, 307–310 (1998) (cit. on p. 78)

[151] J. Wenger, J. Fiurášek, R. Tualle-Brouri, N. J. Cerf, and P. Grangier: "Pulsed squeezed vacuum measurements without homodyning". *Phys. Rev. A* 70, 053812 (2004) (cit. on p. 87)

[152] W. Vogel: "Nonclassical Correlation Properties of Radiation Fields". *Phys. Rev. Lett.* 100, 013605 (2008) (cit. on p. 92)

[153] A. Miranowicz, M. Bartkowiak, X. Wang, Y.-x. Liu, and F. Nori: "Testing nonclassicality in multimode fields: A unified derivation of classical inequalities". *Phys. Rev. A* 82, 013824 (2010) (cit. on p. 92)

[154] J. Sperling, W. Vogel, and G. S. Agarwal: "True photocounting statistics of multiple on-off detectors". *Phys. Rev. A* 85, 023820 (2012) (cit. on pp. 95, 96)

[155] J. Sperling, W. Vogel, and G. S. Agarwal: "Correlation measurements with on-off detectors". *Phys. Rev. A* 88, 043821 (2013) (cit. on pp. 96, 97)

[156] T. Nitsche, F. Elster, J. Novotný, A. Gábris, I. Jex, S. Barkhofen, and C. Silberhorn: "Quantum Walks with Dynamical Control: Graph Engineering, Initial State Preparation and State Transfer". *arXiv:1601.08204 [quant-ph]* (2016) (cit. on p. 102)

[157] F. Elster, S. Barkhofen, T. Nitsche, J. Novotný, A. Gábris, I. Jex, and C. Silberhorn: "Quantum walk coherences on a dynamical percolation graph". *Scientific Reports* 5, 13495 (2015) (cit. on p. 102)

[158] B. Brecht, A. Eckstein, R. Ricken, V. Quiring, H. Suche, L. Sansoni, and C. Silberhorn: "Demonstration of coherent time-frequency Schmidt mode selection using dispersion-engineered frequency conversion". *Phys. Rev. A* 90, 030302 (2014) (cit. on p. 103)

[159] B. Brecht, D. V. Reddy, C. Silberhorn, and M. G. Raymer: "Photon temporal modes: a complete framework for quantum information science". *arXiv:1504.06251 [quant-ph]* (2015) (cit. on p. 103)

[160] J. Sperling, W. Vogel, and G. S. Agarwal: "Balanced homodyne detection with on-off detector systems: Observable nonclassicality criteria". *EPL* 109, 34001 (2015) (cit. on p. 103)

[161] A. Luis, J. Sperling, and W. Vogel: "Nonclassicality Phase-Space Functions: More Insight with Fewer Detectors". *Phys. Rev. Lett.* 114, 103602 (2015) (cit. on p. 103)

[162] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley: "Tomography of quantum detectors". *Nat Phys* 5, 27–30 (2009) (cit. on p. 118)

[163] L. Zhang, H. B. Coldenstrodt-Ronge, A. Datta, G. Puentes, J. S. Lundeen, X.-M. Jin, B. J. Smith, M. B. Plenio, and I. A. Walmsley: "Mapping coherence in measurement via full quantum tomography of a hybrid optical detector". *Nature Photonics* 6, 364–368 (2012) (cit. on p. 118)