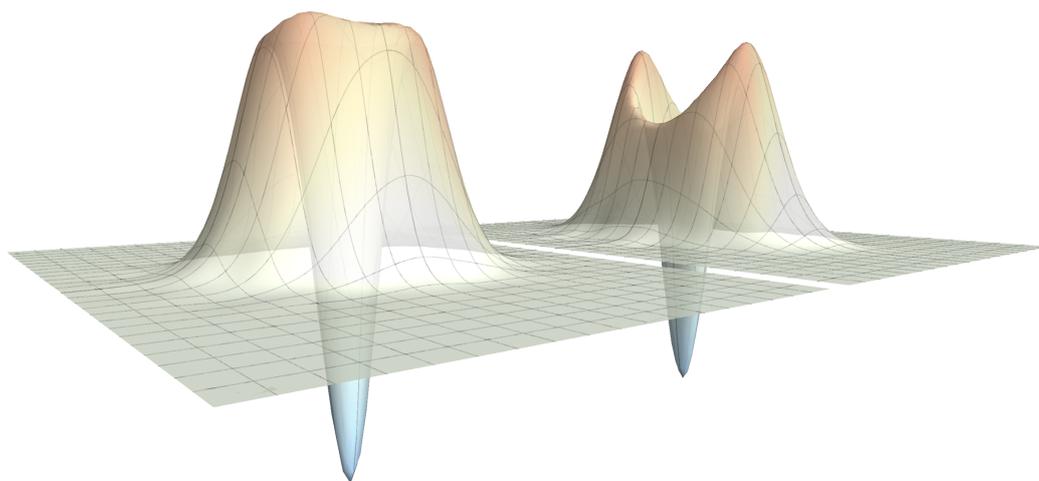




Generation of single photons and Schrödinger kitten states of light



PhD thesis
Jonas Schou Neergaard-Nielsen

Danish National Research Foundation
Center for Quantum Optics – Quantop
Niels Bohr Institute

Academic supervisor: Eugene S. Polzik

July, 2008

Generation of single photons and Schrödinger kitten states of light

Jonas Schou Neergaard-Nielsen

This thesis is submitted in partial fulfilment of the requirements
for the Ph.D. degree at University of Copenhagen.

Copyright © 2008 Jonas Schou Neergaard-Nielsen

Acknowledgments

The work presented in this thesis was carried out at QUANTOP – the Danish National Research Foundation Center for Quantum Optics – at the Niels Bohr Institute, University of Copenhagen. Under the supervision of group leader Professor Eugene S. Polzik, I started the PhD project in May 2005, following the conclusion of my master project in the same lab. In total, I have spent almost 5 years in the Quantop lab – 5 years that have been very rewarding at both a professional and a personal level. Despite the various disappointments, doubts, and annoyances that invariably come with such a project (they have been few!), I truly feel that I have been lucky to work at this lab, which – in my opinion – must be close to the ideal place to do a doctor course. The front-line research carried out in the group, the inspiring environment with lots of visitors and international meetings, the highly international constellation of the lab members, where more than 15 nationalities have been represented (and the German population at times has outnumbered the native Danish!), and, not the least, the extremely friendly atmosphere that has been maintained in spite of the growth from 7-8 members when I started to the current ca. 25 members, including the theory branches – all these things add up to the feeling of a fairly unique research group, that I am happy to have been part of for so long.

My greatest thanks therefore go to Eugene for providing the foundation for this environment, and for giving me the opportunity to do research here, as well as for his continued inspiration. As for all my great colleagues – present and past, experimentalists and theoreticians – it is impossible for me to thank any individual over another. Naturally, some persons have meant more to me than others, but everyone has contributed in his or her own way to the special Quantop atmosphere, and I have enjoyed the company of all. I do not believe that I will find the same level of familiarity at any other workplace – nor the same amount of cake at coffee breaks!

I should, however, mention those that have been directly involved in my experiment. I am particularly grateful to my main co-experimentalists, Bo Melholt Nielsen and Christian Hettich. Bo has been working with me throughout the time of my PhD, first as a master student, later as a PhD student himself. The experiment has been our common project, and all of the experimental work described in this thesis is as much due to his effort as to mine. Christian joined the project as a post doc during my time as master student and the first few months of my PhD. He was really a driving force in the initial stages of the experiment, and he taught me a lot about working in a quantum optics labora-

tory. It has been great to work with these two guys, but also the other people who have contributed to our experiment for shorter periods of time: Arnt Inge Vistnes, Hiroki Takahashi, Olafur Sigurðsson, and Ana Predojevic. I have also had interesting theoretical discussions with Klaus Mølmer, Anne E. B. Nielsen, and Thomas Kiesel. In 2006, I visited the lab of Professor Akira Furusawa at Tokyo University for 3 months. It was an eye-opening experience, and I would like to thank Akira, secretary Yumiko Yoshikawa, and all the students and researchers at the lab for their kindness and welcoming attitude towards me. I also acknowledge the assistance of the technical and administrative staff at NBI, especially that of group secretary Charlotte Hviid.

Although life at Quantop at times is like being in one big family, I would not have missed all the great times with my friends from “real life”. I will extend my warmest thanks to all of them, but perhaps in particular to those of my former fellow students with whom I have had valuable discussions about life in physics and life in general. Finally, I send my deepest love and appreciations to my mother, father, and sister for always being there for me, and to Hiroko for waiting for me across the sea.

List of publications

The main results of this thesis have been published in the following journal papers:

- J. S. Neergaard-Nielsen, B. Melholt Nielsen, C. Hettich, K. Mølmer, and E. S. Polzik, "Generation of a superposition of odd photon number states for quantum information networks," *Physical Review Letters* **97**, 083604 (2006)
- J. S. Neergaard-Nielsen, B. Melholt Nielsen, H. Takahashi, A. I. Vistnes, and E. S. Polzik, "High purity bright single photon source," *Optics Express* **15**, 7940 (2007)

Results achieved at Tokyo University (briefly mentioned in Chapter 6) were published in

- N. Takei, N. Lee, D. Moriyama, J. S. Neergaard-Nielsen, and A. Furusawa, "Time-gated Einstein-Podolsky-Rosen correlation," *Physical Review A* **74**, 060101(R) (2006)

I have presented the results at the following conferences and schools:

- QCMC 2006, International Conference on Quantum Communication, Measurement and Computing, Tsukuba, Japan, November 2006
- QUROPE winter PhD school, Obergurgl, Austria, February 2007 (poster)
- New Frontiers in Optical Technologies, summer PhD school, Tampere, Finland, August 2007 (poster)
- QIPC 2007, International Conference on Quantum Information Processing and Communication, Barcelona, Spain, October 2007 (poster)

The results have also been included by the thesis supervisor in the following invited and keynote talks:

- Quantum and Nonlinear Optics, PhD school, Danish Technical University, August 2006
- Quantum Computation, Communication and Complexity, Lectures at the Institut Henri Poincare, Paris, January 2006
- Few Qubit Applications, summer PhD school, Wales, August 2006
- International Committee on Optics Topical Meeting on Optoinformatics/ Information Photonics, St. Petersburg, September 2006
- Quantum Optics III, Pucon, Chile, November 2006
- Gordon Conference on Quantum Information Science, Lucca, Italy, April 2007
- QELS 2007, Quantum Electronics and Laser Science conference, Baltimore, May 2007
- IQEC 2007, International Quantum Electronics Conference, Munich, May 2007
- ICOLS 2007, International Conference on Laser Spectroscopy, Colorado, June 2007
- Photons, Atoms and Quanta, Royal Society, London, September 2007
- Quantum Information Processing and Communication Conference, Barcelona, October 2007
- Quantum Optics, Quantum Nanophysics & Quantum Information, European Science Foundation Workshop, Vienna, October 2007
- Quantum-Atom Optics Downunder, Wollongong, Australia, December 2007
- Lorentz Workshop: Applied Quantum Measurement (AQM07), Leiden, November 2007
- Quantum Optics: from photons and atoms to molecules and solid state systems, Obergurgl, February 2007
- Nonlinear and Quantum Optics, PhD school, Sao Paolo, February 2008
- Solvay workshop on Bits, Quanta, and Complex systems: modern approaches to photonic information processing, Brussels, May 2008

Contents

Acknowledgments	iii
List of publications	v
1 Introduction	1
2 Basics	5
2.1 Quantum state representations	5
2.1.1 The density matrix	5
2.1.2 Quadrature operators of light fields	7
2.1.3 The Wigner function	8
2.2 Quantum states of light	12
2.2.1 Gaussian states	12
2.2.2 Non-Gaussian states	18
2.3 Detecting light	22
2.3.1 Intensity detection	22
2.3.2 Homodyne detection	22
2.3.3 Losses, inefficient detection	24
2.4 Tomographic state reconstruction	25
2.4.1 Maximum likelihood estimation	25
2.4.2 Application to homodyne measurements	26
2.4.3 A short history	30
3 Theory of state generation	31
3.1 Single mode photon subtracted squeezed vacuum	31
3.1.1 Number-resolving detector	31
3.1.2 Realistic photon detection	36
3.2 Gaussian formalism	40
3.3 OPO theory	43
3.3.1 OPO modeling	44
3.3.2 DOPO	46
3.3.3 NDOPO	51
3.4 Photon-subtracted state generation	54
3.4.1 Temporal mode selection	55
3.4.2 Frequency filtering of trigger mode	57
3.4.3 Plugging in	59
3.4.4 Discussion	61
3.4.5 Single photon states from NDOPO output	64

4	Experimental methods	69
4.1	Overview	69
4.2	Making squeezed vacuum	73
4.2.1	Second harmonic generation	73
4.2.2	OPO	79
4.3	Trigger system	84
4.3.1	The number of phase matched OPO modes	84
4.3.2	Filter specifications	85
4.3.3	Cavity locking	86
4.4	Data acquisition	92
4.4.1	Homodyne detector	92
4.4.2	Data sampling	94
4.4.3	Extracting quadrature data	95
4.5	Phase information	99
4.5.1	Calibrated, continuously-scanned phase	100
4.5.2	Phase-locked local oscillator	102
5	Data presentation and analysis	107
5.1	Presentation of the data sets	107
5.2	State reconstruction	110
5.2.1	Introductory remarks	110
5.2.2	Investigations of the reconstruction method	111
5.2.3	Full reconstruction with simple exponential mode function	115
5.2.4	The choice of mode function	118
5.2.5	Detector compensated reconstruction	122
5.3	Additional analysis	127
5.3.1	Error estimation	127
5.3.2	Simple curve fitting to single photon data	130
5.3.3	Production rates, spectral brightness	130
6	Outlook	135
6.1	Time gated state generation	135
6.1.1	Towards storage in atomic memory	135
6.1.2	Results of the pumped pulse experiment	138
6.2	Other paths forward	143
A	Various calculations	147
A.1	Gaussian integrals	147
A.2	Fidelity between 1-PSSqV and odd cat state	148
A.3	Equivalency of two detection models in the low efficiency limit	149
A.4	DOPO temporal correlation functions after filtering through an optical cavity	149
A.5	Full expressions for the OPO state generation model	151
	Bibliography	153

Introduction

The discovery in 1956 by Hanbury Brown and Twiss of photon bunching in thermal light spurred the rise of quantum optics as a rich field of fundamental research [55]. Although the Hanbury Brown-Twiss effect could be understood entirely in terms of the classical theory of light as electromagnetic waves governed by Maxwell's equations, many experiments followed that showed behaviour of light that was only explainable with the help of quantum theory. According to quantum theory, an electromagnetic field consists of discrete quanta of energy – photons – that can be considered as bosonic particles. Depending on the way it is observed, light can show characteristics of either waves or particle. This wave-particle duality is one of many counter-intuitive notions of quantum mechanics. Another one is the concept of superposition – that a system can exist in two or more distinct states simultaneously. The superposition principle and wave-particle duality permeates all of quantum mechanics, and not least quantum optics. Schrödinger [138] voiced some concern over the implications of quantum superposition (and the measurement problem) with his famous cat paradox where the state of a cat, dead or alive, is in a superposition with that of a weakly radioactive material – they are entangled, as he coined it. The eventual spontaneous decay of a single atom of the material will trigger an intricate device that leads to instantaneous death of the creature. If unobserved and unperturbed, the cat itself will thus be in a superposition state of dead and alive. The deeper significance of the paradox remains a subject of discussion among proponents of different interpretations of quantum mechanics. Since then, though, the name “Schrödinger cat” has also come to denote a range of nonclassical states of various physical systems, with the common quality that they can be considered superpositions of macroscopically distinct states. A large part of fundamental research in quantum optics consists of generating new kinds of light displaying explicitly nonclassical (i.e. quantum) effects, such as the superposition in a Schrödinger cat state. Other very important effects that have been demonstrated include, from the particle point of view, photon anti-bunching [70] and polarization entangled photons [5, 117], and effects more closely related to the wave description such as field fluctuations below the classical level (squeezing) [143] and quadrature entanglement [118].

Within the last few decades, alongside research into fundamental concepts of quantum mechanics, quantum optics has also provided an important setting for the ever-growing field of quantum information processing, where the peculiarities of quantum mechanics is put to practical use in computation and communication protocols that might outperform classical counterparts. In turn, the promises of quantum information have provided an impetus for much of the quantum optics research of recent years, such as the landmark demonstrations of teleportation of quantum states [15, 45]. The traditional description of quantum information is in terms of qubits, which emphasizes the particle aspect of light, as photons are natural carriers of a single qubit. However, an alternative approach is becoming increasingly popular, in which continuous variables (cv) play the key role for encoding of information [20, 23]. In the case of light, the continuous variables are typically (although not limited to) the quadratures of the field, which are more strongly linked to the wave aspect of light. To a large extent, cv quantum information protocols make use of Gaussian states and operations, for example in quantum key distribution with coherent states [53] whose quadrature distributions are Gaussian. The realization that entanglement of Gaussian states cannot be distilled by Gaussian operations [37, 39, 49], however, leads to the necessity of non-Gaussian operations such as photon counting for long-distance communication [21]. Furthermore, it has been shown that universal quantum computation is possible with a certain type of Schrödinger cat states as resources – states which are highly non-Gaussian [131].

The work presented in this thesis is a primarily experimental endeavour into the generation and characterization of pure non-classical and, in particular, non-Gaussian states. As with many other projects within this field, it has one foot in the “fundamental research” camp and another in the “applications” camp – it is a truly fascinating side of modern quantum optics that these two overlap to such a great extent. The generation of a novel state of light featuring highly non-classical characteristics, namely a negative probability distribution in phase space, adds to the already large range of demonstrations of quantum features of light. This “Schrödinger kitten” state, although not a fully-fledged macroscopic superposition (a cat), still has some properties in common with its larger siblings. On the other hand, a longer-term goal of the project is to employ this kitten – and just as importantly our single photon states – in a quantum information network involving atomic ensembles as quantum memories. Although all-optical computation schemes have been devised, as mentioned above (also with discrete variables [71]), it still seems reasonable to assume that future quantum communication networks or computers will need a stationary long-term storage for quantum information. With that in mind, there is a huge research effort going into the development of efficient interactions between light and matter, in particular in the form of individual atoms or larger ensembles. From a technical point of view, the states we generate are very well suited to interactions with atoms due to their narrow spectral widths relative to comparable sources of high-purity non-Gaussian states.

The experimental techniques we have developed could find other applications in quantum state engineering and control, for example for entanglement distillation, but they are also fundamentally interesting in the context of wave-particle

duality. The basic concept is conditional state preparation, where a photon, detected in one mode of a correlated two-mode state, heralds the generation (or existence) of a certain state in the second mode. This is an idea familiar from many single photon sources. What is special in our case is that the initial two-mode state is a continuous wave, as opposed to the more common short pulses. The discrete detection of a single photon at a random time from this continuous wave induces the preparation of the conditional state within a temporally localized mode. The state characterization is done by homodyne tomography, which measures the continuous quadrature variables of the field (wave), but in this case conditioned on the detection of a photon (particle)! Homodyne measurements on continuous wave fields are usually analyzed in the frequency domain, but the conditioning means that we have to go to the time domain, which leads us to consider the precise temporal “wave packet” extent of the heralded single photon and kitten states. Previously, conditional state preparation with continuous waves was performed for Gaussian bright squeezed states [73] and for distillation of non-Gaussian states [57]; however, these were not conditioned on photon detections but rather threshold measurements of intensity or amplitude. The experiments closest to ours are probably those of Foster et al. [41] where homodyne detection of a non-classical continuous field (from a cavity-QED system) was conditioned on the detection of a photon in part of the field. They observed non-classical correlations in the recorded photocurrent, but no quantum state was extracted from the data.

Although the relationship between the wave and the particle description of our experiment is fascinating, the more philosophical considerations will not play a major role in this thesis. The main emphasis will be on a precise theoretical modeling of the generated states in Chapter 3, based on the theory of squeezed states from an optical parametric oscillator and a Gaussian mode reduction formalism; on the description and characterization of the experimental tools and methods in Chapter 4; and on a detailed analysis of the obtained data, in particular the full tomographic state reconstruction and investigations into the temporal state selection in Chapter 5. An introduction to the basics of quantum optical states and homodyne tomography is given in Chapter 2, while Chapter 6 describes some extensions of the experiment already carried out and some possible directions for future investigations.

Basics

2.1 Quantum state representations

2.1.1 The density matrix

A pure state of a quantum system is represented by its state vector or ket, $|\psi\rangle$, which can be expanded into basis kets of the given Hilbert space: $|\psi\rangle = \sum_k c_k |\phi_k\rangle$. The expectation value of any observable of the system represented by the operator \hat{A} acting on the state is $\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$. If the Hamiltonian, \hat{H} , of the system is time-independent, the state ket evolves according to $|\psi(t)\rangle = \hat{U}(t) |\psi(0)\rangle$, with the unitary operator $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$. A system consisting of several modes A, B , etc. which are in pure states $|\psi_A\rangle_A, |\psi_B\rangle_B, \dots$, is given by the tensor product of its constituent states, $|\psi_{AB\dots}\rangle = |\psi_A\rangle_A \otimes |\psi_B\rangle_B \otimes \dots$, often abbreviated $|\psi_A\rangle |\psi_B\rangle \dots$ or $|\psi_A, \psi_B, \dots\rangle$. However, a multimode system can also be in a state which cannot be separated into a product of the constituent states – then it is entangled.

It is possible that a system is in a statistical mixture (or ensemble) of two or more pure states, meaning that each state ket $|\psi_i\rangle$ appears with a probability w_i . Such a mixed state cannot be described as a single state ket but requires the density operator formulation,

$$\hat{\rho} = \sum_i w_i |\psi_i\rangle \langle \psi_i|.$$

The fractional populations w_i add up to 1. For a pure state $|\psi\rangle$, $\hat{\rho} = |\psi\rangle \langle \psi|$. In a basis $\{|\phi_k\rangle\}$, the density operator has a matrix representation

$$\hat{\rho} = \sum_{m,n} \rho_{mn} |\phi_m\rangle \langle \phi_n|,$$

$$\rho_{mn} = \langle \phi_m | \hat{\rho} | \phi_n \rangle = \sum_i w_i \langle \phi_m | \psi_i \rangle \langle \psi_i | \phi_n \rangle.$$

The density matrix is Hermitian and normalized to trace 1. The diagonal elements are $\rho_{nn} = \sum_i w_i |\langle \phi_n | \psi_i \rangle|^2$ and are thus all non-negative. It will be infinite for an infinitely dimensional Hilbert space, but otherwise it will in general consist of $2M$ real parameters for an M -dimensional Hilbert space, but only $2M - 1$

are independent due to the normalization constraint. The expectation value of any observable of the system can be obtained from the density matrix as

$$\langle \hat{A} \rangle = \text{tr}[\hat{A}\hat{\rho}] = \sum_{n,i} w_i \langle \phi_n | \hat{A} | \psi_i \rangle \langle \psi_i | \phi_n \rangle = \sum_i w_i \langle \psi_i | \hat{A} | \psi_i \rangle.$$

In that sense, the density matrix contains the full information about the given state. For a multimode system, say $\hat{\rho}_{AB}$, the density matrix of the mode A is given by the partial trace over mode B ,

$$\hat{\rho}_A = \text{tr}_B[\hat{\rho}_{AB}].$$

This is really a definition of $\hat{\rho}_A$, but it is sensible (and the unique choice, [113, §2.4.3]) since it results in correct expectation values of any operator \hat{A} acting only on mode A . The extension of this operator to the multimode Hilbert space is $\hat{A} \otimes \hat{I}_B$, where \hat{I}_B is the identity operator on mode B , and $\text{tr}_{AB}[(\hat{A} \otimes \hat{I}_B)\hat{\rho}_{AB}] = \text{tr}[\hat{A}\hat{\rho}_A]$; the expectation values are the same whether using $\hat{\rho}_{AB}$ or $\hat{\rho}_A$. Even if the multimode system is in a pure state, the individual modes need not be – if two modes are entangled, the individual modes obtained by tracing over the other will be mixed states.

Measurements on a quantum system

An observable \hat{A} can be diagonalized with eigenvalues λ corresponding to the possible outcomes of a measurement of \hat{A} :

$$\hat{A} = \sum_{\lambda} \lambda \hat{\Pi}_{\lambda}.$$

$\hat{\Pi}_{\lambda}$ is the projector $\sum_i |\phi_{\lambda,i}\rangle \langle \phi_{\lambda,i}|$ onto the space of eigenstates $\{|\phi_{\lambda,i}\rangle\}$ of \hat{A} with eigenvalue λ (if λ is a nondegenerate eigenvalue, $\hat{\Pi}_{\lambda} = |\phi_{\lambda}\rangle \langle \phi_{\lambda}|$). After a measurement yielding the outcome λ , the state of the system $\hat{\rho}$ is transformed according to [113, §2.4.2]

$$\hat{\rho}' = \frac{\hat{\Pi}_{\lambda} \hat{\rho} \hat{\Pi}_{\lambda}}{\text{pr}(\lambda)},$$

where $\text{pr}(\lambda) = \text{tr}[\hat{\rho} \hat{\Pi}_{\lambda}]$ is the probability of getting the outcome λ . The projectors are idempotent, $\hat{\Pi}_{\lambda}^2 = \hat{\Pi}_{\lambda}$, and mutually orthogonal, $\hat{\Pi}_{\lambda} \hat{\Pi}_{\lambda'} = 0$ for $\lambda \neq \lambda'$. Thus, subsequent measurements of \hat{A} will all give the same result. Projective measurements are special cases of the more general class of measurements, POVMs (positive operator-valued measure). A POVM is a set $\{\hat{E}_m\}$ of positive operators¹ fulfilling the completeness relation $\sum_m \hat{E}_m = \hat{I}$. A positive operator can be decomposed as $\hat{E}_m = \hat{M}_m^{\dagger} \hat{M}_m$. When measuring a POVM with outcome m , the state of the system is transformed as

$$\hat{\rho}' = \frac{\hat{M}_m \hat{\rho} \hat{M}_m^{\dagger}}{\text{pr}(m)}, \quad (2.1)$$

¹A positive operator \hat{A} has $\langle \psi | \hat{A} | \psi \rangle \geq 0$ for any state ket $|\psi\rangle$. Equivalently, all eigenvalues are non-negative (a positive operator is Hermitian and therefore diagonalizable).

with the probability of obtaining outcome m

$$\text{pr}(m) = \text{tr}[\hat{M}_m \hat{\rho} \hat{M}_m^\dagger] = \text{tr}[\hat{\rho} \hat{E}_m].$$

The elements \hat{E}_m of a POVM *can* be projectors, but in general they are not idempotent, so subsequent measurements do not necessarily yield the same result.

2.1.2 Quadrature operators of light fields

The description of electromagnetic (or light) fields makes heavy use of the notion of modes. A mode refers to a single degree of freedom of the electromagnetic field such as a polarization, spatial, temporal, frequency mode, etc. For a given mode decomposition, a state of the light field is called single mode if all but one mode show vacuum statistics of the operators [114]. A state can be single mode in one decomposition or degree of freedom and multimode in another. A single mode of the quantized electromagnetic (or light) field is described as a harmonic oscillator with the Hamiltonian

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \quad (2.2)$$

where \hat{a}^\dagger and \hat{a} are the creation and annihilation operators of field excitations. The energy eigenstates are denoted $|n\rangle$, and consist of n excitations of the mode, i.e. n photons. They are termed number states or Fock states interchangeably, while specifically the ground state $|0\rangle$ is called the vacuum state. The number states are orthogonal, $\langle m|n\rangle = \delta_{mn}$, and it is common and convenient to choose them as the basis for the infinite-dimensional Hilbert space. For the remainder of this thesis, all density matrices are written in the number state basis:

$$\rho_{mn} = \langle m|\hat{\rho}|n\rangle.$$

The diagonal of the density matrix is then the photon number distribution of the state; $\text{pr}(n) = \rho_{nn}$.

The creation and annihilation operators are not Hermitian and are thus not observables. However, they can be combined in the following way to create the so-called quadrature operators, which are Hermitian²:

$$\hat{x} = \frac{\hat{a}^\dagger + \hat{a}}{\sqrt{2}} \quad , \quad \hat{p} = \frac{i(\hat{a}^\dagger - \hat{a})}{\sqrt{2}}.$$

From the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$, we get $[\hat{x}, \hat{p}] = i$. The notation x, p gives reference to the standard position and momentum operators of quantum mechanics, which obey the same algebra, but also to the position and momentum of a classical harmonic oscillator, while the name ‘quadrature’ refers to the

²Three different normalization conventions are in common use in the definition of the quadrature operators: 1, $\sqrt{2}$, or 2 in the denominator. Our choice of $\sqrt{2}$ gives a reverse transformation from \hat{x}, \hat{p} to \hat{a}^\dagger, \hat{a} with the same normalization factor: $\hat{a}^\dagger = (\hat{x} - i\hat{p})/\sqrt{2}$, $\hat{a} = (\hat{x} + i\hat{p})/\sqrt{2}$. Furthermore, it corresponds to a choice of $\hbar = 1$ when comparing with the standard position-momentum commutator, $[\hat{x}, \hat{p}] = i\hbar$.

decomposition of an electromagnetic field into two 90° -out of phase components (cos and sin). The Hamiltonian in terms of the quadrature operators is also similar to the classical harmonic oscillator:

$$\hat{H} = \frac{1}{2}\hbar\omega(\hat{x}^2 + \hat{p}^2).$$

Since the quadrature operators do not commute – they are conjugate variables – they cannot be simultaneously determined. Rather, they must obey Heisenberg's uncertainty relation,

$$\text{Var } x \text{ Var } p \geq \frac{|[\hat{x}, \hat{p}]|}{4} = \frac{1}{4}. \quad (2.3)$$

A state fulfilling the equality is called a minimum uncertainty state.

The quadrature operators can be generalized to any rotated superposition of \hat{x} and \hat{p} ,

$$\hat{q}_\theta \equiv \cos\theta \hat{x} + \sin\theta \hat{p} = \frac{\hat{a}^\dagger e^{i\theta} + \hat{a} e^{-i\theta}}{\sqrt{2}}.$$

When we discuss general quadrature observables, we will mostly refer to them as q -values, rather than x and p , and think of them as values along a real axis rotated by the phase θ from the x -axis.

The overlap between the eigenstates of the quadrature operators, $|x\rangle$, and a state $|\psi\rangle$ is the wave function $\psi(x) = \langle x|\psi\rangle$, whose absolute square is the quadrature probability distribution of the given state. For the general phase quadrature variable, \hat{q}_θ , we write the probability distribution

$$\text{pr}_\theta(q) = |\langle q_\theta|\psi\rangle|^2, \quad (2.4)$$

or, for a general, possibly mixed state,

$$\text{pr}_\theta(q) = \langle q_\theta|\hat{\rho}|q_\theta\rangle = \text{tr}[|q_\theta\rangle\langle q_\theta|\hat{\rho}]. \quad (2.5)$$

The wave functions of the harmonic oscillator eigenstates (number states) are important and useful. For the general rotated quadrature, they are given by [78, 89]

$$\langle q_\theta|n\rangle = \frac{e^{-in\theta}}{\sqrt{2^n n! \sqrt{\pi}}} H_n(q) e^{-q^2/2}, \quad (2.6)$$

where $H_n(q)$ is the n 'th Hermite polynomial. For the vacuum state,

$$|\langle q_\theta|0\rangle|^2 = \frac{1}{\sqrt{\pi}} e^{-q^2},$$

which is a Gaussian with variance $1/2$, the minimum allowed.

2.1.3 The Wigner function

The classical harmonic oscillator has well-defined position and momentum, and its state can be represented as a point in the phase space spanned by the x and p variables. This is not possible for the quantum harmonic oscillator, due to the

uncertainties or quantum noise associated with it due to (2.3). The state of a classical harmonic oscillator can of course also be uncertain to some degree because of e.g. insufficient knowledge of the system. In that case one could describe the state by a probability distribution over phase space, $P(x, p)$, giving the probability $\iint_A P(x, p) dx dp$ of observing the system's variables within any area A . One could think of doing the same for the quantum system; however, the minimum quantum uncertainty imposed by Heisenberg's relation is fundamentally different from any source of uncertainty for a classical system. It will not make sense to assign a probability of observation to an area in phase space smaller than that of a 2-D Gaussian distribution with variance $1/2$, since a measurement of x with a precision Δx would yield an uncertainty in p of at least $1/2\Delta x$.

There are very good reasons, though, for introducing to the quantum theory a phase space function with some of the same properties as an ordinary probability distribution. Indeed, the distribution of a state over phase space is not just a matter of the quantum noise, and it is quite handy to be able to describe and visualize the state as a phase space distribution. The Wigner function, named after E. P. Wigner, who introduced it [167], is in many ways similar to a classical distribution function. It can be defined in several different ways. Originally,

$$W(x, p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iyp} \langle x - \frac{y}{2} | \hat{\rho} | x + \frac{y}{2} \rangle dy. \quad (2.7)$$

The Wigner function has a one-to-one correspondence with the density matrix, hence it contains the full information about the state. It is normalized,

$$\iint_{-\infty}^{\infty} W(x, p) dx dp = 1,$$

and its marginal distributions are the correct probability distributions for x and p ,

$$\begin{aligned} \int_{-\infty}^{\infty} W(x, p) dp &= \text{pr}(x) \\ \int_{-\infty}^{\infty} W(x, p) dx &= \text{pr}(p), \end{aligned} \quad (2.8)$$

or, in general [13],

$$\int_{-\infty}^{\infty} W(q \cos \theta - p \sin \theta, q \sin \theta + p \cos \theta) dp = \text{pr}_\theta(q). \quad (2.9)$$

These properties of the Wigner function are similar to those of a classical distribution function [20]. Furthermore, the expectation value of an operator \hat{A} can be evaluated from the Wigner function by

$$\langle \hat{A} \rangle = \text{tr}[\hat{\rho} \hat{A}] = 2\pi \iint_{-\infty}^{\infty} W(x, p) W_A(x, p) dx dp, \quad (2.10)$$

where $W_A(x, p)$ is the Wigner function $\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iyp} \langle x - \frac{y}{2} | \hat{A} | x + \frac{y}{2} \rangle dy$ of the operator \hat{A} .³ When the operator \hat{A} is symmetric in \hat{x} and \hat{p} , the expectation value

³This formula for calculating expectation values is a special case of the overlap formula [78, §3.1.2], $\text{tr}[\hat{A}\hat{B}] = 2\pi \iint_{-\infty}^{\infty} W_A(x, p) W_B(x, p) dx dp$.

formula becomes particularly classical-looking, since the Wigner function of $\hat{A} = \mathcal{S}(\hat{x}^m \hat{p}^n)$ is simply $W_A(x, p) = x^m p^n$, where $\mathcal{S}(\cdot)$ means symmetrization of the operators in the argument [20]. For example, $\mathcal{S}(\hat{x}^2 \hat{p}) = \frac{1}{3}(\hat{x}^2 \hat{p} + \hat{x} \hat{p} \hat{x} + \hat{p} \hat{x}^2)$. This property is called Weyl correspondence.

The key feature that differentiates the Wigner function from a classical distribution is that it can take negative values. For example, the Wigner function of the 1-photon number state $|1\rangle$ is, using (2.6),

$$\begin{aligned} W_1(x, p) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iyp} \langle x - \frac{y}{2} | 1 \rangle \langle 1 | x + \frac{y}{2} \rangle dy \\ &= \frac{1}{\pi^{3/2}} \int_{-\infty}^{\infty} e^{iyp} (x - \frac{y}{2})(x + \frac{y}{2}) \exp \left[-\frac{1}{2}(x - \frac{y}{2})^2 - \frac{1}{2}(x + \frac{y}{2})^2 \right] dy \\ &= \frac{1}{\pi} e^{-x^2 - p^2} (2x^2 + 2p^2 - 1), \end{aligned}$$

which becomes negative within a circle of radius $1/\sqrt{2}$ around the origin, in which it is $-1/\pi$. This value is also the minimum value the Wigner function can take, while the maximum is $1/\pi$ – another property that distinguishes it from a proper probability distribution [78, §3.1.2]. Hence, it is called a quasi-probability distribution. However, it does not really limit our use of the function as a probability measure, since we cannot assign any probability to a single point anyway, due to the Heisenberg uncertainty. By integrating the Wigner function over a Gaussian with variances fulfilling (2.3), such as the vacuum state function,

$$W_0(x, p) = \frac{1}{\pi} e^{-x^2 - p^2}, \quad (2.11)$$

the result will always become non-negative [30]. The possibility of negative Wigner functions is a clearly quantum feature, and the negativity of a state is often considered as a strong signature of a non-classical state [88]. The first observation of a negative Wigner function was done by Leibfried et al. [76] for the motional state of a trapped ion.

Computational formulas using number state basis

It is often quite useful to calculate a state's density matrix in the number state representation from a given Wigner function and vice versa. From (2.10), we see that

$$\rho_{mn} = \text{tr}[\hat{\rho} |n\rangle \langle m|] = 2\pi \iint_{-\infty}^{\infty} W(x, p) W_{mn}(x, p) dx dp, \quad (2.12)$$

with

$$W_{mn}(x, p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iyp} \langle x - \frac{y}{2} | m \rangle \langle n | x + \frac{y}{2} \rangle dy.$$

From Leonhardt [78, §5.2.6], this function is

$$\begin{aligned} W_{mn}(x, p) &= \\ &\begin{cases} \frac{1}{\pi} e^{-x^2 - p^2} (-1)^m (-x + ip)^{m-n} \sqrt{2^{m-n} \frac{n!}{m!}} L_n^{m-n}(2x^2 + 2p^2) & , m \geq n \\ \frac{1}{\pi} e^{-x^2 - p^2} (-1)^m (x + ip)^{n-m} \sqrt{2^{n-m} \frac{m!}{n!}} L_m^{n-m}(2x^2 + 2p^2) & , m < n \end{cases} \end{aligned} \quad (2.13)$$

where L_k^a are generalized Laguerre polynomials. Conversely, from the definition (2.7), the Wigner function, given a density matrix ρ_{mn} , is

$$\begin{aligned} W(x, p) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iyp} \langle x - \frac{y}{2} | \left(\sum_{m,n} \rho_{mn} |m\rangle \langle n| \right) |x + \frac{y}{2}\rangle dy \\ &= \sum_{m,n} \rho_{mn} W_{mn}(x, p). \end{aligned}$$

The value at the origin has a particularly simple expression – it depends only on the diagonal matrix elements, since $W_{mn}(0, 0) = \delta_{mn} (-1)^m / \pi$:

$$W(0, 0) = \frac{1}{\pi} \sum_n (-1)^n \rho_{nn}.$$

This is related to the fact that all off-diagonal density matrix elements contain phase information about the state, but the phase space origin has no defined phase.

Characteristic function

Yet another way to represent quantum states, often used in theoretical texts, is the (Wigner-Weyl) characteristic function. It is defined as [36]

$$\chi(\beta) = \langle \hat{D}(-\beta) \rangle = \text{tr}[\hat{\rho} \hat{D}(-\beta)],$$

where $\hat{D}(\alpha)$ is the standard displacement operator, $\hat{D}(\alpha) = \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a})$, and $\beta = (u + iv) / \sqrt{2}$, such that

$$\hat{D}(-\beta) = \exp\left(-\frac{u + iv}{\sqrt{2}} \hat{a}^\dagger + \frac{u - iv}{\sqrt{2}} \hat{a}\right) = e^{i(u\hat{p} - v\hat{x})}.$$

The density matrix is conversely obtained from the characteristic function by

$$\hat{\rho} = \frac{1}{2\pi} \int \chi(\beta) \hat{D}(\beta) d^2\beta.$$

The Wigner function is related simply to the characteristic function by a 2-dimensional Fourier transform:

$$W(x, p) = \frac{1}{(2\pi)^2} \iint e^{i(vx - up)} \chi(u, v) dudv.$$

Multimode states

The characteristic and Wigner functions are readily generalized to multimode states. The vector and matrix notation from Simon [142], Eisert and Plenio [36] is introduced; for an N -mode system,

$$\begin{aligned} \hat{\xi} &= (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_N, \hat{p}_N)^T \\ \xi &= (x_1, p_1, \dots, x_N, p_N)^T \\ \zeta &= (u_1, v_1, \dots, u_N, v_N)^T, \end{aligned}$$

and the block diagonal $2N \times 2N$ matrix (the symplectic matrix)

$$\sigma = \begin{pmatrix} 0 & 1 & & & & \\ -1 & 0 & & & & \\ & & \ddots & & & \\ & & & 0 & 1 & \\ & & & -1 & 0 & \end{pmatrix}.$$

Since the field operators for different modes commute, the symplectic matrix allows to write all commutation relations of the system as

$$[\hat{\xi}_j, \hat{\xi}_k] = i\sigma_{jk}.$$

The characteristic function for an N -mode state $\hat{\rho}$ is written neatly as

$$\chi(\zeta) = \langle e^{i\zeta^T \sigma \hat{\xi}} \rangle = \text{tr}[\hat{\rho} e^{i\zeta^T \sigma \hat{\xi}}],$$

with the following Fourier relations with the Wigner function:

$$\begin{aligned} W(\xi) &= \frac{1}{(2\pi)^{2N}} \int e^{i\xi^T \sigma \zeta} \chi(\zeta) d^{2N}\zeta, \\ \chi(\zeta) &= \int e^{-i\zeta^T \sigma \xi} W(\xi) d^{2N}\xi. \end{aligned} \tag{2.14}$$

2.2 Quantum states of light

2.2.1 Gaussian states

A natural classification criteria for optical quantum states is whether they are Gaussian or not. It turns out that the states most readily available in experiments are Gaussian, whereas non-Gaussian states take some more effort to produce. Gaussian states are defined by having a Gaussian characteristic function, or, equivalently, a Gaussian Wigner function (the Fourier transform of a Gaussian function is again Gaussian). A Gaussian function is fully determined by its first and second moments. Hence it can be described using only the quadrature mean values, $\langle \hat{\xi}_j \rangle = \text{tr}[\hat{\rho} \hat{\xi}_j]$, and the covariances collected in the real, symmetric, and positive $2N \times 2N$ covariance matrix Γ with elements

$$\Gamma_{jk} = 2 \text{Re} \langle (\hat{\xi}_j - \langle \hat{\xi}_j \rangle) (\hat{\xi}_k - \langle \hat{\xi}_k \rangle) \rangle.$$

We will be almost exclusively concerned with states with zero mean values, i.e. states centred around the origin in phase space. In that case, the covariance matrix is

$$\Gamma_{jk} = 2 \text{Re} \langle \hat{\xi}_j \hat{\xi}_k \rangle = \langle \hat{\xi}_j \hat{\xi}_k \rangle + \langle \hat{\xi}_j \hat{\xi}_k \rangle^* = \langle \hat{\xi}_j \hat{\xi}_k \rangle + \langle \hat{\xi}_k \hat{\xi}_j \rangle = \langle \{\hat{\xi}_j, \hat{\xi}_k\} \rangle,$$

where $\{\hat{\xi}_j, \hat{\xi}_k\}$ is the anti-commutator. The diagonal entries are twice the variances of the quadratures, while the off-diagonal entries describe correlations or

covariances between different quadratures and modes (which are zero for uncorrelated modes)⁴. Hence, the covariance matrix of a vacuum state is the identity matrix. Heisenberg's uncertainty relation can be formulated very generally as $\Gamma + i\sigma \geq 0$ (the ≥ 0 meaning positive semidefiniteness of the matrix, i.e. all eigenvalues ≥ 0) [36].

For a covariance matrix Γ , the Wigner function is [20, 142]

$$W(\xi) = \frac{1}{\pi^N \sqrt{\det \Gamma}} e^{-\xi^T \Gamma^{-1} \xi}, \quad (2.15)$$

with the corresponding characteristic function (see Appendix A.1)

$$\chi(\zeta) = e^{-\frac{1}{4} \zeta^T \sigma^T \Gamma \sigma \zeta}. \quad (2.16)$$

A single mode Gaussian has a covariance matrix

$$\Gamma^{(1)} = \begin{pmatrix} a & c \\ c & b \end{pmatrix},$$

but through a rotation of the state in phase space (or a rotated coordinate system), it can be brought on the simple form

$$\Gamma^{(1)} = \begin{pmatrix} a' & 0 \\ 0 & b' \end{pmatrix}, \quad (2.17)$$

which then corresponds to a Wigner function with the axes of its elliptical shape along the x - and p -axes. The uncertainty relation for these parameters is $a'b' \geq 1$.

Coherent states

Within quantum theory, the coherent state, which is defined as the eigenstate of the annihilation operator, $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$, is the closest thing to a single point in phase space and hence to the classical harmonic oscillator. It has a number state expansion given by

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (2.18)$$

so the photon number distribution is Poissonian with mean and variance $|\alpha|^2$. The Wigner function is

$$W_\alpha(x, p) = \frac{1}{\pi} e^{-(x-x_\alpha)^2 - (p-p_\alpha)^2}, \quad \alpha = \frac{x_\alpha + ip_\alpha}{\sqrt{2}},$$

which is a Gaussian with minimum uncertainty (variance 1/2), equally distributed in x and p . In fact it is just the vacuum state with the identity matrix as covariance, displaced to have its center in $\alpha = (x_\alpha + ip_\alpha)/\sqrt{2}$. No matter how large the amplitude $|\alpha|$ becomes, the quadrature noise is the same, and so, for

⁴In several references, f.x. Simon [142], Braunstein and van Loock [20], the covariance matrix is defined with a factor 1/2 relative to the definition used here.

classical (large) values of $|\alpha|$, the noise becomes negligibly small and the state is the classical phase space point. The time evolution of an initial coherent state $|\alpha(0)\rangle$ is [95]

$$\begin{aligned} |\alpha(t)\rangle &= e^{-i\hat{H}t/\hbar} |\alpha(0)\rangle \\ &= e^{-i\omega t(\hat{n}+1/2)} e^{-|\alpha(0)|^2/2} \sum_{n=0}^{\infty} \frac{\alpha(0)^n}{\sqrt{n!}} |n\rangle \\ &= e^{-i\omega t/2} e^{-|\alpha(0)|^2/2} \sum_{n=0}^{\infty} \frac{(\alpha(0)e^{-i\omega t})^n}{\sqrt{n!}} |n\rangle \\ &= e^{-i\omega t/2} |\alpha(0)e^{-i\omega t}\rangle, \end{aligned}$$

using that $\hat{a}^\dagger \hat{a}$ in the Hamiltonian (2.2) is the number operator with $\hat{n} |n\rangle = n |n\rangle$. That means it just evolves into another coherent state of the same amplitude - it is circling around in the complex phase plane like a classical harmonic oscillator, keeping the same minimal noise distribution. The coherent state is indeed very similar to the classical description of light as an oscillating field; therefore, the coherent state and statistical mixtures thereof are called classical states.

The coherent states form an over-complete basis of the Hilbert space; they are non-orthogonal, with an overlap of

$$\langle \alpha | \beta \rangle = e^{-|\alpha|^2/2 - |\beta|^2/2 + \alpha^* \beta},$$

but they are complete, $\frac{1}{\pi} \int |\alpha\rangle \langle \alpha| d^2\alpha = 1$, and hence span the entire Hilbert space, so they can be used to decompose any state:

$$|\psi\rangle = \frac{1}{\pi} \int |\alpha\rangle \langle \alpha| \psi\rangle d^2\alpha.$$

More generally, a density operator can be expanded into coherent states as well:

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

The distribution function $P(\alpha)$ is called the Glauber-Sudarshan P-representation or coherent state representation. For a given density matrix, the distribution can be obtained from

$$P(\alpha) = \frac{e^{|\alpha|^2}}{\pi^2} \int \langle -\beta | \hat{\rho} | \beta \rangle e^{|\beta|^2} e^{-\beta\alpha^* + \beta^*\alpha} d^2\beta.$$

The P-representation of a coherent state $|\gamma\rangle$ is just the two-dimensional Dirac delta function, $P_\gamma(\alpha) = \delta^{(2)}(\alpha - \gamma)$. $P(\alpha)$ is real and normalized to 1 (due to $\hat{\rho}$ being Hermitian with trace 1), so if it is otherwise non-negative and not ill-behaved⁵ it acts as a probability density distribution. The state represented by such a P-function is then an ensemble of classical coherent states and hence a classical state itself. However, the P-representation does not always behave so nicely; hence it is another example of a quasi-probability distribution. If a state has a P-function which is ill-behaved or becomes negative, it means that it cannot be described from a classical point of view as a statistical ensemble of harmonic oscillators - it is a nonclassical state. This is at least the most common definition of nonclassicality [95, 139].

⁵In this context, ill-behaved means more singular than a delta function.

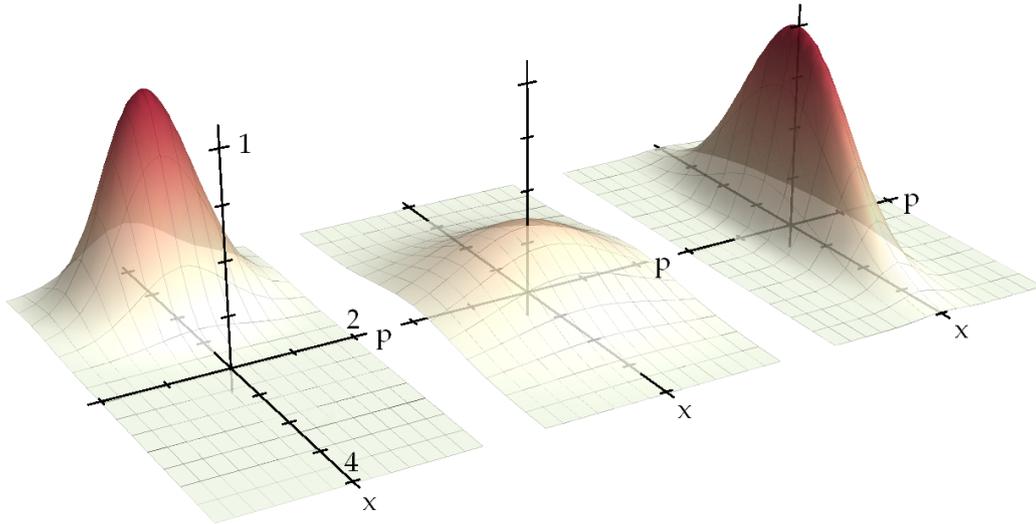


FIGURE 2.1: Wigner functions of, from left to right, a coherent state with amplitude $\alpha = -2$, a thermal state with mean photon number $n_{\text{th}} = 1$, and a squeezed vacuum state with squeezing level $r = \ln 2$ (squeezed variance $1/8$). The function values have been multiplied by π , which makes 1 and -1 the maximum and minimum values attainable. We will be doing this scaling in all plots and graphs throughout the thesis.

Thermal state

The radiation from a black body of temperature T is in a thermal state. This is a mixed state which must be formulated as a density matrix:

$$\hat{\rho}_{\text{th}} = \frac{1}{n_{\text{th}} + 1} \sum_{n=0}^{\infty} \left(\frac{n_{\text{th}}}{n_{\text{th}} + 1} \right)^n |n\rangle\langle n|,$$

where n_{th} is the mean photon number occupation, related to the temperature through $n_{\text{th}} = (\exp(\hbar\omega/k_B T) - 1)^{-1}$. Black body radiation is a quite classical phenomenon, and the state can indeed be written as an ensemble of coherent states [135]:

$$P_{\text{th}}(\alpha) = \frac{1}{\pi n_{\text{th}}} e^{-|\alpha|^2/n_{\text{th}}}.$$

It is actually a Gaussian distribution of the Gaussian coherent states. There is clearly no phase dependence, so the Wigner function has zero displacement, and there must be the same uncertainty in \hat{x} and \hat{p} . From this, and the operator identity $\hat{n} = (\hat{x}^2 + \hat{p}^2 - 1)/2$, the covariance matrix is seen to be

$$\Gamma_{\text{th}} = \begin{pmatrix} 1 + 2n_{\text{th}} & 0 \\ 0 & 1 + 2n_{\text{th}} \end{pmatrix}.$$

Squeezed states

Loosely speaking, squeezed states are coherent states which have had their noise distribution squeezed, such that one quadrature is less noisy than for the coherent state while the conjugate quadrature is correspondingly more noisy – in the

ideal case, the uncertainty product is still the minimum allowed⁶. The Wigner function is still a 2-D Gaussian, but it has become elongated along some direction. The possibility of squeezed states of light was discussed in the 1970's by Stoler [145] and Yuen [171]. The first experimental demonstration of squeezed light was carried out in 1985 by Slusher et al. [143] using four-wave mixing in a beam of Na-atoms inside an optical cavity, and soon after, with better results, by Wu et al. [169] using an optical parametric oscillator, which soon became the standard tool for producing highly squeezed light – more on that in §3.3.

The direction in phase space along which the noise is minimal (the squeezing angle) could possibly be aligned with the general angle of the state in phase space (the state's phasor), in which case the amplitude is squeezed while the phase is anti-squeezed. The opposite will be the case if the squeezing angle is perpendicular to the state's phasor. In the case of zero amplitude of the field, we talk about a squeezed vacuum state, and it makes no sense to distinguish between amplitude- or phase-squeezing. Quadrature squeezing is a wide subject area; for the sake of brevity, the discussion will be restricted to squeezed vacuum. With the squeezing direction along one of the x or p axes in phase space, the covariance matrix is as in (2.17) with either a' or b' less than the vacuum/coherent state level of 1. A pure squeezed state is a minimum uncertainty state with $a' = b'^{-1} = e^{2r}$.

Mathematically, a squeezed vacuum (with arbitrary squeezing angle) can be generated from vacuum by applying the squeezing operator,

$$\hat{S}(\zeta) = \exp\left(\frac{1}{2}\zeta(\hat{a}^\dagger)^2 - \frac{1}{2}\zeta^*\hat{a}^2\right), \quad (2.19)$$

with $\zeta = re^{i\phi}$, such that a squeezed vacuum state is $|\zeta\rangle = \hat{S}(\zeta)|0\rangle$ (the ζ here is not related to the previous ζ , describing phase space variables). The effect of the squeezing operation is most easily investigated in the Heisenberg picture [20]. The creation and annihilation operators are transformed according to

$$\begin{aligned} \hat{S}^\dagger(\zeta)\hat{a}^\dagger\hat{S}(\zeta) &= \hat{a}^\dagger \cosh r + \hat{a}e^{-i\phi} \sinh r \\ \hat{S}^\dagger(\zeta)\hat{a}\hat{S}(\zeta) &= \hat{a} \cosh r + \hat{a}^\dagger e^{i\phi} \sinh r, \end{aligned}$$

as is easily shown using the operator identity

$$e^{\hat{X}}\hat{Y}e^{-\hat{X}} = \hat{Y} + [\hat{X}, \hat{Y}] + \frac{1}{2!}[\hat{X}, [\hat{X}, \hat{Y}]] + \dots$$

From these transformations, the rotated quadrature operator transforms into

$$\hat{q}_\theta^S \equiv \hat{S}^\dagger(\zeta)\hat{q}_\theta\hat{S}(\zeta) = \hat{q}_\theta \cosh r + \hat{q}_{\phi-\theta} \sinh r.$$

Now, for the quadrature at the special angle $\phi/2$ and its conjugate at $\phi/2 + \pi/2$,

$$\begin{aligned} \hat{q}_{\phi/2}^S &= \hat{q}_{\phi/2}(\cosh r + \sinh r) \\ \hat{q}_{\phi/2+\pi/2}^S &= \hat{q}_{\phi/2+\pi/2} \cosh r + \hat{q}_{\phi/2-\pi/2} \sinh r = \hat{q}_{\phi/2+\pi/2}(\cosh r - \sinh r). \end{aligned}$$

⁶The word squeezing has been used for other effects as well, such as number squeezing. We will reserve it for quadrature squeezing.

With an initial vacuum state, the variances of these transformed rotated quadratures are thus

$$\begin{aligned}\text{Var } q_{\phi/2}^S &= \text{Var } q_{\phi/2} (\cosh r + \sinh r)^2 = \frac{1}{2} e^{2r} \\ \text{Var } q_{\phi/2+\pi/2}^S &= \text{Var } q_{\phi/2+\pi/2} (\cosh r - \sinh r)^2 = \frac{1}{2} e^{-2r},\end{aligned}$$

giving the covariance matrix stated above, after a rotation in phase space of $-\phi/2$. The corresponding Wigner function, with $\phi = 0$, is

$$W_{\text{sq}}(x, p) = \frac{1}{\pi} \exp\left(-\frac{x^2}{e^{2r}} - \frac{p^2}{e^{-2r}}\right).$$

This function fulfills all requirements for being a classical probability distribution – there are no negativities. However, the P-representation does not exist as a well-behaved function. It is formally expressed as [135, §12.4.4]

$$P_{\zeta}(\alpha) = \exp\left[\frac{e^{2r} - 1}{4} \frac{\partial^2}{\partial x^2} - \frac{1 - e^{-2r}}{4} \frac{\partial^2}{\partial p^2}\right] \delta(\alpha),$$

which is more singular than the delta function. Hence, it cannot be expressed as an ensemble of coherent states, and by the standard criterium it is a nonclassical state. An example of a squeezed vacuum Wigner function is illustrated in Figure 2.1, together with a coherent and a thermal state.

The squeezed vacuum state has the following number state expansion [95, §21.5]:

$$|\zeta\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{n!}}{(n/2)!} \left(\frac{e^{i\phi} \tanh r}{2}\right)^{n/2} \delta_{0n}^{[2]} |n\rangle. \quad (2.20)$$

Here we introduced a special notation which will make some formulas in the next chapter tidier; a modulo-2 Kronecker delta function which we can define

$$\delta_{mn}^{[2]} \equiv \frac{1 + (-1)^{m+n}}{2},$$

for integers m and n such that $\delta_{mn}^{[2]}$ is 1 if both m and n are even or both are odd, and 0 otherwise. Used as above with one index being either 0 or 1, it will act as an “odd/even” delta function for the other index. Thus, the squeezed vacuum state contains only even photon numbers.

Two-mode squeezed state

By combining two squeezed vacuum states with orthogonal squeezing phases, or by applying the two-mode squeezing operator

$$\hat{S}^{(2)}(\zeta) = \exp\left(\zeta \hat{a}_1^\dagger \hat{a}_2^\dagger - \zeta^* \hat{a}_1 \hat{a}_2\right)$$

to a two-mode vacuum $|00\rangle$, a two-mode squeezed state is created. This state is also called an EPR state, since the two modes are entangled in the quadratures (“position” and “momentum”), very much like in the original Einstein–Podolsky–Rosen-paper [34]. In a properly rotated phase space, the covariance

matrix of the state is [20]

$$\Gamma_{\text{sq}}^{(2)} = \begin{pmatrix} \cosh 2r & 0 & -\sinh 2r & 0 \\ 0 & \cosh 2r & 0 & \sinh 2r \\ -\sinh 2r & 0 & \cosh 2r & 0 \\ 0 & \sinh 2r & 0 & \cosh 2r \end{pmatrix}$$

Thus, the individual modes, obtained by tracing over the other mode, are just thermal states with mean photon number $\langle \hat{n} \rangle = (\langle \hat{x}^2 \rangle + \langle \hat{p}^2 \rangle - 1)/2 = (\cosh 2r - 1)/2 = \sinh^2 r$. The entanglement of the two modes is apparent, though, from the existence of off-diagonal elements in the matrix. While individual mode quadratures are not squeezed, combinations of the two are:

$$\begin{aligned} \text{Var}(x_1 - x_2) &= \text{Var}(p_1 + p_2) = e^{-2r} \\ \text{Var}(x_1 + x_2) &= \text{Var}(p_1 - p_2) = e^{2r}. \end{aligned}$$

This continuous variable entanglement is essential for cv quantum information protocols such as teleportation or dense coding. It was first demonstrated by Ou et al. [118] using an optical parametric oscillator, and later by Silberhorn et al. [141] by the Kerr effect in a fiber. Related examples of continuous variable entanglement are the Stokes (polarization) entanglement of Bowen et al. [17] and the entanglement between the collective spin states of two macroscopic atomic ensembles by Julsgaard et al. [65].

2.2.2 Non-Gaussian states

In a quantum optics experiment, Gaussian states are quite hard to “de-Gaussify”. Most operations that can change a state does not change its Gaussian character. Such Gaussian operations include those implemented by linear optics (e.g. beam splitters, phase shifters), squeezing and homodyne detection. In order to de-Gaussify a state, non-Gaussian operations such as photon counting/detection is needed. Therefore, fairly few non-Gaussian states have been demonstrated experimentally so far. A non-Gaussian state will per definition have a non-Gaussian Wigner function. Moreover, it is likely to have negativities. In fact, any pure non-Gaussian state will have some negativity of the Wigner function [88], but for example linear losses will drive the state towards the Gaussian vacuum state, so the negativities may disappear. For our purpose, the two most interesting classes of pure non-Gaussian states are the Fock or number states, and the Schrödinger cat states. We should, however, also mention that a range of other non-Gaussian optical states have been generated and characterized by their Wigner functions by the groups in Konstanz (displaced Fock state [90], two-mode Fock state [6]) and in Florence (single photon added coherent states [175, 177] and thermal states [173]).

Fock states

The whole concept of photons is of course deeply quantum, so a state of the light field with a well-defined number of photons (a Fock or number state) is highly

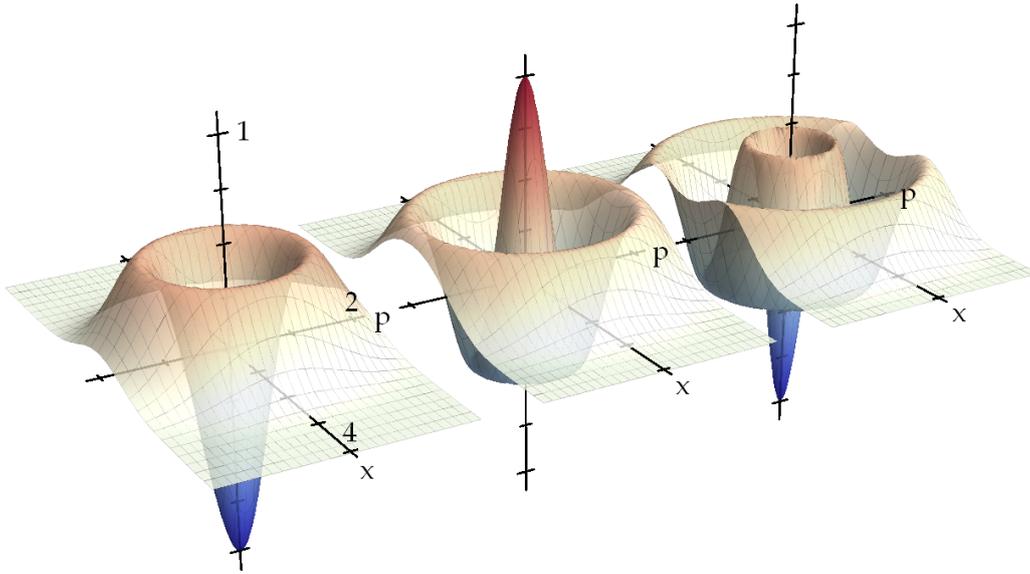


FIGURE 2.2: $\pi W(x, p)$ for the 1-, 2-, and 3-photon Fock states, from left to right. Notice their alternating negative and positive values in the origin.

nonclassical. This is clear from the fact that the Wigner functions of all Fock states except the vacuum exhibit negative values. Furthermore, their P-functions contain derivatives of the delta function, and are thus highly singular and do not correspond to classical probability distributions. The Wigner functions are, from (2.13),

$$W_n(x, p) = \frac{1}{\pi} e^{-x^2 - p^2} (-1)^n L_n^0(2x^2 + 2p^2).$$

Hence, even-number states have $W(0,0) = 1/\pi$, but for $n > 0$ they do have negativities in other regions of phase space.

The complete certainty of the photon number of a Fock state should lead to infinite uncertainty in the operator conjugate to \hat{n} . Although there does not exist a well-defined phase operator [95, §10.7], it can still be argued that phase is conjugate to photon number, so the “uncertainty relation” is satisfied since there is no phase dependence of the Fock states. This is reflected in their Wigner functions being rotationally invariant about the origin. Compare this with the coherent state, whose photon number uncertainty grows with increasing amplitude while the phase uncertainty decreases.

$n = 1$ number states, i.e. single photons, have been generated and characterized in many different ways in a host of different physical systems – for a (quite incomplete) overview, see e.g. the paper reporting our own single photon source [104]. However, the first full state characterizations showing negative Wigner function values were performed in 2001 by Lvovsky et al. [91] for a running field and in 2002 by Bertet et al. [12] for a cavity field. Recently, the same feat was accomplished by Ourjoumtsev et al. [121] for a two-photon state. The high purities of the field states that are needed in order to observe negative Wigner functions are only attained for a small fraction of the developed single photon

sources – most sources are characterized by coincidence counting, where the vacuum part of the field is eliminated.

Schrödinger cat states

The name “Schrödinger cat state” is a somewhat loose term, used in relation to many quite different quantum systems, each being more or less analogous to Schrödinger’s original cat paradox and its concept of quantum superposition of macroscopic objects [138]. Examples are the superposition of spatial states of a trapped ion [100], multi-particle entangled states [75], and superposition of flux states in a SQUID [43]. In the context of harmonic oscillators such as light fields, the “Schrödinger cat” is usually a nickname for the more precisely termed ‘coherent state superposition’ (CSS), which in its standard incarnation is defined as

$$|\mathfrak{C}\rangle = \mathcal{N}(|\alpha\rangle + e^{i\beta} |-\alpha\rangle), \quad (2.21)$$

that is, a superposition of two coherent states of same amplitude but opposite phase. \mathcal{N} is a normalization factor. The special cases where the coherence factor $e^{i\beta}$ is 1 or -1 are called even and odd cats, respectively. The argument for naming these states Schrödinger cats is that coherent states (especially ones with large amplitude) are seen as macroscopic, classical entities, and they are distinguishable by a macroscopic measurement (homodyne detection) [85].⁷ Yurke and Stoler [172] suggested a way to directly generate such a coherent state superposition using a Kerr-like Hamiltonian, but it has not been realized yet. With the help of cavity-QED interactions with Rydberg atoms, Brune et al. [22] generated a coherent state superposition, but it was trapped inside the cavity and could only be indirectly probed via the atoms.

If, instead of the superposition in (2.21), the state was just a statistical mixture (ensemble) of the two coherent states, its P-representation would be the sum of the two corresponding delta functions, but since it is a coherent superposition, it cannot be described by a well-behaved P-function. In Mandel and Wolf [95] its formal expression is derived; it involves derivatives of the delta function of infinitely high order! A statistical mixture of the two coherent states would have a density operator $\frac{1}{2} |\alpha\rangle\langle\alpha| + \frac{1}{2} |-\alpha\rangle\langle-\alpha|$ and the Wigner function would just be the sum of the two separate Gaussians. For the coherent superposition, the density matrix is $|\mathfrak{C}\rangle\langle\mathfrak{C}|$ and contains interference terms between $|\alpha\rangle$ and $|-\alpha\rangle$.

⁷A quotation from the amusing satirical review by Bergou and Englert [11]: *The original, very complicated set-up that consisted of a radioactive sample, a Geiger counter, a hammer, a cell of poisonous gas, etc. – in addition to the cat itself – gradually evolved into arrangements that are considerably simpler and experimentally more accessible. At the present state of cat evolution it has become reduced to the superposition of two slightly excited coherent states of opposite phase. The amount of theoretical work behind this reduction cannot be overestimated!*

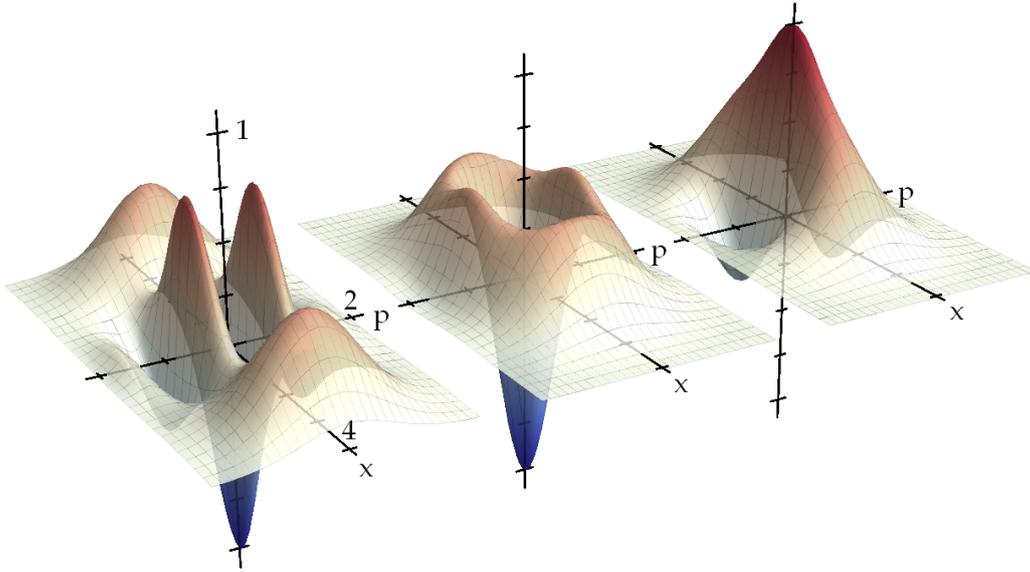


FIGURE 2.3: Examples of Wigner functions for cat states, or coherent state superposition states. From left to right, odd cat with $\alpha = 2$, odd cat with $\alpha = 1$, even cat with $\alpha = 1$. The interference fringes between the two coherent state peaks at α and $-\alpha$ quickly become more pronounced with increasing amplitude.

In the Wigner function, expressed by [48, 112]

$$W_{\pm}(x, p) = \frac{1}{2 \pm 2e^{-2\alpha^2}} \left(W_0(x - \sqrt{2}\alpha, p) + W_0(x + \sqrt{2}\alpha, p) \right. \\ \left. \pm 2W_0(x, p) \cos(2\sqrt{2}p\alpha) \right),$$

this is reflected as a wavy pattern with negativities between the two Gaussian peaks, as seen in Figure 2.3. In the above expression, plus and minus denote even and odd cat states, respectively, while $W_0(x, p)$ is the vacuum Wigner function (2.11). α is assumed real.

The photon number distribution of a cat state oscillates between odd and even photon numbers. In particular, for the odd and even cats, their number state expansions are, using (2.18),

$$|\tilde{\mathcal{C}}_{-}^{\otimes}\rangle = \mathcal{N}_{-}(|\alpha\rangle - |-\alpha\rangle) = \frac{1}{\sqrt{\sinh|a|^2}} \sum_{n=0}^{\infty} \delta_{1n}^{[2]} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \\ |\tilde{\mathcal{C}}_{+}^{\otimes}\rangle = \mathcal{N}_{+}(|\alpha\rangle + |-\alpha\rangle) = \frac{1}{\sqrt{\cosh|a|^2}} \sum_{n=0}^{\infty} \delta_{0n}^{[2]} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (2.22)$$

that is, the odd cat contains only odd photon number contributions, while there are only even photon numbers in the even cat.

2.3 Detecting light

2.3.1 Intensity detection

The essential measurement device in quantum optics is obviously the photodetector, converting an optical signal into an electronic signal. In particular photodiodes are important, as they are probably the most widely used photodetectors in experiments. Photodiodes are usually very linear; the generated current is directly proportional to the incident light power. The proportionality constant can be expressed as either the responsivity, stated in units of A/W, or as the quantum efficiency, which is the number of electrons per incident photon. At near-infrared, good silicon PIN photodiodes have quantum efficiencies in the 90% range, sometimes as high as 99%. With such high efficiencies, an optical signal can be faithfully converted into an electronic signal, processable by the subsequent detector circuitry. Very weak fields, however, can not be detected, since they become covered in the noise of the electronics. Commonly, optical powers on the order of microwatts is needed to get a decent signal-to-noise ratio. Very weak fields, at intensity levels which are more conveniently measured in photons per second than in watts, can be detected using avalanche photo diodes (APDs) that feature internal gain. When operated with a reverse bias above the diode breakdown voltage (Geiger mode), each photon absorption event results in an electron avalanche large enough to be detected. That way, it is actually possible to observe individual photons with quite good timing precision. Unfortunately, the quantum efficiencies of APDs are not quite as high as for standard photodiodes – usually 60% or below. Furthermore, they are not capable of distinguishing between events in which one or several photons are absorbed; such efficient photon number-resolving detectors are in high demand for several quantum information protocols, but are technically demanding to construct.

An essential aspect of light detection in quantum optics is that of light noise. As already mentioned, coherent states, which closely describe a laser's output, have a Poissonian photon number (intensity) distribution with a variance proportional to the intensity, $\text{Var } n = |\alpha|^2$. This intrinsic intensity noise is called shot noise and is an effect of the quantized nature of light. The linear scaling of the shot noise variance with light power is in contrast to that of classical noise or signals, which are proportional to the field amplitude and thus scale quadratically with the intensity. This makes it possible to distinguish classical and shot noise in a laser beam by observing the power scaling. In experiments it is often an essential necessity to be shot noise limited, that is, getting rid of all classical sources of noise [7, §8].

2.3.2 Homodyne detection

Since all varieties of photodetectors measure the light intensity and not the field, phase information is not directly accessible; only through interference of two or more phase-coherent fields can any knowledge of the phase or phase-dependent properties of the interfering fields be revealed. A very powerful technique for

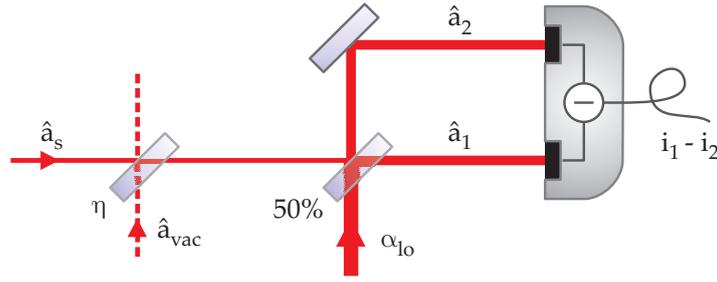


FIGURE 2.4: The basic scheme of a homodyne detector. The difference current output of the detector is proportional to the quadrature variables of the incoming field \hat{a}_s . Non-ideal detection efficiency is modeled by a fictitious beamsplitter in the signal beam.

phase-sensitive measurement is balanced homodyne detection. Employing two standard photodiode detectors and interference with an auxiliary beam (the local oscillator), it is capable of measuring the quadrature variables q_θ at any given phase of the field under inspection. Not only that; it can do so, even for fields of vanishingly small intensity, something which – as explained above – is completely infeasible with normal photodetection. Homodyne detection is a standard technique in quantum optics and is covered extensively in several textbooks [7, 78, 135], so the following will be just a brief recapitulation.

First, let us briefly recall how a beamsplitter works. Two incoming modes, \hat{a}_1 and \hat{a}_2 , are each transmitted with intensity transmittance T and reflected with reflectance $R = 1 - T$. The mode operators are transformed into the output modes \hat{a}'_1 and \hat{a}'_2 according to

$$\begin{pmatrix} \hat{a}'_1 \\ \hat{a}'_2 \end{pmatrix} = \begin{pmatrix} \sqrt{T} & \sqrt{1-T} \\ -\sqrt{1-T} & \sqrt{T} \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix},$$

or, in an equivalent operator description [20],[78, §4.1.2],

$$\begin{pmatrix} \hat{a}'_1 \\ \hat{a}'_2 \end{pmatrix} = \hat{B}^\dagger(T) \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \hat{B}(T),$$

$$\hat{B}(T) = \exp[\Theta(\hat{a}_1\hat{a}_2^\dagger - \hat{a}_1^\dagger\hat{a}_2)] \quad \text{with } \cos\Theta = \sqrt{T}. \quad (2.23)$$

In a homodyne detector, the weak signal beam to be measured is mixed on a 50-50 beamsplitter ($T = 1/2$) with a comparatively very strong local oscillator beam (LO), as illustrated in Figure 2.4. The signal beam is described by the operator \hat{a}_s while the strong LO can be treated classically as $|\alpha_{lo}|e^{i\theta}$ with an adjustable phase relative to the signal. The two modes after the beamsplitter are

$$\hat{a}_1 = \frac{1}{\sqrt{2}}(|\alpha_{lo}|e^{i\theta} + \hat{a}_s),$$

$$\hat{a}_2 = \frac{1}{\sqrt{2}}(|\alpha_{lo}|e^{i\theta} - \hat{a}_s).$$

The photocurrents produced in the detectors are proportional to the number of photons (or the photon flux) in the detected fields:

$$i_1 \propto \hat{n}_1 = \hat{a}_1^\dagger \hat{a}_1, \quad i_2 \propto \hat{n}_2 = \hat{a}_2^\dagger \hat{a}_2.$$

By electronically subtracting the two currents, a difference current i_- proportional to the quadrature operator at phase θ of the signal field is obtained:

$$i_- = i_1 - i_2 \propto \hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2 = |\alpha_{10}| (\hat{a}_s e^{-i\theta} + \hat{a}_s^\dagger e^{i\theta}) = \sqrt{2} |\alpha_{10}| \hat{q}_{s,\theta}. \quad (2.24)$$

The fluctuations of the homodyne signal are proportional to both the local oscillator amplitude and the signal field quadrature. With vacuum as the signal input, the noise is entirely due to the LO, and the variance (noise power) should scale linearly with the intensity. However, for a constant LO intensity, the noise variance can also be seen as the level of the quantum fluctuations of the vacuum test state. If the vacuum is replaced with a different quantum state, the fluctuations of this state relative to the vacuum will be reflected in the total homodyne noise level. This is exemplified in the classic squeezed vacuum noise trace, which goes above and below the vacuum level for varying LO phases. At maximum squeezing, the noise of the homodyne signal is still high, as determined by the LO intensity, but it is lower than when a vacuum state is measured.

It is quite important that the two detection arms of the homodyning be balanced. As can be seen from (2.24), the method hinges on the cancelation of the $|\alpha|^2$ terms in the difference photocurrent. Any classical noise of the LO will be correlated in the two arms and will hence cancel out – but only if the balancing is perfect. If not, the classical noise will “pollute” the quantum noise of the signal beam.

2.3.3 Losses, inefficient detection

Quantum states are fragile creatures which easily degrade through interaction with the environment. For a light mode, degradation or losses are usually described by the beamsplitter model. A field impinging upon a beamsplitter with transmittance η loses a fraction $1 - \eta$ of its intensity while getting admixed with a fraction $1 - \eta$ of the input from the other port. Even if this input is vacuum, it has an effect due to its vacuum fluctuations. The homodyne signal in (2.24) will change to

$$i_- \propto \sqrt{2} |\alpha_{10}| (\sqrt{\eta} \hat{q}_{s,\theta} + \sqrt{1 - \eta} \hat{q}_v),$$

where v denotes the admixed vacuum mode. Most other sources of loss for the field, such as absorption in a medium or undetected photons due to an inefficient detector, can be treated in the same way as the beamsplitter, since they have the same effect on the homodyne outcome [78, §4.1.4]. An imperfect overlap (mode-matching) between LO and signal field, quantified by the interference visibility V enters as a loss that must be modeled by a beamsplitter with transmittance $\eta = V^2$ [7].

Later, we will need to know the effect of a beamsplitter on an impinging field with a specified photon number n . With vacuum in the other input port, the beamsplitter operator (2.23) transforms the input state $|n, 0\rangle$ into [78, §4.1.3]

$$\hat{B}(\eta) |n, 0\rangle = \sum_{k=0}^n \sqrt{B_k^n(\eta)} |k, n - k\rangle, \quad (2.25)$$

where

$$B_k^n(\eta) = \binom{n}{k} \eta^k (1 - \eta)^{n-k}$$

is exactly the binomial distribution, expressing the probability that out of n indistinguishable photons, k are transmitted and $n - k$ are reflected. If the beam-splitter represents a loss of an incoming signal state $\hat{\rho}_s = \sum_{\mu, \nu} \rho_{\mu\nu} |\mu\rangle\langle\nu|$, the reflected output mode (ν) must be traced over after operation of $\hat{B}(\eta)$ to obtain the lossy state $\hat{\rho}_s(\eta)$:

$$\begin{aligned} \langle m | \hat{\rho}_s(\eta) | n \rangle &= \langle m | \text{tr}_\nu [\hat{B}(\eta) (\hat{\rho}_s \otimes |0\rangle_\nu \langle 0|) \hat{B}^\dagger(\eta)] | n \rangle \\ &= \sum_j \sum_{\mu, \nu} \rho_{\mu\nu} \langle m, j | \hat{B}(\eta) | \mu, 0 \rangle \langle \nu, 0 | \hat{B}^\dagger(\eta) | n, j \rangle \\ &= \sum_j \sum_{\mu, \nu} \rho_{\mu\nu} \sqrt{B_m^\mu(\eta)} \langle j | \mu - m \rangle_\nu \sqrt{B_n^\nu(\eta)} \langle \nu - n | j \rangle_\nu \quad (2.26) \\ &= \sum_j \rho_{m+j, n+j} \sqrt{B_m^{m+j}(\eta) B_n^{n+j}(\eta)}. \end{aligned}$$

This transformation is called a generalized Bernoulli transformation.

2.4 Tomographic state reconstruction

2.4.1 Maximum likelihood estimation

An essential problem in experimental science is that of model estimation based on observational data – given some observation of a system, which model can best describe the underlying structure of the system? Commonly, one knows or makes a qualified assumption of a functional relationship between the independent and dependent variables of the system; $y = f(x_1, x_2, \dots)$, where y is the variable being measured. The model will depend on one or more parameters \mathbf{s} ; $y = f(x_1, x_2, \dots; \mathbf{s})$, and these parameters should be estimated based on the observations $\{y_i\}$. One can calculate the probability $\text{pr}(\{y_i\} | \mathbf{s})$ for the observed outcome to have happened, given specified model and parameters. For parameter estimation, though, the data is given and the parameters are unknown, in which case the probability is considered a function of the parameters and is called the likelihood function:

$$L(\mathbf{s}) \equiv \text{pr}(\{y_i\} | \mathbf{s}).$$

Intuitively, with no other a priori information than the observed data, the parameters of the model that are most likely to be correct are the ones that maximize the likelihood function. This is the basis of the maximum likelihood method. The parameters \mathbf{s}_0 which maximize the likelihood are called the maximum likelihood estimator. In fact, the familiar least-squares method of parameter estimation is a special case of maximum likelihood estimation, which is relevant under the assumption that the measurement outcomes are independent and normally distributed [130, §15.1].

If the data samples $\{y_i\}$ are independent and identically distributed, the total outcome probability is just the product of the probabilities of the individual samples. Usually the logarithm of the likelihood is used, since the product turns into a sum, while the location of the maximum is unchanged:

$$\ln L(\mathbf{s}) = \ln \prod_i \text{pr}(y_i|\mathbf{s}) = \sum_i \ln \text{pr}(y_i|\mathbf{s}). \quad (2.27)$$

In standard maximum likelihood estimation, the maximization of the log-likelihood can be carried out in various ways, one of which is the iterative expectation–maximization algorithm [161]. So far, most quantum state estimations with maximum likelihood have been done with a variation of this algorithm.

2.4.2 Application to homodyne measurements

We are going to use the maximum likelihood method to estimate which quantum states of light are produced in our experiment. We can run the state generation a large number of times under the same experimental conditions, while sampling a given quadrature variable once for each state by homodyne detection. These samples will then be independent and identically distributed - a distribution determined by the density matrix of the state through (2.5). Based on the observed samples, the density matrix (and, along with that, the Wigner function) can be estimated in various ways, one of which is by maximum likelihood, in which case the matrix entries play the role as the parameters \mathbf{s} . This estimation is often referred to as quantum state reconstruction. The idea of reconstructing a 2D-distribution (Wigner function) from observations of its 1D projections (marginal quadrature distributions) at various angles is similar to that of medical tomographic imaging, which has lend this technique the name homodyne tomography.

Since the homodyne measurement gives continuous-valued outcomes that have infinitely small probability of being exactly observed, the results must be binned. For a density matrix $\hat{\rho}$ (true or estimated), the probability $\text{pr}_{\theta,j}$ of observing the quadrature variable \hat{q}_θ within the j th bin $]q_j; q_{j+1}[$ is

$$\begin{aligned} \text{pr}_{\theta,j} &= \int_{q_j}^{q_{j+1}} \text{pr}_\theta(q) dq \\ &= \int_{q_j}^{q_{j+1}} \text{tr}[|q_\theta\rangle\langle q_\theta| \hat{\rho}] dq \\ &= \text{tr}[\hat{I}_{\theta,j} \hat{\rho}], \end{aligned}$$

where

$$\hat{I}_{\theta,j} = \int_{q_j}^{q_{j+1}} |q_\theta\rangle\langle q_\theta| dq$$

is the projection operator for the j th bin. If the bins cover the entire range of q -values, the probabilities and the projectors for a specific angle add to unity/identity: $\sum_j \text{pr}_{\theta,j} = 1$, $\sum_j \hat{I}_{\theta,j} = \hat{I}$. With $N_{\theta,j}$ observations in the j th bin, the likelihood as a

function of $\hat{\rho}$ is

$$\begin{aligned}\ln L(\hat{\rho}) &= \ln \prod_{\theta,j} \text{pr}_{\theta,j}^{N_{\theta,j}} \\ &= \sum_{\theta,j} N_{\theta,j} \ln \text{pr}_{\theta,j}.\end{aligned}$$

Here we have assumed a measurement procedure where the quadrature phase angle is kept constant for a while as the corresponding quadrature variable \hat{q}_θ is being sampled a number of times; afterwards, the LO phase is adjusted to a new value of θ , and the measurement continues. The \sum_θ denotes summation over all the sampled phase angles. The phase angle plays a somewhat special role in the maximum likelihood method for homodyne measurements. It is usually not a random variable but rather has a value under the control of the experimenter⁸. It can be set to a fixed value or be scanned continuously; either way, for each set phase angle there is an entire corresponding quadrature probability distribution $\text{pr}_\theta(q)$ and set of projectors $\{\hat{\Pi}_j\}$. In the reconstruction, each phase angle distribution ought to be weighted by the number of samples obtained at that phase, $N_\theta = \sum_j N_{\theta,j}$. We will see below that this also comes out naturally from the maximization of the likelihood.

Following [10, 61, 158], we decompose $\hat{\rho}$ into $\hat{\sigma}^\dagger \hat{\sigma}$, where $\hat{\sigma}$ is a complex lower triangular matrix with real diagonal elements. This Cholesky decomposition ensures that $\hat{\rho}$ is Hermitian and positive definite. We are seeking the conditions, under which $\ln L(\hat{\rho})$ is maximal with respect to variations of $\hat{\rho}$ or, equivalently, $\hat{\sigma}$. Furthermore, the requirement of unity trace of a density matrix adds a constraint which can be incorporated using the method of Lagrange multipliers; for some constant λ , the following must hold at maximum likelihood:

$$\begin{aligned}\frac{\partial}{\partial \hat{\sigma}} (\ln L(\hat{\rho}) - \lambda(\text{tr } \hat{\rho} - 1)) &= 0 \\ \Rightarrow \sum_{\theta,j} \frac{N_{\theta,j}}{\text{pr}_{\theta,j}} \frac{\partial}{\partial \hat{\sigma}} \text{tr}[\hat{\Pi}_{\theta,j} \hat{\sigma}^\dagger \hat{\sigma}] - \lambda \frac{\partial}{\partial \hat{\sigma}} \text{tr}[\hat{\sigma}^\dagger \hat{\sigma}] &= 0.\end{aligned}$$

The matrix derivative of a trace is [124] $\frac{\partial}{\partial \mathbf{X}} \text{tr}[\mathbf{A}\mathbf{X}] = \mathbf{A}^\text{T}$. Carrying out the differentiations and transposing the equation gives

$$\begin{aligned}\sum_{\theta,j} \frac{N_{\theta,j}}{\text{pr}_{\theta,j}} \hat{\Pi}_{\theta,j} \hat{\sigma}^\dagger - \lambda \hat{\sigma}^\dagger &= 0 \\ \Rightarrow \sum_{\theta,j} \frac{N_{\theta,j}}{\text{pr}_{\theta,j}} \hat{\Pi}_{\theta,j} \hat{\rho} &= \lambda \hat{\rho},\end{aligned}\tag{2.28}$$

where we multiplied by $\hat{\sigma}$ from the right. To determine λ we take the trace on both sides of (2.28). The right side becomes λ , while on the left side, $\text{tr}[\hat{\Pi}_{\theta,j} \hat{\rho}]$ is nothing else but $\text{pr}_{\theta,j}$, leaving

$$\lambda = \sum_{\theta,j} N_{\theta,j} = N.$$

⁸Some works do make use of phase-randomized data, either intentionally so or out of necessity [91, 101, 176].

N denotes the total number of samples over all phase angles. Inserting this and a couple of N_θ 's in (2.28):

$$\sum_{\theta} \frac{N_{\theta}}{N} \sum_j \frac{N_{\theta,j}}{N_{\theta} \text{pr}_{\theta,j}} \hat{I}_{\theta,j} \hat{\rho} = \sum_{\theta} \frac{N_{\theta}}{N} \sum_j \frac{f_{\theta,j}}{\text{pr}_{\theta,j}} \hat{I}_{\theta,j} \hat{\rho} = \hat{\rho},$$

with $f_{\theta,j}$ being the frequency of observations in the j th bin within the θ phase angle measurement. Denoting the operator on the left hand sides by $\hat{R}(\hat{\rho})$,

$$\hat{R}(\hat{\rho}) \equiv \sum_{\theta} \frac{N_{\theta}}{N} \sum_j \frac{f_{\theta,j}}{\text{pr}_{\theta,j}} \hat{I}_{\theta,j}, \quad (2.29)$$

this extremal equation says that for the maximum likelihood estimator $\hat{\rho}_0$, $\hat{R}(\hat{\rho}_0)$ must be the identity operator. This can indeed also be seen from the definition of \hat{R} ; for a data set $\{f_{\theta,j}\}$, the distribution which is the most likely must have its probabilities equal to the measured frequencies in all bins. When this is the case, \hat{R} turns into a simple sum of the projection operators, which we earlier assumed to be the identity.

As outlined in several papers [38, 89, 132], the solution to the extremal equation $\hat{R}(\hat{\rho}_0)\hat{\rho}_0 = \hat{\rho}_0$, which can equivalently be written

$$\hat{R}(\hat{\rho}_0)\hat{\rho}_0\hat{R}(\hat{\rho}_0) = \hat{\rho}_0,$$

can be found by an iterative method closely related to the classical expectation–maximization algorithm. Starting from an initial guess $\hat{\rho}^{(0)}$, the iteration continues as

$$\hat{\rho}^{(k+1)} = \hat{R}(\hat{\rho}^{(k)})\hat{\rho}^{(k)}\hat{R}(\hat{\rho}^{(k)})$$

followed by normalization to trace 1 for each step. The iteration is supposed to monotonically increase the likelihood and converge towards the fixed point $\hat{\rho}_0$. Řeháček et al. [132] demonstrates that this convergence is not always guaranteed, but in our experience, with our type of experimental data, the likelihood is always increased. By studying the operator \hat{R} , we can see qualitatively why applying it to an estimated $\hat{\rho}^{(k)}$ will generate a new estimate closer to $\hat{\rho}_0$: If $\hat{\rho}^{(k)}$ over-emphasizes the (θ, j) bin compared to the observation, $f_{\theta,j}/\text{pr}_{\theta,j}$ will be less than 1, while an under-emphasis will make this fraction larger than 1. Thus, application of $\hat{R}(\hat{\rho}^{(k)})$ will change the estimate towards a better balance through the different weights on the projectors $\hat{I}_{\theta,j}$.

The density matrix reconstruction is conveniently carried out in the number state basis, where the matrix representation of the projection operator is

$$(\hat{I}_{\theta,j})_{mn} = \int_{q_j}^{q_{j+1}} \langle m|q_{\theta}\rangle \langle q_{\theta}|n\rangle dq,$$

with $\langle m|q_{\theta}\rangle$ from (2.6). Since the Hilbert space is infinite, in the reconstruction it is necessary to truncate it to dimension $M+1$ at some maximal photon number M . To some degree, this truncation limits the objectivity of the reconstruction, since a choice of M has to be made. It can be chosen large enough, though, that in practice there is no doubt that all photon numbers contributing to the state has been included.

Coming back to the issue of the treatment of the different phase angles under which the quadratures have been sampled, we see from (2.29) that the iteration operator consists of a sum of individual operators for each angle, weighted by the relative number of obtained data points at each angle. Just as expected. Another important question is how many different phase angles should be probed? Only phases between 0 and π must be probed, since the other half-circle of phases give the same distributions with a sign difference ($\hat{q}_{\theta+\pi} = -\hat{q}_\theta$). But if the quadrature distributions are only sampled for very few phase angles, say 2 or 3, one can easily imagine that features of the Wigner function in between those phases go unnoticed. The maximum likelihood reconstruction only estimates states that fit well with the observed data. Hence, a sufficient number of phase angles must be observed in order to increase the azimuthal resolution of the reconstructed Wigner function. Leonhardt [78] shows that the required number of phase angles is approximately equal to the dimension of the state, that is, the highest excited photon number. This was shown for the reconstruction methods mentioned in the next section, not for maximum likelihood, but there is no reason to assume that it is fundamentally different for the different methods. An alternative approach is to use a continuously scanned LO phase and record each sample as a (θ, q_θ) pair. As in Lvovsky [89], the binning in both θ and q can be made so small that each bin contains only 0 or 1 data point. In that case, the iteration operator (2.29) becomes

$$\hat{R}(\hat{\rho}) = \frac{1}{N} \sum_i \frac{1}{\text{pr}_i} \hat{\Pi}_i, \quad (2.30)$$

where the sum is over all data samples indexed by i .

Finally, another very useful feature of the maximum likelihood method is the possibility to explicitly include imperfect detection efficiency in the algorithm [10, 89]. As discussed in §2.3.3, a detector efficiency η can be modeled by transmission through a beamsplitter with transmittance η . To estimate the density matrix $\hat{\rho}_{\eta=1}$ of the state as it was before the deterioration due to non-unity detector efficiency, one could simply reconstruct the detected lossy $\hat{\rho}_\eta$ and subsequently perform the inverse of the Bernoulli transformation (2.26). However, statistical errors on the low photon number density matrix elements have a tendency to turn into unrealistically large or possibly negative high photon number elements. This can be avoided by instead incorporating the non-unity efficiency into the reconstruction algorithm. Thus, the projection operator $\hat{\Pi}_{\theta,j}$ above is exchanged with an element of a POVM, which, from [10, 89], is

$$\hat{E}_{\theta,j}(\eta) = \sum_{m,n,k} \sqrt{B_m^{m+k}(\eta) B_n^{n+k}(\eta)} \int_{q_j}^{q_{j+1}} dq \langle n|q_\theta\rangle \langle q_\theta|m\rangle |n+k\rangle \langle m+k|, \quad (2.31)$$

where the summation of m and n runs from 0 to the truncation number M , and k from 0 to $M - \max(m, n)$. This operator now represents a detection event in the (θ, j) bin by an η -efficiency detector, and the iteration algorithm will directly reconstruct the state $\hat{\rho}_{\eta=1}$ before the detector.

2.4.3 A short history

The application of maximum likelihood methods to quantum state reconstruction was developed around 1997 by Hradil [60], Mogilevtsev et al. [98], and Banaszek [8–10]. Before then, a couple of other methods were in use for homodyne tomography. The very first experimental demonstration of homodyne tomography was done by Smithey et al. [144] who also developed the idea based on the realization by Vogel and Risken [163] of the relationship between homodyne measurements and the Wigner function. The reconstruction method employed there consists in basically inverting the Radon transformation (2.9) to construct the Wigner function directly from observed quadrature distributions. The inverse transformation integrates the marginal quadrature distributions with a kernel function [78]

$$K(q) = \frac{1}{2} \int_{-\infty}^{\infty} |\xi| \exp(i\xi q) d\xi.$$

This kernel function is not regular, so for numerical applications a regularization must be done. A typical choice is to cut off the integration over ξ at some specific frequency k_c . This corresponds to low-pass filtering the shape of the Wigner function, which is quite undesirable; if k_c is chosen too low, sharp radial features of the Wigner function will be filtered and smoothed, while a high k_c introduces large oscillations in the shape. This problem is clearly demonstrated in some of our earlier work [105].

Another method introduced by d'Ariano et al. [27] and further developed by Leonhardt et al. [79] avoids this need for unphysical filtering. Instead of first calculating the Wigner function, each density matrix element is directly sampled from the data by

$$\rho_{mn} = \left\langle \left\langle \frac{1}{\pi} f_{mn}(q) e^{i(m-n)\theta} \right\rangle \right\rangle_{q,\theta},$$

where the averaging is over the quadrature observations. The $f_{mn}(q)$ are so-called pattern functions, independent of the data. Instead of the somewhat arbitrary filtering necessary with the inverse Radon method, this pattern function technique requires the more physically meaningful truncation of the photon numbers in the density matrix – just as the maximum likelihood method. The pattern function do however suffer from another major drawback, namely unphysical resulting density matrices. Since each matrix element is reconstructed individually, there is no mechanism which forces the total density matrix to be normalized. Furthermore, due to statistical errors, small diagonal elements can even become negative. The maximum likelihood method overcomes these limitations by specifically including the formal constraints on density matrices into the algorithm and reconstructing the entire matrix as an entity.

See also the reviews by D'Ariano et al. [28] and Lvovsky and Raymer [92] for more details about quantum state tomography.

Theory of state generation

In this chapter, a theoretical model describing our experiments with non-Gaussian state generation will be developed. At first, in §3.1, the theory for photon subtraction from a single mode squeezed vacuum will be reviewed, as it provides some useful insights into the scheme, although it fails to describe the outcomes of our experiments correctly. The rest of the chapter explores the multimode output of an OPO, the reduction to two modes (signal and trigger), and the state resulting from a photon detection, based on the Gaussian state formalism.

3.1 Single mode photon subtracted squeezed vacuum

3.1.1 Number-resolving detector

Schrödinger cat states in the shape of coherent state superpositions have proven quite difficult to generate experimentally [85]. In a 1997 paper, Dakna et al. [26] suggested a rather simple procedure for generating states very similar to Schrödinger cats with small amplitude α (Schrödinger kittens). As sketched in Figure 3.1, an initial squeezed vacuum state is split into two modes on a tapping beamsplitter with transmittance T and vacuum on the other input port. If a photon number-resolving detector in the reflected (trigger) mode records m' photons, then the resulting state in the transmitted (signal) mode will be a squeezed vacuum with m' photons subtracted. The first experimental demonstration of this idea was done by Wenger et al. [166], who observed a non-Gaussian state, which was not pure enough to have negative Wigner function values, though. In the following, we will calculate the exact form of the output state in a slightly simpler way than [26], and show that these photon subtracted squeezed vacuum states (PSSqV) resemble proper kitten states to a quite good approximation.

In the Schrödinger picture, the two-mode state incident on the beamsplitter is

$$|\psi_{\text{in}}\rangle = |\zeta\rangle_s |0\rangle_t,$$

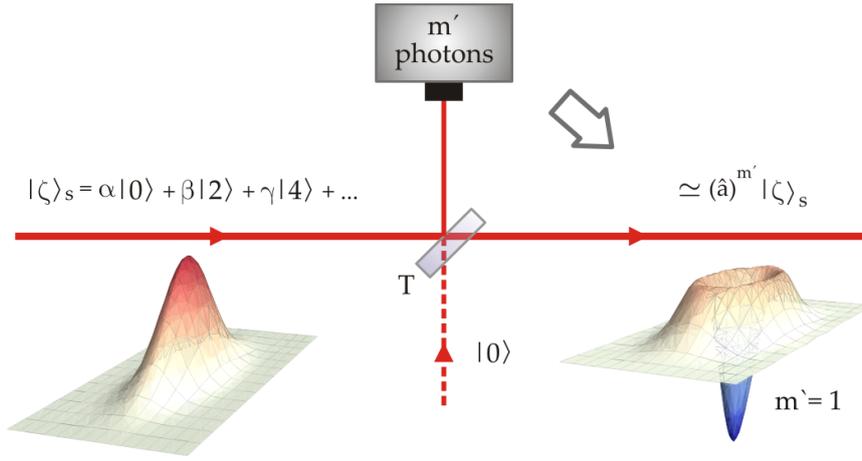


FIGURE 3.1: The simple model of Dakna et al. [26] for the generation of Schrödinger cat-like states by conditional subtraction of an exact number of photons from a single-mode squeezed vacuum. The scheme assumes an ideal photon-number resolving detector.

where $|\zeta\rangle_s$ with $\zeta = re^{i\phi}$ is the squeezed vacuum state with (2.20)

$$|\zeta\rangle_s = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{n!}}{(n/2)!} \left(\frac{e^{i\phi} \tanh r}{2} \right)^{n/2} \delta_{0n}^{[2]} |n\rangle_s.$$

The output state after the beamsplitter is, using (2.25),

$$\begin{aligned} |\psi_{bs}\rangle &= \hat{B}(T) |\psi_{in}\rangle \\ &= \sum_{\mu=0}^{\infty} {}_s\langle \mu | \zeta \rangle_s \hat{B}(T) |\mu, 0\rangle \\ &= \sum_{\mu=0}^{\infty} \sum_{n=0}^{\mu} {}_s\langle \mu | \zeta \rangle_s \sqrt{B_n^{\mu}(T)} |n, \mu - n\rangle \\ &= \sum_{n,m=0}^{\infty} {}_s\langle n + m | \zeta \rangle_s \sqrt{B_n^{n+m}(T)} |n, m\rangle. \end{aligned}$$

The expansion coefficients are explicitly

$$\begin{aligned} \langle n, m | \psi_{bs} \rangle &= {}_s\langle n + m | \zeta \rangle_s \sqrt{B_n^{n+m}(T)} \\ &= \frac{1}{\sqrt{\cosh r}} \frac{(n+m)!}{\left(\frac{n+m}{2}\right)! \sqrt{n!m!}} \left(\frac{e^{i\phi} \tanh r}{2} \right)^{\frac{n+m}{2}} \sqrt{T^n (1-T)^m} \delta_{nm}^{[2]}. \end{aligned}$$

Now, if the detector in the reflected trigger mode detects m' photons, the state is projected onto $|m'\rangle_t \langle m'|$, and the transmitted signal mode collapses into the conditional photon-subtracted state

$$|\psi_{m'}\rangle_s = \frac{1}{\sqrt{\text{pr}_t(m')}} \sum_{n=0}^{\infty} \langle n, m' | \psi_{bs} \rangle |n\rangle_s, \quad (3.1)$$

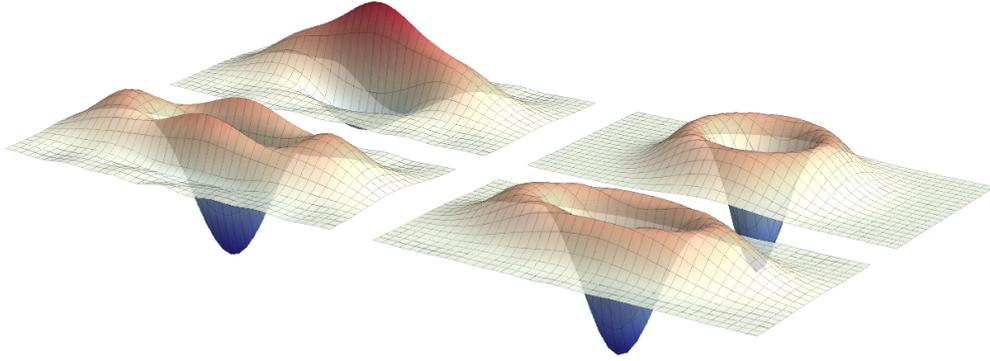


FIGURE 3.2: Wigner functions of photon subtracted squeezed vacuum states. From the front frame and counter-clockwise, the number of subtracted photons are 1, 1, 2, and 3, respectively. $r = 0.5$ for all except the rightmost which has $r = 0.1$. $T = 0.95$ for all.

where the normalization factor $\text{pr}_t(m')$ is the probability of detecting m' photons in the trigger mode, given by

$$\text{pr}_t(m') = \sum_{n=0}^{\infty} |\langle n, m' | \psi_{\text{bs}} \rangle|^2.$$

We can calculate an expression for this probability (by Mathematica):

$$\begin{aligned} \text{pr}_t(m') = & \frac{1}{\cosh r} \frac{1}{m'!} \left(\frac{1-T}{T} \right)^{m'} (\tanh s)^{m'+\delta_{1m'}^{[2]}} ((m' - \delta_{0m'}^{[2]})!!)^2 \\ & \times {}_2F_1 \left(\frac{m'+1+\delta_{1m'}^{[2]}}{2}, \frac{m'+1+\delta_{1m'}^{[2]}}{2}, \frac{1}{2} + \delta_{1m'}^{[2]}, \tanh^2 s \right), \end{aligned}$$

where ${}_2F_1$ is the hypergeometric function, $!!$ is the double factorial¹, and s is a rescaled squeezing parameter defined by

$$\tanh s = T \tanh r.$$

Figure 3.2 shows the Wigner functions of some examples of the PSSqV states $|\psi_{m'}\rangle_s$, for $m' = 1, 2$ and 3 . They are clearly non-Gaussian, and while the 1-PSSqV states have the look of a simple squeezed volcano, the shapes become more complex for higher numbers of subtracted photons.

We notice from (3.1) the fairly obvious fact that, with an odd (even) number of photons detected in the trigger mode, the signal mode consists of a superposition of only odd (even) Fock states – see Figure 3.3. More interestingly, we can decompose $|\psi_{m'}\rangle_s$ into a superposition of two states opposite in phase space. Abbreviating

$$\begin{aligned} F_{n,m'} &= \frac{(n+m')!}{\left(\frac{n+m'}{2}\right)! \sqrt{n!}}, \\ v &= \sqrt{\frac{e^{i\phi} \tanh s}{2}}, \end{aligned}$$

¹The double factorial is defined recursively by $n!! = n \cdot (n-2)!!$ with $(-1)!! = 0!! = 1!! = 1$. For example, $7!! = 7 \cdot 5 \cdot 3 \cdot 1$. [165]

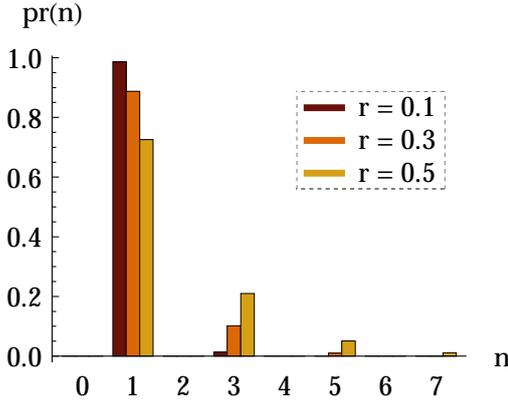


FIGURE 3.3: Photon number distributions for 1-PSSqV states with different squeezing levels. $T = 0.95$. Because single-mode squeezed vacuum is a superposition of even-number states, only odd numbers appear in the state after subtraction of a photon. For very low squeezing levels, the squeezed vacuum is primarily vacuum with a small 2-photon component, so conditioned on the detection of a trigger photon, the signal state is essentially a single photon.

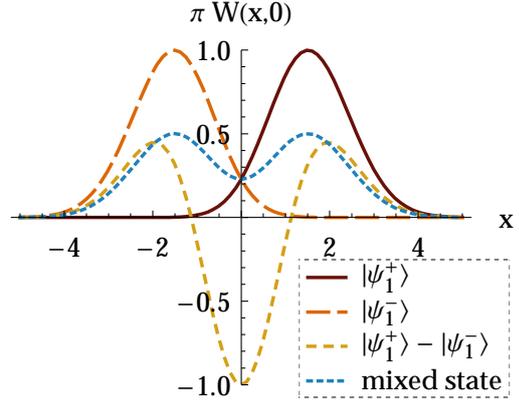


FIGURE 3.4: Wigner function slices of the 1-PSSqV state $|\psi_1\rangle \propto |\psi_1^+\rangle - |\psi_1^-\rangle$ and its component states (normalized) for $T = 0.95$, $r = 0.5$, $\phi = 0$. The blue function is what the state would look like if it was an incoherent mixture of the two states, $|\psi_1^+\rangle\langle\psi_1^+| + |\psi_1^-\rangle\langle\psi_1^-|$. The dip in the Wigner function of $|\psi_1\rangle$ is due to interference in the coherent superposition of the two component states.

$$\mathcal{N}_{m'} = (\tanh s)^{m'+\delta_{1m'}^{[2]}} ((m' - \delta_{0m'}^{[2]})!)^2 \times {}_2F_1 \left(\frac{m'+1+\delta_{1m'}^{[2]}}{2}, \frac{m'+1+\delta_{1m'}^{[2]}}{2}, \frac{1}{2} + \delta_{1m'}^{[2]}, \tanh^2 s \right),$$

(3.1) can be rewritten (skipping the s index)

$$\begin{aligned} |\psi_{m'}\rangle &= \frac{1}{\sqrt{\mathcal{N}_{m'}}} \sum_{n=0}^{\infty} F_{n,m'} v^{n+m'} \delta_{nm'}^{[2]} |n\rangle \\ &= \frac{v^{m'}}{2\sqrt{\mathcal{N}_{m'}}} \left[\sum_{n=0}^{\infty} F_{n,m'} v^n |n\rangle + (-1)^{m'} \sum_{n=0}^{\infty} F_{n,m'} (-v)^n |n\rangle \right] \\ &\equiv \frac{v^{m'}}{2\sqrt{\mathcal{N}_{m'}}} \left[|\psi_{m'}^+\rangle + (-1)^{m'} |\psi_{m'}^-\rangle \right]. \end{aligned}$$

This superposition of two distinct states already bears some resemblance with the odd/even cat states in (2.22). Figure 3.4 shows that the Wigner functions of these states look somewhat like two weak coherent states opposite in phase space.

Fidelity

To quantify how close two states are to each other, the fidelity is a useful and commonly accepted measure [19, 64]. The fidelity between a pure state $|\psi_1\rangle$ and

a mixed state $\hat{\rho}_2$ is defined as

$$\begin{aligned}\mathcal{F}(|\psi_1\rangle, \hat{\rho}_2) &\equiv \langle \psi_1 | \hat{\rho}_2 | \psi_1 \rangle = \text{tr}[|\psi_1\rangle\langle\psi_1| \hat{\rho}_2] \\ &= 2\pi \iint_{-\infty}^{\infty} W_1(x, p) W_2(x, p) dx dp,\end{aligned}$$

where W_1, W_2 are the Wigner functions of the two states. \mathcal{F} is 1, only if the two states are identical, and 0 if they are orthogonal. For two pure states, $\mathcal{F}(|\psi_1\rangle, |\psi_2\rangle) = |\langle \psi_1 | \psi_2 \rangle|^2$. For the fidelity (or overlap) between the m' -PSSqV $|\psi_{m'}\rangle$ and an odd cat state $|\sim_{\odot-}^{\otimes}\rangle$ (2.22), we get

$$\mathcal{F}(|\psi_{m'}\rangle, |\sim_{\odot-}^{\otimes}\rangle) = |\langle \psi_{m'} | \sim_{\odot-}^{\otimes} \rangle|^2 = 0$$

if m' is even, since the m' -PSSqV then has only even photon number components while the odd cat has only odd. If, on the other hand, m' is odd,

$$\begin{aligned}\mathcal{F}(|\psi_{m'}\rangle, |\sim_{\odot-}^{\otimes}\rangle) &= \left| \sum_{n=0}^{\infty} \langle \psi_{m'} | n \rangle \langle n | \sim_{\odot-}^{\otimes} \rangle \right|^2 \\ &= \left| \frac{1}{\sqrt{\mathcal{N}_{m'}} \sqrt{\sinh|\alpha|^2}} \sum_{n=0}^{\infty} \delta_{m'n}^{[2]} F_{n,m'}(v^*)^{n+m'} \frac{\alpha^n}{\sqrt{n!}} \right|^2 \quad (3.2) \\ &= \frac{|v|^{2m'}}{\mathcal{N}_{m'} \sinh|\alpha|^2} \left| \sum_{n=0}^{\infty} \delta_{1n}^{[2]} \frac{(n+m')!}{\left(\frac{n+m'}{2}\right)! n!} (\alpha v^*)^n \right|^2.\end{aligned}$$

For an even cat state, the fidelity is 0 for odd m' ; for even m' it is the same expression as above, only with $\cosh|\alpha|^2$ instead of $\sinh|\alpha|^2$, and with $\delta_{0n}^{[2]}$ instead of $\delta_{1n}^{[2]}$.

In the special case of one subtracted photon, $m' = 1$, the fidelity becomes (§A.2)

$$\mathcal{F}(|\psi_1\rangle, |\sim_{\odot-}^{\otimes}\rangle) = \frac{|\alpha|^2}{\cosh^3 s \sinh|\alpha|^2} \exp(|\alpha|^2 \tanh s \cos(2 \arg \alpha - \phi)).$$

This expression was also obtained by Lund, Jeong et al. [62, 85, 86] for the fidelity between the odd cat and a squeezed single photon, that is, a single photon state passed through an ideal squeezer. The reason is of course that the two states, 1-PSSqV ($m' = 1$) and squeezed single photon, are equivalent; from [86]

$$\hat{S}(s) |1\rangle = \frac{1}{\sinh s} \hat{a} \hat{S}(s) |0\rangle,$$

which shows this equivalency (the $\sinh s$ factor just normalizes the photon subtracted state on the right hand side). It is not the same amount of squeezing required, though, to generate a 1-PSSqV and a squeezed photon state. A 1-PSSqV generated with a squeezing parameter r and a beamsplitter transmittance of T corresponds to a single photon squeezed with a parameter $s = \text{arctanh}(T \tanh r)$.

The fidelities between different m' -PSSqV's and cat states (odd/even as appropriate) are plotted in Figure 3.5 a) as a function of the squeezing parameter r . For each r , the cat state amplitude α has been optimized to yield the highest possible fidelity. This optimal α is plotted in Figure 3.5 b). The fidelity is almost perfect for squeezing parameters below 0.2, but for $r \gtrsim 0.6$, it rapidly degrades. In order to keep the fidelities higher than, say 0.95, the cat amplitude can not get higher than 1.5 – 2.

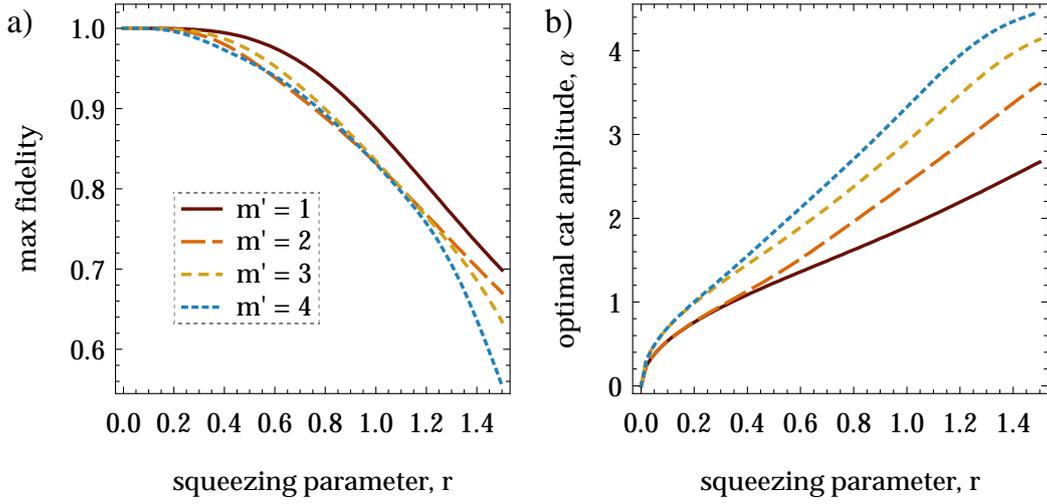


FIGURE 3.5: For different numbers of detected trigger photons m' , a) shows the maximum attainable fidelity between a m' -PSSqV with given squeezing parameter r , and any cat state. b) shows the corresponding amplitudes of the optimal cat state.

3.1.2 Realistic photon detection

The discussion so far has assumed that the trigger detector signalling the subtraction of one or more photons from the squeezed field can resolve photon numbers perfectly. No such detector exists in reality. As mentioned in §2.3.1, commercially available photon counting detectors (like APDs) cannot distinguish between one or several photons impinging on it at the same time – they can only signal whether some photons were absorbed or not. Furthermore, they have a certain dead time after a detection event, within which they cannot accept new detections. These kinds of detectors are commonly called on/off detectors, as opposed to the idealized photon number resolving detectors. Whereas the detection of m photons on a number resolving detector is described theoretically by the projection onto a number state, $|m\rangle_t\langle m|$, the on/off detector's on or “click” event effectively projects the state of the signal mode into a statistical mixture of 1-PSSqV, 2-PSSqV, 3-PSSqV, etc., and can be modeled by a summation over all projection operators for $m \geq 1$,

$$\hat{\Pi}_t^{\text{on}} = \sum_{m \geq 1} |m\rangle_t\langle m|,$$

or, since the projectors are complete,

$$\hat{\Pi}_t^{\text{on}} = \hat{I}_t - \hat{\Pi}_t^{\text{off}} = \hat{I}_t - |0\rangle_t\langle 0|.$$

The efficiency of the detector must also be taken into account. Apart from the quantum efficiency of the APD itself, there will be propagation losses on the way from the beamsplitter to the detector, yielding a total efficiency η_t for the reflected photons to be detected. This efficiency can be modeled by a fictitious beamsplitter before a 100% efficiency detector, or, alternatively, it can be modeled

by replacing the on- and off-projection operators with POVMs [147]

$$\hat{E}_t^{\text{on}}(\eta_t) = \hat{I}_t - \hat{E}_t^{\text{off}}(\eta_t), \quad \hat{E}_t^{\text{off}}(\eta_t) = \sum_{m=0}^{\infty} (1 - \eta_t)^m |m\rangle_t \langle m|,$$

because $(1 - \eta_t)^m$ is the probability that m photons will *not* be detected. For unit efficiency, $\hat{E}_t^{\text{off}}(1)$ is again $|0\rangle_t \langle 0|$, with the convention $0^0 = 1$. The two-mode state resulting from a click event in the trigger mode is obtained, according to (2.1), from operating with $\hat{E}_t^{\text{on}}(\eta_t)$ on the initial two-mode state after the beam-splitter², $|\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|$. Tracing over the trigger mode (whereby the POVM element can be collected on one side of the density matrix, due to cyclic invariance of the trace) then gives the state of the signal mode:

$$\begin{aligned} \hat{\rho}_{\text{on},s}(\eta_t) &= \frac{\text{tr}_t [\hat{E}_t^{\text{on}}(\eta_t) |\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|]}{\text{pr}_{\text{on}}} \\ &= \frac{\text{tr}_t [(\hat{I}_t - \sum_m (1 - \eta_t)^m |m\rangle_t \langle m|) |\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|]}{\text{tr}_{s,t} [(\hat{I}_t - \sum_m (1 - \eta_t)^m |m\rangle_t \langle m|) |\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|]} \\ &= \frac{\sum_m (1 - (1 - \eta_t)^m) {}_t \langle m | \psi_{\text{bs}} \rangle \langle \psi_{\text{bs}} | m \rangle_t}{1 - \sum_{n,m} (1 - \eta_t)^m \langle n, m | \psi_{\text{bs}} \rangle \langle \psi_{\text{bs}} | n, m \rangle}. \end{aligned}$$

Yet another imperfection in the APD to be considered is the existence of dark counts; due to thermal noise in the diode, the detector will sometimes give a click without any impinging photons. Such a dark count gives no information at all about the state of the trigger mode – the corresponding action on the trigger mode is just the identity operator; $\hat{E}_t^{\text{dark}} = \hat{I}_t$. The resulting signal state is then simply the original squeezed vacuum (that has suffered a loss of $1 - T$ from the beamsplitter):

$$\hat{\rho}_{\text{dark},s} = \frac{\text{tr}_t [\hat{I}_t |\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|]}{\text{tr}_{s,t} [\hat{I}_t |\psi_{\text{bs}}\rangle \langle \psi_{\text{bs}}|]} = \sum_m {}_t \langle m | \psi_{\text{bs}} \rangle \langle \psi_{\text{bs}} | m \rangle_t.$$

In the event of a dark count, there might indeed have been photons split off from the state, but this is no different from just observing the unconditioned signal state, that is, regardless of whether the trigger detector gives a click or not.

A trigger detector click will be either a genuine photon detection, $\hat{E}_t^{\text{on}}(\eta_t)$, or a dark count, \hat{E}_t^{dark} . Hence, an individual signal state conditioned on a click will be either a PSSqV or a squeezed vacuum state. Following Wenger et al. [166] we call the fraction of genuine counts the ‘modal purity’ and assign it the symbol Ξ . The fraction of dark counts is then $1 - \Xi$ and the ensemble averaged density matrix is

$$\hat{\rho}_{\text{on},s}(\eta_t, \Xi) = \Xi \hat{\rho}_{\text{on},s}(\eta_t) + (1 - \Xi) \hat{\rho}_{\text{dark},s}. \quad (3.3)$$

Dark counts are actually not the only source of false clicks. If any photons occupying modes other than the trigger mode are detected by the APD, they will, as well, project the signal mode onto a squeezed vacuum state, and they can be incorporated into the modal purity Ξ on equal footing with dark counts. An important aspect of our experiment is to avoid such uncorrelated photons from reaching the trigger APD.

²Implicitly, the full two-mode operator is $\hat{I}_s \otimes \hat{E}_t^{\text{on}}(\eta_t)$.

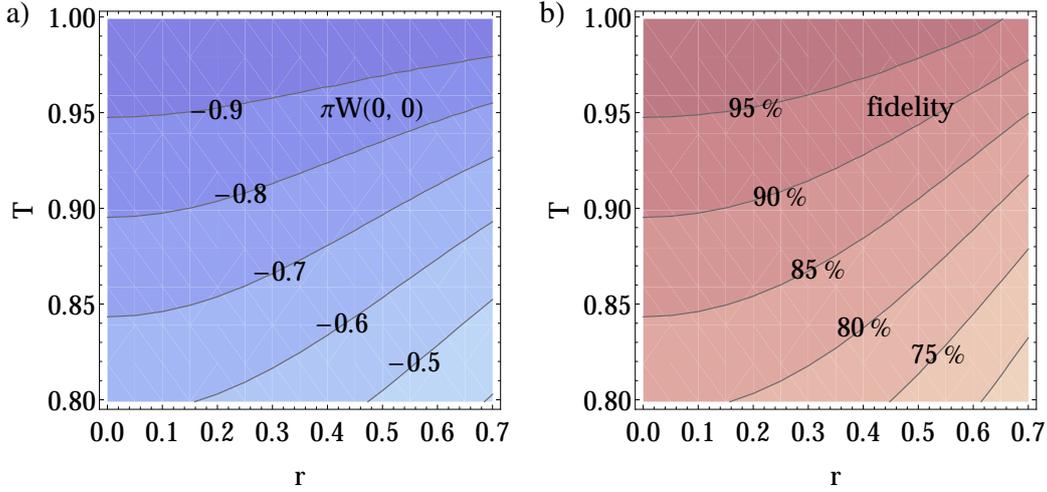


FIGURE 3.6: The r - and T -dependence of **a)** $\pi W(0,0)$, the value of the negative dip in the Wigner function of the PSSqV state which is ideally -1 , but is reduced when using an on/off detector, and **b)** $\mathcal{F}(|\tilde{\alpha}_-^{\text{odd}}\rangle, \hat{\rho}_{\text{on},s})$, the fidelity with an odd kitten state with optimized amplitude. The parameters of the on/off detector are in both plots $\eta_t = 0.1$ and $\zeta = 1$. The summation in the POVM for the on/off detector has been truncated at $M = 15$.

Effect on the states

The step from an ideal photon number detector to the more realistic on/off detector model naturally has some consequences for the quality of the generated PSSqV states. This is clearly exemplified in the value of the negative Wigner function dip at the origin of phase space, $W(0,0)$. For 1 subtracted photon (or 3 or 5 or...), the value in the ideal detector case is always $-1/\pi$, no matter what the squeezing parameter r and the beamsplitter transmittance T are. Once we lose the information about exactly how many photons were reflected off the beamsplitter, it starts to matter whether there is a high or a low probability of having more than one photon in the trigger arm. Thus, if the squeezing level is higher, there will be more photons in the initial field and it is more likely that two or more will be reflected; the same goes if the beamsplitter transmittance is relatively low. Then there will be some admixture in the signal state of the 2-PSSqV state which has $W(0,0) = 1/\pi$ as seen in Figure 3.2. The dependence on r and T of $W(0,0)$ as well as the fidelity with an odd Schrödinger kitten is sketched in Figure 3.6. Especially the beamsplitter transmittance has a dramatic effect on both numbers – it clearly has to be kept as close to unity as possible. The disadvantage of a high transmittance (and low reflectance) is of course to lower the production rate, since fewer photons are detected by the trigger detector.

T and r are parameters of the initial state preparation scheme that become highly relevant due to the limited capabilities of the realistic photon detector. The properties of the trigger detection apparatus itself will also influence the quality of the generated signal state. While the detector efficiency η_t has a small, though noticeable, effect, the modal purity \mathcal{E} has a quite direct and severe impact on the quality of the generated state, as shown in Figure 3.7. A low efficiency definitely

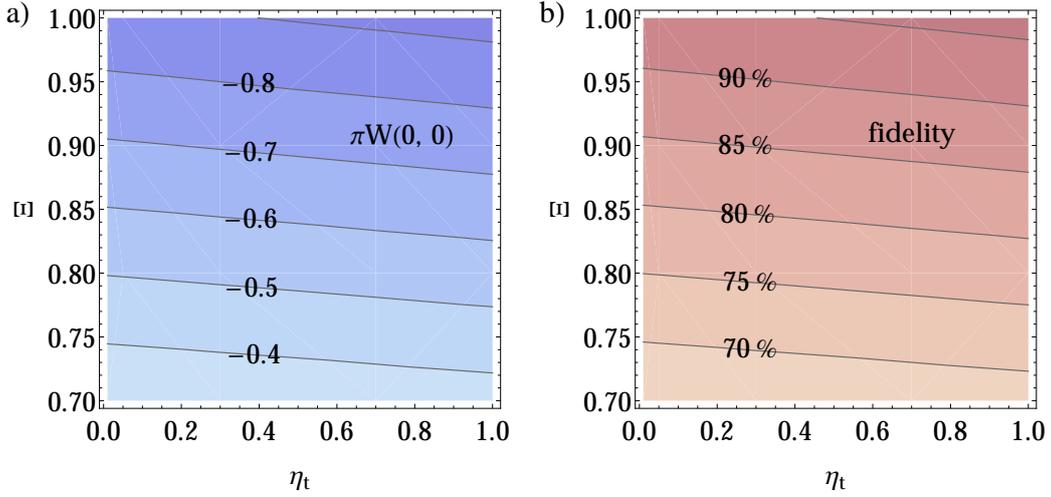


FIGURE 3.7: The η_t - and \mathcal{E} -dependence of **a)** $\pi W(0,0)$, and **b)** $\mathcal{F}(|\tilde{\mathcal{O}}_-\rangle, \hat{\rho}_{\text{on},s})$. The squeezing parameter is $r = 0.5$ and the beamsplitter transmittance $T = 0.95$.

means that the production rate will go down, but the quality of the state is only weakly degraded, at least for realistic values of T and r . It is more likely that two photons in the trigger arm will give rise to a click, than if there is only one – hence, the non-unity efficiency leads to relatively higher proportion of 2-PSSqV states (or higher) in the signal. But as long as the initial probability of having more than one photon in the trigger arm is kept low, the effect will be only moderate.

Whereas the three previously discussed parameters T , r , and η_t all serve to mix the 1-PSSqV state with higher number subtracted states to varying degrees, the influence of a non-unity modal purity is to admix plain squeezed vacuum with the amount $1 - \mathcal{E}$. Both fidelity and Wigner function dip will naturally suffer heavily from that. Finally, there will also be losses in the signal arm. These will alter the state according to the Bernoulli transformation (2.26) with an efficiency η_s . In the experimental situation, these signal losses along with dark counts/uncorrelated clicks are the major limitations to be beaten.

The emphasis given here to the Wigner function value in the origin, $W(0,0)$, will continue throughout the thesis – it will be the main figure of merit for our generated states. There are several good reasons for emphasizing this particular parameter: Our target states, odd kittens and single photons, both have a large negativity in the center of phase space; however, experimental imperfections will quickly reduce the negativity, so it is a very sensitive gauge of the performance of the experiment. Furthermore, $W(0,0)$ is easy to calculate and visualize, and is a parameter commonly encountered in the literature (e.g. [69]). For more complex states, it might not be the optimal figure of merit, though – the 2-photon state, for example, has a positive peak in the center which is only weakly affected by losses, admixture with squeezed vacuum etc.

3.2 Gaussian formalism

The theory developed so far has assumed an ideal input squeezed state, i.e. a pure minimum uncertainty state, and furthermore any mentioning of relevant physical modes such as frequency, temporal, and spatial modes have been neglected. Several illustrative results have been derived, showing the influence on such an ideal system of different photon detector models and characteristics, squeezing level, and beamsplitting ratio. However, a more realistic model – in particular for the input squeezed state – is necessary to properly describe the experiments. As we will see in §3.3, the OPO output consists of a continuum of frequency modes with different squeezing levels, so it should in principle be represented by a multimode density matrix. The insufficiency of single mode theories were realized by Sasaki and Suzuki [133] and by Mølmer [99], who developed different formalisms to take account of the multimode character of the system. The following theory of the state generation scheme is based on the latter of these works. An essential ingredient of the formalism is the temporal mode selection (§3.4), which reduces the system to two modes, signal and trigger, such that a two-mode density matrix suffices to describe it. The states within these modes will be mixed, even for perfect experimental conditions with no losses.

Once we depart from the idealized model, the explicit formulation in terms of number states, as done in the previous sections, becomes rather cumbersome. It is still possible to perform state simulations using the density matrix; however, since the state of the system is Gaussian until the de-Gaussifying photon detection takes place, a Gaussian formalism is actually more appropriate and more transparent as well. The state is then described simply by its covariance matrix (there is no displacement in phase space), which can encompass any initial squeezed vacuum state, not only minimum uncertainty states. The further steps in the modeling can then go via the characteristic function or the Wigner function – as they are Fourier transforms of each other, there is no significant argument for using one or the other. Kim et al. [69] went via the characteristic function to calculate results for the fidelity and $W(0,0)$ for kittens prepared with the on/off detector model. They did not consider non-unity photon detector efficiency, though. On the other hand, they could get results for non-ideal initial squeezed states due to the covariance matrix formulation. In the part of parameter space where their theory and that developed in the previous sections overlap, the two models agree completely. In the theory paper written by Mølmer [99] in connection with our experimental demonstration of the kitten state preparation (and in subsequent related works by Nielsen and Mølmer, [107–110]), the Wigner function is used, and we will do the same in the following.

Very generally, we assume that the trigger and signal modes are in a two-mode Gaussian state just prior to the click event of the trigger detector. This state can be the pure squeezed state after a beamsplitter, as obtained from the single mode theory of the previous section, or it can be a mixed state, for example from our reduced multimode system. From the covariance matrix Γ of the two modes, the Wigner function $W_{\Gamma}(\xi)$ is as in (2.15). We saw before, that to describe in the density matrix representation the effect of a trigger detection on the signal

mode, we operate on the state $\hat{\rho}_\Gamma$ with the on-POVM and trace over the trigger mode:

$$\hat{\rho}_{\text{on},s} = \frac{\text{tr}_t [\hat{E}_t^{\text{on}} \hat{\rho}_\Gamma]}{\text{tr}_{s,t} [\hat{E}_t^{\text{on}} \hat{\rho}_\Gamma]}. \quad (3.4)$$

In the Wigner function representation, this corresponds to integrating over $W_\Gamma(\xi)$ multiplied by the Wigner function of the on-POVM, cf. (2.10),

$$W_{\text{on}}(x_s, p_s) = \frac{2\pi \iint W_\Gamma(\xi) W_E(x_t, p_t) dx_t dp_t}{2\pi \int W_\Gamma(\xi) W_E(x_t, p_t) d^4\xi}. \quad (3.5)$$

With the 100% efficiency on-POVM, $\hat{E}_t^{\text{on}} = \hat{I}_t - |0\rangle_t\langle 0|$, the operator Wigner function becomes

$$W_E(x_t, p_t) = \frac{1}{2\pi} - \frac{1}{\pi} e^{-x_t^2 - p_t^2}. \quad (3.6)$$

The last term is the vacuum, while the first is the Wigner function of the identity operator, as obtained directly from (2.7). For the detector with efficiency η_t , the vacuum part of the Wigner function should be exchanged with the corresponding sum of number state Wigner functions. However, the non-unity efficiency is more easily included by inserting a fictitious beamsplitter before the detector, thus altering the covariance matrix, and then use (3.6). Through rotations in phase space, a two-mode Gaussian state can in general be put on the generic form

$$\Gamma = \begin{pmatrix} a & 0 & e & 0 \\ 0 & b & 0 & f \\ e & 0 & c & 0 \\ 0 & f & 0 & d \end{pmatrix}.$$

With these covariance parameters, the numerator and denominator of the conditioned signal state expression (3.5) becomes, using W_E from (3.6),

$$2\pi \iint W_\Gamma(\xi) W_E(x_t, p_t) dx_t dp_t = \frac{\exp\left(-\frac{x_s^2}{a} - \frac{p_s^2}{b}\right)}{\pi\sqrt{ab}} - \frac{2 \exp\left(-\frac{(1+c)x_s^2}{a(1+c)-e^2} - \frac{(1+d)p_s^2}{b(1+d)-f^2}\right)}{\pi\sqrt{(a(1+c)-e^2)(b(1+d)-f^2)}}, \quad (3.7)$$

$$\text{pr}_{\text{on}} = 2\pi \int W_\Gamma(\xi) W_E(x_t, p_t) d^4\xi = 1 - \frac{2}{\sqrt{(1+c)(1+d)}}. \quad (3.8)$$

The resulting state is thus the difference of two Gaussian functions, one of which is the unconditioned signal state obtained by simply tracing over the trigger mode. Non-unity modal purity Ξ is included, as in (3.3), by admixing with the unconditioned state:

$$W_{\text{on}}(x_s, p_s; \Xi) = \Xi W_{\text{on}}(x_s, p_s) + (1 - \Xi) W_{\text{uncond}}(x_s, p_s), \quad (3.9)$$

with

$$W_{\text{uncond}}(x_s, p_s) = \frac{\exp\left(-\frac{x_s^2}{a} - \frac{p_s^2}{b}\right)}{\pi\sqrt{ab}}. \quad (3.10)$$

Sometimes – e.g. in Mølmer [99] – a different model of the photon detection event is chosen, in which the action on the trigger mode of the density matrix is not a projection, but rather the subtraction of a single photon, described by the annihilation operator:

$$\hat{\rho}_{\text{on},s} = \frac{\text{tr}_t [\hat{a}_t \hat{\rho}_\Gamma \hat{a}_t^\dagger]}{\text{tr}_{s,t} [\hat{a}_t \hat{\rho}_\Gamma \hat{a}_t^\dagger]}. \quad (3.11)$$

Through some correspondences between application of \hat{a} and \hat{a}^\dagger to the density matrix and differentiation of the Wigner function [47], this amounts to a transformation in the Wigner representation of

$$W_{\text{on}}(x_s, p_s) = \frac{1}{\text{pr}_{\text{on}}} \int \frac{1}{2} \left(x_t^2 + p_t^2 + \frac{1}{4} \left(\frac{\partial^2}{\partial x_t^2} + \frac{\partial^2}{\partial p_t^2} \right) + x_t \frac{\partial}{\partial x_t} + p_t \frac{\partial}{\partial p_t} + 1 \right) W_\Gamma(\xi) dx_t dp_t.$$

One could start arguing whether this model of the detection process is more realistic, since photons are indeed absorbed by the detector. However, when tracing over the trigger mode, it does not really matter what actually happened to the photons in that mode. In fact, one can easily show (see Appendix A.3) that this model is equivalent to the projection operator model described above in the limit of low detector efficiency, $\eta_t \rightarrow 0$. So the projection model is more general, but for experimentally relevant efficiencies ($\eta_t \approx 0.1$), the two models produce almost identical predictions.

Example

The modeling of the conditioning process that turns a Gaussian state into a non-Gaussian state in the signal mode is thus quite simple. Most of the specifics of a given physical setup is handled prior to the photon detection, while staying within the Gaussian formalism. We briefly demonstrate how the idealized system, treated in the previous section using state ket formalism, can be easily treated in terms of the covariance matrix. With an initial squeezed vacuum state (not necessarily pure) as input in the signal mode, and vacuum input in the trigger mode, the covariance matrix before the beamsplitter is

$$\Gamma^i = \begin{pmatrix} 2\langle \hat{x}_{i,s}^2 \rangle & 0 & 0 & 0 \\ 0 & 2\langle \hat{p}_{i,s}^2 \rangle & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The beamsplitter transforms the coordinates according to $\hat{\xi}^i \rightarrow \hat{\zeta} = B(T)\hat{\xi}^i$ with the unitary matrix

$$B(T) = \begin{pmatrix} \sqrt{T} & 0 & \sqrt{1-T} & 0 \\ 0 & \sqrt{T} & 0 & \sqrt{1-T} \\ -\sqrt{1-T} & 0 & \sqrt{T} & 0 \\ 0 & -\sqrt{1-T} & 0 & \sqrt{T} \end{pmatrix}. \quad (3.12)$$

The corresponding covariance matrix transformation is $\Gamma^i \rightarrow \Gamma = B(T)\Gamma^i B(T)^T$. The efficiencies of the trigger channel as well as the signal channel are now modeled by a fictitious beamsplitter in the trigger mode with transmittance η_t , and in the signal mode with η_s . Using a matrix $B(\eta_t)$ like above on the trigger mode and an auxiliary vacuum mode, doing the same with $B(\eta_s)$ on the signal mode, and then disregarding the auxiliary output modes results in a transformation of the covariance matrix of

$$\Gamma \rightarrow \Gamma' = \underline{\eta}\Gamma\underline{\eta} + \underline{V},$$

with the efficiency matrix $\underline{\eta} = \text{diag}(\sqrt{\eta_s}, \sqrt{\eta_s}, \sqrt{\eta_t}, \sqrt{\eta_t})$ and the vacuum contribution matrix $\underline{V} = \text{diag}(1 - \eta_s, 1 - \eta_s, 1 - \eta_t, 1 - \eta_t)$. The final covariance matrix of the two-mode system just before the trigger detection event is thus

$$\Gamma' = \underline{\eta}B(T)\Gamma^i B(T)^T \underline{\eta} + \underline{V},$$

and the signal state conditioned on a click is then obtained from (3.5)-(3.9).

The results of this analysis are exactly the same as those obtained from the state ket formalism – but this Gaussian approach is somewhat more elegant.

3.3 OPO theory

Since the first experiment showing noise reduction below the vacuum level, performed by Slusher et al. [143] using four-wave mixing in sodium atoms, most experiments demonstrating or utilizing squeezed light have employed nonlinear crystals, fibers or waveguides. Nonlinear crystals have a large second-order susceptibility, $\chi^{(2)}$, which permits frequency-mixing between incoming and outgoing optical fields. One possible effect is spontaneous parametric down-conversion (SPDC) in which a small part of an incident pump field at frequency ω_p is converted into two fields at ω_+ and ω_- , with $\omega_+ + \omega_- = \omega_p$. Apart from frequency, the down-converted twin fields can be correlated in the spatial, temporal, and polarization degrees of freedom, and SPDC is thus often used for entangled photon pair production. Furthermore, the two fields will be quadrature-correlated due to their common dependence on the pump field phase, so they can be in a two-mode squeezed state. In the case of mode degeneracy of the two fields, the common mode will be in a normal squeezed state. We will be employing a degenerate OPO for kitten state generation and a (frequency) non-degenerate OPO to make single photons. The polarization is always degenerate, that is, we use Type-I down-conversion.

The Hamiltonian describing the SPDC process is, assuming that the pump field amplitude $\beta = \langle \hat{a}_{\omega_p} \rangle$ is large enough to be treated classically [136],

$$\hat{H} = \frac{i\hbar g}{2} (\beta e^{-i\omega_p t} \hat{a}_+^\dagger \hat{a}_-^\dagger - \beta^* e^{i\omega_p t} \hat{a}_+ \hat{a}_-), \quad (3.13)$$

where g is the interaction strength. To produce a reasonable amount of squeezing, the product of the interaction strength, which is proportional to the $\chi^{(2)}$ and the length of the crystal, and the pump amplitude should be fairly high. There

are two quite different ways of running the down-conversion process; either by pulsed or by continuous-wave (cw) pumping, and each way has its merits.

With pulsed pumping, very high peak powers can be obtained when using picosecond, or even femtosecond lasers, on the order of several hundred kW. However, for such short pulses, the crystal must be very short to avoid effects of group velocity mismatch [122]. The down-converted squeezed light will be ultrashort pulses as well, and will thus populate a very wide bandwidth. Usual squeezing levels obtained are around 3 dB noise reduction [50, 122]. These relatively low levels, as well as the ultrashort pulsed nature itself, limit the usability for many squeezed light applications. For conditional state engineering as presented in the previous sections, the pulsed scheme clearly has its justification, since high squeezing levels are not necessary.

By cw pumping, the pump is a steady state field with powers typically between a few mW and a few hundred mW – considerably lower than the powers involved in pulsed lasers. However, the problems associated with short pulses are not present and the nonlinear crystal can be made considerably longer, usually around a cm. In order to further enhance the interaction strength, the crystal can be placed inside an optical cavity resonant on the down-converted frequency. Such an arrangement is called an optical parametric oscillator (OPO). The squeezed light generated from an OPO is continuous, with a bandwidth matching that of the cavity. The best squeezing results so far are obtained from OPOs; 9 dB below vacuum noise by Takeno et al. [151] using a long traveling-wave cavity, and -10 dB in a monolithic cavity by Vahlbruch et al. [159]. It should be noted, that although pulsed pump experiments almost always are single-pass configurations, Zavatta et al. [174] recently implemented cavity enhancement of pulsed SPDC using a cavity with a round-trip time equal to the inverse of the laser repetition rate.

3.3.1 OPO modeling

The following derivation of the OPO's cavity and output fields is based on the formalism in Collett and Gardiner [25]. A few more details are given in my master thesis [105].

Cavity spectrum

We model the cavity as in Figure 3.8. The singly resonant OPO is kept at resonance with the down-converted field at $\omega_0 = \omega_p/2$, but not with the pump. This is possible by having the mirrors transmitting the pump wavelength, but being highly reflective at the down-converted wavelength - they are assumed to be perfectly reflecting. One mirror, the output coupler, is transmitting a small fraction, described by the decay rate γ_1 . There will inevitably be losses in the cavity, which are modeled by a fictitious beamsplitter with a decay rate of γ_2 . The relation between the decay rates and the output coupler transmittance T_1

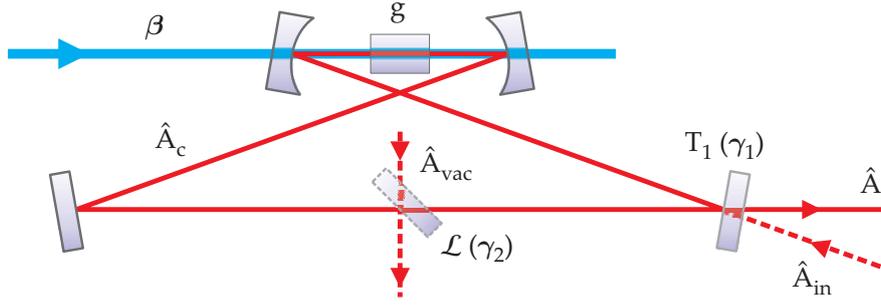


FIGURE 3.8: Schematic model of the OPO. The blue pump is not resonant in the cavity. A fictitious beamsplitter models the intra-cavity losses.

and intra-cavity (intensity) losses \mathcal{L} is

$$\gamma_1 = \frac{1 - \sqrt{1 - T_1}}{\tau_{\text{cav}}} \approx \frac{T_1}{2\tau_{\text{cav}}}, \quad \gamma_2 = \frac{1 - \sqrt{1 - \mathcal{L}}}{\tau_{\text{cav}}} \approx \frac{\mathcal{L}}{2\tau_{\text{cav}}},$$

where $\tau_{\text{cav}} = l_{\text{cav}}/c$ is the cavity round-trip time, l_{cav} the cavity length, and the approximations hold for $T_1, \mathcal{L} \ll 1$. The total decay rate $\gamma = \gamma_1 + \gamma_2$ is the HWHM (half-width half-maximum) bandwidth of the approximately Lorentzian shaped resonances. Around the central resonance at ω_0 , there is a comb of other resonances at frequencies separated by an integer times the cavity's free spectral range $\Delta\omega = 2\pi\tau_{\text{cav}}^{-1}$,

$$\omega_n = \omega_0 + n\Delta\omega.$$

For an OPO, the field operators in the Hamiltonian (3.13) refers to the intra-cavity field, which can be separated into the different resonant modes. These frequency modes are pairwise correlated, the mode at ω_n with the one at ω_{-n} , and the operators \hat{a}_+ and \hat{a}_- refer then to these modes. When considering such distinct resonance pairs and their correlations, we talk about the non-degenerate OPO (NDOPO). As opposed to this, the degenerate OPO (DOPO) considers the description of just the central mode at ω_0 , in which $\hat{a}_+ = \hat{a}_-$. The DOPO and NDOPO are not separate modes of operation of the OPO – there are correlations in the degenerate and non-degenerate modes at the same time. Rather, since the different resonances can be distinguished in the subsequent processing and analysis of the OPO output (for example in a homodyne detector with limited spectral range), the distinction between DOPO and NDOPO is merely a shift of emphasis on the different parts of the output. Dunlop et al. [33] have actually expanded the present theoretical formalism such that it in principle covers the entire spectrum of the OPO output. However, we choose to keep the DOPO and NDOPO descriptions separate, keeping the formulas a bit simpler. Furthermore, we assume that the cavity is exactly on resonance with the central frequency ω_0 . The theory in ref. [33] also included finite detuning (although it seems that some of their results are erroneous).

3.3.2 DOPO

The field operators are needed in both the time and frequency domains. We use the symmetric Fourier transform for the relation between the two domains:

$$\begin{aligned}\hat{a}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} \hat{a}(t) dt, \\ \hat{a}(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \hat{a}(\omega) d\omega.\end{aligned}$$

It is useful to change the operators to a frame rotating with the central optical frequency ω_0 by writing $\omega = \omega_0 + \Omega$:

$$\begin{aligned}\hat{A}(\Omega) &\equiv \hat{a}(\omega_0 + \Omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\Omega t} \hat{A}(t) dt, \\ \hat{A}(t) &\equiv \hat{a}(t) e^{i\omega_0 t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\Omega t} \hat{A}(\Omega) d\Omega.\end{aligned}\tag{3.14}$$

For the daggered operators, the Fourier transform takes $\hat{A}^\dagger(t) \xleftrightarrow{\mathcal{F}} \hat{A}^\dagger(-\Omega)$. The standard time domain commutation relation for the field operators (which holds for both \hat{a} and \hat{A} operators) is

$$[\hat{A}(t), \hat{A}^\dagger(t')] = \delta(t - t'),$$

which in the frequency domain becomes

$$[\hat{A}(\Omega), \hat{A}^\dagger(-\Omega')] = \delta(\Omega + \Omega').$$

Assuming the detuning of the cavity resonance from $\omega_p/2$ to be zero, the Heisenberg picture equation of motion for the intra-cavity field \hat{A}_c of the degenerate OPO is, in the rotating frame [25],

$$\frac{d\hat{A}_c(t)}{dt} = \epsilon e^{i\phi} \hat{A}_c^\dagger(t) - \gamma \hat{A}_c(t) + \sqrt{2\gamma_1} \hat{A}_{\text{in}}(t) + \sqrt{2\gamma_2} \hat{A}_{\text{vac}}(t),\tag{3.15}$$

where $\epsilon = g|\beta|$ (with units s^{-1}), and ϕ is the phase offset of the pump field β . The intra-cavity losses introduce vacuum noise, described by \hat{A}_{vac} , and unless an input field is specifically injected through the output coupler, \hat{A}_{in} is also a vacuum operator. If $\epsilon > \gamma$, the cavity mean field builds up faster than it decays and the leaking field will acquire a coherent amplitude; this condition is the oscillation threshold. Let $|\beta_{\text{thr}}|$ be the pump amplitude necessary to reach threshold; $g|\beta_{\text{thr}}| = \gamma$. Then there is the following experimentally useful relation between the various parameters:

$$\frac{\epsilon}{\gamma} = \frac{|\beta|}{|\beta_{\text{thr}}|} = \sqrt{\frac{P_p}{P_{\text{thr}}}},\tag{3.16}$$

where the P_p and P_{thr} denote pump power and pump threshold power, respectively. We are only interested in the sub-threshold behaviour of the OPO, so in the following, $\epsilon < \gamma$ can be assumed.

The output field \hat{A} is related to the cavity field through $\sqrt{2\gamma_1}\hat{A}_c = \hat{A}_{\text{in}} + \hat{A}$. In frequency space, the solution to the equation of motion for the output field is

$$\hat{A}(\Omega) = \frac{1}{D(\Omega)} [B_1(\Omega)\hat{A}_{\text{in}}(\Omega) + C_1e^{i\phi}\hat{A}_{\text{in}}^\dagger(-\Omega) + B_2(\Omega)\hat{A}_{\text{vac}}(\Omega) + C_2e^{i\phi}\hat{A}_{\text{vac}}^\dagger(-\Omega)], \quad (3.17)$$

with the abbreviations

$$\begin{aligned} B_1(\Omega) &= \gamma_1^2 - (\gamma_2 - i\Omega)^2 + \epsilon^2, \\ B_2(\Omega) &= 2\sqrt{\gamma_1\gamma_2}(\gamma - i\Omega), \\ C_1 &= 2\gamma_1\epsilon, \\ C_2 &= 2\sqrt{\gamma_1\gamma_2}\epsilon, \\ D(\Omega) &= (\gamma - i\Omega)^2 - \epsilon^2. \end{aligned}$$

The expression for the daggered operator is identical, except of course for the general exchange $\hat{A}(\Omega) \rightarrow \hat{A}^\dagger(-\Omega)$, and then a change of C_1 and C_2 to their complex conjugates (a sign change of the pump phase ϕ).

When calculating the correlations in the output field, most terms vanish, since for the input vacuum field \hat{A}_{in} ,

$$\langle \hat{A}_{\text{in}}(\Omega)\hat{A}_{\text{in}}(\Omega') \rangle = \langle \hat{A}_{\text{in}}^\dagger(-\Omega)\hat{A}_{\text{in}}^\dagger(-\Omega') \rangle = \langle \hat{A}_{\text{in}}^\dagger(-\Omega)\hat{A}_{\text{in}}(\Omega') \rangle = 0,$$

while the only non-zero term is the anti-normal ordered

$$\langle \hat{A}_{\text{in}}(\Omega)\hat{A}_{\text{in}}^\dagger(-\Omega') \rangle = \langle [\hat{A}_{\text{in}}(\Omega), \hat{A}_{\text{in}}^\dagger(-\Omega')] + \hat{A}_{\text{in}}^\dagger(-\Omega')\hat{A}_{\text{in}}(\Omega) \rangle = \delta(\Omega + \Omega').$$

The same holds for \hat{A}_{vac} . The correlations in the output become

$$\begin{aligned} \langle \hat{A}(\Omega)\hat{A}(\Omega') \rangle &= \frac{B_1(\Omega)C_1 + B_2(\Omega)C_2}{D(\Omega)D(\Omega')} e^{i\phi} \delta(\Omega + \Omega') \\ &= G_1(\Omega) e^{i\phi} \delta(\Omega + \Omega'), \\ \langle \hat{A}^\dagger(-\Omega)\hat{A}^\dagger(-\Omega') \rangle &= \frac{B_1(\Omega')C_1 + B_2(\Omega')C_2}{D(\Omega)D(\Omega')} e^{-i\phi} \delta(\Omega + \Omega') \\ &= G_1(\Omega) e^{-i\phi} \delta(\Omega + \Omega'), \\ \langle \hat{A}^\dagger(-\Omega)\hat{A}(\Omega') \rangle &= \frac{C_1^2 + C_2^2}{D(\Omega)D(\Omega')} \delta(\Omega + \Omega') \\ &= G_2(\Omega) \delta(\Omega + \Omega'), \\ \langle \hat{A}(\Omega)\hat{A}^\dagger(-\Omega') \rangle &= \frac{B_1(\Omega)B_1(\Omega') + B_2(\Omega)B_2(\Omega')}{D(\Omega)D(\Omega')} \delta(\Omega + \Omega') \\ &= (1 + G_2(\Omega)) \delta(\Omega + \Omega'), \end{aligned}$$

with the real-valued expressions

$$\begin{aligned} G_1(\Omega) &= \frac{2\gamma_1\epsilon(\gamma^2 + \Omega^2 + \epsilon^2)}{D(\Omega)D(-\Omega)}, \\ G_2(\Omega) &= \frac{4\gamma_1\gamma\epsilon^2}{D(\Omega)D(-\Omega)}, \\ D(\Omega)D(-\Omega) &= [(\gamma - \epsilon)^2 + \Omega^2][(\gamma + \epsilon)^2 + \Omega^2], \end{aligned}$$

where Ω' was set equal to Ω , since the delta-function will not allow any other values anyway. From this, the time domain correlation functions can be calculated, using (3.14):

$$\begin{aligned}
\langle \hat{A}(t)\hat{A}(t') \rangle &= \langle \hat{A}^\dagger(t)\hat{A}^\dagger(t') \rangle^* \\
&= \frac{1}{2\pi} \iint_{-\infty}^{\infty} e^{-i(\Omega t + \Omega' t')} \langle \hat{A}(\Omega)\hat{A}(\Omega') \rangle d\Omega d\Omega' \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\Omega(t-t')} G_1(\Omega) d\Omega \\
&= \frac{\gamma_1 \epsilon e^{i\phi}}{2} \left(\frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma-\epsilon} + \frac{e^{-(\gamma+\epsilon)|t-t'|}}{\gamma+\epsilon} \right),
\end{aligned} \tag{3.18}$$

and similarly

$$\begin{aligned}
\langle \hat{A}^\dagger(t)\hat{A}(t') \rangle &= \langle \hat{A}(t)\hat{A}^\dagger(t') \rangle - \delta(t-t') \\
&= \frac{\gamma_1 \epsilon}{2} \left(\frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma-\epsilon} - \frac{e^{-(\gamma+\epsilon)|t-t'|}}{\gamma+\epsilon} \right).
\end{aligned} \tag{3.19}$$

This last expression gives us some additional important information: For $t' = t$, it is just the expectation value of the number operator for the field at time t . Therefore, we have the (time-independent) photon number of the OPO output in the degenerate mode:

$$\langle \hat{n} \rangle = \langle \hat{A}^\dagger(t)\hat{A}(t) \rangle = \frac{\gamma_1 \epsilon^2}{\gamma^2 - \epsilon^2}. \tag{3.20}$$

Notice the units of inverse time; since the field is not a restricted cavity mode, but rather a continuously traveling wave, the photon number is actually the number of photons per unit time (second). Therefore, it is the photon production rate of the OPO.

Quadrature correlation

Without loss of generality, we can set $\phi = 0$. It will simplify the following expressions for the quadrature correlations, and make the covariance matrix diagonal (no $\hat{X}-\hat{P}$ correlation). To keep the notation consistent, we use capital letters for the rotating frame quadrature operators:

$$\begin{aligned}
\langle \hat{X}(t)\hat{X}(t') \rangle &= \langle \hat{A}(t)\hat{A}(t') \rangle + \langle \hat{A}^\dagger(t)\hat{A}(t') \rangle + \frac{1}{2}\delta(t-t') \\
&= \frac{1}{2}\delta(t-t') + \frac{\gamma_1 \epsilon}{\gamma-\epsilon} e^{-(\gamma-\epsilon)|t-t'|}, \\
\langle \hat{P}(t)\hat{P}(t') \rangle &= -\langle \hat{A}(t)\hat{A}(t') \rangle + \langle \hat{A}^\dagger(t)\hat{A}(t') \rangle + \frac{1}{2}\delta(t-t') \\
&= \frac{1}{2}\delta(t-t') - \frac{\gamma_1 \epsilon}{\gamma+\epsilon} e^{-(\gamma+\epsilon)|t-t'|}.
\end{aligned} \tag{3.21}$$

Thus, the output field at a given time is correlated with the same field at earlier and later times within a range of roughly the inverse cavity bandwidth. The

$\delta(t - t')/2$ terms look a little odd, but they are essential to the quantum description of the light field and would not appear in a classical theory. They represent the vacuum fluctuations which have a constant level of $1/2$ in the frequency domain. In the time-domain it becomes a delta function, meaning that these vacuum fluctuations are completely independent and uncorrelated. The delta functions are often gotten rid of by rewriting the operators in normal-ordered form, denoted by $:\cdot:$, where all creation operators are moved to the left of any annihilation operators. In the normal-ordered equivalents of the expressions above, the commutator $[\hat{A}(t), \hat{A}^\dagger(t')]$, which gives rise to the delta-function, does not appear. The relationship between the full correlations above and the normal ordered ones is

$$\langle \hat{X}(t)\hat{X}(t') \rangle = \frac{1}{2}\delta(t - t') + \langle : \hat{X}(t)\hat{X}(t') : \rangle. \quad (3.22)$$

Compare this with the similar relation for the standard single mode quadrature operators:

$$\langle \hat{x}^2 \rangle = \frac{1}{2}\langle (\hat{a}^\dagger)^2 + \hat{a}^2 + \hat{a}^\dagger\hat{a} + ([\hat{a}, \hat{a}^\dagger] + \hat{a}^\dagger\hat{a}) \rangle = \frac{1}{2} + \langle : \hat{x}^2 : \rangle.$$

It is clear that the $1/2$ and $\delta(t - t')/2$ terms originating in the commutator is the vacuum variance, while the normal-ordered correlation is the additional contribution which makes the total variance larger or smaller than for vacuum. In a covariance between two different modes, the commutator becomes zero and leaves no vacuum term – the vacuum states in two modes are uncorrelated. We can write this in a general way, valid for all second-order moments of quadrature operators $\hat{\xi}_j, \hat{\xi}_k$:

$$\langle \hat{\xi}_j\hat{\xi}_k \rangle = [\hat{a}_j, \hat{a}_k^\dagger] + \langle : \hat{\xi}_j\hat{\xi}_k : \rangle,$$

where \hat{a}_j is the mode operator corresponding to the quadrature operator $\hat{\xi}_j$ which might be time- or frequency-dependent. The normal-ordered second-order moments become particularly useful when considering the effect of transformations of the fields on the quadrature correlations. For example, a beamsplitting operation which transforms $\hat{X}(t) \rightarrow \hat{X}'(t) = \sqrt{T}\hat{X}(t) + \sqrt{1-T}\hat{X}_{\text{vac}}(t)$ changes the correlation to

$$\begin{aligned} \langle \hat{X}'(t)\hat{X}'(t') \rangle &= T\langle \hat{X}(t)\hat{X}(t') \rangle + (1-T)\langle \hat{X}_{\text{vac}}(t)\hat{X}_{\text{vac}}(t') \rangle \\ &= T\left(\frac{1}{2}\delta(t - t') + \langle : \hat{X}(t)\hat{X}(t') : \rangle\right) + (1-T)\frac{1}{2}\delta(t - t') \\ &= \frac{1}{2}\delta(t - t') + T\langle : \hat{X}(t)\hat{X}(t') : \rangle. \end{aligned}$$

We see that it is not necessary to keep track of the additional vacuum mode when considering second-order moments. It is enough to transform the normal-ordered moments and add the vacuum term afterwards. If the input to the second port was not vacuum, a term $(1-T)\langle : \hat{X}_2(t)\hat{X}_2(t') : \rangle$ would be added to the above expression.

One reason, why the rotating frame operators are so useful, comes from the way a homodyne detector analyzes the signal field. When the signal is mixed with the local oscillator at ω_0 , the optical sidebands at $\omega_0 \pm \Omega$ are partly converted

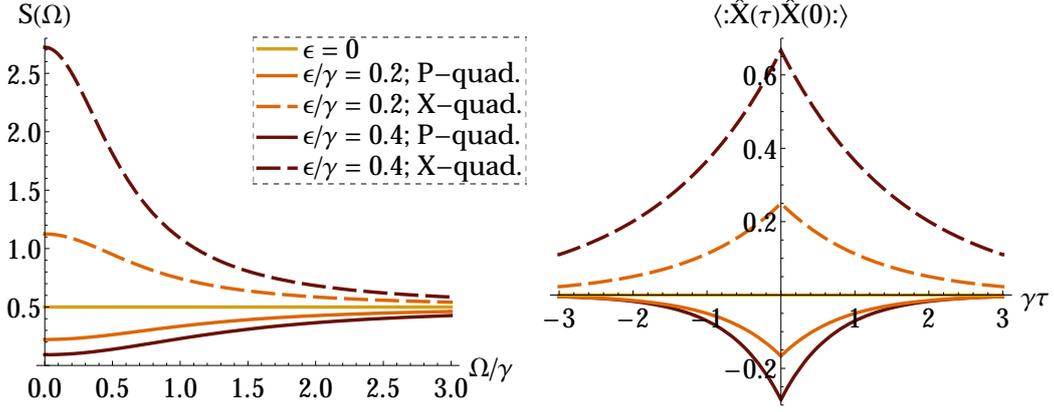


FIGURE 3.9: Squeezing and anti-squeezing spectra $S(\Omega)$ and temporal correlations $\Gamma(\tau)$ of the DOPO output. Here, $\gamma_1 = \gamma$, meaning no intra-cavity losses. $\epsilon = 0$ corresponds to the vacuum state.

into low-frequency components at Ω which are reflected in the photocurrent (as long as Ω is within the detector's electronic bandwidth). Thus, the spectrum of the photocurrent is a direct measurement of the $\hat{X}(\Omega)$ and $\hat{P}(\Omega)$ operators. For example, the spectral density (or power spectrum) of the photocurrent, as measured on a spectrum analyzer, is proportional to the spectral density $S(\Omega)$ of the light field quadrature which is in phase with the LO. This spectral density corresponds to the variance at different frequencies and is defined by [95, §2.4.1]

$$S_x(\Omega)\delta(\Omega + \Omega') \equiv \langle \hat{X}(\Omega)\hat{X}(\Omega') \rangle,$$

or, equivalently, by the Wiener-Khinchine theorem, as the Fourier transform of the auto-correlation function,³

$$S_x(\Omega) = \int_{-\infty}^{\infty} \langle \hat{X}(\tau)\hat{X}(0) \rangle e^{i\Omega\tau} d\tau,$$

where $\langle \hat{X}(\tau)\hat{X}(0) \rangle = \langle \hat{X}(t)\hat{X}(t') \rangle$ for $\tau = t - t'$, since the OPO output field fluctuations are stationary, and hence only depend on the time difference τ , not the specific times t and t' . Either way, the spectral densities of the X and P quadratures for the DOPO output become

$$\begin{aligned} S_x(\Omega) &= \frac{1}{2} + \frac{2\epsilon\gamma_1}{(\gamma - \epsilon)^2 + \Omega^2}, \\ S_p(\Omega) &= \frac{1}{2} - \frac{2\epsilon\gamma_1}{(\gamma + \epsilon)^2 + \Omega^2}. \end{aligned} \quad (3.23)$$

We see that the P quadrature is squeezed below the vacuum level, while X is anti-squeezed. If $\gamma_1 = \gamma$, i.e. no intra-cavity losses, the squeezed state is pure, since $S_x(\Omega)S_p(\Omega) = 1/4$ then. The spectral densities and temporal correlations are illustrated in Figure 3.9.

³There is no factor of $1/\sqrt{2\pi}$ missing in this expression; when using the symmetric Fourier transform for the field operators, Fourier transforms of derived expressions such as the auto-correlation will not necessarily have the same form.

3.3.3 NDOPO

For the non-degenerate OPO, we study the two resonance modes separated symmetrically n free spectral ranges away from the central frequency, that is, at $\omega_{\pm} = \omega_0 \pm n\Delta_{\omega}$, and consider them as separate commuting field modes. We then introduce the rotating frame operators (two different frames):

$$\begin{aligned}\hat{A}_+(\Omega) &= \hat{a}(\omega_+ + \Omega), \\ \hat{A}_+(t) &= \hat{a}(t)e^{i\omega_+ t}, \\ \hat{A}_-(\Omega) &= \hat{a}(\omega_- + \Omega), \\ \hat{A}_-(t) &= \hat{a}(t)e^{i\omega_- t}.\end{aligned}$$

A homodyne detector with local oscillator frequency ω_+ will observe the quadrature operators derived from $\hat{A}_+(\Omega)$, while the fluctuations described by $\hat{A}_-(\Omega)$ can be observed with a local oscillator at ω_- . In most practical settings, the frequency separation $2n\Delta_{\omega}$ between the resonances will by far exceed the detection bandwidth of the homodyne detectors, in which case the detectors will not see the other modes. We use this argument to justify the separation of the field into two distinct modes; it is not strictly correct, since they have a frequency domain commutator

$$\begin{aligned}[\hat{A}_+(\Omega), \hat{A}_-^\dagger(-\Omega')] &= [\hat{a}(\omega_0 + n\Delta_{\omega} + \Omega), \hat{a}^\dagger(\omega_0 - n\Delta_{\omega} - \Omega')] \\ &= \delta(2n\Delta_{\omega} + \Omega + \Omega'),\end{aligned}$$

but it is effectively zero when the detector does not resolve the full range. We will see that the Ω -sideband of the \hat{A}_+ mode is correlated with the $-\Omega$ -sideband of the \hat{A}_- mode, which is no surprise: in a naive photon picture, a photon down-converted at frequency $\omega_+ + \Omega$ will be accompanied by its twin at the symmetric frequency $\omega_- - \Omega$. If the detector bandwidth was wide enough, a homodyne detector with LO at ω_0 would observe squeezing at the $n\Delta_{\omega} + \Omega$ sideband.

The equations of motion for the two cavity fields are

$$\frac{d\hat{A}_{c,\pm}(t)}{dt} = \epsilon e^{i\phi} \hat{A}_{c,\mp}^\dagger(t) - \gamma \hat{A}_{c,\pm}(t) + \sqrt{2\gamma_1} \hat{A}_{\text{in},\pm}(t) + \sqrt{2\gamma_2} \hat{A}_{\text{vac},\pm}(t).$$

The solutions in frequency space, propagated to the output fields $\hat{A}_{\pm}(\Omega)$, are very similar to those for the DOPO in (3.17):

$$\begin{aligned}\hat{A}_{\pm}(\Omega) &= \frac{1}{D(\Omega)} [B_1(\Omega) \hat{A}_{\text{in},\pm}(\Omega) + C_1 e^{i\phi} \hat{A}_{\text{in},\mp}^\dagger(-\Omega) \\ &\quad + B_2(\Omega) \hat{A}_{\text{vac},\pm}(\Omega) + C_2 e^{i\phi} \hat{A}_{\text{vac},\mp}^\dagger(-\Omega)].\end{aligned}$$

Using the assumption mentioned above, that the inter-mode commutator is zero, the correlations in the following output operator combinations vanish:

$$\begin{aligned}\langle \hat{A}_{\pm}(\Omega) \hat{A}_{\pm}(\Omega') \rangle &= \langle \hat{A}_{\pm}^\dagger(-\Omega) \hat{A}_{\pm}^\dagger(-\Omega') \rangle \\ &= \langle \hat{A}_{\pm}^\dagger(-\Omega) \hat{A}_{\mp}(\Omega') \rangle = \langle \hat{A}_{\pm}(\Omega) \hat{A}_{\mp}^\dagger(-\Omega') \rangle = 0.\end{aligned}$$

The quadrature variances all become

$$\begin{aligned} \langle \hat{X}_{\pm}(\Omega) \hat{X}_{\pm}(\Omega') \rangle &= \langle \hat{P}_{\pm}(\Omega) \hat{P}_{\pm}(\Omega') \rangle = \\ &= \frac{1}{2} \frac{C_1^2 + C_2^2 + B_1(\Omega)B_1(-\Omega) + B_2(\Omega)B_2(-\Omega)}{D(\Omega)D(-\Omega)} \delta(\Omega + \Omega') \\ &= \left(\frac{1}{2} + G_2(\Omega) \right) \delta(\Omega + \Omega'), \end{aligned} \quad (3.24)$$

Thus, the fluctuations are larger than for vacuum (when $\epsilon > 0$) and phase-insensitive. That means that each of the two output modes are in thermal states when considered separately. However, there are also correlations between the two modes:

$$\begin{aligned} \langle \hat{X}_{\pm}(\Omega) \hat{X}_{\mp}(\Omega') \rangle &= -\langle \hat{P}_{\pm}(\Omega) \hat{P}_{\mp}(\Omega') \rangle = \\ &= \frac{1}{2} \frac{(B_1(\Omega)C_1 + B_2(\Omega)C_2)(e^{i\phi} + e^{-i\phi})}{D(\Omega)D(-\Omega)} \delta(\Omega + \Omega') \\ &= \cos \phi G_1(\Omega) \delta(\Omega + \Omega'), \end{aligned} \quad (3.25)$$

and

$$\langle \hat{X}_{\pm}(\Omega) \hat{P}_{\mp}(\Omega') \rangle = \langle \hat{P}_{\pm}(\Omega) \hat{X}_{\mp}(\Omega') \rangle = \sin \phi G_1(\Omega) \delta(\Omega + \Omega').$$

The cross-correlations become particularly meaningful when considering the variances of sums and differences of the two modes' quadratures, where it becomes clear that they are in a two-mode squeezed state. For $\phi = 0$,

$$\begin{aligned} \langle (\hat{X}_{+}(\Omega) \pm \hat{X}_{-}(\Omega'))^2 \rangle &= 2\langle \hat{X}_{+}(\Omega) \hat{X}_{+}(\Omega') \rangle \pm 2\langle \hat{X}_{+}(\Omega) \hat{X}_{-}(\Omega') \rangle \\ &= (1 + 2G_2(\Omega) \pm 2G_1(\Omega)) \delta(\Omega + \Omega') \\ &= \left(1 \pm \frac{4\gamma_1\epsilon}{(\gamma \mp \epsilon)^2 + \Omega^2} \right) \delta(\Omega + \Omega'), \end{aligned}$$

and oppositely for the P -quadrature:

$$\langle (\hat{P}_{+}(\Omega) \pm \hat{P}_{-}(\Omega'))^2 \rangle = \left(1 \mp \frac{4\gamma_1\epsilon}{(\gamma \pm \epsilon)^2 + \Omega^2} \right) \delta(\Omega + \Omega').$$

The vacuum level is $\delta(\Omega + \Omega')$, so the X -difference and P -sum are squeezed for $\epsilon > 0$, and in the limit of infinite squeezing, $\epsilon \rightarrow \gamma$, the two modes are entangled in the original EPR-sense [34].

There is a whole range of these pairwise correlated, spatially degenerate⁴, but spectrally separated modes in the OPO output – only limited by the phase matching bandwidth of the down-conversion process, as will be discussed in §4.3.1. Lu and Ou [84] pointed out, that all these non-degenerate modes must be filtered out before reaching a photon detector (APD), if the purpose of the APD is to detect only photons from the degenerate mode. In contrast to a homodyne

⁴In the case of type-I down-conversion they are also polarization degenerate. It is possible to generate entangled beams with orthogonal polarization from type-II down-conversion in OPO's, but that requires some additional experimental effort to keep both polarizations resonant in the cavity - see e.g. [72]

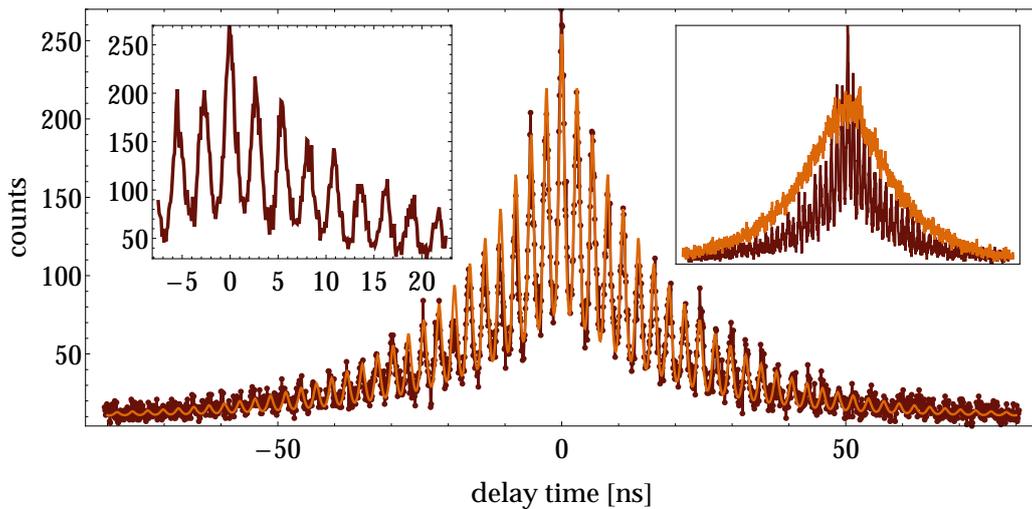


FIGURE 3.10: Intensity correlation function of the unfiltered (and unseparated) OPO output, obtained from time-delay coincidence counting on two APDs. The oscillations due to beating between the different frequency modes are clearly observed, and they are well fitted by the theory of refs. [52, 83, 84]. The left inset is a closeup of the central region. The right inset shows (in orange) the correlations after the implementations of the filters, such that only the degenerate frequency mode reaches the APDs. The unfiltered correlations are shown for reference only - the vertical normalization is not adjusted properly.

detector, a photon detector does basically not discriminate by frequency, so photons from all the different modes will be detected, but most of them will not be correlated with the mode of interest⁵, so in effect they induce “dark counts” in the detector. The existence of the frequency comb of down-converted modes can be inferred from a measurement of the intensity auto-correlation of the unfiltered NDOPO output. Since all modes are born from the same pump field, they have a fixed phase relationship with each other, which leads to a beating effect in the total output intensity very similar to a mode-locked laser. This was discussed by Lu et al. [83], Lu and Ou [84], and the predicted oscillating intensity correlation was observed by Goto et al. [52]. We measured the same kind of oscillations in our OPO output before implementing any trigger filtering. These intensity correlations, acquired by coincidence counts on two APDs, are plotted in Figure 3.10. We clearly see the double-sided exponential envelope and the oscillations in between with a period of exactly the cavity round-trip time $\tau_{\text{cav}} = 2.7$ ns. This can be understood if considering that two entangled photons are born simultaneously, while they will not necessarily leave the OPO at the same time. Since they can only leave the cavity at a certain time of each round trip, the coincidence counts will appear in peaks with the spacing τ_{cav} . The pattern is partly blurred by the finite resolution t_{res} of the APDs. The fitted curve

⁵For single photon state generation, this will be one of the non-degenerate modes rather than the degenerate.

in the figure is given by

$$g(\tau) = C_1 \left[C_2 + e^{-2\gamma|\tau|} \sum_{n=-\infty}^{\infty} \left(1 + \frac{2|\tau - n\tau_{\text{cav}}| \ln 2}{t_{\text{res}}} \right) \times \exp \left(-\frac{2|\tau - n\tau_{\text{cav}}| \ln 2}{t_{\text{res}}} \right) \right], \quad (3.26)$$

which is a slightly rough model obtained in ref. [52] in the limit of low gain, $\epsilon \ll \gamma$. C_1 and C_2 are intensity-dependent scaling factors. The best fit to the data is achieved for $t_{\text{res}} = 0.55$ ns, which is a bit higher than the APD maker's specification of 0.35 ns. After implementing the filtering (see §4.3.2), the peaks disappear, and only a smoothed double-exponential is left, as seen in the inset of Figure 3.10.

Getting back to the formal description, the quadrature temporal correlations, which will be used for the conditioned state analysis, are obtained from (3.24) and (3.25) with $\phi = 0$:

$$\begin{aligned} \langle \hat{X}_{\pm}(t) \hat{X}_{\pm}(t') \rangle &= \langle \hat{P}_{\pm}(t) \hat{P}_{\pm}(t') \rangle \\ &= \frac{1}{2} \delta(t - t') + \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\Omega(t-t')} G_2(\Omega) d\Omega \\ &= \frac{1}{2} \delta(t - t') + \frac{\gamma_1 \epsilon}{2} \left(\frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma - \epsilon} - \frac{e^{-(\gamma+\epsilon)|t-t'|}}{\gamma + \epsilon} \right), \end{aligned} \quad (3.27)$$

$$\begin{aligned} \langle \hat{X}_{\pm}(t) \hat{X}_{\mp}(t') \rangle &= -\langle \hat{P}_{\pm}(t) \hat{P}_{\mp}(t') \rangle \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\Omega(t-t')} G_1(\Omega) d\Omega \\ &= \frac{\gamma_1 \epsilon}{2} \left(\frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma - \epsilon} + \frac{e^{-(\gamma+\epsilon)|t-t'|}}{\gamma + \epsilon} \right). \end{aligned} \quad (3.28)$$

With the continuous-time correlation functions in place for the DOPO and ND-OPO, the next step is to reduce to two modes, which will turn these functions into simple number-valued entries in the correlation matrix that represents the observed trigger and signal modes. We will see that there is a freedom in the definition of the signal mode, which should be used to optimize the non-Gaussian character of the conditioned signal state.

3.4 Photon-subtracted state generation

The following is based on the scheme for producing kitten states from an initial squeezed vacuum state, that is, the output from the DOPO. The single photon state generation using the NDOPO is briefly discussed in the end.

3.4.1 Temporal mode selection

The continuous wave squeezed vacuum generated by a standard OPO is a stationary field in the sense mentioned before, that the two-time correlation functions only depend on the time difference, and not the absolute time. It makes perfect sense to describe such a continuous field in the frequency domain, and to assign a quantum state to individual sidebands. In experiments on cw squeezed light generation and related cw protocols such as EPR-entanglement and teleportation, measurements are usually performed at a given narrow sideband well within the squeezing bandwidth, but above the possibly noisy low frequencies. If no other parameters of the system change, the state defined at a sideband does not change with time. However, when we introduce a time-dependent element to the state generation, such as a photon detector which clicks at discrete times, it is no longer entirely useful to consider the field only at individual sidebands. In the kitten experiment, the detection of a trigger photon at time t_c obviously alters the correlated signal field in the vicinity of this moment of time. The field is then no longer stationary and a single sideband is not sufficient to describe the state of the field at the interesting time around t_c . Instead of a formulation in terms of sidebands, we need to utilize the entire spectrum. Alternatively – and preferably – we can work in the time domain which is more intuitive.

The OPO output field correlations in the last section are defined for a continuum of temporal modes. We will need to assign a single temporal mode, which our conditioned quantum state occupies. That is done by selecting a mode function to integrate over, i.e. we perform a temporal filtering. In sideband measurements, the temporal mode is implicitly chosen as the transform of the spectral shape of the sideband. In pulsed experiments, the pulse itself is the temporal mode. We define the filtered field mode operator \hat{A}_s as

$$\hat{A}_s = \int_{-\infty}^{\infty} f_s(t) \hat{A}_s(t) dt.$$

$\hat{A}_s(t)$ is the mode operator describing the signal field before the photon detection, i.e. the OPO output field after beamsplitting into trigger and signal mode, and other subsequent transformations (primarily additional losses). $f_s(t)$ is the (amplitude) filter function corresponding to the chosen temporal mode. To preserve commutation relations, the filter function must satisfy $\int |f_s(t)|^2 dt = 1$. In a homodyne detection of the conditioned signal state, the temporal selection can be done by shaping the local oscillator amplitude according to $f_s(t)$. When the LO is off, the homodyne detector measures nothing. A much easier alternative is to record the detector output continuously, and subsequently filter the sampled noise electronically. There is a large freedom in choosing an appropriate temporal mode, but it is of course sensible to choose one that optimizes the state with respect to a certain characteristic, such as the fidelity or Wigner function dip size – this is the temporal equivalent of spatial mode-matching. The importance of such temporal mode-matching was pointed out by Ou and Kimble [116].

The optimal choice of a filter function for the signal mode is not obvious. Intuitively, though, it should be similar to the double-sided exponentially decaying correlation functions of the OPO output: When the trigger APD clicks, it signals

the subtraction of a photon in the part of the signal mode which is correlated with the trigger mode that gives the click. The photon-subtracted state is distributed over the time of the correlations, which is both in the past and in the future relative to the time of the trigger click. If we choose a too narrow filter function, not all of the PSSqV state is included in the mode; if it is too wide, some of the surrounding non-photon subtracted squeezed vacuum is included, decreasing the purity of the state. For low gains and no trigger filtering, Nielsen and Mølmer [110] showed that the double-sided exponential function with same decay constant as the OPO,

$$f_s(t) = \sqrt{\gamma}e^{-\gamma|t|},$$

is in fact the optimal choice. For larger gains, the picture becomes more complicated – they found some optimal mode functions using numerical methods, but it does not seem possible to find closed analytical expressions.

A temporal mode $f_t(t)$ should also be defined for the trigger. In contrast to the homodyne detector, which measures the signal beam continuously, the trigger APD detector is highly time selective, so the mode of the trigger is defined by the detector itself. The timing resolution of commercial APDs is quite good (~ 350 ps specified for the model in our setup). The extent of the correlation functions are several orders of magnitude longer (~ 100 ns in our setup), so the shape of the trigger mode is not as important as the exact time of the click. It can be modeled as a delta function or a narrow normalized top-hat function centred on t_c . For both the trigger and signal modes, we will consider only real mode functions $f_t(t)$ and $f_s(t)$, which makes sense since the correlation functions (3.18), (3.19) do not have a time-dependent phase factor. With real mode functions, the quadrature operators are filtered in the same way as the mode operators:

$$\hat{X}_s = \int_{-\infty}^{\infty} f_s(t)\hat{X}_s(t)dt,$$

and equivalently for the trigger mode and the P quadrature. The filtering of the fields is a linear operation, so it does not change their Gaussian character. The covariance matrix built from $\{\hat{X}_s, \hat{P}_s, \hat{X}_t, \hat{P}_t\}$ will then describe the two-mode state within the given temporal modes just before the trigger detection. One can take advantage of the normal-ordering of the correlations when calculating the entries in the covariance matrix. For example, for the Γ_{11} entry:

$$\begin{aligned} \Gamma_{11} &= 2\langle(\hat{X}_s)^2\rangle = 2\left\langle\int f_s(t)\hat{X}_s(t)dt\int f_s(t')\hat{X}_s(t')dt'\right\rangle \\ &= 2\iint f_s(t)f_s(t')\left(\frac{1}{2}\delta(t-t') + \langle:\hat{X}_s(t)\hat{X}_s(t'):\rangle\right)dtdt' \\ &= \int f_s(t)^2dt + 2\iint f_s(t)f_s(t')\langle:\hat{X}_s(t)\hat{X}_s(t'):\rangledtdt' \\ &= 1 + 2\iint f_s(t)f_s(t')\langle:\hat{X}_s(t)\hat{X}_s(t'):\rangledtdt'. \end{aligned}$$

A covariance term like Γ_{24} is obtained by filtering once with each of the mode functions:

$$\Gamma_{24} = \langle\{\hat{P}_s, \hat{P}_t\}\rangle = \iint f_s(t)f_t(t')\langle:\{\hat{P}_s(t), \hat{P}_t(t')\}:\rangledtdt'.$$

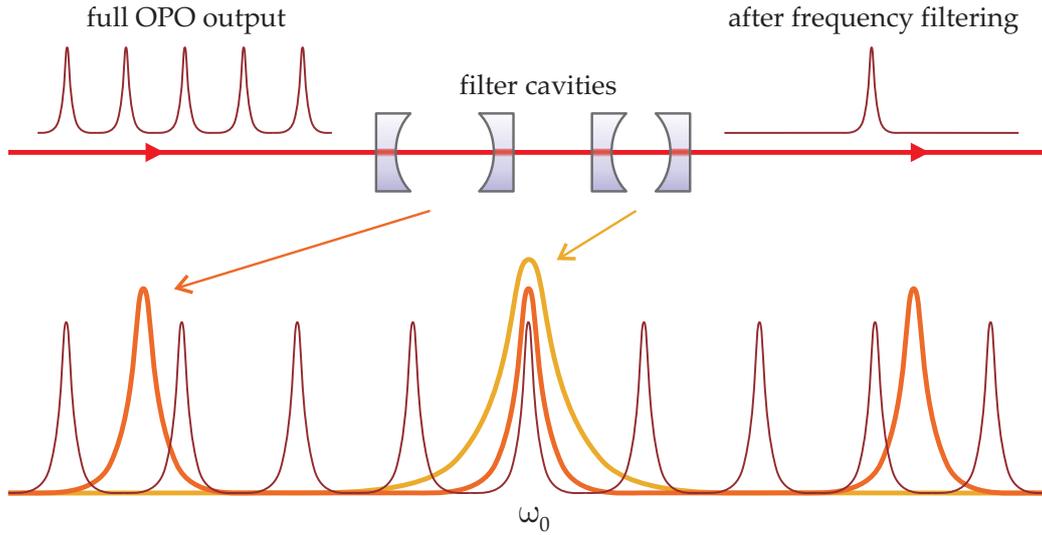


FIGURE 3.11: Top: The frequency filtering of the trigger mode using empty optical resonators. Bottom: The spectral structure of the OPO output (brown) and the two filtering cavities. The filters have different free spectral ranges in order to filter out more non-degenerate modes.

3.4.2 Frequency filtering of trigger mode

As mentioned previously, photons arriving at the trigger detector, that are uncorrelated with the signal mode, will effectively be dark counts, leading to a non-photon subtracted state in the conditioned signal mode. Such are the photons down-converted into all the non-degenerate frequency modes of the OPO. In order that only photons from the degenerate mode reaches the APD, all other modes should be filtered out. This can be done by a serial arrangement of several empty cavities that have different free spectral ranges and are all resonant on the central frequency. Then, frequencies that are not simultaneously resonant in all the cavities will be filtered out. Hopefully, the degenerate mode is the only one meeting the resonance criteria. The filtering scheme is illustrated in Figure 3.11. The filtering cavities also alter the degenerate mode in the trigger arm, convoluting the field with the temporal response of the cavities. This in turn alters the correlation functions of the two-mode field, which may change the optimal temporal mode function for the signal mode.

In order to keep the following calculations reasonably simple, we assume that there is only one cavity filter, or at least that one cavity dominates the filtering by having a much narrower bandwidth than the others. In our experiment this is not really the case, since two cavities have the same bandwidth while the third is 3 times as wide. The combined action of the filters can be approximated by a single cavity with a narrower bandwidth, though. Assuming this cavity to have symmetric input and output coupler, each with a decay rate of $\kappa/2$, giving a total decay rate of κ , the response function of the cavity is $\kappa e^{-\kappa t}$, so the field $\hat{A}(t)$ entering the cavity will be transformed according to the convolution [99]

$$\hat{A}(t) \rightarrow \hat{A}'(t) = \kappa \int_{-\infty}^t \hat{A}(y) e^{-\kappa(t-y)} dy + v(t) \hat{A}_{\text{vac}}(t), \quad (3.29)$$

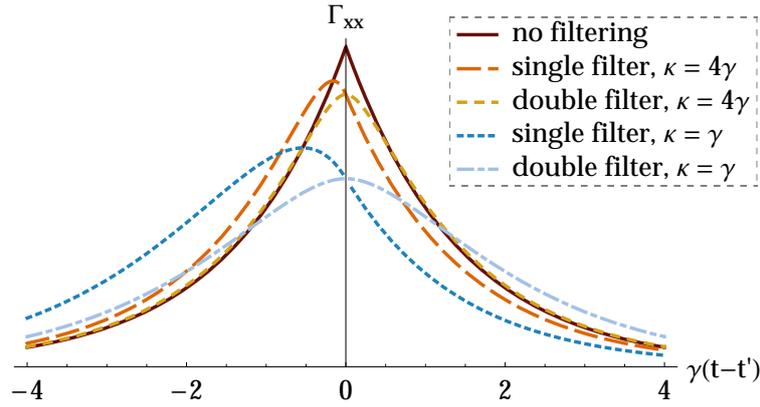


FIGURE 3.12: The normal-ordered correlation function $\Gamma_{xx}(t-t') \equiv \langle : \hat{X}(t)\hat{X}(t') : \rangle$ where either none (brown), one (orange, dark blue), or two (yellow, light blue) of the \hat{X} operators has been filtered through a cavity. The narrower the filter cavity bandwidth κ is, the more the correlation functions become smoothed out.

where the last term represents the vacuum component added wherever the incoming field is filtered out (the cavity essentially acts as a frequency-dependent beamsplitter). The details of this vacuum term are irrelevant, since it is not needed in the normal-ordered formulation. The convolution stops at time t due to causality – the output field at t cannot depend on the input field at later times. The filtering is a real transformation (at least when the cavity is not detuned from resonance), so the quadrature operators transform in the same way. The covariance matrix after the filtering is obtained in a similar manner as we saw for the temporal mode selection before: The normal-ordered correlation functions involving the trigger mode are convoluted once or twice with the cavity response function, and the vacuum state covariance matrix with diagonal entries $\delta(t-t')$ is added.

We can calculate explicit expressions for the normal ordered DOPO output correlation functions as given by (3.21) and (3.22) when filtered either once or twice. The results for the X quadrature are, for a single filtering (again with the prime on \hat{X} denoting a filtered field):

$$\begin{aligned} \langle : \hat{X}(t)\hat{X}'(t') : \rangle &= \kappa \int_{-\infty}^{t'} \langle : \hat{X}(t)\hat{X}(y') : \rangle e^{-\kappa(t'-y')} dy' \\ &= \begin{cases} \frac{\kappa\gamma_1\epsilon}{(\gamma-\epsilon)(\gamma-\epsilon+\kappa)} e^{-(\gamma-\epsilon)(t-t')} & , t \geq t' \\ \frac{\kappa\gamma_1\epsilon}{\gamma-\epsilon-\kappa} \left(\frac{2e^{\kappa(t-t')}}{\gamma-\epsilon+\kappa} - \frac{e^{(\gamma-\epsilon)(t-t')}}{\gamma-\epsilon} \right) & , t < t', \end{cases} \end{aligned} \quad (3.30)$$

and for the double filtering:

$$\begin{aligned} \langle : \hat{X}'(t)\hat{X}'(t') : \rangle &= \kappa^2 \int_{-\infty}^t \int_{-\infty}^{t'} \langle : \hat{X}'(y)\hat{X}'(y') : \rangle e^{-\kappa(t+t'-y-y')} dy dy' \\ &= \frac{\kappa^2\gamma_1\epsilon}{(\gamma-\epsilon)^2 - \kappa^2} \left(\frac{e^{-\kappa|t-t'|}}{\kappa} - \frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma-\epsilon} \right). \end{aligned} \quad (3.31)$$

The results for the P quadrature are similar; the functions get a sign change, and all occurrences of $\gamma - \epsilon$ should be exchanged with $\gamma + \epsilon$. The outcome of the filtering is to smoothen out the otherwise sharply peaked exponential correlation function, as illustrated in Figure 3.12. The singly filtered correlation is particularly interesting: Since the cavity delays the field that is filtered through, the correlations are shifted to earlier times. That is, when a photon is detected in the filtered trigger arm, most of the correlated part of the signal arm is already in the past. With no filtering of the trigger, the correlations are equally in the past and in the future.

3.4.3 Plugging in

The kitten state generation experiment consists of the following main steps:

1. Generating squeezed vacuum in the OPO.
2. Tapping off a part of the squeezed vacuum on a beamsplitter with high transmittance.
3. Frequency filtering the tapped-off part, the trigger mode, through a series of cavities.
4. Detecting trigger photons by a photon counter.
5. Measuring a quadrature of the signal field on a homodyne detector.
6. Temporally filtering the homodyne current to select the part of the signal field that is correlated with the detected trigger mode.

The measurement is then repeated many times at different local oscillator phases to build statistics about the quadrature distributions of the conditioned signal field. The last two steps are applicable for the state characterization we have performed so far. For application of the kitten state in subsequent quantum information protocols, it should of course not be measured (destructively) at once. The theoretical modeling of the photon-subtracted state in the given temporal mode is however the same. In the model, step 6 above comes before the quadrature measurement in step 5 – we imagine then that the homodyne detector measures exactly one quadrature value of the field within the selected mode function.

The theoretical prediction of the outcome of the state generation is obtained from the final covariance matrix by the method described in §3.2. We summarize the preceding sections by following the transformation of the covariance matrix through the outlined steps above. As we have seen, it suffices to transform the normal-ordered covariance matrix and add the vacuum matrix in the end. Initially, the covariance matrix is a function of time, or rather of the two-time difference $t - t'$.

1. Output from OPO The initial state of the trigger mode is vacuum, so most of the normal-ordered covariance matrix will be zero, except the signal mode sub-matrix which has the diagonal entries given by (3.21) and (3.22):

$$\begin{aligned} : \Gamma_{xx}^i(t-t') : &= 2 \langle : \hat{X}(t) \hat{X}(t') : \rangle, \\ : \Gamma_{pp}^i(t-t') : &= 2 \langle : \hat{P}(t) \hat{P}(t') : \rangle, \end{aligned}$$

with the index i for initial variance. The initial matrix is thus

$$: \Gamma^i(t-t') : = \begin{pmatrix} : \Gamma_{xx}^i(t-t') : & 0 & 0 & 0 \\ 0 & : \Gamma_{pp}^i(t-t') : & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

2. After beamsplitter The beamsplitter transforms the covariance matrix to $: \Gamma^{\text{bs}}(t-t') : = B : \Gamma^i(t-t') : B^T$ with the beamsplitter matrix (3.12). So far, the matrix still has a simple expression, so we can write it out (skipping the $(t-t')$ parts to make room):

$$: \Gamma^{\text{bs}}(t-t') : = \begin{pmatrix} T : \Gamma_{xx}^i & 0 & -\sqrt{T(1-T)} : \Gamma_{xx}^i & 0 \\ 0 & T : \Gamma_{pp}^i & 0 & -\sqrt{T(1-T)} : \Gamma_{pp}^i \\ -\sqrt{T(1-T)} : \Gamma_{xx}^i & 0 & (1-T) : \Gamma_{xx}^i & 0 \\ 0 & -\sqrt{T(1-T)} : \Gamma_{pp}^i & 0 & (1-T) : \Gamma_{pp}^i \end{pmatrix}$$

3. Trigger filtering The frequency filtering of the trigger mode does not affect the signal mode sub-matrix (the top left 2x2 block). The trigger mode sub-matrix is filtered twice, while the covariances are filtered only once. We denote the trigger filtered covariance matrix $\Gamma^{\text{tf}}(t-t')$:

$$: \Gamma_{mn}^{\text{tf}}(t-t') : = \begin{cases} : \Gamma_{mn}^{\text{bs}}(t-t') : & , ss \\ \kappa^2 \int_{-\infty}^t \int_{-\infty}^{t'} : \Gamma_{mn}^{\text{bs}}(y-y') : e^{-\kappa(t+t'-y-y')} dy dy' & , tt \\ \kappa \int_{-\infty}^{t'} : \Gamma_{mn}^{\text{bs}}(t-y) : e^{-\kappa(t'-y)} dy & , st, \end{cases}$$

where ss means the signal variances with $mn \in \{11, 22\}$, tt the trigger variances with $mn \in \{33, 44\}$, and st the covariances with $mn \in \{13, 24, 31, 42\}$. The explicit expressions are obtained trivially using (3.30) and (3.31).

4+6. Temporal mode selection The two-time correlation matrix $: \Gamma^{\text{tf}}(t-t') :$ is turned into the temporal mode selected matrix $: \Gamma^{\text{tms}} :$ by carrying out the filtering with the chosen mode functions $f_s(t)$ and $f_t(t)$, as described in §3.4.1.

Final steps $: \Gamma^{\text{tms}} :$ could now be the final correlation matrix describing the two-mode state just before a trigger click. We are still missing one thing, though – the losses of the trigger, represented by the efficiency parameter η_t , and of the

signal, η_s . As noted before, these can be easily implemented by the diagonal efficiency matrix $\underline{\eta} = \text{diag}(\sqrt{\eta_s}, \sqrt{\eta_s}, \sqrt{\eta_t}, \sqrt{\eta_t})$, so that the final normal-ordered covariance matrix is

$$:\Gamma^{\text{final}}: = \underline{\eta} : \Gamma^{\text{tms}} : \underline{\eta}.$$

The OPO output coupler decay rate γ_1 actually just acts as an additional loss if it is less than γ , the full decay rate. γ_1/γ is the proportion of the cavity field escaping into the desired output mode – this ratio is called the escape efficiency η_{esc} of the OPO. This efficiency term is common for the trigger and signal modes, and just to slightly clean up the theory it makes sense to get rid of the γ_1 symbol by exchanging it with γ and incorporating η_{esc} into η_t and η_s .

To get the full covariance matrix, the vacuum noise given by the identity matrix is added to the normal-ordered matrix:

$$\Gamma^{\text{final}} = \underline{I} + :\Gamma^{\text{final}}:.$$

The signal state conditioned on a trigger click is then obtained by plugging the entries of Γ^{final} into the formulas (3.5)–(3.8). The effect of dark counts, represented by the modal purity Ξ , can be included by mixing an amount Ξ of the conditioned state with an amount $1 - \Xi$ of the unconditioned state (squeezed vacuum) given by (3.10) with $a = \Gamma_{11}^{\text{final}}$, $b = \Gamma_{22}^{\text{final}}$.

3.4.4 Discussion

The model developed here gives different predictions from those of the simple single mode model in the beginning of the chapter. This is no surprise, since the squeezed vacuum state produced in the OPO with its frequency-dependent squeezing and anti-squeezing is somewhat more complicated than the “text-book” single mode squeezed state. We will take a look at the dependence of the generated state on the various parameters, and make comparisons with the single mode theory.

We must choose some temporal mode functions for the trigger and signal modes. The natural choice for the trigger is a delta-function centred at $t_c = 0$. For ease of calculation, we choose for the signal the double-exponential function $f_s(t) = \sqrt{\gamma_s} e^{-\gamma_s |t|}$, where we let the decay constant γ_s be yet another free parameter of the model, to allow for at least a rudimentary optimization for high gains. Following the model recipe given before and using these mode functions, it is possible (with the aid of Mathematica) to achieve analytic expressions for the 6 different entries in the final covariance matrix, and hence for the conditioned and unconditioned Wigner functions. The expressions are not particularly illuminating, but for completeness they are listed in Appendix §A.5. The analytical expressions make simulations with varying parameters very fast in Mathematica.

The first important thing to note is, that the unconditioned squeezed vacuum is not a minimum uncertainty state, as is case in the single mode theory, where the squeezing and anti-squeezing levels are just $e^{\mp 2r}/2$. The reason is the temporal mode selection, which mixes different frequency components into the final state,

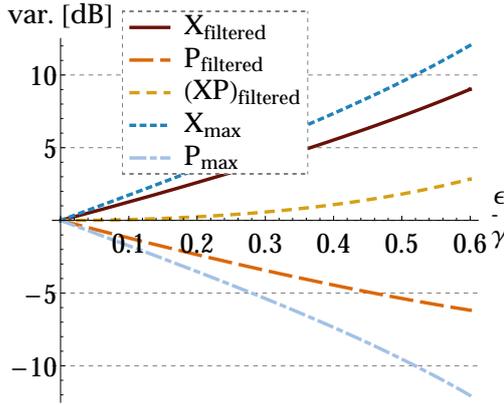


FIGURE 3.13: Squeezing (P) and anti-squeezing (X) variances as a function of ϵ . The levels are given on a logarithmic dB scale relative to the vacuum level. Blue curves are at DC ($\Omega = 0$), while the brown and orange curves are for the temporally filtered mode. The yellow curve is the product of the X and P variances relative to the minimum uncertainty level of $1/4$. T and η_s are set to 1, while the mode function used for filtering has $\gamma_s = \gamma$.

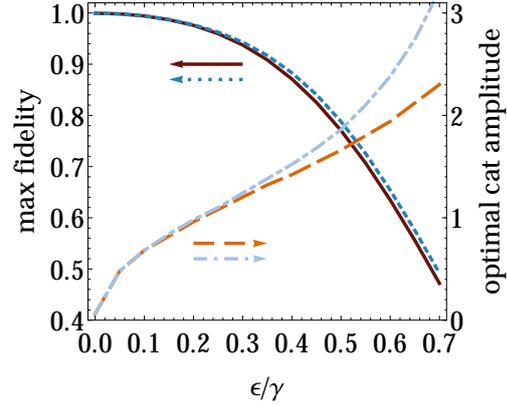


FIGURE 3.14: Fidelity between the kitten states and real odd cat states, and the amplitude of the cat which attains the optimal fidelity for a given ϵ . Efficiencies η_s and η_t are set to 1, $\kappa = \infty$ (no trigger filtering), and $T = 99.9\%$. For the lower curves (brown and orange), the mode function has $\gamma_s = \gamma$, while for the upper ones (blue), γ_s has been optimized dependent on ϵ .

such that the squeezing and anti-squeezing spectra in Figure 3.9 are averaged somewhat out. Figure 3.13 shows the effect. The squeezing and anti-squeezing in the temporally filtered state are compared to the DC-component of the squeezed state ($\Omega = 0$) as a function of ϵ . Both squeezing and anti-squeezing decrease when the field is filtered, which is clear since the higher frequency sidebands are closer to the vacuum state. However, the squeezing decreases more rapidly than the anti-squeezing with increasing pump strength (ϵ), which means that their uncertainty product becomes larger than the Heisenberg limit of $1/4$. One could say that the state has become thermalized. From this observation, we already get a hint that the quality of the photon-subtracted state will degrade for higher pump levels (since a photon-subtracted thermal state is still a Gaussian thermal state).

We see this degradation in the following figures. Figure 3.14 shows the obtainable fidelities between the generated states and the true odd cat states, as well as the corresponding amplitudes of the cat states that optimize the fidelities. This figure can be compared with Figure 3.5 for the single mode theory, although the squeezing parameter r and the pump parameter ϵ can not be directly related to each other. The two sets of curves represent different choices of the mode function width γ_s . For the brown/orange curves it is set equal to γ , while for the blue curves, γ_s has been optimized for each ϵ -value.⁶ The cat amplitudes

⁶The optimization is done with respect to the Wigner function depth, rather than the fidelity, since a combined optimization of γ_s and α is a very slow process. A test run shows that fidelity-aimed optimization only gives about 1% higher fidelity compared to the Wigner depth-aimed optimization.

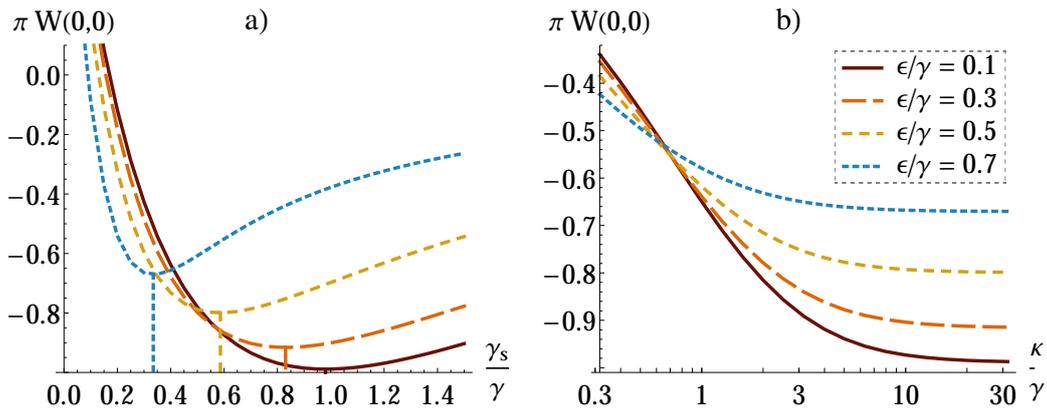


FIGURE 3.15: Wigner function values in the origin as a function of **a)** the mode function width γ_s , and **b)** the trigger filter cavity bandwidth κ , for 4 different pump levels (the legend is for both graphs). The other parameter values are $T = 99.9\%$, $\eta_s = \eta_t = 1$. In **a)**, $\kappa = \infty$, while in **b)**, γ_s is optimized for lowest $W(0,0)$ -value.

obtainable with a reasonable fidelity are more limited than for the single mode theory. For example, whereas the latter can achieve an $\alpha = 2$ cat state with a fidelity of 85%, only 75% is possible with the current model.

Figure 3.15a) and b) shows the influences on the Wigner depth of γ_s and the trigger filter bandwidth κ , respectively. It is clear from a) that the mode function should be chosen carefully; that the state is degraded if the mode is too wide or too narrow. It is also interesting that the optimal mode function width (marked by the vertical lines) decreases with increasing ϵ . An explanation for this is the decrease of the anti-squeezing bandwidth with higher pump levels – from (3.23) it is $\gamma - \epsilon$, while the squeezed quadrature’s bandwidth is $\gamma + \epsilon$. Since the amount of quadrature phase angles with squeezing is less than the amount with anti-squeezing, the overall bandwidth of the state will decrease with increasing ϵ , although not as fast as $\gamma - \epsilon$. It is also obvious from this graph, that higher pump levels decrease the Wigner function depth. This was not the case in the single mode theory, where the cat-fidelity decreased but the Wigner depth remained essentially $-1/\pi$ (for the $\sim 100\%$ T used here). One might in fact argue that the disparity between the bandwidths of the squeezed and anti-squeezed quadratures is the reason for the state degradation; the spectrum of the state is “pulled apart”, and it becomes impossible for a single mode function to encompass fully the aspects of all the quadratures. Another way of understanding the degradation for high pump levels is, that it is caused by the increasing density of temporally neighbouring, uncorrelated “wave packet tents” containing squeezed vacuum. It is perhaps not quite correct to imagine the continuous squeezed OPO output as a series of “tents” scattered randomly along the time axis, but it is a convenient intuitive picture – see Figure 3.16.

While the mode function is an easy subject of optimization in the experiment, and the pump level can be set at will, the trigger filter bandwidth is a more fixed parameter once the cavity has been set down in the setup. It is still illustrating, though, to see its effect on the state, as in Figure 3.15b). There, the mode func-

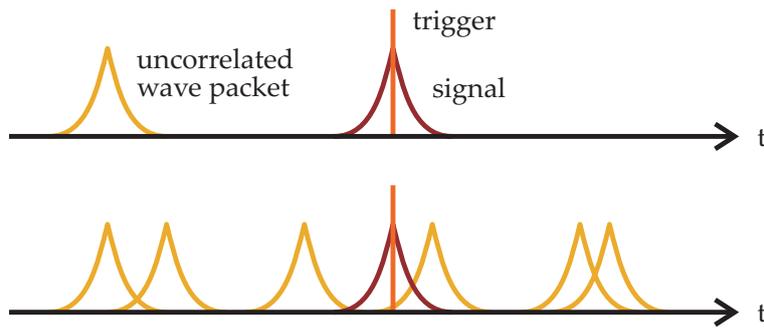


FIGURE 3.16: The stylized picture of the OPO output as consisting of a series of uncorrelated wave packets, each with a length of the inverse OPO bandwidth. For low pump levels (top), the “tents” are scattered far apart, so a state conditioned on the subtraction of a trigger photon can be quite pure. For stronger pumps (bottom), the spontaneous down-conversion happens more frequently, and the wave packets start to have a significant overlap, which reduces the purity of the conditioned states.

tion width is optimized in all points of the graph. For a filter bandwidth that is narrower or a few factors wider than the OPO bandwidth, the state quality is clearly reduced. The reason for that is the smoothing of the temporal correlation functions caused by the filter cavity, which makes the sharp exponential mode function a bad choice. Thus, most of the degradation can be avoided by choosing a more suitable (and complicated) mode function. For the remaining parameters of the model, T , η_s , η_t , Ξ , their effect on the state is straight-forward: when they decrease, the state degrades. For η_t , the degradation is only slight, while for the other three it is severe. η_s - or Ξ - values of roughly 50% leads to the disappearance of the negative dip in the Wigner function. These two parameters, the signal efficiency and the modal purity, are the most important ones to optimize in the experiment – they simply need to be as high as possible. For T there is a trade-off, since higher values close to 1 leads to lower production rates. The value of the optimal pump level, ϵ , is of course also a trade-off between high squeezing/kitten amplitude, and high fidelity and Wigner function depth.

3.4.5 Single photon states from NDOPO output

The scheme for producing single photons from the non-degenerate OPO is very similar to the standard SPDC (pulsed) photon sources: Pairs of photons are generated by down-conversion into different modes, and detection of a photon in one mode heralds the existence of its twin photon in the other mode. The photon pairs are produced at random times, and for such probabilistic sources, there is always a small risk that two pairs appear instead of one. The result is that the heralded state sometimes will contain two photons instead of one, which is detrimental to protocols for secure quantum communication. To keep the two-photon contribution down, the pump power must be kept low, which in turn limits the rate of the photon source. For very low pump levels, the NDOPO single photon source can be understood in the same way; photon pairs are produced at random times, but with a long average interval between successive pairs. The detection of a photon in, say, the ω_- mode then heralds the existence of a single photon

within the correlated part of the ω_+ mode. The resulting state is very simple in comparison to the kitten state; it is rotationally invariant in phase space, and the density matrix consist almost entirely of the 1-photon and vacuum entries (vacuum due to the inevitable losses), plus perhaps a small 2- and higher photon number contributions. It would be easy to develop a simple theory to describe that experiment, but the framework developed in the previous sections for the DOPO works equally well for the NDOPO, so there is no reason not to use that.

In contrast to the DOPO scheme in §3.4.3, the output of the NDOPO is already in a two-mode correlated state. The tapping beamsplitter is not part of the single photon scheme. Instead, the two frequency modes need to be spatially separated, since they are co-propagating out of the OPO. They are also polarization degenerate, so it is not possible to use a polarizing beam splitter, but they can be separated based on their frequency difference. By letting the first of the frequency filter cavities in the trigger arm be resonant on the ω_- frequency, the ω_- mode will be transmitted towards the APDs, while the ω_+ mode is reflected, as long as it is not resonant as well. For the theoretical description, this means that the ω_- mode, as the trigger, is frequency filtered just the same as the trigger mode in the kitten scheme, while nothing happens to the ω_+ mode, which is directed towards the homodyne detector. Thus, the trigger filtering, temporal mode selection of signal, and the final steps in the recipe of §3.4.3 apply equally well here. However, instead of having the $\Gamma^{\text{bs}}(t - t')$ as the preliminary covariance matrix before these steps, it is now the matrix that represents the two-mode squeezed state directly from the NDOPO:

$$:\Gamma^{\text{i,nd}}(t - t') := \begin{pmatrix} :\Gamma_a^{\text{i,nd}}: & 0 & :\Gamma_c^{\text{i,nd}}: & 0 \\ 0 & :\Gamma_a^{\text{i,nd}}: & 0 & -:\Gamma_c^{\text{i,nd}}: \\ :\Gamma_c^{\text{i,nd}}: & 0 & :\Gamma_a^{\text{i,nd}}: & 0 \\ 0 & -:\Gamma_c^{\text{i,nd}}: & 0 & :\Gamma_a^{\text{i,nd}}: \end{pmatrix}.$$

For lack of space, the $(t - t')$ parts were again left out. The auto- and cross-correlations are

$$\begin{aligned} :\Gamma_a^{\text{i,nd}}: &\equiv :\Gamma_a^{\text{i,nd}}(t - t') := 2\langle :\hat{X}_\pm(t)\hat{X}_\pm(t') : \rangle = 2\langle :\hat{P}_\pm(t)\hat{P}_\pm(t') : \rangle, \\ :\Gamma_c^{\text{i,nd}}: &\equiv :\Gamma_c^{\text{i,nd}}(t - t') := 2\langle :\hat{X}_\pm(t)\hat{X}_\mp(t') : \rangle = -2\langle :\hat{P}_\pm(t)\hat{P}_\mp(t') : \rangle, \end{aligned}$$

as given in (3.27)-(3.28).

While for the DOPO scheme it is appropriate to consider the state conditioned on a trigger click as a photon-subtracted state, this is not really the case for the NDOPO scheme. The photon that gives a click in the trigger APD was never a part of the thermal state of the signal mode, so the signal field has not lost a photon. Rather, one can say that it has lost its vacuum component; if there was a photon in the trigger mode, it is apparent that there must also be photons – whether one or more – in the correlated part of the signal mode. Thus, the conditioned state is basically the renormalized thermal state without its vacuum part. We can call it a vacuum-subtracted thermal state, especially if the contents of higher photon numbers do not warrant the name single photon state. The photon number distributions are charted in Figure 3.17, where the

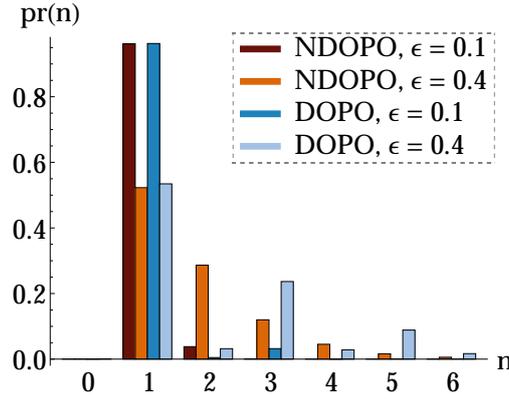


FIGURE 3.17: Photon number distribution for the vacuum-subtracted thermal state (NDOPO) and the kitten state (DOPO) for a low and a high(er) pump level. The other parameters are set to $\gamma_s = \gamma$, $\kappa = \infty$, $T = 99.9\%$, $\eta_s = \eta_t = 1$. Note that the DOPO output (kitten state) has oscillating numbers, while the NDOPO's photon numbers are monotonically decreasing from $n = 1$.

lack of vacuum is apparent for both the DOPO and the NDOPO originating states. Whereas the kitten state from the DOPO has very low even-number contributions (as it should), the vacuum-subtracted thermal state has the same monotonically decreasing distribution (from $n = 1$) as a thermal state. Another thing to note here is, that even a relatively weak pump level of $\epsilon = 0.1$ still leads to a noticeable 2-photon contribution of 4%, so to be used as a single photon source it is really essential to keep the pumping at a minimum. It can not be lowered too much, though; if the trigger rate becomes too low, the unavoidable dark counts of the APD might start to become significant and contaminate the state.

Figure 3.18 highlights some other interesting differences between the vacuum-subtracted thermal state and the kitten state. The influence of the mode function width on a) the Wigner function in the origin and b) the 1- and 2-photon contributions to the state are plotted for different pump levels. The a) figure can be directly compared to the equivalent graph for the kitten state, Figure 3.15a). The depth of the Wigner function dip decreases much more rapidly with increasing ϵ here than for the kitten, and the optimal mode functions have a larger decay rate than the OPO cavity, where for the kitten state they were smaller. Both effects can be attributed to the influence of the 2-photon contribution. The 2-photon number state has $W(0,0) = +1/\pi$, so the increase of this contribution with higher pump levels means a fast shrinking of the Wigner dip. The kitten state has very little 2-photon content, so this effect is quite small. The high γ_s -values for the optimal mode function is partly explained by the b) figure, which shows that the 2-photon content decreases for higher γ_s . The 1-photon part also decreases, though, and what is not shown here is that most of the remaining part of the photon number distribution goes to the vacuum, $n = 0$. This is the case on both sides of $\gamma_s = \gamma$: For low γ_s , the temporal mode function covers more than just the correlated field, and therefore includes some normal thermal state which is mostly vacuum. For high γ_s , the spectrum of the mode function

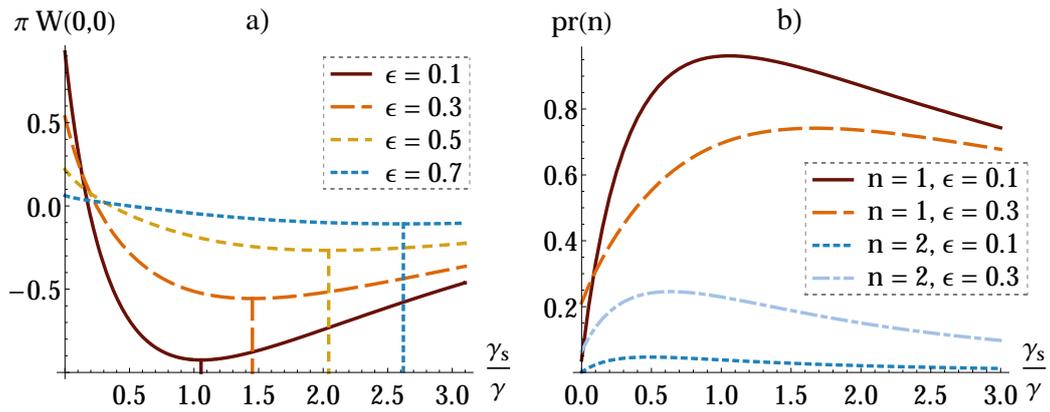


FIGURE 3.18: The dependence on the mode function width γ_s of **a)** the $\pi W(0,0)$ value for the vacuum-subtracted thermal state, and **b)** the 1- and 2-photon components of the state (brown/orange and blue, respectively), for different values of the pump level ϵ . Other parameters are $\kappa = \infty$, $\eta_s = \eta_t = 1$.

covers a wider range than the bandwidth of the field, so the vacuum state at the higher frequencies gets included. These results, that higher pump levels require a temporally narrower mode function, agree well with the observations of Nielsen and Mølmer [110]. In any case, the question of the optimal mode function is not quite as clear as it is for the kitten: Is it better to have a deep Wigner function, or to have a low 2-photon probability? For a single photon source, the latter would probably be the answer.

Experimental methods

4.1 Overview

The setup for the kitten and single photon state generation experiments is outlined in Figure 4.1. It is a quite linear setup, so in this section we go briefly through it step-by-step, from laser to data acquisition. The remaining sections contain more in-depth descriptions of the important aspects of the experiment.

1. Laser Our main (and only) laser is a home-built titanium sapphire laser (Ti:Sapph) designed for continuous single-mode operation around the wavelength range 852 nm - 860 nm. In our earlier experiments we operated at 860 nm, since the KNbO_3 crystals we used at the time were mode-matched at room temperature for this wavelength, while in later experiments we changed to 852 nm. This is the wavelength of the D2 transition in atomic caesium, the element used by two other experiments in the Quantop group. It is pumped by a Coherent Verdi V10 frequency-doubled Nd:YAG laser (532 nm), usually at a power of 9 W. The Ti:Sapph has many quirks and its output power can vary quite a bit. On a good day it can give 800 mW, but more often it is around 500-600 mW. An etalon keeps the lasing single-mode. The laser is prepared for stable frequency locking via an intra-cavity Brewster plate and a piezo-mounted mirror, but for the experiments described in this thesis we did not use the locking, as the passive stability was sufficient. When we operate at 852 nm, we do however visually monitor the fluorescence in a cell of caesium illuminated by a weak branch of the laser, as a guideline to the correct wavelength. The output beam is transmitted through an optical isolator to limit the amount of light reflected back into the laser cavity; too much light will cause it to flip between several longitudinal modes. The laser was constructed by Lasse Leick, whose thesis contains more details [77].

2. Beam splitting Most of the laser power is used for frequency doubling. Before the doubling stage, a small part is split off for use in 2-3 auxiliary beams: the local oscillator, and two beams injected into the OPO, primarily for cavity locking. The local oscillator is transmitted through a mode cleaning cavity

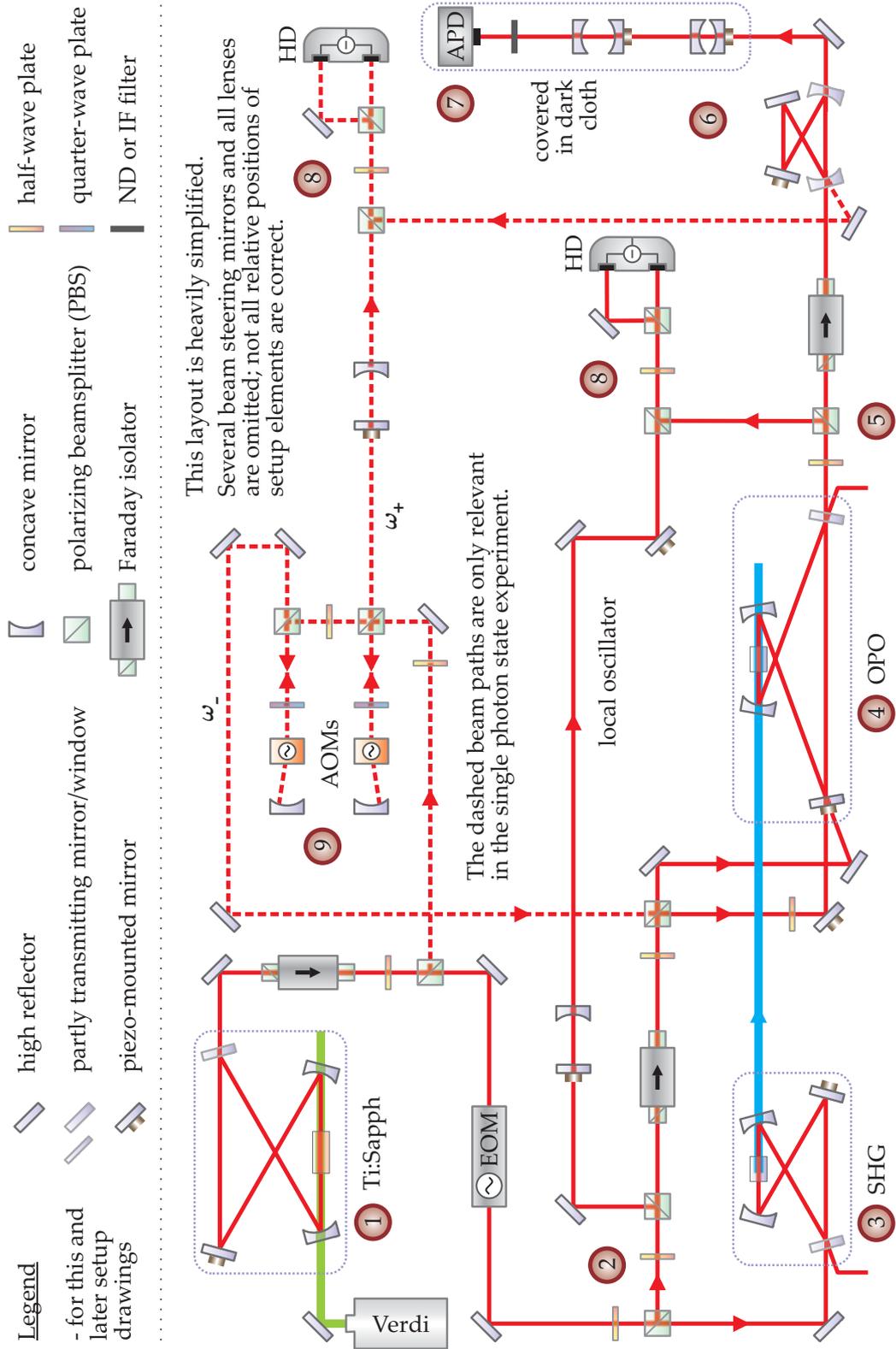


FIGURE 4.1: Setup for kitten and single photon generation.

(Fabry-Perot resonator) locked on the side of the resonance. The mode quality of the uncleaned beam is actually pretty neat already, but the cavity also serves to keep the LO power constant, irrespectively of drifts in the Ti:Sapph power.

3. Frequency doubling The infrared light is frequency doubled to 426 nm or 430 nm in the process of second harmonic generation (SHG). A moderate-finesse bow-tie cavity enhances the incident field to several watts, and a temperature controlled non-linear crystal is positioned in a tight waist between two concave mirrors of the cavity. All mirrors except the input coupler are high-reflectors for the IR, but the concave mirror through which the frequency doubled light escapes has $\sim 90\%$ transmittance for the blue. The cavity is kept on resonance with a Pound-Drever-Hall lock [31]. An electro-optic modulator in the main beam provides a phase modulation at 18 MHz for the Pound-Drever-Hall error signal which is obtained from the light reflected off the input coupler.

4. OPO The OPO cavity configuration is very similar to that of the SHG; a non-linear crystal in a tight waist between two concave mirrors of a 4-mirror bow-tie configuration. The blue light from the SHG is mode-matched to fit the TEM_{00} mode of the OPO cavity. The way this is done is detailed later. Whereas the SHG cavity has one partly transmitting mirror used for the input, the OPO has a similar mirror used to let the squeezed down-converted light escape from the cavity. It can be locked on resonance with a Pound-Drever-Hall lock as well, by injecting one of the auxiliary beams through one of the high-reflector mirrors in the opposite direction of the down-converted field. That way, this lock beam will escape the cavity into a different direction than the squeezed light and can be monitored with an 18 MHz detector. This lock method is preferred in situations where we are only interested in getting a squeezed vacuum. However, there will always be a tiny part of the lock beam that is reflected or scattered from the non-linear crystal into the output mode. This is no problem when measuring sideband squeezing on a homodyne detector, but when employing single photon detectors it is detrimental, since light from the lock beam will cause "dark counts", i.e. clicks in the APD not originating from the down-converted field. Hence, we have developed other methods to lock the OPO, as well as the subsequent filter cavities.

5. Tapping beamsplitter A small part of the generated squeezed vacuum is tapped off and send towards the APD to act as conditioning trigger photons. The remaining main part of the field goes to the homodyne detector. The tapping off is done by a half-wave plate (HWP) + polarizing beam splitter (PBS) combination in order to be able to control the tapping (reflection) ratio.

6. Filtering line The frequency filtering in the trigger arm, which gets rid of most of the nondegenerate OPO modes, is performed by a series of three empty cavities and a 0.5 nm interference filter. The interference filter is specified for transmission of 852 nm light, but by tilting it at an angle with the incident beam

it can be tuned to 860 nm as well. The first filter cavity is in a bow-tie (traveling wave) configuration and thus relatively long compared with the following two, which are of the linear Fabry-Perot (standing wave) type. When they are all resonant on the central frequency, the degenerate mode is transmitted with a typical efficiency of 10-15%, while the nondegenerate modes ought to be almost completely rejected. The variation of free spectral ranges and bandwidths of the different cavities works towards this end. The common resonance is found by injecting one of the auxiliary beams into the OPO in the forward direction through a high-reflector mirror. This seed beam then simulates the mode of the down-converted light, with the same central frequency, so when all cavities are resonant on the seed, they will also be set correctly for the down-converted light. The injected beam (the seed) is blocked when taking measurements.

7. APD The trigger photon detector is a Perkin-Elmer SPCM-AQR-13 (single photon counting module) which has built in all the electronic circuits necessary for operation of the APD in Geiger mode. It has a specified quantum efficiency at 852 nm of 44%, a dark count rate of about 100 c/s, a specified timing jitter of 350 ps, and a dead time after a click of about 35 ns. To avoid stray light impinging on the diode, the module is shielded within a plastic box, a cardboard cover, and a lightproof cloth, where the latter – apart from the APD module – also covers the last two filter cavities with only a small opening to allow through the light coming from the first filter cavity. Due to the careful light shielding, we are able to run the experiment with full laboratory lighting without seeing an increase in the count rate. The output of the module is connected to both the oscilloscope as a trigger and to a pc where the count rate is monitored and used as input for a cavity locking program.

8. Homodyne detection In the balanced homodyne detection, the signal field to be probed should be spatially overlapped with the local oscillator (LO) with near 100% efficiency. The overlapping can be done on a non-polarizing 50/50 beam splitter – which we did for our initial measurements – or on a PBS. After spatially mixing the two beams on a PBS they are actually not overlapped, since they are in orthogonal polarizations. By rotating the polarization of each beam by 45° with a half wave plate, each beam will be split 50/50 on a subsequent PBS – the p-polarized portion will be transmitted while the s-polarized portion is reflected. In the outputs of this second PBS, the two beams will thus have the same polarization and can interfere. The two PBS outputs are directed onto two photodiodes in a balanced detector, where the photocurrents from the two diodes are subtracted before further electronic amplification. The amplified difference photocurrent is then sampled by the fast digital oscilloscope conditioned on the triggering by an APD signal.

9. Shifted beams for photon state generation When running the experiment with the purpose of generating single photon states, a few changes and additional elements are needed compared to the kitten setup outlined above. The different way of splitting the OPO output into a trigger and a signal mode when

operating in the NDOPO mode was already described in §3.4.5: The two co-propagating EPR beams are separated on the first filter cavity based on their frequency difference, so that all of the ω_- mode goes towards the APD.

The frequency mode pair chosen is the two nearest neighbours of the degenerate mode, that is, the modes separated from ω_0 by plus/minus one free spectral range of the OPO cavity. Since their frequencies are shifted from the central frequency of the main laser, it is also necessary to shift the auxiliary beams' frequencies to match. In particular the local oscillator must match the signal mode frequency ω_+ , but also the seed beam used to align all trigger filtering cavities should be shifted to ω_- . The shifted frequencies are obtained by double passes through acousto-optical modulators driven at half the OPO free spectral range frequency.

Apart from these changes to the setup, another important difference to the kitten experiment is the blue pump power, which should be kept at a minimum to avoid two-photon contributions, whereas for the kittens the pump should be sufficiently large to give large state amplitudes.

4.2 Making squeezed vacuum

4.2.1 Second harmonic generation

Most experiments involving squeezed light generation and characterization employ a frequency doubling stage for the production of the OPO pump, such that the down-converted squeezed light appears at the same wavelength as the main laser before doubling. When a certain wavelength of the squeezed light is required, such as the 852 nm for us, a suitable laser at half the wavelength is often not available. Another compelling reason for employing the frequency doubling is the necessity of phase coherence between the local oscillator and the probed beam in a homodyne detection, in order for them to interfere. This phase coherence is most easily obtained by using the same main laser beam as the source for both beams.

Phase matching

During the last few years we have tried out different configurations for the SHG. Most of the time we have used a 10 mm long potassium niobate (KN) crystal as the nonlinear medium, but at some point we changed to 15 or 20 mm long periodically poled potassium titanyl phosphate (PPKTP). Both kinds of crystals are positioned in a temperature-controlled brass oven, and the crystal temperature is tuned to obtain the optimal phase matching condition for the SHG process. To achieve the full potential of the conversion, the wave vector (or phase) mismatch between the fundamental (IR) field at ω and the harmonic (blue) field at 2ω should be zero. If not, a phase shift will accumulate along the length of the crystal, causing the blue light generated in different regions of the crystal

to interfere destructively, shifting the energy back into the IR.¹ The wave vector mismatch is

$$\Delta k = k_{2\omega} - 2k_{\omega} = \frac{4\pi(n_{2\omega} - n_{\omega})}{\lambda_{\omega}}.$$

Due to the wavelength dependence of the indices of refraction, n_{ω} and $n_{2\omega}$, the phase matching condition $\Delta k \approx 0$ is usually not fulfilled. However, if the crystal is birefringent, the index of refraction for one polarization at a certain wavelength can become equal to the index for the orthogonal polarization at an entirely different wavelength. So if the fundamental and harmonic fields have orthogonal polarizations, the phase matching condition can be fulfilled at a specific wavelength. In non-critical phase matching, which we employ in the KN crystal, the crystal is cut in a special way so that both fundamental and harmonic beams propagate together along an optical axis. That way there is no walk-off between the two polarizations, and the exact phase matching wavelength is not critically dependent on the incident angle, as is the case in critical angle-tuned phase matching. On the other hand, the phase matching depends on the temperature, which should be well controlled and stabilized for the desired wavelength. At room temperature the KN is phase matched around 860 nm, while it must be cooled to 5°C for the 852 nm.

The periodically poled KTP makes use of an entirely different phase matching method, namely quasi-phase matching, in which the wave vector mismatch is made up for by a periodical pattern of domains in the crystal with opposite signs of the nonlinearity. The length of each domain should be equal to the coherence length in the crystal, $\pi/\Delta k$ – the distance over which a phase difference of π is accumulated between newly and previously generated blue light. With the sign inversion of the crystal structure, the blue light generated right inside the flipped domain will again be in phase with the light generated in the beginning of the previous domain. In effect, the grating pattern imposes additional momentum to the second harmonic field, with a wave vector $-2\pi/\Lambda$, where Λ is the grating period (twice the domain length). This additional momentum makes the overall phase mismatch close to zero. The quasi-phase matching method gives a lower efficiency of the conversion from fundamental to harmonic than perfect phase matching, by a factor $2/\pi$ [123]. However, the effective nonlinearity of nonlinear crystals is very different depending on the propagation direction through the crystal, and only certain directions – which might not be the ones with highest nonlinearity – are viable for standard non-critical phase matching. By quasi phase matching, on the other hand, any propagation direction can be chosen, so the highest nonlinearity of the given material can be utilized. The most important feature of quasi-phase matching, though, is that the poling period can (in principle) be designed for phase matching for any desired wavelength at a certain temperature. Thus, the PPKTP crystals we purchased from Raicol Crystals Ltd. were designed for phase matching of 852/426 nm at room temperature with the same polarization for both fundamental and harmonic, which is achieved with a grating period of $\Lambda = 4.2 \mu\text{m}$. In practice we obtain the

¹The wave vector mismatch should only be exactly zero in the case of plane waves. For focused Gaussian beams, the Boyd-Kleinman theory [18] gives more precise predictions for the optimal wave vector mismatch, as well as prescriptions for ideal focusing in the crystal.

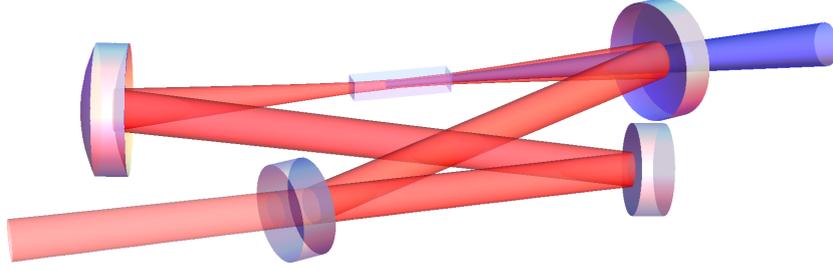


FIGURE 4.2: The latest layout of the second harmonic generation cavity. The curved mirrors around the crystal has a 100 mm radius of curvature, and the total cavity length is 375 mm. That gives a waist of 59 μm . The input coupler (lower left mirror) has a transmittance of 5.7%.

optimal conversion at a temperature of about 45° for the operating wavelength at 852.3 nm.

Theoretical conversion efficiency

An essential parameter in the description of both second harmonic generation and down-conversion is the single-pass nonlinear conversion efficiency E_{nl} which quantifies the power conversion from the fundamental to the harmonic frequencies in a single pass of the crystal:

$$P_{2\omega} = E_{\text{nl}} P_{\omega}^2.$$

From Boyd and Kleinman [18], with a Gaussian beam focused in the center of the crystal, the nonlinear conversion efficiency is determined by

$$E_{\text{nl}} = \frac{16\pi^2 d_{\text{eff}}^2 L_c}{\epsilon_0 c \lambda_{\omega} n_{\omega} n_{2\omega}} h(\alpha, L'_c, \sigma) e^{-(\alpha_{\omega} + \alpha_{2\omega}/2)L_c}, \quad (4.1)$$

where $h(\alpha, L'_c, \sigma)$ is the focusing function

$$h(\alpha, L'_c, \sigma) = \frac{1}{2L'_c} \iint_{-L'_c/2}^{L'_c/2} \frac{\exp[-\alpha(\tau + \tau') - i\sigma(\tau - \tau')]}{(1 + i\tau)(1 - i\tau')} d\tau d\tau'. \quad (4.2)$$

The various parameters and variables are described in Table 4.1. The σ parameter in the focusing function is a scaled wave vector mismatch, and is optimized by the temperature tuning. For the optimal value of the focusing function, which is 1.068, the focus parameter L'_c , which depends on the waist size w_0 , should be 5.68. For the 10 mm long KN crystal, the optimal waist is then 14.5 μm , while for the 15 mm and 20 mm PPKTP crystals, the optimal waists are 20 μm and 23 μm , respectively.

A resonant cavity has a power build-up of the intra-cavity field relative to the mode-matched injected field by a factor [7]

$$\frac{P_{\omega}^c}{P_{\omega}^{\text{in}}} = \frac{T_1}{(1 - r)^2}, \quad (4.3)$$

	$\alpha = (\alpha_\omega - \alpha_{2\omega}/2)z_R$ $L'_c = L_c/z_R$ $\sigma = \Delta k z_R$
d_{eff}	The effective nonlinear coefficient for the PPKTP. Measurements by the manufacturer [4] and Le Targat et al. [74] suggest a value around 9.5 pm/V.
L_c	Crystal length.
λ_ω	The fundamental wavelength, 852 nm.
$n_\omega, n_{2\omega}$	The refractive index at the fundamental and harmonic wavelengths, respectively, at room temperature along the crystal's z -direction. From Kato and Takaoka [67], they are $n_\omega = 1.84$, $n_{2\omega} = 1.94$.
$\alpha_\omega, \alpha_{2\omega}$	Linear absorption coefficients. While the infrared absorption is negligible ($\alpha_\omega < .002 \text{ cm}^{-1}$ as judged from cavity round-trip losses), the absorption at 426 nm is considerable since it is close to the bandgap edge around 400 nm. We measure $\alpha_{2\omega} \approx 0.2 \text{ cm}^{-1}$, but our measurement was not extremely precise, and more detailed investigations show a figure closer to $\alpha_{2\omega} = 0.1 \text{ cm}^{-1}$ [56].
z_R	Rayleigh range of the focused beam inside the crystal, related to the waist size through $z_R = \pi w_0^2 n_\omega / \lambda_\omega$.

TABLE 4.1: Parameters of the Boyd-Kleinman theory for the second harmonic generator with PPKTP.

where T_1 is the input coupler transmittance. r is the round-trip amplitude attenuation caused by outcoupling, linear losses \mathcal{L} (in the other mirrors and absorption/reflection in the crystal), and the nonlinear losses from frequency conversion, which is proportional to the power of the cavity field. Hence,

$$r = \sqrt{1 - T_1} \sqrt{1 - \mathcal{L}} \sqrt{1 - E_{\text{nl}}^{\alpha=0} P_\omega^c}. \quad (4.4)$$

The meaning of $E_{\text{nl}}^{\alpha=0}$ is that for the calculation of the round-trip attenuation, a modified E_{nl} factor should be used, where $\alpha_\omega = \alpha_{2\omega} = 0$ in (4.1). The reason is that Boyd and Kleinman consider the effect of absorption upon the amount of generated second harmonic outside the crystal, and it is clear that it will decrease for a nonzero $\alpha_{2\omega}$. However, the amount of the fundamental field that is lost to frequency conversion does not depend on the second harmonic absorption, so from the perspective of lost fundamental rather than created harmonic field, the E_{nl} term should not contain the absorption coefficient for the second harmonic. As for the absorption at the fundamental wavelength it is most conveniently incorporated in the overall linear loss term \mathcal{L} , since it contributes to the measured value of these losses². When α_ω is small, it hardly makes any difference anyway whether the absorption is modeled accurately as being distributed throughout

²The intra-cavity losses can e.g. be measured by observing the size of the dip in the reflection off the scanned cavity [137].

the length of the crystal, or as just decreasing the average steady state cavity field amplitude through \mathcal{L} . For the same reason we can also neglect α_ω when calculating the generated second harmonic power, whereas it is essential to include $\alpha_{2\omega}$ here:

$$P_{2\omega}^{\text{out}} = E_{\text{nl}}(P_\omega^c)^2 \quad \text{with } \alpha_\omega = 0.$$

$P_{2\omega}^c$ can be obtained numerically from the transcendental equation (4.3), with r from (4.4).

Performance

The first generation of the SHG was designed by Christian Schori and Jens Lykke Sørensen in Århus, and it is described in the thesis of the former [137]. It was a 70 cm long bow-tie cavity with the 3x3x10 mm KN crystal positioned between two 50 mm radius-of-curvature mirrors. This configuration gave a beam waist size in the center of the crystal of 16 μm . This is a little larger than the optimal 14.5 μm , but due to thermal lensing effects caused by absorption of both IR and blue light (as described in e.g. Polzik and Kimble [129]), the waist size was increased to improve cavity stability. The highest output power achieved during the experiment's time in the Copenhagen lab was 270 mW (320 mW inside the cavity) from a 550 mW pump. This result lies several years back, though, and both the available pump power and the quality of the crystals used degraded since then. Later we would typically have 150 mW blue from 350-400 mW IR. Another problem, we had with the KN crystals, was some ugly mode profiles of the generated blue beam. The crystals were very inhomogeneous, and by translating the crystal in the transverse directions, the blue mode profile could form many interesting shapes. A clean TEM₀₀ mode was hard to obtain, though. When translating the crystal it was also clear that the losses of the IR beam from either the crystal surface or bulk varied a lot. These annoyances were factors in the decision to change to a PPKTP crystal for frequency doubling as well (apart from the use in the OPO).

Initially, when we started using PPKTP in the SHG, we tried to keep the same cavity configuration as we had used for the KN crystal, although the waist size of 16 μm was in fact smaller than the ideal size. It soon turned out, though, that the crystal would get permanent damage after a while of producing high power blue light. With an input power of 250 mW, it could produce perhaps 130 mW at first, but after a minute or so it would suddenly jump to 105 mW instead. Later we could see that the losses for the infrared had increased in the part of the crystal where the beams had passed. This so-called gray tracking is a well known, but complex phenomenon in KTP [14], [80], [102]. It is highly intensity dependent, so by making the focusing in the crystal looser we could reduce the effect substantially. A waist size twice as large as the optimal, for example, reduces the fundamental intensity by roughly a factor 3 (taking account of lower nonlinear losses and hence a larger intra-cavity field), while the second harmonic conversion efficiency suffers a decrease of less than 10%, as estimated from the above formulas. Promising results of high-efficiency and high-power second harmonic generation with PPKTP at similar wavelengths using larger-than-optimal waist sizes had already been achieved by Le Targat et al. [74] and

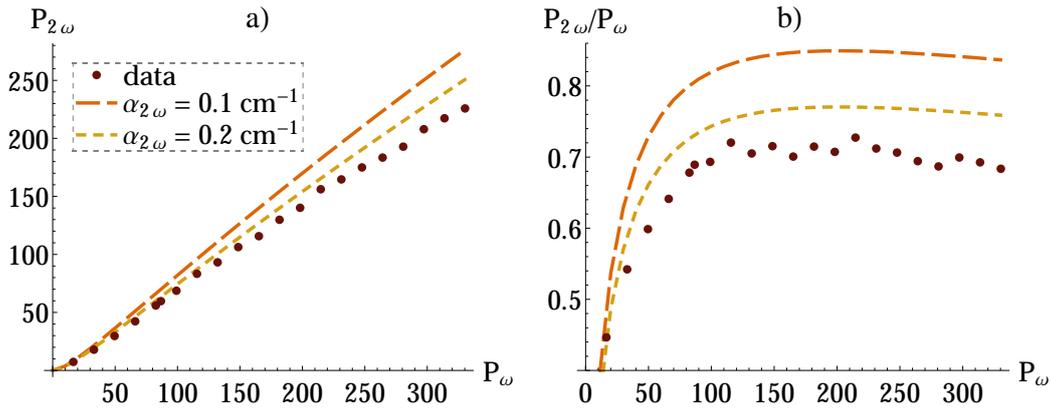


FIGURE 4.3: Frequency doubler performance in its latest incarnation. Generated blue powers in **a)** and conversion efficiency in **b)** versus the mode-matched infrared power. The curves are theoretical expectations, based on various measured and estimated parameters, as described in the text.

Villa et al. [162], and recently with a 4 times optimal waist by Lundeman et al. [87].

The modified SHG cavity has a waist size of $59 \mu\text{m}$, the result of changing to 100 mm radius of curvature mirrors and making the cavity shorter (375 mm). With this much larger waist size, the gray tracking problems have basically disappeared, although some thermal issues still seem to limit the attainable output power a bit. The expected nonlinear conversion efficiencies for the 20 mm crystal, with two different estimates for the 426 nm absorption coefficient, are then:

	$\alpha_{2\omega} = 0.1 \text{ cm}^{-1}$	$\alpha_{2\omega} = 0.2 \text{ cm}^{-1}$	$\alpha_{2\omega} = 0$
E_{nl} :	1.76%/W	1.60%/W	1.95%/W

Figure 4.3 shows a measurement of the generated blue power for different pump powers and the corresponding conversion efficiencies $P_{2\omega}/P_{\omega}$. The blue powers stated are right outside the crystal, inferred from measurements outside the cavity, while the infrared powers are the mode-matched part (mode-matching into the cavity was 93%). Two curves show the expected performance calculated from (4.3) and the above E_{nl} values. The linear losses were measured to be $\mathcal{L} = 0.4\%$, but there might be additional infrared absorption in the presence of the blue light which is not included in that number (see the discussion about BLIIRA in the following section). The measured performance does not quite live up to the theoretical expectation. However, the theoretical model contains many parameters which are not known very well, and some of them can alter the outcome quite a bit. For example, increasing the linear loss term \mathcal{L} to 0.8% would make the lower curve fit the data much more nicely. Anyway, the data demonstrate a reasonably high conversion efficiency, but the amount of blue light attainable is limited by the available laser power. Another important improvement from the KN-based SHG is the spatial mode, which is now a pretty, clean TEM_{00} mode.

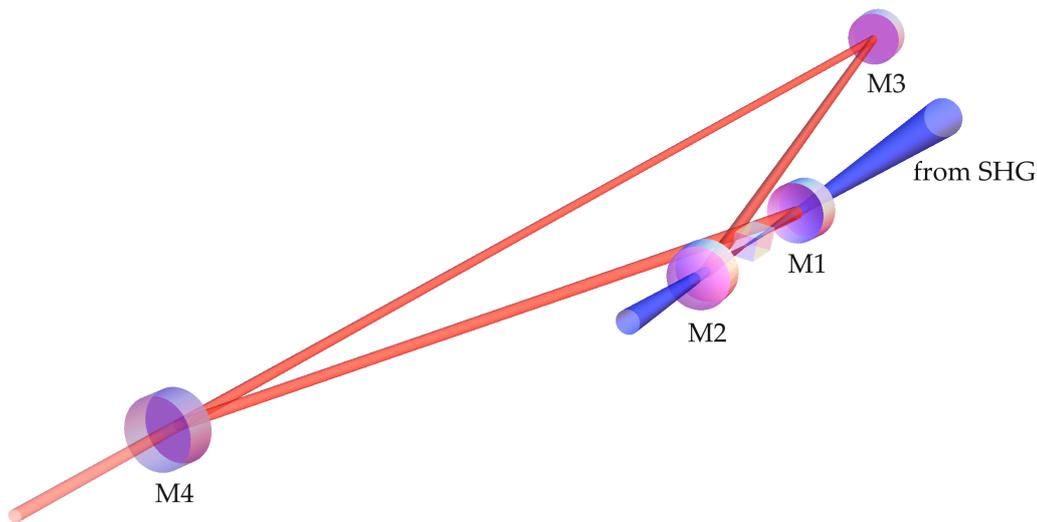


FIGURE 4.4: Layout of the OPO, as described in the text.

4.2.2 OPO

Design

The OPO cavity has maintained its layout, as shown in Figure 4.4, throughout the various iterations of the experimental setup. Like the SHG, it was designed by Christian Schori and Jens Lykke Sørensen. It is 81 cm long, yielding a free spectral range of 370 MHz. The two concave mirrors M1 and M2 surrounding the 10 mm long crystal has radii of curvature of 50 mm. Stability of the resonator mode requires the distance between the concave mirrors to be between 54.5 mm and 58 mm for PPKTP and 1 mm more for KN with a higher index of refraction (2.3). The distance is usually adjusted to be in the central part of the stability range, but not too close to the center since that would lead to frequency degeneracy of the TEM_{00} and higher-order spatial modes. The waist size in the position of the crystal is $15 \mu\text{m}$ with this configuration, close to the optimal waist sizes of $16 \mu\text{m}$ for PPKTP and $14.5 \mu\text{m}$ for KN. The output coupler M4 has a transmittance of $T_1 = 12.5\%$ for the infrared ($\gamma_1/2\pi = 3.7 \text{ MHz}$). The M3 mirror is piezo-mounted for cavity length control.

The blue pump light coming from the SHG should be mode-matched into the OPO cavity mode. Since the cavity mirrors are transparent for the blue light, no resonant build-up occurs and the standard mode-matching method of observing the cavity length-scanned transmission spectrum is not an option. Instead, we have used two other methods to do this, both involving the operation of the OPO as a frequency doubler. The schemes are illustrated in Figure 4.5. The first method, as described in Polzik et al. [128], employs an intermediate cavity between the SHG and OPO with high-reflectors for blue. Light is coupled into the cavity via a weakly reflecting beamsplitter. The OPO is made to produce blue light propagating backwards towards the SHG by pumping IR through M4 in the opposite direction of the squeezed light output. The mode of this beam then represents the OPO cavity mode. The blue beams from the SHG and the

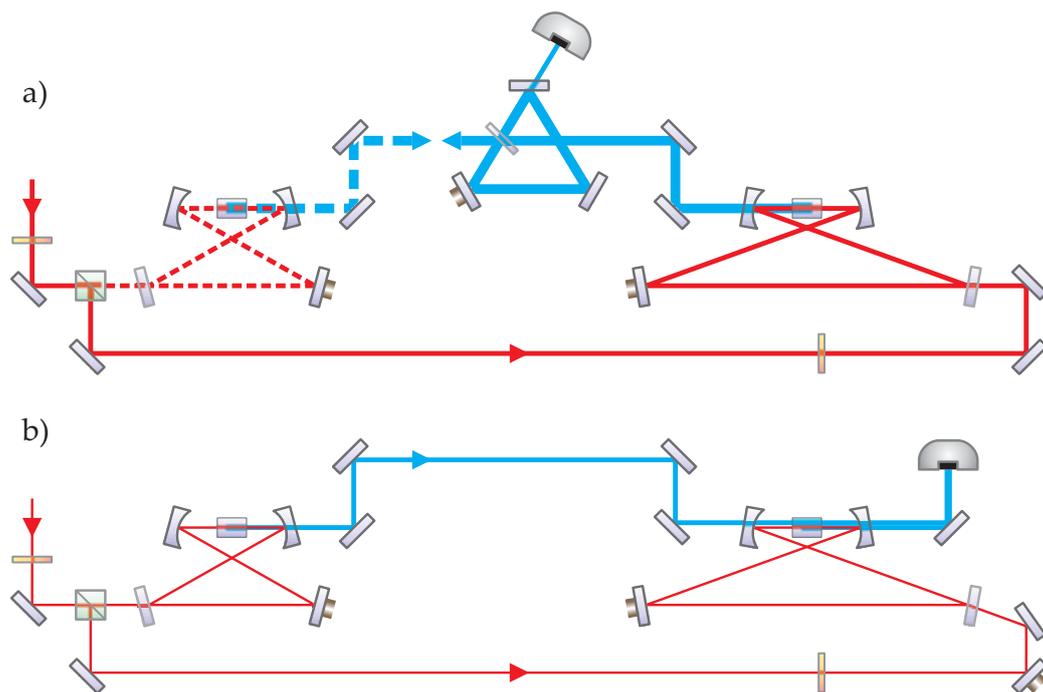


FIGURE 4.5: Blue mode-matching schemes. In **a)**, the SHG and OPO cavities produce blue light in turns, and are mode-matched individually to the intermediate cavity. In **b)**, the two cavities are simultaneously producing very weak blue light, propagating in the same direction. The two beams are then matched by observing their interference after the OPO.

OPO are then, one at a time, mode-matched to the intermediate cavity with the result that the SHG and OPO are mode-matched. This method works well but is somewhat cumbersome. The second method is more direct, but the actual reason why we changed was to get rid of the intermediate cavity to make room on the table for a pulse shaping device sitting in a big box (more on this in §6.1.2). The OPO is pumped with infrared propagating in the forward direction instead of backward, so that the generated blue light escapes through M2. By producing blue light in the SHG and OPO simultaneously, the two beams can be overlapped directly by monitoring their interference behind mirror M2. The rough alignment can conveniently be done by the naked eye, while the fine tuning can be carried out by scanning the phase of the infrared pump to the OPO and observing the interference pattern on a scope. It is important that the blue power levels are as low as possible while still retaining a visual pattern; if the blue light from the SHG has a considerable amplitude, the infrared field in the OPO will undergo phase-sensitive parametric amplification (see page 82), which will again alter the amplitude of the OPO-produced blue light and make the mode-matching unreliable. With this method we have achieved up to 96% mode-matching efficiency.

Potassium niobate

When we made the first kitten measurements, the nonlinear crystal was potassium niobate. This type of crystal has an unfortunate property called BLIIRA (blue light induced infrared absorption), by which the illumination of the crystal by blue light increases the absorption losses for light at infrared wavelengths [93]. This effect is detrimental to the generation of high purity strongly squeezed states; for higher losses, the escape efficiency $\eta_{\text{esc}} = T_1 / (T_1 + \mathcal{L})$ goes down, and as a consequence so do the amount of squeezing and the state purity. For our measurement series that were published in ref. [103], the BLIIRA losses were between 1.0% and 1.5%, depending on the blue power. Without blue light we measured the total passive intra-cavity losses to be 1.0%. For the combined losses, the escape efficiency was then 83-86%. The losses also influence the threshold pump power of the OPO as defined in (3.16). The threshold power P_{thr} can be related to the cavity losses and the single-pass conversion factor by [128]

$$P_{\text{thr}} = \frac{(T + \mathcal{L})^2}{4E_{\text{nl}}}, \quad (4.5)$$

so higher losses increase the threshold, and therefore decrease the interaction strength for a given pump power. The crystal had a nonlinearity of $E_{\text{nl}} = 1.5\%/W$, yielding a threshold power (for 1.5% BLIIRA) of 375 mW.

PPKTP

For the purpose of making highly nonclassical states with high purity and negative Wigner functions in the case of non-Gaussian states, a high overall signal efficiency is needed, as discussed in §3.4.4. There are plenty of potential sources of loss outside the OPO, so it is a bad start to have only 83% escape efficiency. Therefore we switched to PPKTP which should be free of the BLIIRA effect. The first crystal we used was kindly gifted to us by Prof. Furusawa of Tokyo University. It was a 1x1x10 mm crystal quasi-phase matched for 860 nm fundamental wavelength at room temperature. Later we purchased a small stock of similar crystals, but with a phase matching wavelength of 852 nm. The end faces are coated with a high quality dual band anti-reflection coating specified for $R < 0.05\%$ @ 852 nm and $R < 0.1\%$ @ 426 nm. With the crystal in place in the cavity, we measure a total intra-cavity loss of 0.4%, which includes crystal bulk and surface losses and losses at the mirrors (except the output coupler). With such a low loss, the escape efficiency becomes as much as 97%, a huge improvement from the KN. Unfortunately, the crystals are not completely free from additional light-induced losses.

We observe that the intra-cavity losses increase after illumination with the pump light. The amount of increase depends on the optical power and the duration of the illumination. Immediately after turning on the pump (about 100 mW) there is a rapid increase of about 0.1% additional losses during the first second. Afterwards the losses continue rising but at a slower pace. After several minutes they seem to flatten out at a level of 0.3-0.4% extra. The light induced losses persist after turning off the light, but they slowly disappear – after 10-20 minutes the

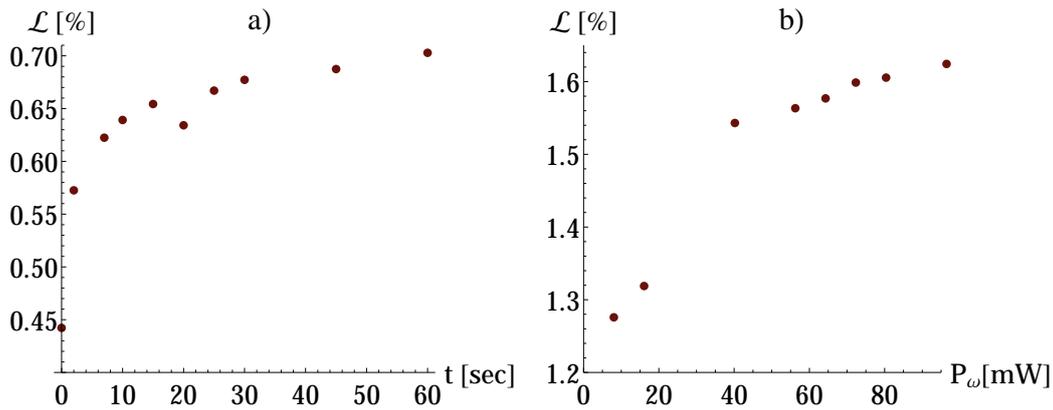


FIGURE 4.6: Increased total linear cavity losses due to BLIIRA. In **a)**, 100 mW of blue light was shone on the crystal for different durations and the losses were recorded immediately afterwards. In **b)**, different powers were used for a 2 minute duration. The initial losses in b) were higher than usual because of dirty mirrors.

total losses have recovered to the value before illumination, irrespectively of the duration of the exposure. The data presented in Figure 4.6 were all measured at the phase matching temperature. However, due to parametric gain of the IR probe beam, the reflection dip method for measuring cavity loss is not applicable while pump light is on. Therefore, the measurements were done by illuminating the crystal for a set duration and measuring the losses immediately after blocking the pump light. We made a check at a temperature far away from phase matching, where the loss effect was still present (also at the same magnitude) – here, we confirmed that the losses are basically identical while the blue light is still on and immediately after turning it off. In the Furusawa group, similar measurements were performed, showing the same characteristics and magnitude of the effect. They measured all the losses far from the phase matching temperature, so they were able to get a better time-resolved picture. This showed in greater detail the two different time-scales of the effect that we see hints of: An initial small ms-scale increase followed by a larger minute-scale increase [152]. We do not have a good understanding of the cause for these effects. They could be some sort of photo refractive effect or thermal lensing, but we have not yet made the effort of a detailed study which might reveal some clues about their nature and possible ways to overcome the extra losses. Due to these additional light-induced losses, the escape efficiency of the OPO decreases to 94-96%, depending on the pump power level. With these losses, the cavity bandwidth is just above 4 MHz.

Parametric gain and squeezing

In the day-to-day characterization of the OPO performance, an essential observable parameter is the parametric gain G (or simply gain). This is the phase-sensitive amplification or deamplification of a degenerate frequency probe (seed) beam injected into the OPO, caused by the nonlinear interaction with the pump light. The amount by which the probe intensity is amplified when in-phase with

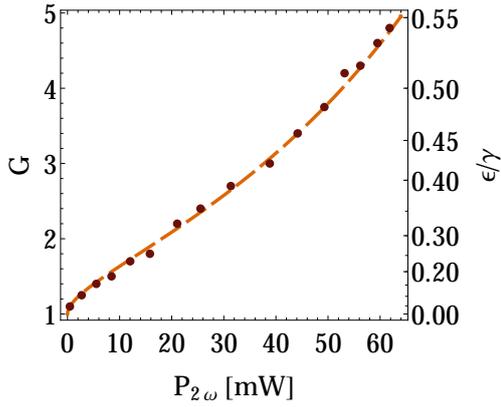


FIGURE 4.7: Observed parametric gain and corresponding pump parameter ϵ for a range of (mode-matched) blue pump powers. The curve shows the best fit to the gain function $(1 - \sqrt{P_{2\omega}/P_{\text{thr}}})^{-2}$, yielding $P_{\text{thr}} = 209 \pm 1$ mW.

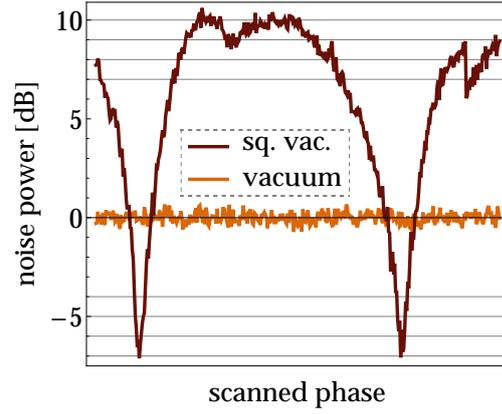


FIGURE 4.8: Almost 7 dB squeezing, measured at the 1.1 MHz sideband with the LO phase scanned. The electronic dark noise (-10 dB relative to the shot noise) has been subtracted from both the vacuum and the squeezed vacuum traces.

the pump, compared to its intensity without pump, is called the gain factor and is a direct measure of the pump parameter ϵ of the OPO theory in §3.3. The classical equation of motion for the cavity field with the probe field $|\alpha_p|e^{i\phi_p}$ injected through a mirror with decay rate γ_3 is, cf. (3.15),

$$\dot{\alpha}_c = \epsilon\alpha_c^* - \gamma\alpha_c + \sqrt{2\gamma_3}|\alpha_p|e^{i\phi_p}.$$

In steady state, $\dot{\alpha}_c = 0$, the solution to this equation and its complex conjugate is

$$\alpha_c = \frac{\sqrt{2\gamma_3}}{\gamma}\alpha_p \left(\frac{\cos\phi_p}{1-x} + i \frac{\sin\phi_p}{1+x} \right),$$

with $x = \epsilon/\gamma = \sqrt{P_{2\omega}/P_{\text{thr}}}$. The phase-sensitive power gain relative to the cavity field with no pump ($x = 0$) is then

$$G(\phi_p) = \frac{|\alpha_c|^2}{|\alpha_c^{x=0}|^2} = \frac{\cos^2\phi_p}{(1-x)^2} + \frac{\sin^2\phi_p}{(1+x)^2}, \quad (4.6)$$

with a maximum value of $G \equiv G(0) = 1/(1-x)^2$. Thus, by observing the cavity field leakage through another mirror, ϵ can be determined. In Figure 4.7, measured gains for various pump powers are plotted, fitting nicely this function with a threshold power of 209 ± 1 mW. The corresponding nonlinear conversion efficiency, from (4.5), is $E_{\text{nl}} = 2.1\%/W$, with the measured losses during that measurement of 0.8%. This value corresponds well with those obtained from single-pass frequency doubling in the OPO, while the theoretically expected value is slightly higher: 2.3%/W for $\alpha_{2\omega} = 0.2 \text{ cm}^{-1}$.

Figure 4.8 shows a trace of some of the best squeezing we have obtained. The measurement was performed at a sideband frequency of 1.1 MHz by two detectors optimized for shot noise limited detection at this frequency. The signal was analyzed on a spectrum analyzer in zero span mode (so the axis is time),

and subsequently the dark noise was subtracted and the level normalized to the vacuum level. The phase of the local oscillator was scanned to show both squeezing and anti-squeezing. We see that we obtain a squeezing of about -6.5 dB and an anti-squeezing of 10 dB. The expected performance is given by (3.23), with the addition of an efficiency factor on the last term (the non-vacuum part) which is sensitive to losses and inefficiencies. In the measurement of Figure 4.8, the total efficiency was estimated to be 88% , with contributions from the OPO escape efficiency, 94% , propagation from OPO to detector, 98% , homodyne visibility, $(99\%)^2$, and detector quantum efficiency, 98% . The maximum parametric gain was measured to be $G = 7.5$, corresponding to a pump parameter $x = 0.63$. Plugging these numbers in, together with $\Omega/2\pi = 1.1$ MHz gives an expected squeezing $S_p = -7.3$ dB and anti-squeezing $S_x = 10.7$ dB relative to the vacuum level. The discrepancy of the observed and expected squeezing can be accounted for by just $2-4\%$ extra unknown losses. The anti-squeezing is also slightly lower than expected, which is most likely due to a weak detector saturation.

4.3 Trigger system

4.3.1 The number of phase matched OPO modes

The output of the OPO is separated into a trigger and a signal mode, either by the tapping beamsplitter (for kittens) or by the first filter cavity (for photons). After the separation, the trigger mode undergoes heavy frequency filtering to suppress the irrelevant OPO longitudinal modes. To understand why as much as three cavities plus an interference filter is needed, we can estimate the number of longitudinal OPO cavity modes that are phase matched for down-conversion simultaneously. The wave vector mismatch for the down-conversion process is

$$\Delta k = k_{2\omega_0} - k_{\omega_+} - k_{\omega_-} - \frac{2\pi}{\Lambda},$$

with $\omega_+ + \omega_- = 2\omega_0$. From published Sellmeier equations for KTP [67], the wavelength-dependence of the indices of refraction, $n(\lambda)$, can be calculated precisely. By inserting $\sigma = \Delta k z_R$ in the Boyd-Kleinman function (4.2) with fixed z_R , α and L'_c , the nonlinear interaction strength as a function of the down-conversion wavelength is obtained. This curve is plotted in Figure 4.9a), which shows a phase matching bandwidth of 30 nm, corresponding to roughly $30,000$ phase matched cavity modes!

The number of populated cavity modes is somewhat reduced, though, when taking account of the fact that both wavelengths of a down-converted pair should be simultaneously resonant in the cavity, as pointed out by Lu and Ou [84]. In terms of wavelength, the resonance condition for the m th non-degenerate mode is

$$(N_0 + m)\lambda_m = l + L_c n(\lambda_m),$$

where N_0 is an integer, l is the cavity length outside the crystal, L_c the crystal length, and the right hand side is the total optical length of the cavity. N_0 is then the number of degenerate mode wavelengths that fit within the cavity length.

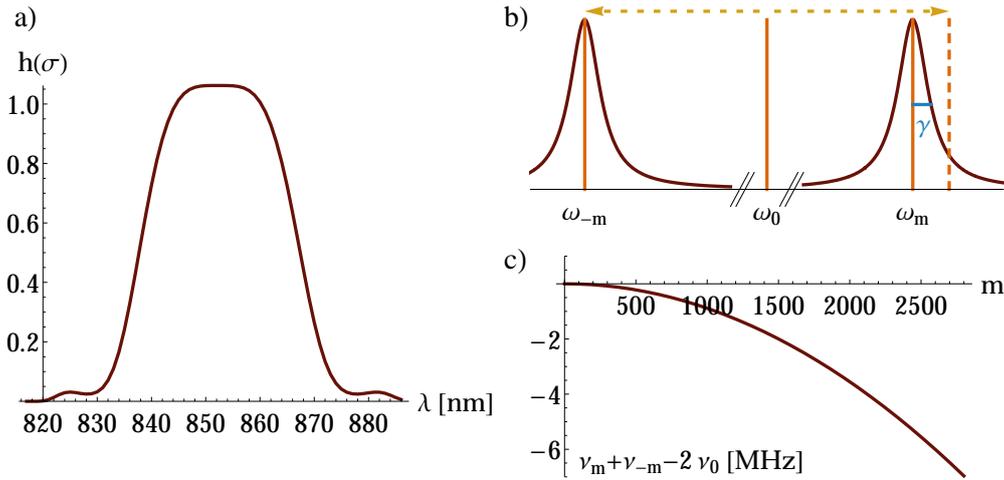


FIGURE 4.9: **a)** The phase matching bandwidth for down-conversion in our PPKTP crystal is ~ 30 nm. $h(\sigma)$ is the Boyd-Kleinman function, calculated from Sellmeier equations. **b)** Due to dispersion in the crystal, the optical cavity length is not the same for high and low frequency sidebands. Therefore the cavity modes become shifted with respect to their “twin” mode, and the down-conversion process – which requires energy conservation – may become suppressed if there is no common resonance for the pair of frequencies symmetric around ω_0 . **c)** The amount of shifting (in MHz) between the m th and $-m$ th mode (shifted $\pm m$ free spectral ranges from the degenerate mode), corresponding to the difference in b) between ω_m and the dashed line. The OPO has a HWHM of 4 MHz, so on the order of 2000 mode pairs fulfill the double resonance condition.

Using the explicit Sellmeier equations, this relation gives the wavelengths λ_m (or frequencies ω_m) for each spectral cavity mode m . In order to have efficient down-conversion into the $m/-m$ mode pair, the difference between their frequency spacing from the degenerate mode should be roughly within a cavity half-linewidth, as illustrated in Figure 4.9b):

$$(\omega_m - \omega_0) - (\omega_0 - \omega_{-m}) \lesssim \gamma.$$

This double-resonance condition is fulfilled for a few thousand modes, as the calculated curve in Figure 4.9c) shows. That is still a very large number of modes that need to be filtered away.

4.3.2 Filter specifications

The trigger path is outlined in Figure 4.11a) on page 89. For convenience we name the filter cavities FC1, FC2, FC3, in the order they appear on the path. Each cavity must be well mode-matched to its neighbours. The mode-matching is done by injecting a seed beam through the locked OPO. FC1 is mode-matched to this beam which simulates the mode of the down-converted field. Then, the length of FC1 is locked and FC2 is mode-matched to the output of FC1, and the procedure is repeated with FC3. Finally, the output of FC3 is focused onto the APD, but no mode-matching is necessary there, so we just adjust the beam position by observing the count rate from the detector. Before the APD is the

filter	length	FSR	finesse	HWHM	OPO modes/FSR
FC1	200 mm	1.5 GHz	30	25 MHz	4
FC2	3.7 mm	40 GHz	300	70 MHz	110
FC3	12 mm	12.5 GHz	250	25 MHz	34
IF	0.5 nm (200 GHz) FWHM — covers 550 OPO modes				
OPO	810 mm	370 MHz	48	3.9 MHz	1

TABLE 4.2: Specifications for the three filter cavities, the interference filter, and the OPO for reference. The numbers for the filters are approximate. The last column, the ratio of the filter FSR to the OPO FSR is a rough lower estimate of the extinction ratio of OPO modes for that filter. It is a lower estimate since the OPO modes not necessarily coincide with the filter resonances.

interference filter (IF) which is tilted to give maximum transmission. The overall transmission of the entire path is 10-15%. Table 4.2 summarizes the specifications of the individual filters, and the spectral layout is illustrated in Figure 4.10. The shortest cavity (FC2) with a free spectral range of 40 GHz has approximately 5 resonances within the bandwidth of the interference filter. To get rid of possible OPO modes transmitted within those resonances, FC3 has a free spectral range such that there are no resonance overlaps between FC2 and FC3 within the coverage of the IF. FC1 with its relatively small FSR and wide resonance provides only a little more filtering, but it is very convenient for separating the photon pairs when running the single photon experiment – it was originally used for the demonstration of strong EPR-correlations in twin beams [136]. In theory, the combination of all four filters should lead to a suppression of the irrelevant OPO modes strong enough that more than 99.9% of the transmitted light is from the relevant mode, that is, basically unity modal purity. In practice, there may be imperfections such as bad mode-matching or unstable locks that do not keep the cavities on resonance, which could impair the suppression. The mode-matching to FC1 and FC3 is quite good; above 95%, while to FC2 it is somewhat harder and we usually have about 80%.

4.3.3 Cavity locking

It is a tricky business to stabilize the lengths of all the cavities in experiments involving photon counting. The usual method of monitoring the transmission or reflection of a beam seeded into the cavity and – one way or the other – extracting an error signal from that does not work. The reason is, that a beam strong enough to give a clean signal on a normal photodetector invariably will contain billions of photons. In sideband measurements like homodyning it usually does not matter, but since the APD is not frequency selective, the lock beam will completely overshadow the relatively few photons from down-conversion. The bow-tie cavities improve the situation a bit by allowing a counter-propagating lock beam, but still there are reflected or scattered photons, especially from the OPO crystal, that enter into the forward-propagating beam direction. Since such lock beam photons will cause false clicks if they reach the APD, other methods

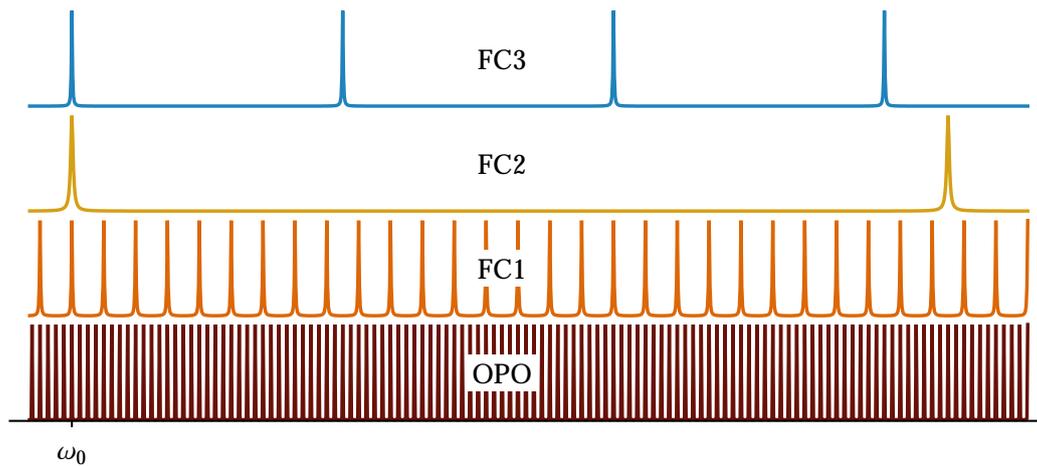


FIGURE 4.10: Spectral layout of the OPO and filter cavities with their considerably different free spectral ranges, arranged to suppress all other modes than the degenerate ω_0 mode (or ω_- for the single photon experiment). Not illustrated is the bandwidth of the interference filter which covers about 5 FC2 modes. The actual experimental layout might be a little different from what is shown here, due to uncertainties in the cavity length measurements.

– or at least an updated variation of the scheme – must be considered. We have developed two different ways of keeping four cavities in series (OPO + FC's) simultaneously locked to resonance; the two are quite similar in terms of the control but different on the optical table.

Using APD signal

The first locking system, which was used for our published results so far, circumvents the problem of a too strong auxiliary lock beam by simply using the down-converted light itself as a lock beam. This light is not detectable on a standard photodiode, but the APD, whose primary objective is to act as a heralding trigger, can double as the lock signal detector. The photon count rate is maximal when all cavities are on resonance; that forms the basis of the error signal. The idea and implementation (LabView programming) of this scheme is mostly due to Christian Hettich.

A LabView program running on a standard Windows PC with National Instruments PCI-6221 and PCI-6722 digital acquisition boards reads from two counter ports and writes to four analog output ports. The inputs are connected to two APD modules and the outputs to a mirror-mounted piezo in each of the four cavities; OPO, FC1, FC2, FC3. The sole purpose of the second APD (APD2) is to aid in the initial simultaneous alignment of all four cavities close to resonance. Since a single detector generates a common error signal for all cavities, it is not possible to bring them into lock one after the other – it must be done in one go. So before turning on the locks, the cavity lengths are scanned at different rates to search for the resonances. A very weak seed beam – attenuated by many orders of magnitude by neutral density filters – is injected into the OPO, while the blue pump beam is shut off. Most of the time no light will reach the APD, but

eventually it will happen that all four cavities are close to resonance at the same time, letting some light through. With four cavities, two of them with finesse around 300, this can take forever. Therefore, APD2 is employed to monitor the transmission of OPO + FC1 + FC2 only. In practice this is done by directing the reflection from the front mirror of FC3 onto APD2 using a PBS + quarter-wave plate combination between FC2 and FC3. Thus, at first only a three-fold simultaneous resonance must be found. An automated algorithm of the program narrows the scan ranges down around the voltage regions where it detected the largest signal. Another four stages of scanning and narrowing down follows, after which each cavity is supposed to be so close to resonance that the scans can be turned off and replaced with constant output voltages. Then the locking part of the program sets in. With the first three cavities locked on the signal from APD2, the last cavity, FC3, is scanned manually until APD1 sees a signal. The locking is then transferred to this signal from APD1, and APD2 is not used anymore. Even with this two-stage resonance search, the procedure is rather slow. The initial search scan rates are ~ 100 Hz for the fastest scanning cavity and 0.05 Hz for the slowest. Because of this, and because the procedure outlined here does not always converge on a common resonance, it can take a considerable amount of time to prepare the cavities for the locking stage. See Figure 4.11a) for an overview of the filter layout on the optical table and Figure 4.11b) for a diagram of the locking stage of the LabView program. The cavity search routine is not illustrated.

To derive individual error signals for each cavity, the four output voltages are modulated slightly with sine waves of different frequencies. For a given cavity, the voltage modulation causes a small modulation of the cavity length and hence of the light transmission as observed by the APD count rate. A digital demodulation of the count rate at the same frequency returns an error signal which has a zero-crossing on the voltage output corresponding to a cavity resonance. A proportional-integral loop then adjusts the voltage offset to keep the error signal close to zero. Since the modulation frequencies are different for the four cavities, individual error signals can be extracted, unaffected by the modulation of the other three cavities. The lock provided in this way is fairly weak. In order to gather enough statistics from the photon counting, the program samples the counts for 60 ms before processing them in a bunch, giving an update rate of a mere 17 Hz. This makes the lock incapable of correcting for possible oscillations faster than ~ 8 Hz. Furthermore, to avoid overshooting, the gain of the PI control must be kept low, meaning that if a cavity somehow falls down from the resonance, it takes on the order of a second to recover. In practice, however, this is not a serious limitation to the experiment. As long as there is no excessive acoustic noise in the lab, all the cavities are quite passively stable. The most unstable one, the OPO, will remain very close to resonance for several seconds after turning off the lock. Therefore, the weak stabilization provided by the APD signal is sufficient for the active long-term locking. The system works well down to a count rate of 10,000 c/s, and 5,000 c/s is still possible, although the lock starts to become fragile.

When all cavities are finally locked to resonance using the seed beam, the blue light is slowly turned on using either an iris or a PBS + half-wave plate com-

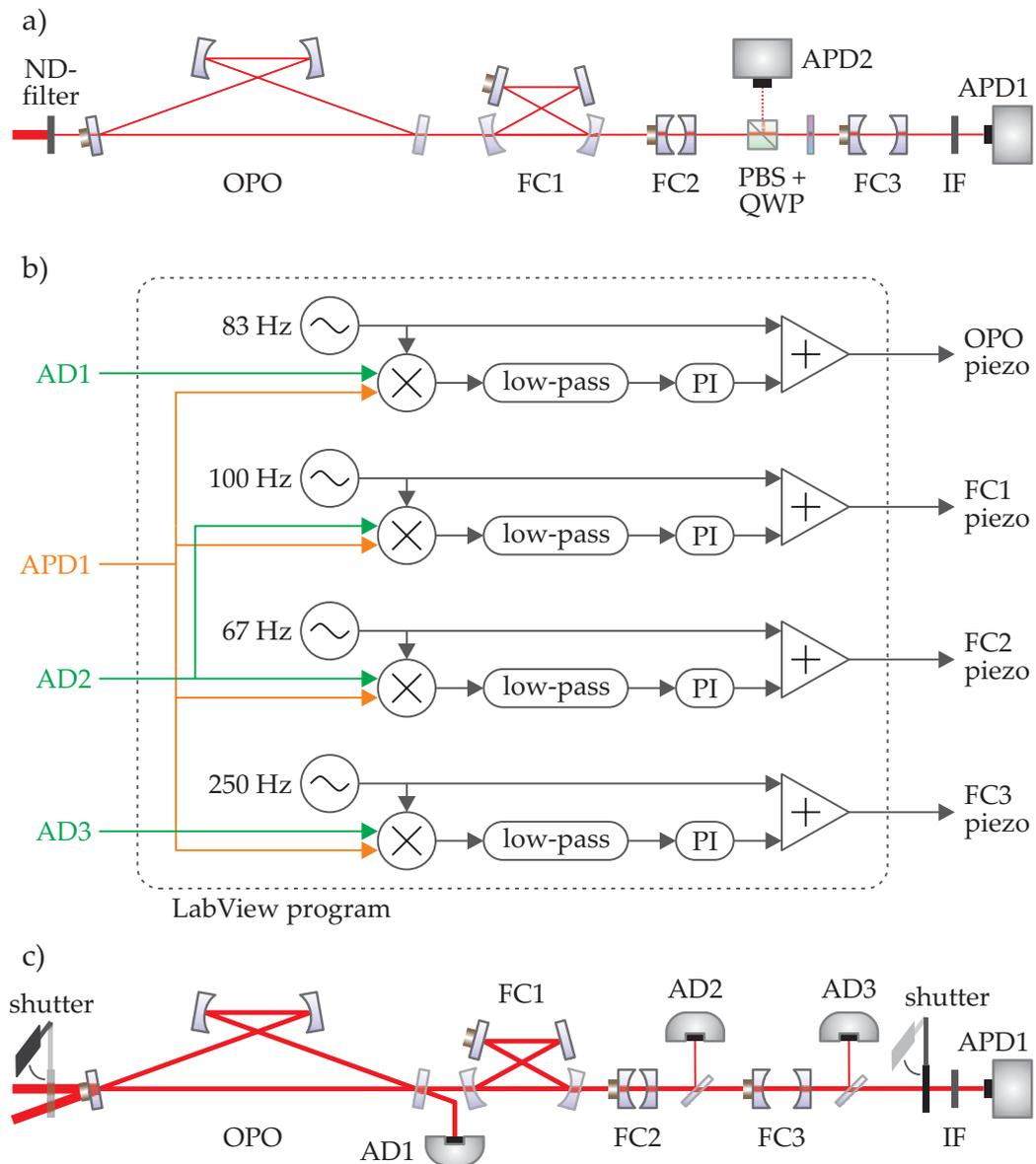


FIGURE 4.11: Overview of the trigger filtering system. **a)** The cavity locking stage in the scheme where APD1 provides the error signal for all for cavity locks. Here, the attenuated seed beam is injected into the OPO. When switching to the state generation, the blue pump light is turned on, after which the seed is blocked. APD2 is used only for alignment of the cavities on resonance. **b)** Diagram of the digital control. The length of each cavity is modulated by different frequencies which are used again for demodulation of the detector signals. These signals can come from either the APD or from the analog photodetectors (AD). The printed frequencies are only an example – other combinations are possible. **c)** The sequential locking scheme where two shutters alternately switches strong seed beams on and off, while always keeping the APD shielded from the strong light. Here the APD only acts as a conditioning trigger detector while the locking is taken care of by the analog detectors.

bination. Turning the pump on too fast disturbs the OPO lock, probably due to the heating of the crystal. With the pump at full strength, the seed beam is blocked completely, after which the down-converted light acts as the lock beam. For the kitten state generation, a usual pump level is $x = \epsilon/\gamma \simeq 0.3$ (for a parametric gain of $G = 2$), which gives a photon production rate of $2.5 \times 10^6 \text{ s}^{-1}$, from (3.20). With about 10% total efficiency of the trigger channel, a tapping beamsplitter reflection ratio of 5% is sufficient for a good locking signal of more than 10,000 counts per second. In the single photon state experiment, an entire OPO mode is sent to the trigger, with no prior beamsplitting, so the gain can be much lower (as it should be to avoid excessive 2-photon contributions). A pump level of only $x = 0.05$ ($G = 1.1$) still gives a $\sim 5,000 \text{ c/s}$ count rate.

Sequential locking with strong beam

As will be explained in §4.5, we needed better control of the phase information obtained during measurements. To this end, the coherent seed beam injection in the OPO became necessary, which in turn required modifying the locking procedure. Partly inspired by Wakui et al. [164], we implemented a sequential scheme where a strong seed beam is switched on and off with mechanical shutters, while the path to the APD is alternately closed and opened as well, to shield the module from the harmful light level. Figure 4.11c) shows the important elements of the altered scheme. A separate LabView program was designed to control the switching sequence, as well as providing the phase lock. It is described in more detail in the next section. Running on a second Windows PC equipped with a NI PCI-6251 board, it outputs TTL digital signals at predefined intervals in a cycle with a rate of 1-4 Hz. During one cycle, each of the two shutters is opened and closed once. When the seed beam is on, the locking program is told to actively control the cavity lengths; when the seed beam is switched off, the active stabilization halts and the output voltages are kept at their last value until the seed and locking is switched on again. When the seed is fully blocked, the shutter in front of the APD is opened, and the scope is set to accept trigger signals from the APD – there is no need to include the dark counts during the period of locking. Figure 4.12 shows an example of such a sequence, with a period of 330 ms.

The locking program is the same as that used in the previous scheme. But now, the error signals are not derived from the APD counts but from analog silicon PIN photodiode detectors monitoring the power level of the strong seed beam. Three detectors are connected to analog inputs of the NI-DAQ card. The first one monitors the transmission of the OPO of a second auxiliary beam, injected in the backwards propagating direction (this beam is switched together with the other seed beam). The second detector monitors the transmission of FC2, via a thin glass plate that reflects about 8% towards the photodiode. The third detector similarly monitors a reflection from a glass plate of the light transmitted through FC3. With three detectors, the resonance search becomes a little easier than with only the two APDs used before. The locking itself can also be done with error signals obtained from different detectors for each cavity (detector 2 for FC1 and FC2 together), although it is not necessary – the threefold demodulation of the signal from the final detector is enough. It is convenient, though, to be able to

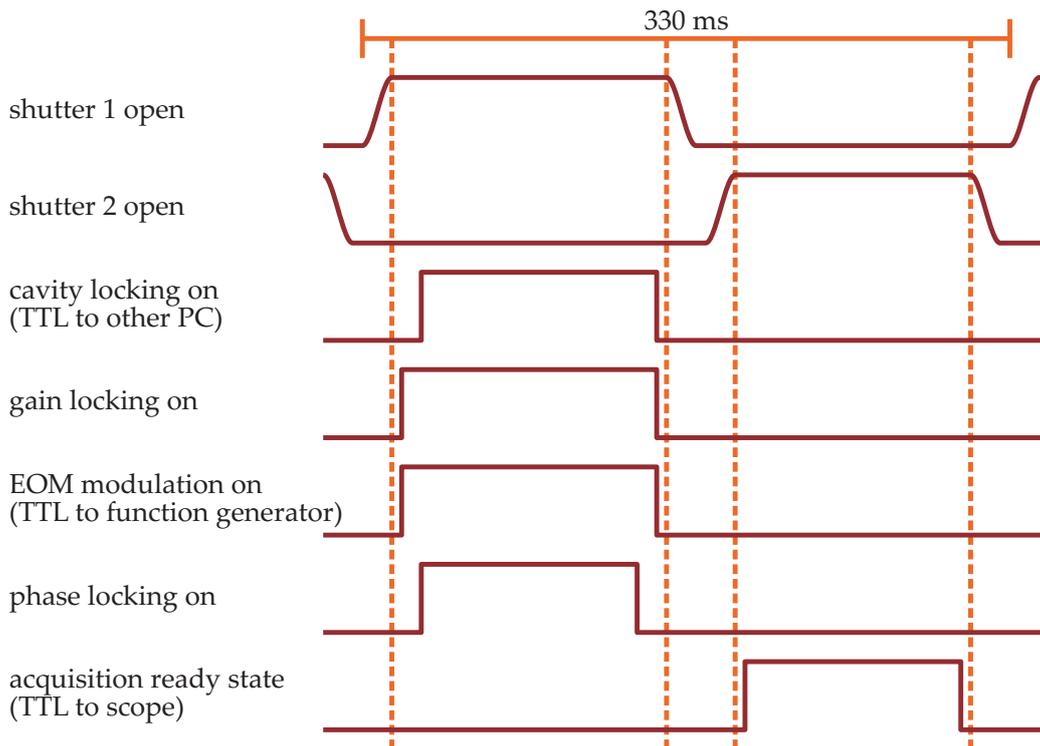


FIGURE 4.12: An example of a shutter sequence and the associated on/off signals sent to other parts of the experiment. The shutters are switched such that no strong light from the seed beam reaches the APD. Cavity locking takes place during the periods when shutter 1 is open (seed beam on), while the scope acquires data during the open time of shutter 2, when the down-converted photons can cause APD clicks. Gain and phase locking is explained in §4.5.2. We attempt to stabilize the gain briefly before turning on the cavity locks, and the EOM modulation should be on before the phase lock will work.

see in realtime, while running measurements, the levels of each detector. If the total transmission goes down, it is easier to see which cavity is causing problems than when only monitoring the total transmission on a single noisy APD.

The mechanical shutters were constructed from old hard disk drives. A hole was drilled through the metal frame and a thin piece of copper was mounted to the pivot arm. On the pivot arm is a coil of copper wire and underneath it two permanent magnets. By sending current through the coil in one or the other direction, it can be made to swing to either of its extreme positions. To avoid vibrations, pads of Sorbothane dampens the swing in each end. We drive the shutters with some quite unsophisticated electronic circuits which simply converts the input 0 V/5 V TTL signal into ± 13 V in an op-amp comparator configuration followed by a buffer amplifier. The HDD coil has a resistance of only 10Ω , so to reduce the current drawn, a 75Ω resistor is inserted in series in the output of the driver, resulting in 150 mA through the coil. The current is kept continuously at +150 mA (-150 mA) when the shutter is open (closed). More complex driver circuits can be built which has a much lower holding current, while providing a larger switching current for faster switching [94], [168]. Our simple driver switches the shutter with a delay of several seconds

after an extended period of being in the same on or off state. In an experimental run, however, we switch the shutters with short intervals, and then the delay between the change of the TTL level and the actual switching of the shutter is considerably shorter. More importantly, the delay is consistent, so it is no problem to compensate for it by just changing the TTL level earlier. We adjust the on/off times of the two shutters relative to each other by observing the APD counts from the seed beam that has been heavily attenuated. When the shutters are well adjusted, no light slips through to the APD.

4.4 Data acquisition

4.4.1 Homodyne detector

An essential component of the various nonclassical state generation experiments is the homodyne photodetector. In order to efficiently resolve the intricate details of quantum optical states in homodyne detection, it has to fulfill several important requirements: It must have a quantum efficiency close to unity, it must be fast enough to detect the high-frequency components of the field, and all classical noise in the local oscillator should be well canceled in the balancing, such that the detector is shot noise limited even for LO powers high enough that the light noise is far above the dark noise of the electronic circuit. Moreover, when the state occupies a wide band all the way from DC to several MHz, the detector must perform in the entire range.

In normal sideband measurements of cw light states, such as in ref. [137], two separate detectors can be used for the two arms, so the difference photocurrent is obtained by subtracting the two signals after the amplification stage. The balancing is then done by applying a strong modulation to the LO at the sideband frequency of interest and adjusting the gain and phase of the two signals until the modulation is maximally canceled out. This approach is less feasible for the broadband detection since all frequencies should be well balanced, which can be difficult to achieve with two separate amplification stages with their possibly different characteristics. Instead, we use a layout where the two diodes are placed close to each other on the same print board, and their photocurrents are subtracted immediately afterwards, before the amplification. There is no way of tuning the two currents with respect to each other, but if the two diodes behave the same, it is not really necessary. If the quantum efficiencies of the diodes should be slightly different, it must be compensated for optically by having a beamsplitting ratio different from 50/50. We have not seen diode asymmetries that required optical powers that were different within the uncertainty of the power meter used. The balancing is easily done by just making the DC level zero with very fine beamsplitter adjustment.

As with several other elements of the setup, there have also been two generations of the homodyne photodetector. The first generation employed two Hamamatsu S3883 silicon PIN diodes with quantum efficiencies of 90%. The electronics were not quite optimal, so the detector response was not flat over the bandwidth of the generated states. As can be seen from the noise power spectrum in Figure 4.13a),

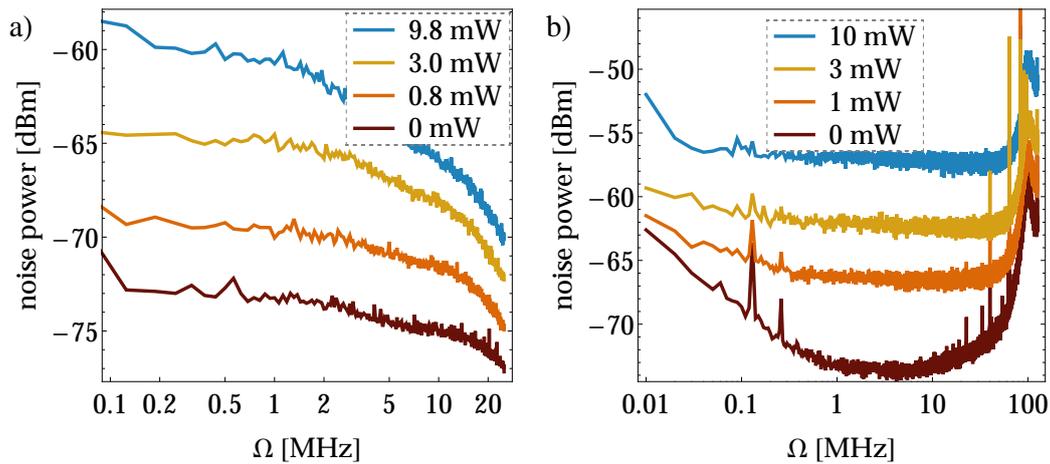


FIGURE 4.13: Spectral characteristics of the homodyne detectors; first generation, **a)**, and second generation, **b)**. The older detector is just barely fast enough to resolve the ~ 4 MHz wide conditioned states, while the newer has an excellently flat performance until the noise peak at 100 MHz, which is filtered out by the temporal mode function filtering of the recorded noise. The data in a) were measured on a spectrum analyzer, while those in b) were obtained by FFT on long traces measured on the scope with a sampling rate of 250 MS/s. Note that the bandwidths of both detectors are sufficiently narrower than the OPO free spectral range of 370 MHz, such that the homodyning detects only one frequency mode of the OPO output.

the 3 dB bandwidth was merely 10 MHz. That of course means that the difference photocurrent is a filtered representation of the measured light state. A new and significantly improved detector was developed by Hiroki Takahashi together with the electrical engineering department. Using faster diodes, faster amplifiers, and optimized circuit design, this detector achieves an almost flat spectrum all the way to 50 MHz (Figure 4.13b)). The strong wide peak appearing around 100 MHz is not a cause of worry since these high frequencies are averaged/filtered out when applying the mode function to the data. The diodes used are from a special production batch based on Hamamatsu S5971, with a specified quantum efficiency of 98%. It is difficult to quantify the amount of classical noise cancelation achieved by especially the new detector. The reason is the direct diode current subtraction layout that makes the output signal extremely sensitive to differences in the light power on the two diodes. Hence, due to detector saturation, it is not possible to block one diode to check the amount of classical noise without and with balancing ($\sim 100 \mu\text{W}$ difference saturates the amplification stage). However, we can still check the power scaling of the light noise and confirm that it is indeed almost shot noise limited up to LO powers of several mW, as demonstrated in Figure 4.14.

Regarding the optical side of the homodyne detector, we found that the polarization homodyning using 2 PBS and a half-wave plate, as described in §4.1, is far superior to using just a non-polarizing 50/50 beamsplitter in terms of stability and convenience of beam alignment and power balancing. With the latter, the angle of incidence on the beamsplitter needed to be precisely adjusted to achieve the 50/50 ratio, meaning that it was a mess to do a precise balancing and get

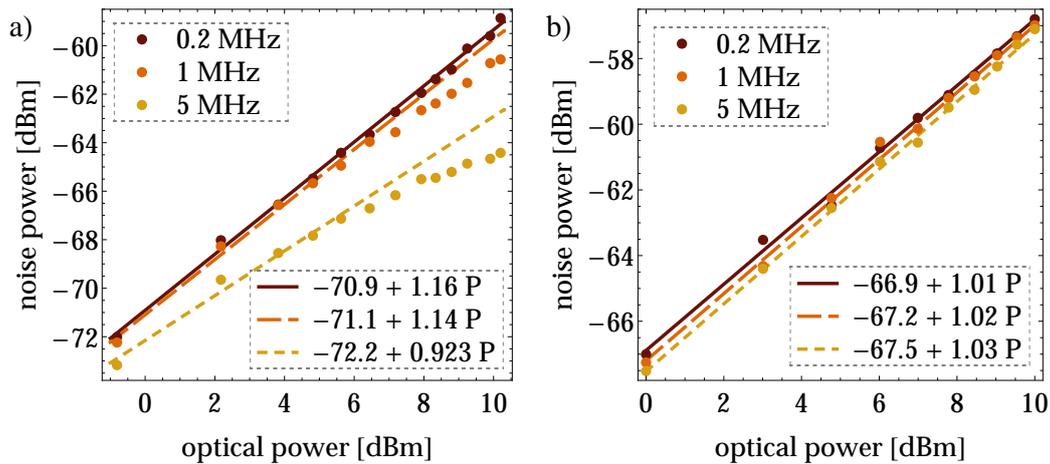


FIGURE 4.14: Power scaling of the light noise for the old, **a**), and new, **b**), homodyne detectors, at three different frequencies – the optical power is measured before the homodyne beamsplitter. The data are from the same series as those in Figure 4.13. The plots are log-log, so a shot-noise limited behaviour would be characterized by a linear trend with a gradient of 1. The new detector is almost fully shot noise limited, while the older detector apparently has not canceled out all classical noise (which would scale with a gradient of 2). We did obtain better scaling at other times, though, but the behaviour depends a lot on how well the beams are focused on the detector’s diodes. The fits in a) are only performed on the first five points, since saturation of the detector obviously limits the noise at higher powers.

a good beam overlap at the same time. During measurements, the balancing could also drift away and was quite sensitive to acoustic noise. Furthermore, the overlap of the two beams needed realignment every day. With the polarizer combination, the angle sensitivity is basically gone, and the balancing (done by the half-wave plate) and the mode-matching are effectively decoupled. The DC output level does not fluctuate as much as before, and interference visibilities as high as 98.5-99% will often remain for more than a week without need for realignment.

4.4.2 Data sampling

The homodyne output signal is recorded on a 1 GHz LeCroy Wavepro 7100 digital oscilloscope with 48 MB fast memory. A useful and important feature of this scope is its ability to perform segmented sampling, where it samples only within a specified interval (including pre-trigger times) around each trigger event and excludes all the dead time in-between – see Figure 4.15. A sequence of 20,000 of such segments can be recorded in a single combined waveform before the data needs to be transferred from memory to a file on the hard disk (which takes a few seconds for a full memory). We usually record intervals of 1 or 2 μs length, so with count rates of 5-20 kc/s, we save between 96% and 99.5% memory space by using segmenting instead of sampling continuously, making possible much longer measurement streaks between file storage. We sample at a rate of 500 MS/s, more than enough to resolve in great detail the variation of

the homodyne photocurrent caused by the detection of the trigger photon. Each segment is centered around the trigger time, that is, with 0.5 or 1 μs before and after the trigger. The temporal width of the conditioned state's mode function is only around 100 ns, but the extra space around the trigger allows us to extract the background squeezed/thermal state as well, by positioning the mode function far away from the trigger time.

When we employ the locking scheme where shutters switch on and off the strong beams, we add an extra gating condition to the triggering of the scope. The LabView program in control of the experimental sequence sends a TTL-high signal to a second input channel of the scope, whenever the shutter in front of the APD is open. By using a logic trigger setting, which demands a simultaneous high value of both this gating signal and the trigger signal from the APD, we make sure that no dark counts, appearing while the APD is blocked, are counted.

We connect the homodyne detector to the scope via a high-pass filter with a bandwidth of 10 kHz to get rid of the DC level of the signal. The reason is that the small drifts in the balanced DC level (which are still present, if at a small scale with the newer detector) will show up as time-varying features of the measured states. This is clearly undesirable. The high-pass filter ensures that the drifts in balancing are removed. Of course it also filters out proper features of the state, but with the 4 MHz bandwidth of the state, it is a negligible loss.

4.4.3 Extracting quadrature data

Every trigger click heralds the production of a single instance of the state under investigation. For each instance, a single quadrature measurement is performed by the homodyne detection, but this measurement is spread out in time to the extent of the correlation function between trigger and signal mode. To extract the single quadrature value from the continuously sampled homodyne signal, the temporal filtering with the relevant mode function must be carried out. In practice, it is done by constructing a numerical array of the same length as the segments and with the shape of the selected mode function. This mode function array is multiplied on each segment, and the sum of the product array is then the (unnormalized) quadrature value for that segment. Normalization is done by recording and extracting the vacuum state in exactly the same way, and calculating the variance of the obtained vacuum values. Figure 4.15 also shows this quadrature extraction procedure.

Some examples of recorded and extracted quadrature data are demonstrated in Figure 4.16. Three states are depicted; a kitten state, a squeezed vacuum (obtained from kitten state data with a temporal mode function shifted far away), and a single photon, all shown together with a vacuum sequence trace for reference. Each sequence consists of 20,000 points obtained with a scanned phase, calibrated as will be explained in the next section. The phase-dependent quadrature distributions of the squeezed and kitten states are clearly seen, while the vacuum and photon states are phase-independent. More interesting are the obvious non-Gaussian distributions of the kitten and single photon points, particularly visible for the kitten with its "white eye" in the center. The reconstruction

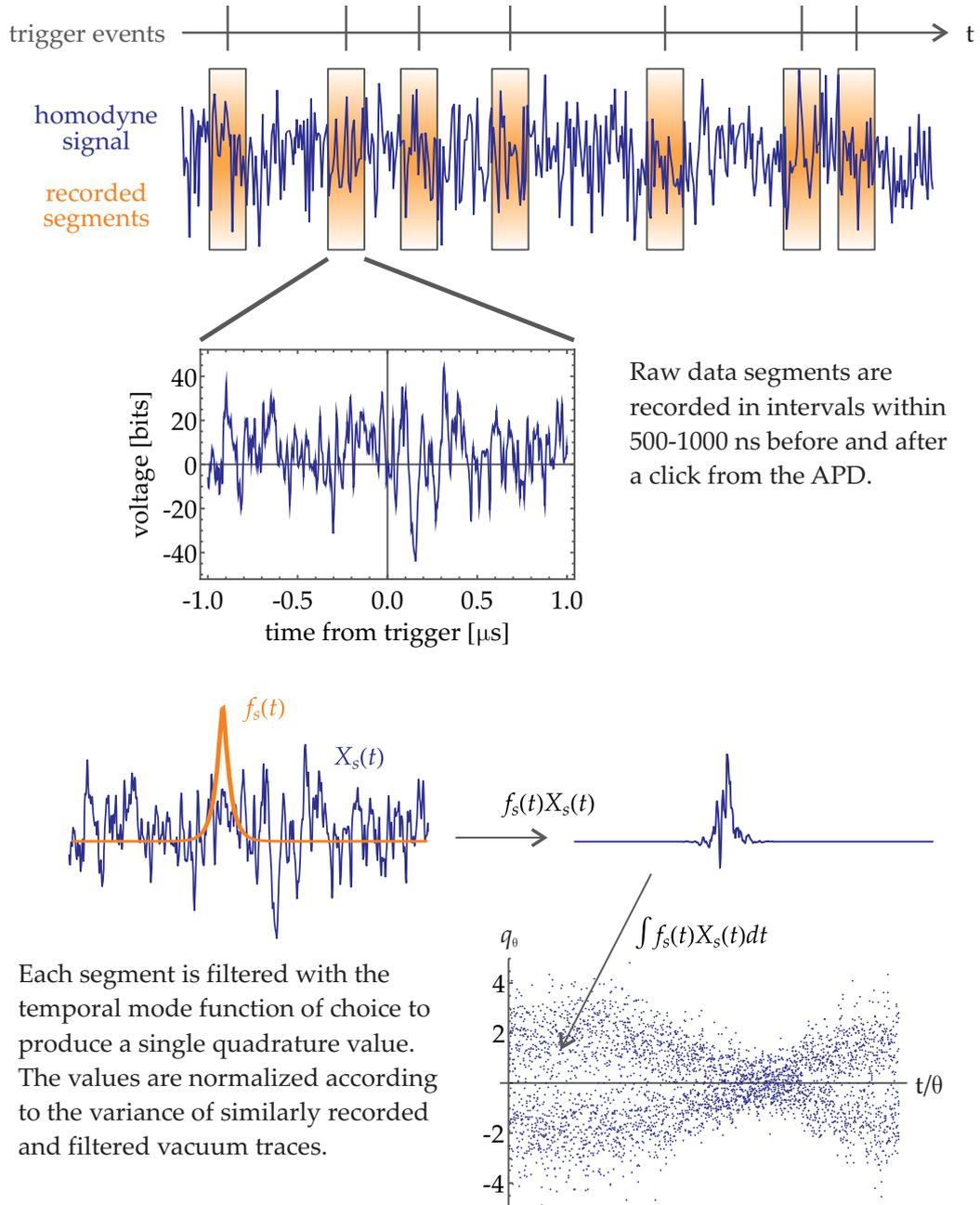


FIGURE 4.15: Illustration of the segmented homodyne sampling, performed in real time by the scope, and the subsequent quadrature data extraction from these segments using the temporal mode filtering. The latter part is done on a PC after the conclusion of the measurement series.

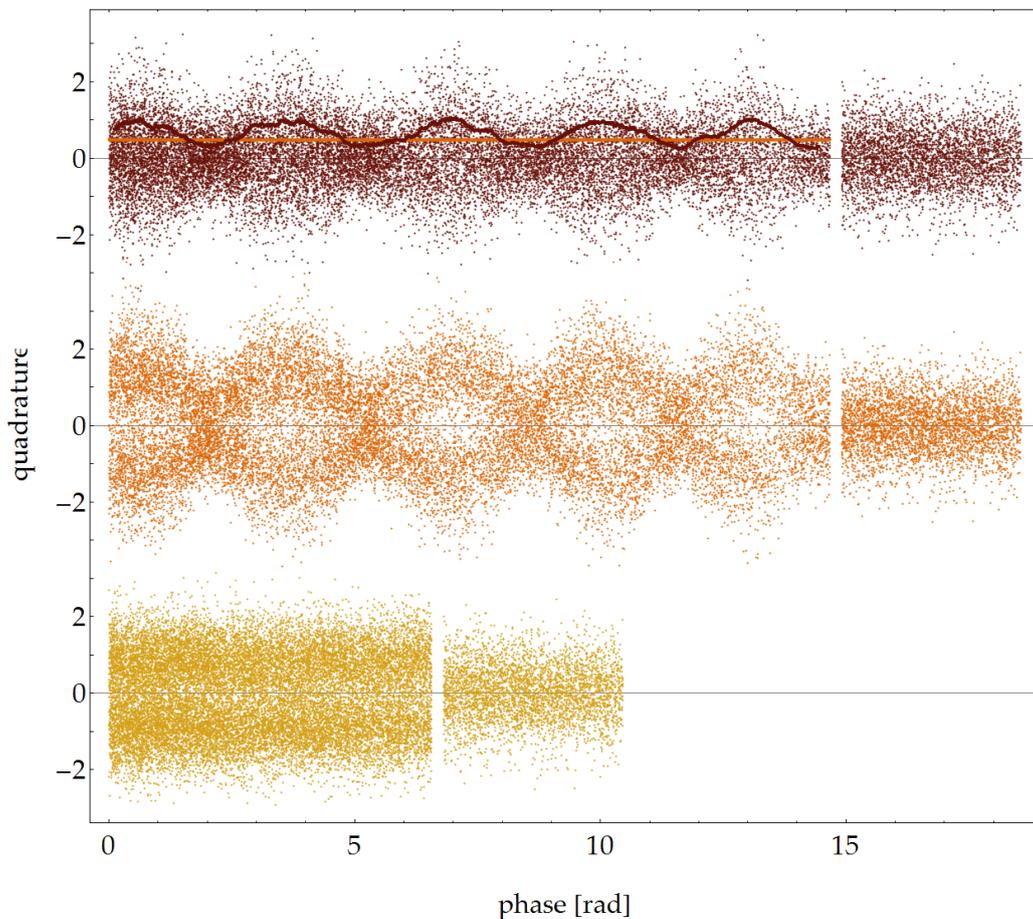


FIGURE 4.16: Examples of sequence traces obtained for three different states; from top to bottom it is a squeezed vacuum, kitten, and single photon state. The small chunks to the right of the traces are reference sequences of the vacuum state. The first two, kitten and squeezed vacuum, are extracted from the same recorded homodyne segments but with different delays of the mode functions. For the squeezed vacuum, it was moved so far away from the trigger time that there was no overlap with the temporal extent of the kitten state. For the squeezed vacuum, the variance is plotted as the brown curve (moving average of 400 points) together with the vacuum level of $1/2$. The minimum and maximum variances are 0.3 and 1.0, respectively (-2 dB/3 dB). The reason for the higher density of points at low phase values is the nonlinear scan of the piezo mounted mirror in the LO path.

of the states from these data will reveal that the regions around 0 are so vacant that their Wigner functions have negative values - even if it may not appear so for the photon sequence.

Mode function and segment noise

As discussed previously, the basic shape of the mode function should be the double-sided exponential corresponding to the OPO decay. For even better matching, the mode function can be adjusted to take account of the effect of the trigger filtering and - in the case of the old homodyne detector - the elec-

tronic detector response. The precise shape is not that critical, though. A wide range of different mode functions result in very similar states, as will be shown in §5.2.4. One thing to take care of is the horizontal offset of the mode function. The trigger TTL pulse from the APD and the homodyne current from the conditioned state do not arrive at the scope simultaneously. After the splitting of the OPO output field into signal and trigger, the signal field has a relatively short path length towards the homodyne detector, which also processes the electronic signal quickly and sends it to the scope. In comparison with that, the trigger has a rather long optical path, and on top of that, the APD module has a delay of some tens of ns. As a consequence, the interesting part of the homodyne signal appears before the trigger time on the scope. This delay must naturally be taken into account by shifting the mode function correspondingly.

To adjust the mode function prior to doing the quadrature extraction, we calculate a trace from the data which we call the segment noise. This is the point-wise variance across all segments in the measurement series. More explicitly: Each segment consists of, say, 1000 points with a mean value around 0. The segment noise trace is then a 1000 points long list where the 1st point is the variance of the 1st point of all the many thousand segments; the 2nd point is the variance of all the 2nd points, and so on. In a single segment there is no clear signature of the photon-subtracted (or vacuum-subtracted) state. However, by taking the segment noise trace, we see that there is an increase of the noise variance in a region close to the trigger, as seen in Figure 4.17. The reason is that the conditional non-Gaussian states have a higher variance than their corresponding “background states”. The shape of the variance peak gives an indication of the correlations between the trigger and the signal fields and can be used as a guide for determining a suitable mode function – most importantly, it shows quite precisely the delay between the arrival to the scope of the homodyne and trigger signals. By taking the variance over all recorded segments, we mix squeezed and anti-squeezed quadratures (in the case of the DOPO/kitten measurements) which is perfectly fine when just seeking some input for the choice of mode function. We can, however, also take the segment noises separately for squeezed and anti-squeezed quadrature segments. In that case we see that the heights of the variance peaks, as well as the background levels are different (Figure 4.17a). The background levels are just the variances of the unconditioned squeezed vacuum state. The reason why they do not depart much from the vacuum level is the wide bandwidth of the homodyne detector; the squeezed vacuum state occupies only a minor region of the detector’s spectrum, and at high frequencies the variance is equal to the vacuum’s. Since the segment noise trace takes the variance in single points (basically delta functions in time), each point includes the entire detector bandwidth and will thus be dominated by the vacuum. In the figure, a) and b) are from data recorded by the old and slow and the new and faster homodyne detectors, respectively. While a) is from a kitten measurement with a pump of $x = 0.28$, b) is from a single photon measurement where the pump was much weaker, $x = 0.12$. This is the main reason for the lower variance peak, but also the different detector bandwidths has a clear influence. For example, the background level in b) is basically equal to the vacuum level of $1/2$, although, with the given gain, the variance of the background thermal state is 0.54 at DC.

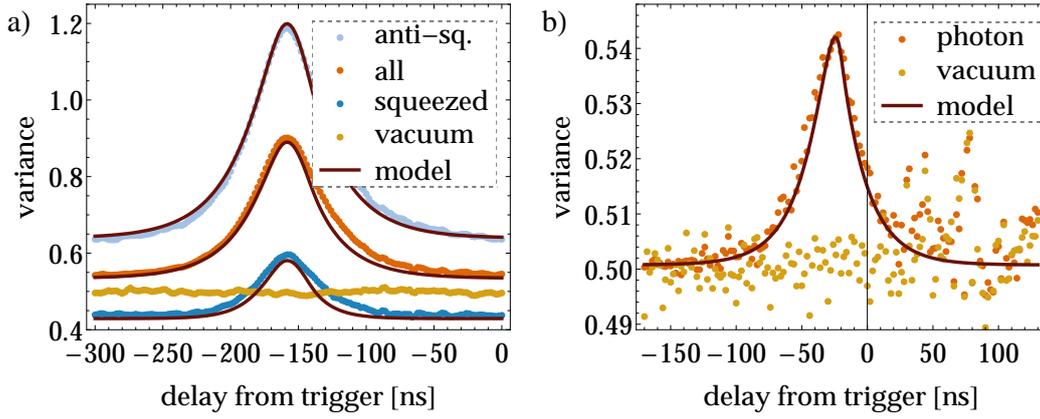


FIGURE 4.17: Segment noise traces of **a)** a kitten state measurement and **b)** a single photon state measurement. The traces show the point-wise noise variance of the recorded segments, as explained in the text. In a), the light blue points show the variance obtained from a selection of the segments recorded at anti-squeezed quadrature phases, while the dark blue points are from squeezed quadrature phases. The orange points uses all recorded segments. For the photon measurement, b), all phases have the same variance, so there is only one trace. The brown curves show the modeled variances, which fit well with the data when assuming a detector bandwidth of 25 MHz in a) and 200 MHz in b). The trigger delay is much larger in a) than the usual 30-40 ns due to some extra delay cables we inserted to try to remove additional ringing that appears after an APD click - a little bit of which can be seen in b) after the trigger at time 0.

The shapes of the squeezed and anti-squeezed segment noise traces can be theoretically modeled from the formalism in §3.4. For each point t_i of the segment noise trace, the signal mode function $f_{s,i}(t)$ should be centered around t_i , and it should have the shape of the impulse response of the homodyne detector – or some approximation thereof, like a Gaussian, a double-sided exponential, or a top-hat function with the width of the inverse detector bandwidth. The variances are obtained by calculating the expectation values $\langle \hat{X}_{s,i}^2 \rangle$, $\langle \hat{P}_{s,i}^2 \rangle$, using the Wigner function obtained in each t_i as the weight function. In Figure 4.17, this modeling was performed, and decent fits were found using a top-hat temporal mode function 25 ns wide (a ~ 25 MHz wide sinc-function in frequency space) for the kitten with the old detector, and a 3 ns wide top-hat function (~ 200 MHz) for the photon with the new detector.

4.5 Phase information

In a full homodyne state tomography measurement, the phase of the local oscillator relative to the signal field is scanned across a range of at least π radians to obtain marginal quadrature distributions at the different angles in phase space. To do a proper state reconstruction, it is necessary to have a way of determining the phase for each recorded quadrature value. We have taken two significantly different approaches to this non-trivial experimental challenge, related to the two different cavity locking methods described in §4.3.3. The phase information is less important for phase independent states such as photon number states

than it is for e.g. the kitten state with its marked phase dependence. Thus, while it is possible to obtain the phase information also for the single photon state – although not without additional complications for the second method described below – it is in practice not really necessary; one can just scan the phase arbitrarily and note that there is no change in the quadrature distributions.

4.5.1 Calibrated, continuously-scanned phase

Most of the time during this project, we have taken the approach of continuously scanning the phase of the local oscillator (by a piezo-mounted mirror) during the course of a measurement. In principle, then, each recorded quadrature point has a different phase value from the others. In the subsequent analysis, the points can then be grouped together in bins of similar phase values, or the point-by-point maximum likelihood method [89] described in §2.4.2 can be used for the reconstruction. We obtain the relative phases between the points from a calibration of the relation between the piezo scan output voltage and the measured phase of a coherent seed beam.

Before the beginning of a measurement, we plan what kind of phase scan we need. The basis is the number of segments to record in a single measurement sequence (usually 20,000), and the expected count rate of the trigger APD. These numbers give the duration of a single sequence, uninterrupted by the writing to files. For the calibration to work properly, each sequence should be contained within a single one-directional sweep of the piezo mirror, that is, it should be finished before the change of direction of the scan. Furthermore, the phase should be scanned at least by π over one sequence, such that a tomographic reconstruction can be done on each 20,000-point sequence. A typical piezo scan setting is a triangular 0.1 Hz sweep from -5 V to 5 V (before further amplification). See Figure 4.18 for an illustration. The calibration is performed by injecting a weak seed beam into the OPO, and recording with the homodyne detector its interference with the LO as the phase is scanned with this setting. The function generator output is monitored on a second channel of the scope, which is set to trigger at -4 V and to record until the turning point of the scan. The response of the piezo transducer is not quite linear, unfortunately, but that just makes a proper calibration that much more important. We record several traces of the interference and fit each of them to a generalized cosine-function model defined by

$$y(t) = y_0 + C \cos(\phi_0 + \omega_1 t + \omega_2 t^2),$$

that takes account of the nonlinear piezo response that is well modeled by the parabolic time-dependence of the phase. We then take the average of the fitted ω_1 and ω_2 values and use them for the calibrated time-phase relationship.

In the state generation measurements following the calibration, we initiate each sequence recording at the same point of the piezo sweep, by setting up an additional trigger constraint on the scope; that the function generator output should be above a certain value, otherwise it should not accept APD triggers. We take this value to be a little later than the -4 V trigger point of the calibration measurement, e.g. at -2 V, since the fit might be a little uncertain in the beginning.

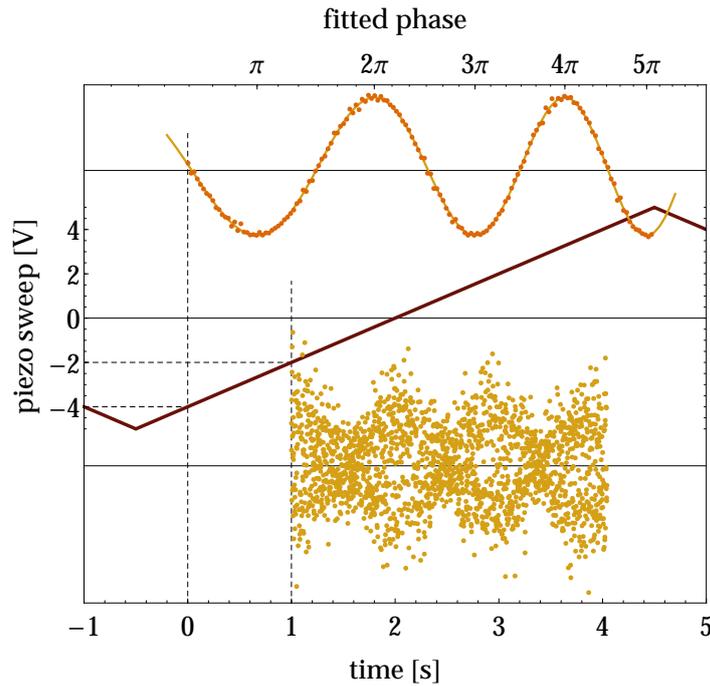


FIGURE 4.18: The scheme for phase calibration, as described in the text. The brown (center) trace is the output from the function generator to the piezo amplifier. The upper data points is a single recording of the interference between the seed beam and the LO. The phase values above it are obtained from the cosine fit to these points, the unbroken curve. The points in the bottom show a single 20,000 points sequence of a kitten state measurement, initiated when the piezo voltage is -2 V.

With the settings mentioned here, there would be 3.5 seconds until the turn of the sweep, so the count rate should be at least $\sim 6,000$ c/s. Each recorded segment of the sequence is accompanied by the time of its trigger event. This time can then be converted into a phase value with the help of the calibrated time-phase conversion function. Only relative phases within one sequence can be faithfully determined with this method. We cannot expect that the phase at the onset of one sequence is the same as that several sweep cycles later. This can be seen from Figure 4.19, where 21 phase calibration traces recorded at consecutive sweeps (i.e. with 10 seconds distance) are plotted. The general shape of the traces are very similar, but their phase offset drifts somewhat. In fact, a reasonable phase stability within a single 2-3 second sequence is also not easy to achieve. Acoustic noise as well as air flow over the optical table can easily disturb the interferometer. Since we do not make direct measurements of the phases but only infer them from the calibration, these disturbances must be limited to a minimum. We did that by covering the entire table in heavy plastic sheets hanging from the overhead shelf – a very cheap but efficient solution. Still, all movement in the laboratory has to be very slow to avoid turbulence. As an example of the achievable phase stability, the average phase value at $t = 4$ s calculated from the fits to the 21 traces in Figure 4.19 is $(3.88 \pm 0.08)\pi$. Note that this is the phase relative to at $t = 0$ s, so it excludes the phase offset drift.

Since we cannot assume all the segments to have the same global phase, we do

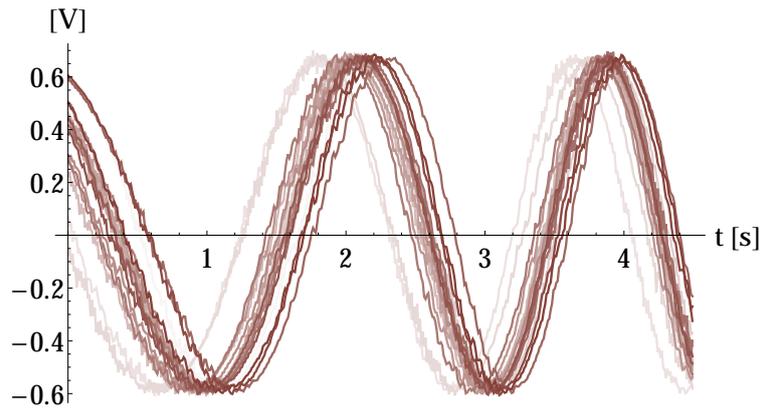


FIGURE 4.19: 21 consecutive phase calibration traces. Their order is indicated by the colour, light to dark. The fairly good short-term stability can be seen from the similarity of the shapes, while the less good longer-term stability gives the drift in phase offset from trace to trace.

individual tomographic reconstructions for each sequence/file. The resulting Wigner functions will then be noisy due to the few points measured, and they will have different directions in phase space. To make an average, all the Wigner functions must be rotated to have the same direction. There is nothing disputable about that procedure; it is just a correction for the fact that the irrelevant global phase drifts over time.

4.5.2 Phase-locked local oscillator

With the installment of the shutter-based cavity locking scheme, we have been able to take advantage of the injected seed beam for real-time phase measurements. On top of that, we can also lock the phase, such that several conditioned homodyne measurements can be performed at (roughly) the same angle in phase space.

Including the seed beam, there are three fields whose phases we must consider: The seed (ϕ_s), the down-converted quantum field (ϕ_q), and the local oscillator (ϕ_{lo}). As usual the global phase is irrelevant and the relative phases between the fields are interesting. The phase difference that we want to control is $\phi_{lo} - \phi_q$, since that defines the direction in phase space of the homodyne measurement. We can not directly observe this relative phase between the quantum field and the LO, though, because the quantum field (in our case at least) does not have a coherent component and hence does not make an interference pattern with the LO.³ That is why the seed beam becomes useful, acting as an intermediate phase reference. First we lock the seed to a fixed relative phase $\phi_s - \phi_q = \varphi$, to be kept constant throughout the measurement, and next we monitor the interference of the seed and LO and extract their relative phase $\phi_{lo} - \phi_s$, which through the first

³One could in principle use the variance as a phase indicator in the case of squeezed vacuum from the OPO. In practice it is a way too weak and noisy signal, though.

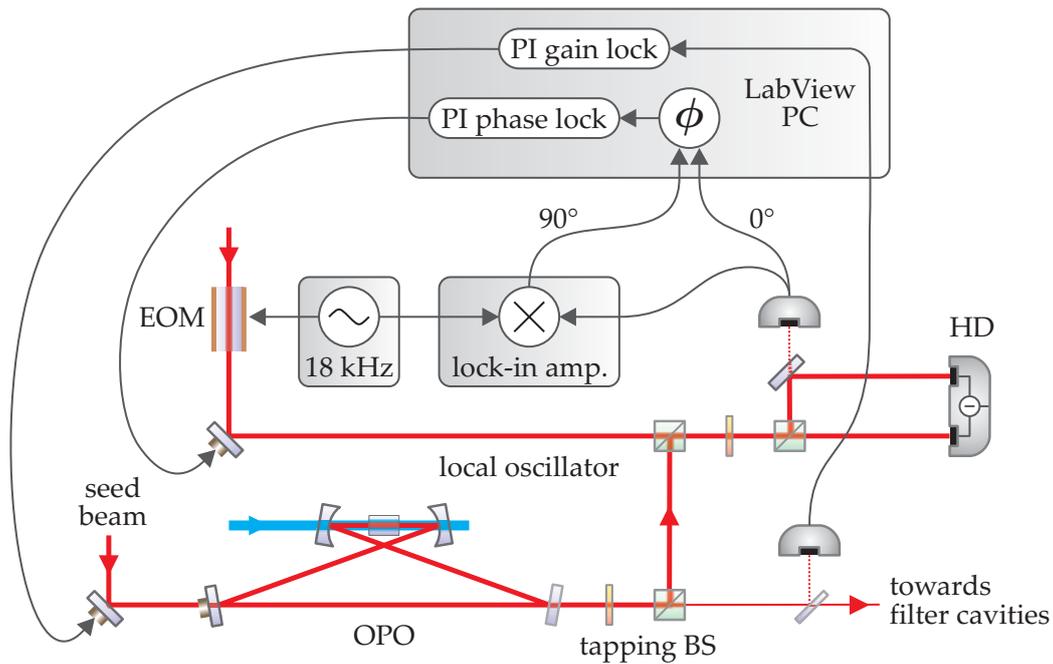


FIGURE 4.20: Scheme for phase locking with injected seed beam.

lock is equal to $\phi_{l_0} - \phi_q - \varphi$. The constant φ is just a global phase which will be responsible for the direction of the measured Wigner function in phase space.

The full phase locking scheme is illustrated in Figure 4.20. The locking of the seed and quantum field phases is done indirectly, via the parametric gain of the seed beam. The method works only when the injected seed is at the degenerate frequency, since the parametric gain of non-degenerate frequencies is phase-invariant. Hence, to make the scheme work for the photon generation experiment (which we did not try so far), we would need to inject beams of both the degenerate and the non-degenerate frequencies; the non-degenerate is necessary for the cavity locking. It is easy to lock the seed beam to a certain gain. We take out the signal for the lock from the trigger arm, after the tapping beamsplitter, to avoid additional losses of the signal field. A glass plate reflects a few percent of the trigger light onto a standard photodetector which thus sees the phase-dependent gain of the seed beam. We choose an arbitrary gain value (lower than the maximum gain) as the set point for the lock, and a simple PI routine in the same LabView program, that also runs the shutter sequence, outputs a feedback voltage to a piezo transducer on a mirror in the seed beam path before the OPO. By locking the seed to a specific gain value, its phase is at the same time fixed relative to the down-converted field, since both the gain and the down-converted phase originates from the pump beam's phase; when the classical parametric gain is maximum, the seed is in phase with the anti-squeezed quadrature, and so on. The set point of the gain lock in turn determines the value of φ .

The seed beam follows the path of the squeezed vacuum towards the homodyne detection arrangement. Here it interferes with the LO to give the overlap

intensity in one arm

$$I_{10+s} = I_{10} + I_s + 2\sqrt{I_{10}I_s} \cos(\phi_{10} - \phi_s).$$

The homodyne detector cannot be used to monitor the interference, since it saturates for the strong seed beam that is required for good signal-to-noise ratios for the other locks. Instead we have a sensitive photodetector monitoring the leakage through a mirror used to steer the beam towards the homodyne detector. The mirror is highly reflecting – about 99.9% – but the small residual leakage is enough to give a clear signal. The size of this signal indicates the phase difference between LO and seed, according to the cosine interference pattern. The phase $\phi_{10} - \phi_s$ is not ambiguously determined, though, from this single number; a given I_{10+s} value corresponds to two different phase values. Furthermore there is a large uncertainty in the phase around the maximum and minimum of the interference fringe. To get an unambiguous phase determination, we create an interference signal shifted 90° relative to the original. An EOM in the LO path before the homodyner modulates the LO phase weakly at $\Omega_{10} = 18$ kHz, which in turn causes a modulation of I_{10+s} . The interference term of I_{10+s} is then proportional to

$$\cos[\phi_{10} - \phi_s + M \cos(\Omega_{10}t)] \approx \cos(\phi_{10} - \phi_s) - \sin(\phi_{10} - \phi_s)M \cos(\Omega_{10}t),$$

where the modulation depth M is assumed small. By demodulating the 18 kHz component of the interference signal in a lock-in amplifier (i.e. multiplying with $\cos(\Omega_{10}t)$, followed by low-pass filtering), we obtain a signal proportional to $-\sin(\phi_{10} - \phi_s)$. With this sine, together with the cosine of the DC component, as inputs to the LabView program, the phase between LO and seed is calculated, and a PI loop can lock the phase to any desired value through feedback to a piezo mirror in the LO path. An example of a kitten state recorded with this phase locking is shown in Figure 4.21.

When the seed beam is continuously on, the phase can be stabilized to within $\pm 2^\circ$. During measurements, the seed must be switched on and off with the shutter, and the program must keep the output voltages to the two piezo mirrors constant during the periods without seed (see Figure 4.12 for the sequence). The lock becomes somewhat more unstable from the switching, and small phase drifts during the unlocked periods add to the increased phase uncertainty of roughly $\pm 5^\circ$.

The method outlined here for measuring and locking the phase difference (up to the global phase φ) between the LO and the down-converted field ought to be further improved for long-term reliability. As it is, it is susceptible to changes of either the blue pump power or the power of the seed beam. Before locking, the gain detector is calibrated by blocking the blue light, such that we get the $G = 1$ level. If the seed power changes or drifts during a measurement, this calibration no longer holds, and we will lock at a different gain than intended. A change in the blue power will not disrupt the calibration but will alter the gain dependence of phase ($G(\phi_p)$ from (4.6)), such that the set gain no longer corresponds to the same phase $\phi_s - \phi_q$. Setting a small value of the gain lock point reduces the effects of blue power drifts, though, since the inverse of the $G(\phi_p)$ function has

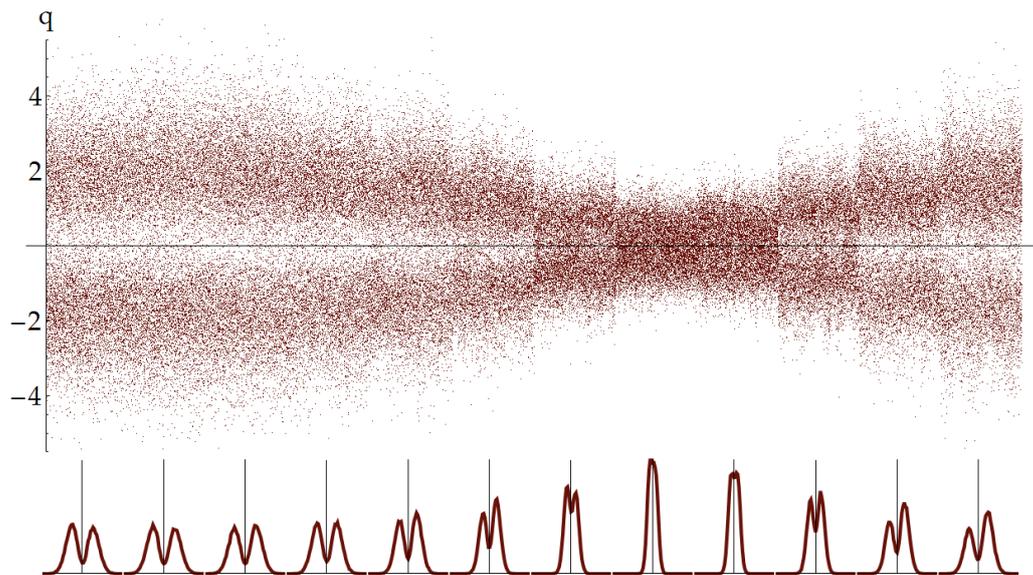


FIGURE 4.21: Kitten state with fairly high gain, $x = 0.46$, measured with phase locked local oscillator in 15° intervals. 10,000 points are shown here for each phase setting, and below each group of points is shown the histogram of the points, that is, the quadrature distribution. The squeezed phase angles are most sensitive to fluctuations and instabilities of the phase, and small glitches and drifts can be seen there if looking carefully. The asymmetry of the distributions could be due to some contamination in the signal by a coherent field – probably originating from the local oscillator.

a smaller x -gradient (blue power dependence) for small G values. Anyhow, for improved long-term phase stability, the optical powers of the seed and pump beams should be stabilized.

Data presentation and analysis

In this chapter we will analyze in detail the data obtained in the kitten state and single photon state generation experiments.¹ I have selected 8 different measurement series to be studied; 6 for the kitten state and 2 for the single photon. These were selected as representatives of different experimental conditions, and are each among the best results we have obtained for the given conditions. They are not in any way particularly “lucky shots”, though, apart from the fact that they were obtained at times when various experimental imperfections plaguing other series had been diminished. Such imperfections could be fluctuating phases, too high propagation or intra-cavity losses, detector saturation due to too tight focusing on the diodes, unstable cavity locking, etc.

5.1 Presentation of the data sets

The 8 representative data series are labeled ks1-6 for the kitten measurements and sp1-2 for the single photon measurements². The most important experimental parameters for each series are listed in Table 5.1. These parameters are the maximal classical parametric gain G (and corresponding pump parameter x), the transmittance of the tapping beamsplitter T , and the directly assessable efficiency factors. These are the OPO escape efficiency η_{esc} , propagation efficiency from OPO to detector η_{pr} , homodyne beamsplitter overlap (visibility squared) η_{hd} , detector quantum efficiency η_{qe} , and an efficiency factor η_{en} related to the electronic noise. The overall signal arm efficiency η_s is the product of these factors. η_{qe} and η_{en} are associated with the detection only and not the state generation. Hence, for a more true estimation of the generated state, these terms can be corrected for in the reconstruction procedure. Finally, the size of the data set is listed as (number of sequences) \times (segments per sequence) \times (samples per segment). In all cases the sampling rate was 500 MS/s.

¹As before, the vacuum-subtracted thermal states from the NDOPO operation will be referred to as “single photon states”, although they also contain vacuum and higher photon numbers.

²For local reference: the data set file locations are; ks1: 20060209/cat2; ks2: 20060215/cat1; ks3: 20060907/cat1; ks4: 20060907/cat5; ks5: 20080402/cat4; ks6: 20080402/cat3; sp1: 20061107/sps2;

label	G/x	η_{esc}	η_{pr}	η_{hd}	η_{qe}	η_{en}	T	data size
ks1	2.3/.34	.86	.91	.96 ²	.90	.90	.97	49×20,000×1000
ks2	1.8/.25	.86	.91	.96 ²	.90	.91	.93	50×20,000×1000
ks3	2.8/.40	.95	.96	.97 ²	.90	.96	.95	10×20,000×1000
ks4	1.9/.27	.95	.96	.97 ²	.90	.96	.93	19×20,000×1000
ks5	3.4/.46	.94	.97	.96 ²	.98	.96	.95	(12×5)×10,000×500
ks6	2.0/.29	.95	.97	.96 ²	.98	.91	.95	(12×5)×10,000×500
sp1	2.0/.29	.95	.92	.97 ²	.98	.92	-	9×20,000×1000
sp2	1.3/.12	.95	.92	.97 ²	.98	.92	-	9×20,000×1000

TABLE 5.1: Experimental parameters for the selected data sets.

The loss/efficiency η_{en} due to the dark (electronic) noise of the homodyne detector circuitry requires some explanation. This dark noise clearly contaminates the measured light noise. In standard sideband squeezing measurements, the noise power (variance of the photocurrent) is observed. Since the light noise and the electronic noise are uncorrelated, they add in quadrature (assuming both are normal distributed), so the electronic noise can be directly subtracted from the overall noise to obtain the noise variance originating from only the light – this was done for the squeezing trace in Figure 4.8. When measuring the noise amplitude instead of the variance, which we do in the time domain detection, the electronic noise can not so easily be deducted, and it becomes an integral part of the state statistics. As shown by Appel et al. [3], the effect is equivalent to the addition of yet another virtual beamsplitter with efficiency (transmittance)

$$\eta_{\text{en}} = 1 - \frac{\langle \Delta N_{\text{el}}^2 \rangle}{\langle \Delta N_{\text{vac}}^2 \rangle},$$

where ΔN_{el}^2 and ΔN_{vac}^2 are the variances of the recorded electronic (dark) noise and vacuum noise, respectively. Here, the recorded vacuum noise includes both the light and electronic noise. This efficiency term can clearly be improved by increasing the local oscillator optical power and thereby the vacuum variance, but, as we have seen in §4.4.1, a too high power will lead to detector saturation and correspondingly decreased efficiency. The noise variance ratio in the formula above can be measured for different frequencies on a spectrum analyzer, but what really matters is the ratio within the chosen temporal mode function. Therefore, we always record a sequence of the electronic noise as well as one or more vacuum sequences, such that we can evaluate this ratio. Different mode functions will include the spectral components with different weights, so the η_{en} values listed in Table 5.1 are only guiding numbers, obtained from a simple double-sided exponential $f_s(t)$ with the decay rate of the OPO (which, by the way, is $\gamma/2\pi = 4.6$ MHz for ks1 and ks2, and $\gamma/2\pi = 4.2$ MHz for the other series). The values are also upper estimates only; in the case that some classical noise is present in the local oscillator, such that the detection is not fully shot noise limited, the $\langle \Delta N_{\text{vac}}^2 \rangle$ term will overestimate the actual shot noise variance.

Notes on the measurements

The following are brief comments on the 8 selected measurement series and explanations of the difference of the parameters in Table 5.1. Pairwise (ks1+2, ks3+4, ks5+6, sp1+2), they represent basically identical experimental conditions, with a high gain and a low gain measurement in each case.

- Series ks1-2 used potassium niobate (KN) as the down-conversion crystal, while the others used PPKTP, giving much better OPO escape efficiencies.
- Series ks1-4 were measured on the old homodyne detector with low quantum efficiency and somewhat slow response, while ks5-6 and sp1-2 used the new improved detector.
- Series ks5-6 were carried out with the new phase-locked acquisition scheme. The phase intervals were 15° over a range of π radians (12 different angles), and for each phase, 12 data sequences of each 10,000 segments were recorded. Each segment was made shorter, spanning only $1 \mu\text{s}$, since the $2 \mu\text{s}$ span of the other series were a little excessive.
- The propagation efficiencies of ks1-2 were low because of an unnecessarily involved path towards the beamsplitter overlap with the local oscillator - a path which was optimized in later measurements. The low propagation efficiencies for the single photon series were also due to a long path towards the homodyne detector. This path was more difficult to make efficient, since the ω_+ signal mode should be reflected on filter cavity 1 first.
- The visibilities of the homodyne beamsplitter overlap are unfortunately all rather low. Within recent times we have regularly achieved 99% or higher visibility, but in the measurements we ran with such high visibilities, there always seemed to be other problems, resulting in non-ideal data.
- The gain values stated for the photon series are the phase sensitive degenerate parametric gains, just as for the kitten series. They were measured by injection of a degenerate frequency probe instead of the usual frequency shifted probe beam for cavity locking.

Expected Wigner functions

To get a first impression of the kind of states we should expect to see from the data, theoretical Wigner functions based on the parameters from Table 5.1 are plotted in Figure 5.1. They are calculated from the formalism in §3.4, with the simple exponential mode function $f_s(t) = \sqrt{\gamma}e^{-\gamma|t|}$ (remembering that this is not necessarily the optimal choice, especially for high gain levels). Apart from the parameters already listed, the value of the trigger path efficiency is set to $\eta_t = 0.1$ – it may not have been exactly this number in all the measurements, but the parameter has only little influence on the outcome, so an approximate number is fine. Furthermore, the trigger filter bandwidth is set to $\kappa/2\pi = 12$

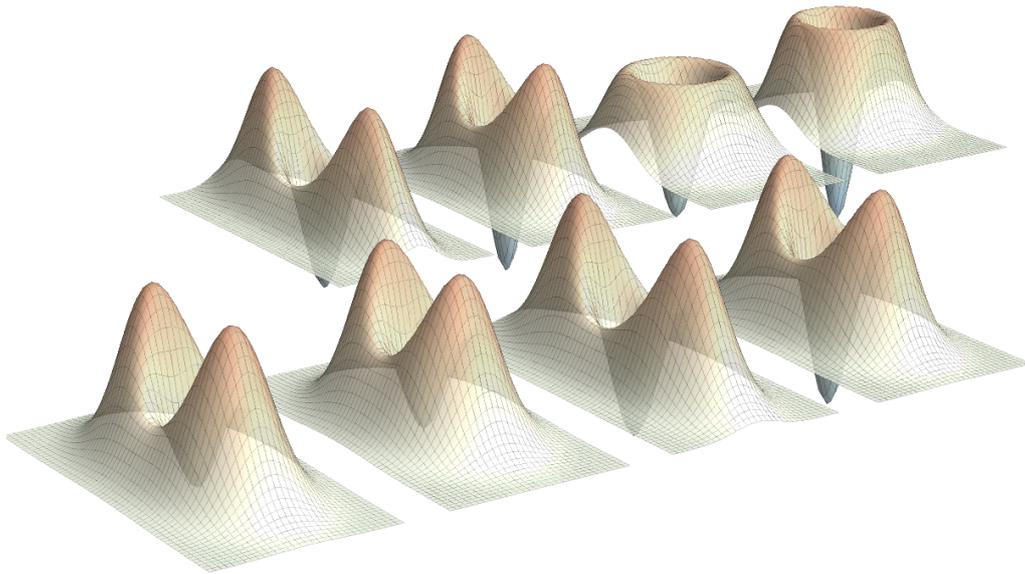


FIGURE 5.1: Theoretically expected Wigner functions for the 8 data series. Front row, left to right: ks1-4; back row, left to right: ks5-6, sp1-2.

MHz. The model developed in §3.4 assumed just a single Lorentzian frequency filter with bandwidth κ , but the combination of the three filter cavities with bandwidths 25/25/75 MHz can be roughly approximated by a single filter with a bandwidth around 12 MHz.

From the plotted functions we see that we should hardly expect any negativity from the raw data of ks1 and ks2, due to their low η_{esc} and η_{qe} values. From the other states, however, regions around the origin of negative Wigner function values should be obtained. The ks3 and ks5 states clearly have a higher amplitude than the other kitten states because of their higher pump levels. For the two single photon states, ps2 should be a bit more pure than ps1, due to the lower gain and therefore lower 2-photon contribution.

5.2 State reconstruction

5.2.1 Introductory remarks

We use the iterative maximum likelihood method, described in §2.4.2, to estimate (reconstruct) the density matrix – and from that, the Wigner function – of the state underlying the observed quadrature distributions. In all calculations we truncate the Hilbert space at a maximum photon number $M = 15$.

Some of the Wigner functions will, for clarity, be presented as contour plots instead of the 3D surface plots. An example is shown in Figure 5.2, together with the colour legend. The plotted values are π times the real Wigner function value. Each contour interval corresponds to a change in $\pi W(x, p)$ of 0.05. The plot range is always $[-4 ; 4]$ in x and $[-2 ; 2]$ in p – the same goes for the 3D surface plots.

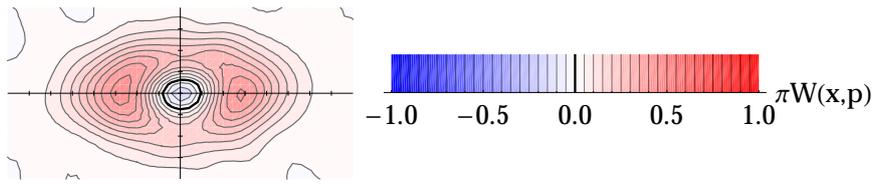


FIGURE 5.2: Example of a Wigner function contour plot, together with the colour legend. The plot range goes from -4 to 4 in x and from -2 to 2 in p . The bold contour defines the zero-crossing around the origin.

5.2.2 Investigations of the reconstruction method

As mentioned previously, the reconstruction procedure can be performed with each individual phase/quadrature data point (θ, q_θ) as input, or the data can be binned into a 2-dimensional histogram of phase and quadrature before starting the iteration. For the phase locked measurements, ks5 and ks6, it only makes sense to use the binned approach, since they are already recorded in phase bins. The resolution of the quadrature binning, dq , can then be made sufficiently small to catch the essential shape of the distributions. For the other measurement series with scanned phase, the point-by-point approach is more natural. It is, however, slow compared to the binned method. For a typical sequence of 20,000 samples, each iteration step of the point-by-point approach requires the evaluation of 20,000 weighted projectors, $(1/\text{pr}_i)\hat{I}_i$, in the expression for the iteration operator $\hat{R}(\hat{\rho}^{(k)})$ (2.30). The projectors themselves are fixed for a given data set, so they are calculated before starting the iteration, but the weight factors, $\text{pr}_i = \text{tr}[\hat{I}_i\hat{\rho}^{(k)}]$, must be evaluated in each step. In contrast, for binned data with $\#\theta$ phase bins and $\#q$ quadrature bins, only $\#\theta\#q$ weighted projectors, $(f_{\theta,j}/\text{pr}_{\theta,j})\hat{I}_{\theta,j}$, must be evaluated for each $\hat{R}(\hat{\rho}^{(k)})$ in (2.29). With a reasonable binning of $\#\theta = 12$, $dq = 0.1$ ($\#q \sim 100$, depending on the size of the state), it is on the order of 10 times faster to use binned than single-point data for a sequence of 20,000 samples. For larger data sets it is even faster – such as ks5 and ks6 where all 60 sequences easily can be lumped together in one big 2D histogram. Although the data set is larger, the dimensions of the histogram are the same, so the iteration step time will not increase.

Convergence and binning

For the sake of speed of the reconstruction, the phase-scanned data series should be binned as well. One can imagine, though, that this binning causes a loss of information and therefore gives less correct results. Another relevant question is how many iterations are necessary before the algorithm has converged to a final estimate of the density matrix. As for the latter, Figure 5.3 plots the change in the density matrix estimate per iteration step. The distance measure used is the trace distance [113, §9.2]

$$\delta(\hat{\rho}_1, \hat{\rho}_2) = \frac{1}{2} \text{tr} |\hat{\rho}_1 - \hat{\rho}_2|,$$

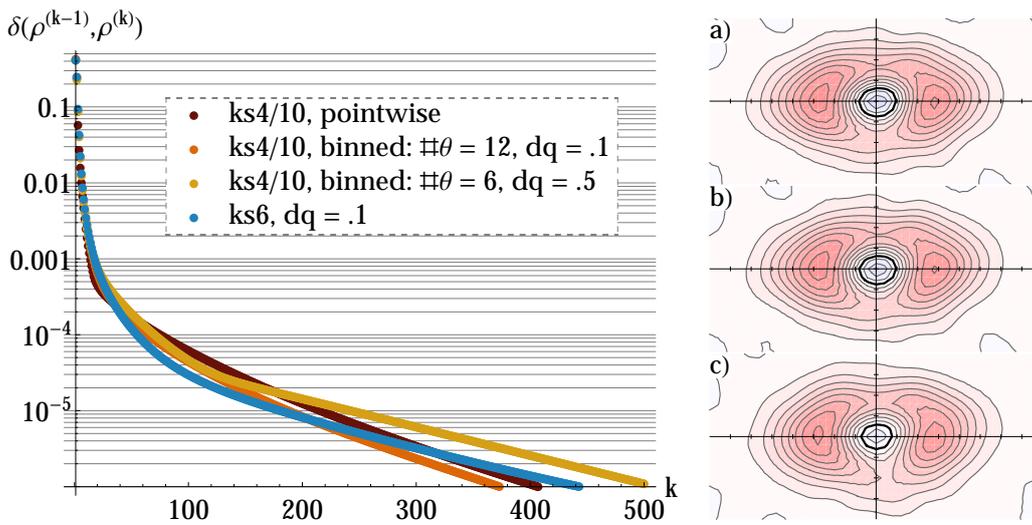


FIGURE 5.3: The trace distance between consecutive state estimates, showing the convergence of the iteration algorithm towards a fixpoint. k is the iteration step number. The brown, orange and yellow values are from a standard 20,000-samples sequence; one reconstructed point-by-point (brown), and the other two with different binnings. The blue values are for a full phase-locked data set of 600,000 samples, but otherwise in bins similar to those for the orange points.

FIGURE 5.4: Wigner function contour plots for a single sequence of ks4, reconstructed with **a)** the point-by-point procedure, **b)** the finest binning from Table 5.2 (48/.1), and **c)** the coarsest binning (6/.5).

where $|\hat{\rho}| = \sqrt{\hat{\rho}^\dagger \hat{\rho}}$, and the square-root can be taken as the Cholesky decomposition. As the number of iterations increases, the change from step to step becomes smaller and smaller, so the reconstruction converges towards a fixpoint. The necessary number of iterations apparently do not depend much on the binning method – for a given 20,000-samples sequence, the plot in Figure 5.3 is essentially the same for the point-by-point and binned data (but remember that the time per iteration is much longer for the point-by-point approach). The convergence of the reconstruction of a full phase-locked data series is also plotted, and there is little difference among these different types of data and histograms.

The question of how finely the phase-scanned data should be binned in order not to lose too much information is addressed in Table 5.2. Here, the same ks4 sequence as before was reconstructed with 12 different binnings, as well as point-by-point. In each case, the iteration was halted when it reached a breakpoint of 10^{-5} in the trace distance between consecutive density matrix estimates. The table shows for each bin setting the trace distance between the density matrix reconstructed with that setting and the density matrix obtained from the point-by-point setting, as well as the $\pi W(0,0)$ values. The Wigner function values are almost immune to changes in the phase binning, but is deteriorated a bit for the large $dq = 0.5$ quadrature bins. This is not strange, since the origin of phase space has no phase – the $W(0,0)$ value depends only on the photon number distributions, not the off-diagonal density matrix elements containing the phase information. On the other hand, the trace distance to the point-by-point result becomes significantly smaller with more phase bins but depends only weakly

dq	$\#\theta$			
	6	12	24	48
0.5	.161	.089	.052	.041
	-.085	-.081	-.082	-.082
0.2	.163	.087	.045	.025
	-.136	-.135	-.137	-.138
0.1	.163	.087	.045	.024
	-.140	-.140	-.140	-.142

TABLE 5.2: The effect of different settings for phase- and quadrature binning. A 20,000-samples sequence of ks4 was reconstructed with iterations until a breakpoint of $\delta(\hat{\rho}^{(k-1)}, \hat{\rho}^{(k)}) = 10^{-5}$. With the point-by-point method, the estimated Wigner function has a value in the origin of $\pi W(0,0) = -.145$. For the various bin settings, the corresponding $\pi W(0,0)$ value is listed as the second number (negative). The first number of each entry is the trace distance to the point-by-point estimated state – which we must assume is the most accurate estimate on basis of the given data.

on the quadrature binning. It is fair to assume that the point-by-point approach gives the most accurate reconstruction based on the available data, and clearly the binning in both θ and q must be rather finely meshed in order to make the binned results very similar to the point-by-point results. We are talking about minor details, though: Figure 5.4 shows three of the reconstructed Wigner functions – the point-by-point one, the one obtained with the finest binning in both θ and q , and the one with the coarsest binning of those in Table 5.2. In particular the first two are practically impossible to recognize from each other, while the last one, with the $\#\theta = 6/dq = 0.5$ setting, is pretty close as well. Hence, as a compromise between speed and information, we choose the setting $\#\theta = 24/dq = 0.2$ for most of the following reconstructions.

Initial density matrix

So far, all the iterations have been started with the identity matrix (normalized) as the initial density matrix estimate $\hat{\rho}^{(0)}$. In fact, the choice of the initial estimator has basically no influence on the outcome of the reconstruction. It seems to be important, though, to have non-zero elements in the diagonal – otherwise some numerical errors accumulate, and after some steps, the iteration procedure becomes unstable (at least for our implementation). It is, for example, impossible to reach a result with the vacuum state as input, but if all the other diagonal elements have just a tiny non-zero component, the procedure converges nicely. This is seen in Figure 5.5 where the convergence from such a vacuum state is plotted together with that from the identity matrix and from a randomly generated density matrix with non-zero entries in all the off-diagonal entries as well. The plot for the vacuum state has a couple of small kinks in the beginning, but otherwise it is very similar to the other two (and most other possible initial states). Furthermore, the resulting estimated states with the three different $\hat{\rho}^{(0)}$ are basically

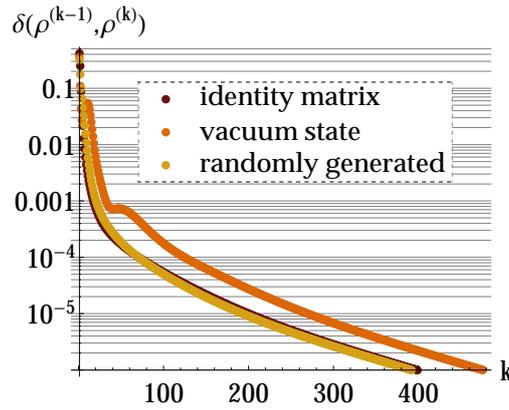


FIGURE 5.5: Convergence of the reconstruction with three different initial density matrix estimates. The vacuum state has a value of 10^{-6} in each diagonal element apart from the ρ_{00} which is 1. The sequence was the same as that used in the previous figures.

identical; the maximum trace distance between any pair of the three is 3×10^{-6} .

Test on artificial data

To further justify the use of the maximum likelihood algorithm and demonstrate its ability to reconstruct to a high precision an observed quantum state, we perform a test on artificially generated quadrature data, where we know a priori the true state. As the test case, we choose an arbitrary density matrix resulting from the reconstruction of one of the phase-scanned sequences. The marginal quadrature distributions are calculated for 100 phase angles between 0 and π , and for each phase angle, a number of simulated data points are generated randomly according to the relevant distribution. The phase angles are so closely spaced that they simulate a phase-scanned measurement.

10 short sequences with 20,000 points each and a single long one with 600,000 points are simulated and reconstructed after binning into histograms with either $\#\theta = 24/dq = 0.2$ or $\#\theta = 48/dq = 0.05$ and using different iteration breakpoint settings (the most detailed of these iterations takes a really long time!). The trace distance to the true state, as well as the $\pi W(0,0)$ values are summarized in the following table, where the numbers given for the 20,000 samples are the mean of the 10 sequences.

		$\#\theta = 24/dq = 0.2$	$\#\theta = 48/dq = 0.05$
		$bp = 5 \times 10^{-6}$	$bp = 5 \times 10^{-7}$
20,000 samples	$\delta(\hat{\rho}_{\text{est}}, \hat{\rho}_{\text{true}})$.061	.054
	$\pi W(0,0)$	$-0.136 \pm .024$	$-0.145 \pm .024$
600,000 samples	$\delta(\hat{\rho}_{\text{est}}, \hat{\rho}_{\text{true}})$.037	.022
	$\pi W(0,0)$	-0.128	-0.136

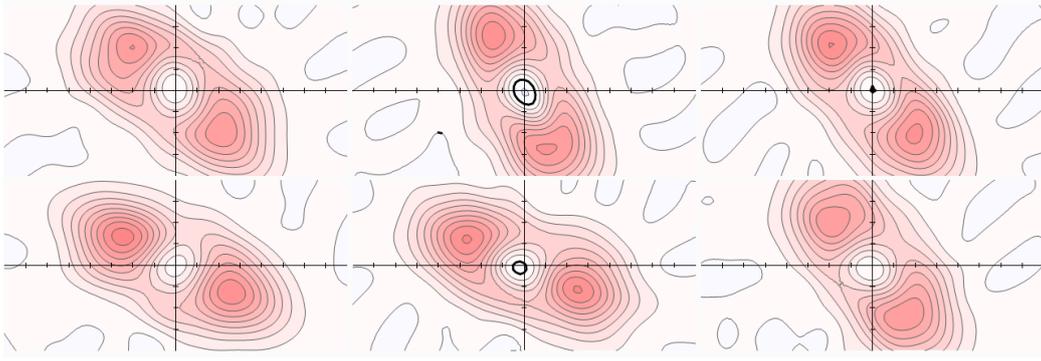


FIGURE 5.6: Reconstructions of the first 6 of the sequences in the ks1 data set. Their directions in phase space vary a bit because of drifts in phase between their respective measurements. To be able to estimate an overall average from the series, the states must be rotated to have the same direction.

For the true test state, $\pi W(0,0) = -.138$. We see that the agreement naturally gets better for the improved binning, but just as importantly, that more samples with better statistics give a more accurate state estimation. This is of course no big surprise, but it reminds us that a possible mismatch between a reconstructed and a true state is not necessarily due to the algorithm, but can also come from the data themselves. The fact that the standard deviation of the Wigner function value for the 10 short sequences is the same for the detailed and the rough estimation shows that this uncertainty is inherent in the data – basing a state estimate on only 20,000 samples introduces more error in the state estimation than choosing a slightly rougher binning for the reconstruction. Finally, another testament to the closeness to the true state of the detailed reconstruction of the 600,000 samples sequence is the fact that the difference of their Wigner functions nowhere in phase space exceeds 0.02 – simply looking at their Wigner functions it is very hard to see any difference.

It is clear that the iterative maximum likelihood method gives a quite good estimate of the observed state. The precision of the reconstructed state is mostly limited by the number of data samples and the time available for the fairly heavy numerical algorithm. A finer binning and more iteration steps allow for a better estimate, but require proportionately longer time.

5.2.3 Full reconstruction with simple exponential mode function

Having established the robustness of the maximum likelihood reconstruction method, it is time to have a closer look at the recorded data and estimate their states. For the most direct comparison to the theoretical expectations from Figure 5.1, we start out with the simple exponentially decaying temporal mode function with decay rate equal to γ . For the iteration, we use the binned approach with $\#\theta = 24/dq = 0.2$ (for ks5-6, $\#\theta$ is naturally 12, from the nature of the data), and halt the iteration at a breakpoint of 5×10^{-6} .

As explained in §4.5.1, the phase offset is not the same from sequence to sequence in the phase-scanned measurements. Therefore they can not immedi-

ately be reconstructed as one big data set, as is the case for the phase-locked series. Figure 5.6 shows individual reconstructions of the first 6 out of the 49 sequences of ks1. They are clearly not oriented in the same direction, so to get a state estimate based on the entire data set, each of the individual reconstructions must be rotated to have the same angle, such that an average can be taken. This rotation can be done either before or after the reconstruction. If done before, the phase values can be shifted by a suitable amount to get the maximum variance at 0 degrees. If done afterwards, the density matrix elements can be multiplied with a phase factor corresponding to the angle of the state in phase space. The fastest way is to rotate first, since the reconstruction then needs to be done, not many times, but only once on the histogram containing all the data.

In Figure 5.7, Wigner functions of all of the eight reconstructed states are plotted, together with the theoretical expectations (same as in Figure 5.1). Also shown are the reconstructed states obtained by shifting the mode function $f_s(t)$ away from the time interval where the trigger and signal modes are correlated. These results are then the “background states”; squeezed vacuum and thermal states. When comparing the measured with the expected states, it is first of all clear that the overall shape and dimensions of the Wigner functions match very well, both for low gain and high gain kittens and for the photon states. A single exception is ks5, which is a bit smaller than expected – most likely, the gain has changed a little from when it was measured to the time of the state generation measurement, or the $G = 3.4$ was a bit off. Another detail is that the two phase-locked states, ks5 and ks6, seem a little wobbly in the phase. They do not have the smooth rounded shapes that ks1-4 do. This indicates that either a) the phase fluctuations while locked are too large, or the phase is not determined precisely enough, or b) 12 different angles in phase space is too little for a high azimuthal resolution of the reconstructed Wigner function. Whether a) or b) is the dominant cause needs to be investigated further, but probably a) is the best guess, since the phase locking system is not yet fine tuned – and on top of that, not so much care was taken in keeping air flow down in the ks5-6 experiments as in the others.

Although the general shapes of the measured states are very close to expectations, the dips of the Wigner functions are not quite as deep as they should be. This particular feature of the Wigner function is extremely sensitive to losses and impurities, as we have seen before, so the discrepancy between data and theory indicates that there are further experimental inefficiencies apart from those already accounted for. The following table lists the additional efficiency factors needed in the theoretical models to bring the expected $W(0,0)$ values on par with those observed. Another adjustable parameter of the theory that has not been included yet is the modal purity \mathcal{E} . The table also lists the \mathcal{E} factor that would lead to agreement between theory and experiment:

	ks1	ks2	ks3	ks4	ks5	ks6	sp1	sp2
extra η_s	.96	.92	.95	.93	.80	.86	.90	.90
\mathcal{E}	.97	.93	.96	.94	.85	.87	.92	.91

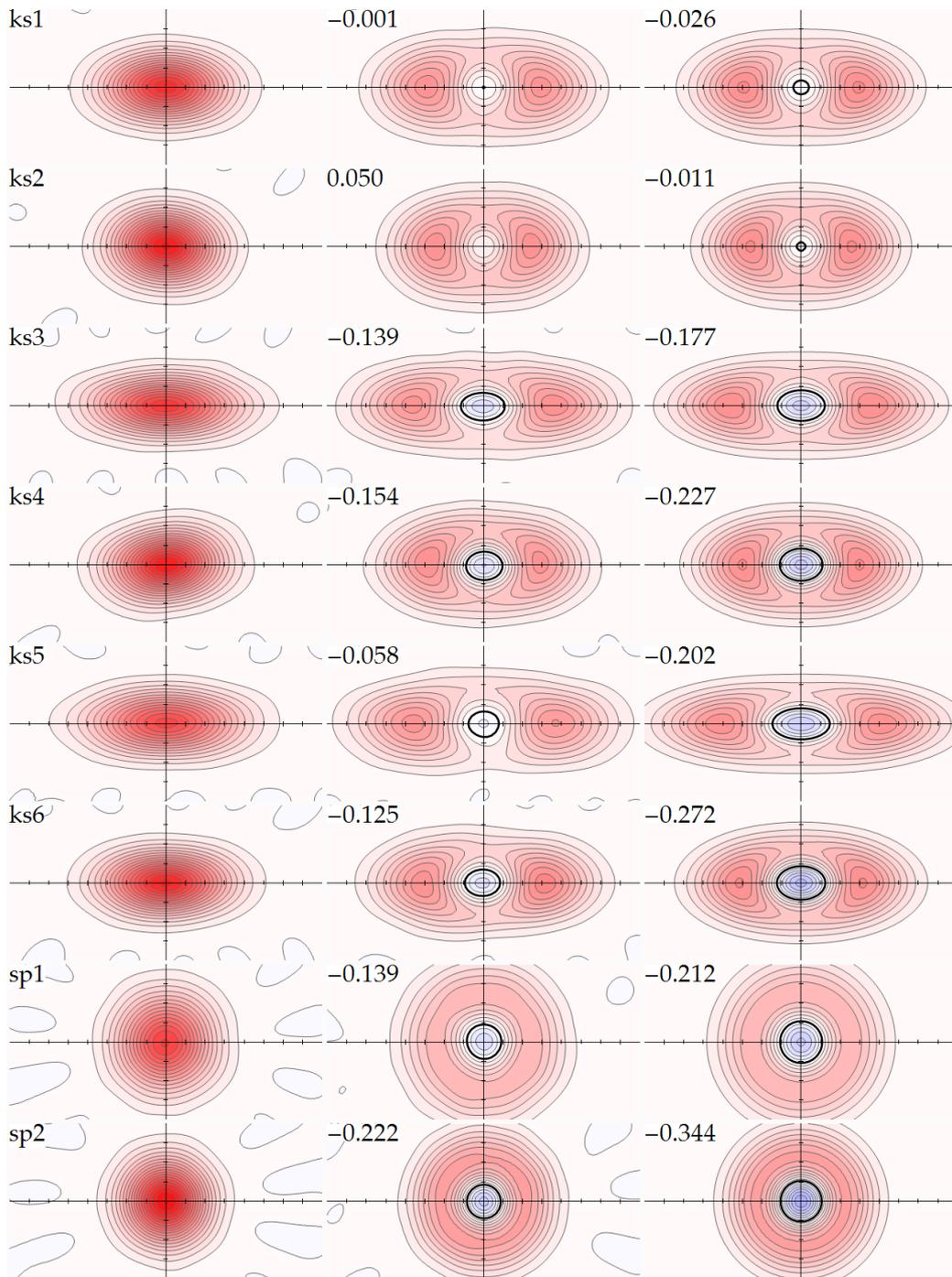


FIGURE 5.7: Wigner functions for the 8 states. In the first column are the reconstructed “background states”, that is, squeezed vacuum and thermal state, while the second column shows the corresponding conditioned states. These should be compared with the theoretical expectations in the third column. The numbers are the $\pi W(0,0)$ values.

It is hard to say with certainty what the sources of the discrepancies are, and also whether they are primarily η_s terms or Ξ terms. The latter describes dark counts and false clicks from uncorrelated photons slipping through the filters. There is certainly a small amount of dark counts in each measurement, corresponding to $\Xi = 0.98 - 0.99$, but we do not know for certain whether some unintended photons manage to make the way through the filters. In theory there should not be, but there might be some unfortunate cavity mode overlaps, e.g. due to imperfect mode-matching. It should be possible, in principle, to estimate the amounts of Ξ versus η_s , since they influence e.g. the segment noise variance or the Wigner function itself in different ways. The difference is rather small, however, so it is almost impossible to distinguish between the few percent that we are talking about – especially due to the uncertainty on other parameters of the model.

For the worst series, ks5 and ks6, there are a couple of possible sources of the additional inefficiency: At the time we made those measurements, the EOM modulating the SHG pump beam for the Pound-Drever-Hall lock was located before the point, where the beam for the local oscillator is split off. Thus, there was also a huge 18 MHz modulation on the local oscillator – too large to be balanced out in the homodyne detection. Most of this classical noise is filtered out by the temporal filtering, but not all, so the detection was not completely shot noise limited. Also for the other measurement series, some uncanceled classical noise could be one cause of additional losses. Another cause could be saturation of the homodyne detector. Especially for ks5, which is more degraded than ks6, this is a plausible bet, since the only difference between ks5 and ks6, apart from the higher gain, was that we increased the power of the local oscillator (as is reflected in the different η_{en} factors in Table 5.1).

5.2.4 The choice of mode function

One of the most interesting aspects of the combination of the continuous-wave light with the discrete-time state conditioning is the localization of the conditioned state – i.e., the question of which signal mode function $f_s(t)$ to employ, in order to extract as much of the state as possible. Something, perhaps more important than finding the optimal mode function, is to investigate different non-optimal shapes, and find to what extent they preserve the non-classical properties of the state. This is particularly important for further processing of the states, for example for storage in atomic memories, where the local oscillator might have to be optically shaped before the overlap with the generated state. It may very well be infeasible to shape it according to an optimal mode function – with regard to both the optical shaping method and the requirements for pulse shaping for the memory. It is outside the scope of this work to theoretically evaluate the optimal function, cf. the discussion in §3.4.1, but in the following we will study the influence of different mode function choices on the outcome of the measured and reconstructed states. The benchmark will be the $W(0,0)$ value. Because of the focus on this number, the data will be finely binned in the quadrature ($dq = 0.05$) and coarsely in the phase ($\#\theta = 6$ or 12). The iterations

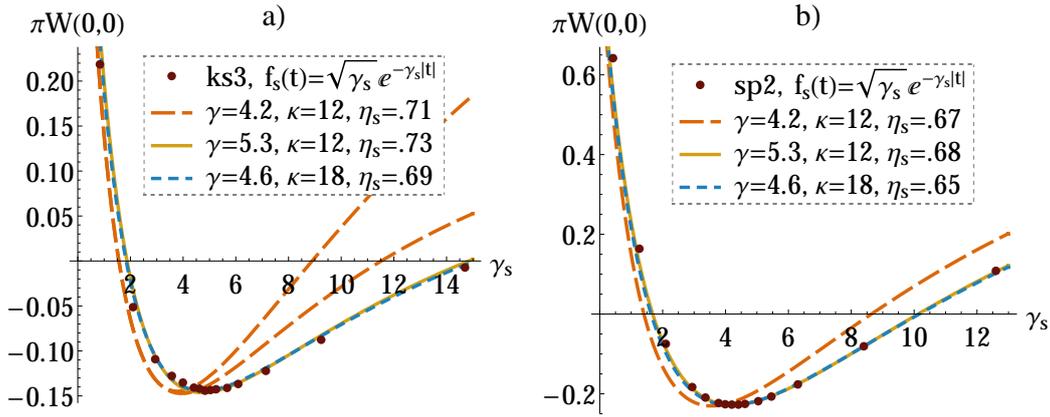


FIGURE 5.8: Wigner function values in the origin, obtained by varying the decay parameter γ_s of the simple exponential mode function in the quadrature extraction of **a)** ks3 and **b)** sp2. The curves are theoretical expectations for different parameters, where the efficiency η_s has been adjusted for the best fit with the data. The orange curves are the expectations based on the independent estimates of the model parameters used previously. In a), the lower of the two orange curves, as well as the yellow and blue, have been corrected for the effect of the slow detector – a reasonable agreement with the data was found by assuming a Lorentzian spectrum with a half-width of 18 MHz. The numbers for γ , κ and γ_s are in units of 2π MHz.

will be halted at a rather early breakpoint of 5×10^{-5} to save time – the trade-off in the Wigner function depth is still rather small at this point.

Variation of the simple exponential mode function

We stay with the simple exponentially decaying mode function that we have used so far for another little while. But now we vary the decay parameter γ_s , and thus the width of the mode function. The modeled dependence of γ_s on the Wigner function is given in Figure 3.15a) (DOPO) and Figure 3.18a) (NDOPO) for unity efficiency. The experimentally obtained dependencies are plotted in Figure 5.8 for two test series, ks3 and sp2, together with the theoretical predictions for different model parameters. As we saw before, the efficiencies are not well determined from the independent estimates, so they have been adjusted in each case to match the data. It is clear that the predictions based on the parameters assumed so far (orange curves) do not fit the data very well. The optimal region for γ_s is at higher values than expected, so either the OPO bandwidth γ or the trigger filtering bandwidth κ , or both, seem to be estimated too low. The yellow and blue curves show two attempts to remedy the situation by adjusting only γ (yellow) or both numbers (blue). They give almost exactly the same prediction, so there is no clue from here about which is more correct. Actually, it is most reasonable that κ is wrong, since it is just an estimated number used to approximate a single Lorentzian frequency filter to the more correct triply Lorentzian from the three cavities. It is not possible, though, to get a fit to the data by only adjusting κ – γ must be increased a little as well. The apparently larger-than-expected OPO bandwidth could be due to a slightly wrong measure-

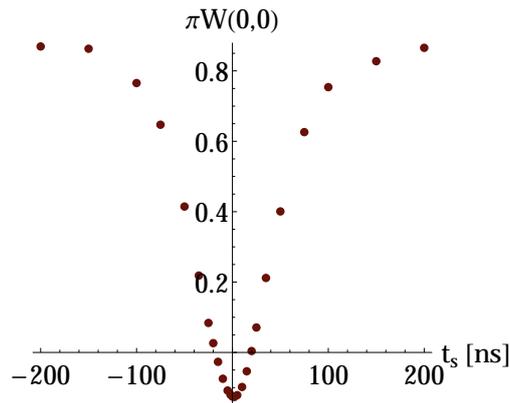


FIGURE 5.9: $\pi W(0,0)$ values of the reconstructed ks6 state, obtained for a simple exponential mode function shifted by t_s relative to the trigger time.

ment of the output coupler transmittance, or possibly because the light-induced losses are higher than estimated. In the latter case, it would also explain part of the discrepancy on the η_s parameter (although for ks3 it is not much off).

As for the optimal width of the mode function, for ks3 it is around $\gamma_s/2\pi = 4.9$ MHz; for ps2 it is indeed at $\gamma_s/2\pi = 4.2$ MHz. But the precise choice is a matter of peanuts – within a range of about 1 MHz, the value of $W(0,0)$ changes only on the third decimal. This is comforting knowledge, meaning that there is some leeway in the selection of mode function.

A note about Figure 5.8a): ks3 was recorded with the old, slow detector, so the points at high γ_s values are seemingly better (lower) than expected. The reason is of course, that the actual mode function becomes a convolution of the selected digital mode function with the detector response. This can be taken into account in the theoretical modeling, so the yellow and blue curves, as well as the lower of the two orange ones, are corrected for a detector with a Lorentzian spectrum with half-width 18 MHz. The upper orange curve is uncorrected.

Another way to vary the exponential mode function is to change its time offset. It is quite obvious that the state will degrade when the mode function is moved away from the time of the trigger click, but it is still interesting to see how much it degrades. This is done in Figure 5.9 for the ks6 series. The decay rate of the mode function was chosen as 4.2 MHz, although the optimal is probably a bit higher, referring to what we saw before. The plot shows that the timing of the mode function (or shaped local oscillator) is rather critical. A shift of only a few ns will significantly decrease the quality of the selected state. After a shift of 200 ns, there is basically only squeezed vacuum left. We can also see from the plot that the trigger delay for this particular series, estimated earlier from the segment noise variance, is a little bit off – the mode function should be shifted by 1-2 ns to gain the optimal Wigner function value.

The time shifted plot can also give a hint about the shape of the optimal mode function, since the procedure of taking a shifted mode function and integrating over the product of that and the data in a way corresponds to taking a convolu-

tion of the “real” shape of the state with the shape of the mode function.³ It is a very resource-expensive way to find the optimal mode function, though, due to the many data extractions and reconstructions needed.

Other shapes

To test the robustness of the generated states against variation in the mode function shape, and to possibly find a better match, we try out a number of different (all real) mode functions. The chosen classes of test functions are the following (without normalization factors), where B is just a general name for the bandwidth parameter:

$$\begin{aligned}
 f_{s,1}(t - t_c) &= e^{-B|t-t_c|} \\
 f_{s,2}(t - t_c) &= \frac{1}{B}e^{-B|t-t_c|} - \frac{1}{B_2}e^{-B_2|t-t_c|} \\
 f_{s,3}(t - t_c) &= e^{-B^2(t-t_c)^2/2} \\
 f_{s,4}(t - t_c) &= \begin{cases} 1 & , -B^{-1} \leq t - t_c \leq B^{-1} \\ 0 & , \text{otherwise,} \end{cases} \\
 f_{s,5}(t - t_c) &= \begin{cases} 1 - \frac{B|t-t_c|}{2} & , -2B^{-1} \leq t - t_c \leq 2B^{-1} \\ 0 & , \text{otherwise,} \end{cases} \\
 f_{s,6}(t - t_c) &= e^{-B^4(t-t_c)^4/2}
 \end{aligned} \tag{5.1}$$

$f_{s,1}$ is just our old simple exponential. $f_{s,2}$ is a convolution of two exponential functions with different bandwidths B and B_2 . To a certain extent, it simulates the smoothing due to the trigger filter cavities. $f_{s,3}$ is a Gaussian, also a bit smoother than the sharp exponential function. $f_{s,4}$ is a completely square top-hat function. The triangular $f_{s,5}$ is a better match than the top-hat to the other functions, although still having sharp edges. Finally, $f_{s,6}$ is somewhere between a square and a Gaussian function.

The low-gain single photon data set sp2 was used as the test case for reconstruction with the different classes of mode functions. The parameters for each type of function that give the lowest $W(0,0)$ value were estimated by brute force trial-and-error. The optimal function within each class is plotted in Figure 5.10. The optimal parameters are listed in Table 5.3 together with the resulting Wigner function value. We see that the smoothed exponential as expected provides slightly better results than the sharp exponential, but the difference is minimal. Also other functions are very close – the Gaussian and triangular functions achieve almost the same result, while the “square-Gaussian” is not too bad either. The small differences between the functions make sense when calculating their mode overlaps, which are quite high – see the last column in the table. The conclusion of this is that the exact shape of the mode function does not matter much, as long as the overall outline is approximately the same as those in Figure 5.10. For fine tuning, a smoothed exponential mode function can

³The “real” shape is in quotation marks because there is no definite mode shape for high gains – only shapes optimized with respect to some parameter, in this case the $W(0,0)$.

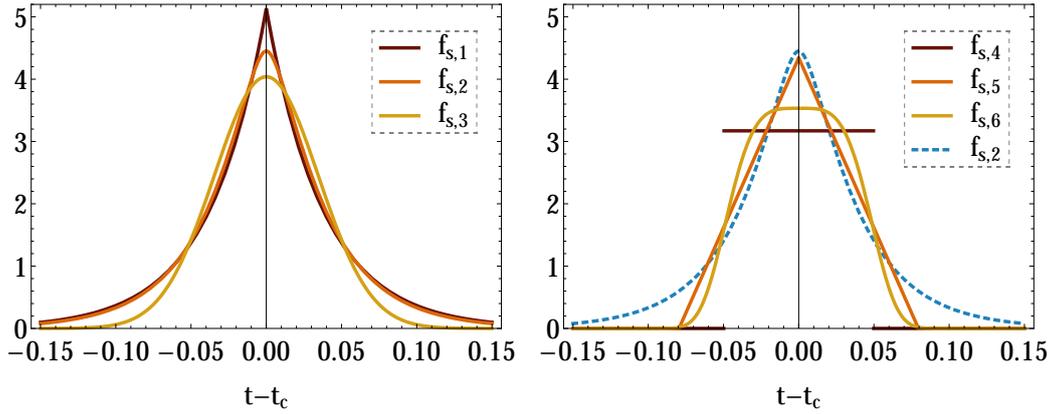


FIGURE 5.10: Six different shapes of signal mode function. The bandwidths of each of them have been optimized for the lowest value of $W(0,0)$ for the data series sp2. The best of the six, $f_{s,2}(t - t_c)$, is plotted in the right frame as well for comparison.

	opt. param. [$2\pi \text{ s}^{-1}$]	$\pi W(0,0)$	η_{relative}	overlap
$f_{s,1}$	$B = 4.2$	-.227	.990	.998
$f_{s,2}$	$B = 6, B_2 = 10$	-.238	1	1
$f_{s,3}$	$B = 4.6$	-.230	.993	.992
$f_{s,4}$	$B = 3.2$	-.082	.863	.915
$f_{s,5}$	$B = 4.0$	-.223	.987	.989
$f_{s,6}$	$B = 3.6$	-.187	.955	.970

TABLE 5.3: Optimal parameters of the six different mode functions with respect to the $W(0,0)$ value of the reconstructed sp2 state. η_{relative} is the additional efficiency factor needed in the theoretical model to explain the degradation relative to $f_{s,2}$, the best among these functions. The last column gives the overlap $\int f_{s,n}(t - t_c) f_{s,2}(t - t_c) dt$ with $f_{s,2}$. It should be noted that a range of different choices of the two parameters of $f_{s,2}$ give essentially the same function and results, e.g. $B = 4.6, B_2 = 18$.

be chosen – there might be other shapes which are even better, but they will most likely be very similar to this. Finally, this analysis was performed for a low gain NDOPO measurement, and the picture will probably be changed a little for some of the higher gains.

5.2.5 Detector compensated reconstruction

We now do a detailed reconstruction of each of the eight series, using the mode function $f_{s,2}$ with its parameters optimized in each case for the deepest Wigner function dip. Furthermore, we incorporate in the reconstruction procedure the compensation for detector efficiency using the POVM in (2.31). That will show us the states as they look, coming out of the generation apparatus, independently of the subsequent characterization apparatus. We do not, however, compensate for the overlap visibility with the local oscillator, since it is important for the generated states to be mode-matchable with another beam for subsequent processing.

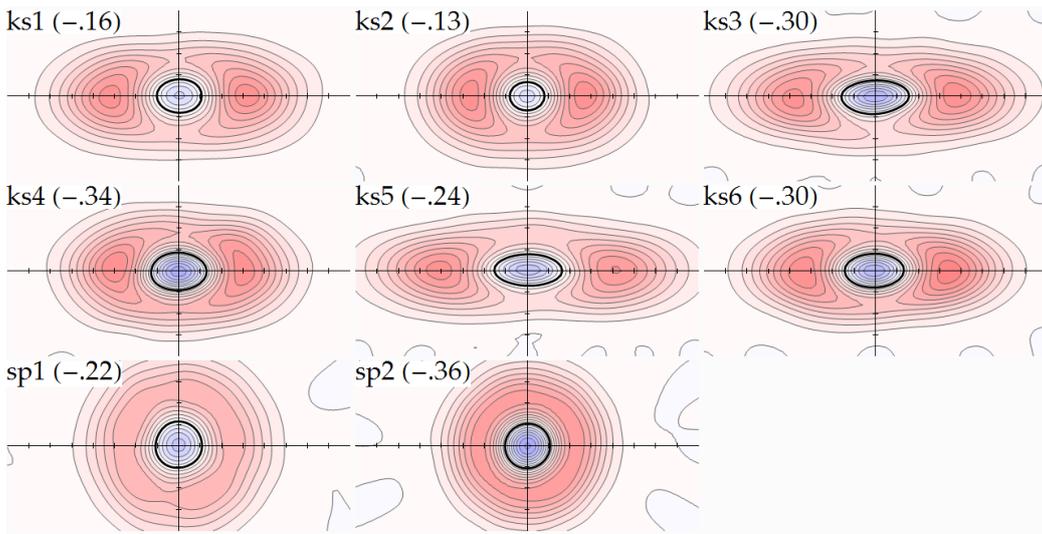


FIGURE 5.11: Wigner functions of the eight states, compensated for the non-unity detector efficiency. The numbers in parentheses are the $\pi W(0,0)$ values.

The propagation efficiency will not be compensated for, either, although part of the losses occur after the LO matching. Only the two factors directly involved in the detection, η_{qe} and η_{en} , will be included in the compensation. The iteration algorithm converges quite slowly when using compensation, so the data binning used is only moderately fine; $dq = 0.1$ for all series, $\#\theta = 24$ for ks1-4 and $\#\theta = 6$ for the phase-independent sp1-2 (and naturally $\#\theta = 12$ for ks5-6). Although the trace distance breakpoint is set rather low, at 2×10^{-6} (corresponding to about 1000 iterations for ks1-4), the $W(0,0)$ is still improving at the time the iteration is halted. The results could thus be improved slightly by running the iteration for several hours, but the difference is small enough to be insignificant.

The resulting Wigner functions, which must be considered as the optimal that can be extracted from our data, are plotted as contours in Figure 5.11, and – just because they look so pretty – as 3D surface plots in Figure 5.12. We see that the ks1 and ks2 states, even after compensation for the low-efficiency detector, are of poorer quality than the other kitten states. The main reason is that they were made with the old KN crystal, while the others used the much better PPKTP.

Fidelity

A parameter that we have neglected a bit throughout this chapter for the sake of simplicity is the fidelity. We make up for that in Figure 5.13 with cat-state fidelities of the kitten states in a) and the photon number distribution of the photon states in b). The kitten-cat fidelities are most easily calculated via the density matrices,

$$\mathcal{F} = \text{tr} [| \sim \hat{0} \rangle \langle \sim \hat{0} | \hat{\rho}_{\text{est}}],$$

where $\hat{\rho}_{\text{est}}$ is the measured and reconstructed state. For the single photon, the fidelity with the target state, $|1\rangle$, is simply the ρ_{11} component. For sp2 it is 68%, for sp1 only 57%. There is only one reason for the lower fidelity of sp1,

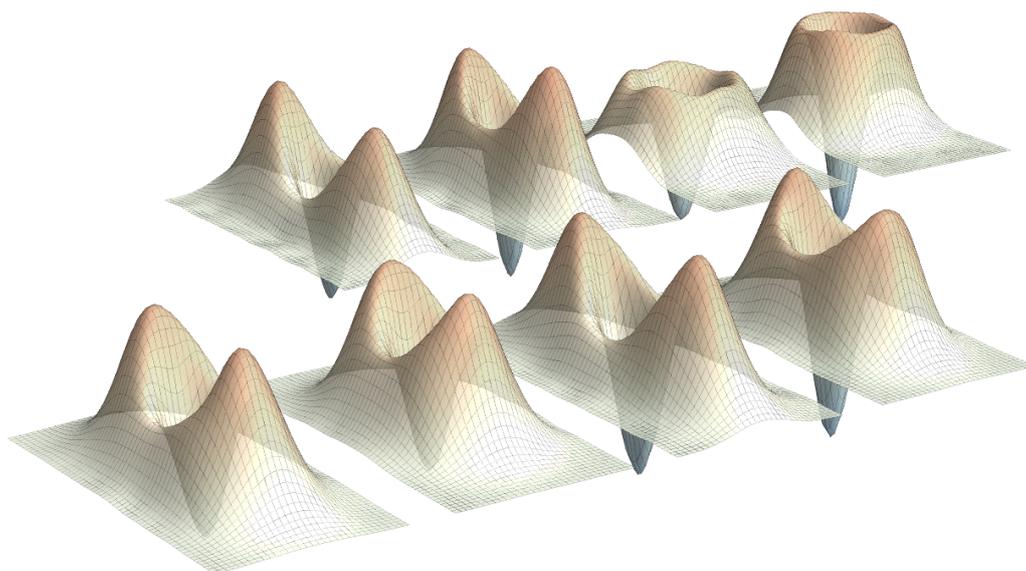


FIGURE 5.12: Same Wigner functions as in Figure 5.11, but with the arrangement like in Figure 5.1 (i.e. ks1-4 in the front row, left to right, and ks5-sp2 in the back row).

and that is the higher gain. As an intended source of single photons, possibly for quantum communication purposes, the high content of 2-, 3-, and even 4-photon components of sp1 is much worse than the lower 1-photon fidelity. The sp2 has a fairly moderate amount of higher photon numbers; only 1.7%. This clearly exemplifies the need for keeping the gain low – and consequently also the production rate.

As for the kitten states, it is interesting that for the four PPKTP-based states (ks3-6), there is a clear pattern of higher gains leading to higher kitten amplitude but lower maximal fidelities (compare with the gain parameters in Table 5.1). That is just what was expected from the theory, Figure 3.14. The amplitudes and fidelities are just somewhat smaller here, due to the experimental inefficiencies. The graph also shows that it is indeed impossible to obtain grown-up cats from this simple generation scheme – although ks5 clearly is a lot more squeezed than ks4, when looking at their Wigner functions, the obtained kitten state amplitude is only increased by about 40%, and with a 10% lower fidelity. Several proposals have been made, though, for ways of generating larger kittens [111, 154] or amplifying existing small kittens [62, 153]. They all involve in some way two or more simultaneous non-deterministic detection events, so the rate of production will be dramatically decreased.

Density matrices

It is enlightening as well to take a glance at the reconstructed density matrices themselves – these are the immediate outputs of the maximum likelihood reconstruction, and they are the basis for the calculation of the Wigner functions. Figure 5.14 shows the eight density matrices (absolute values, $|\rho_{mn}|$) corresponding to the Wigner functions just presented. To make the figures more clear, only

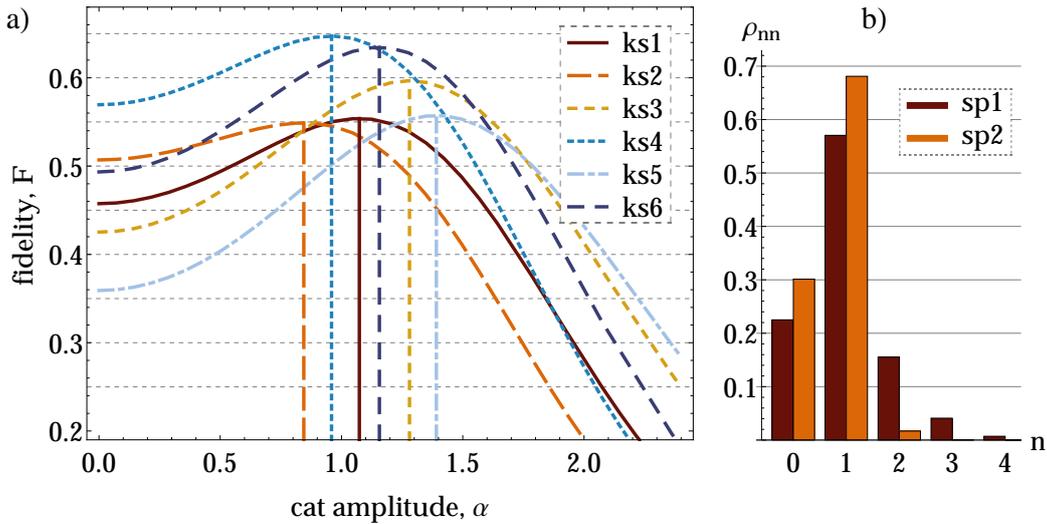


FIGURE 5.13: **a)** The fidelities between each of the six kitten states and a real odd cat state $|\mathcal{C}_{\alpha}^{-}\rangle$ with amplitude α . The vertical lines indicate the α that gives the highest fidelity. **b)** Photon number distribution for the two photon states. The size of the $n = 1$ component is the fidelity with the intended 1-photon state.

photon numbers up to $m = 8$ are included, although the reconstructions used a maximum number of 15. Most of the matrices have essentially vanishing entries at higher numbers than 8, except the largest states ks3 and ks5 – although the entries of these two states’ matrices do almost vanish before reaching $m = 15$. Also, to reduce clutter, only the lower triangles are shown, since the absolute values are symmetric about the diagonal anyway. The density matrices of the single photon states are completely diagonal and therefore offer no more information than was already shown in Figure 5.13b). The kitten state matrices are surely more interesting than those of the single photons, as they have non-zero elements outside of the diagonal as well. This shows that there are – at least in part – coherence between the different photon numbers, as opposed to the mixtures of sp1 and sp2. However, perhaps the most important feature of the kitten state matrices is the prominence of the odd photon number contributions in the diagonal. Apart from the large ρ_{11} elements, for the PPKTP-based states ks3-6, the ρ_{33} elements are also larger than ρ_{22} , and the ρ_{55} ’s are almost as large as ρ_{44} . Indeed it makes sense to call these states “superpositions of odd photon number states” – a perhaps more precise but also more mundane term than “kitten states”⁴. One notices that almost all elements ρ_{mn} that have odd $m - n$ – such as the elements neighbouring the diagonal – are close to zero. This can be seen as a consequence of the 180° rotational symmetry of the Wigner functions, that is, the states approximately fulfill the relation $W(-x, -p) = W(x, p)$. The Wigner functions are thus even, and this property, together with the fact that $W_{mn}(x, p)$ in (2.13) is odd for $m - n$ being odd, means that the integrand in the expression (2.12) for ρ_{mn} is odd, and so the integral becomes zero. The slight wobbliness of the ks5 Wigner function shows up as a small non-zero ρ_{10} element.

⁴Physical Review Letters prefers mundaneness!

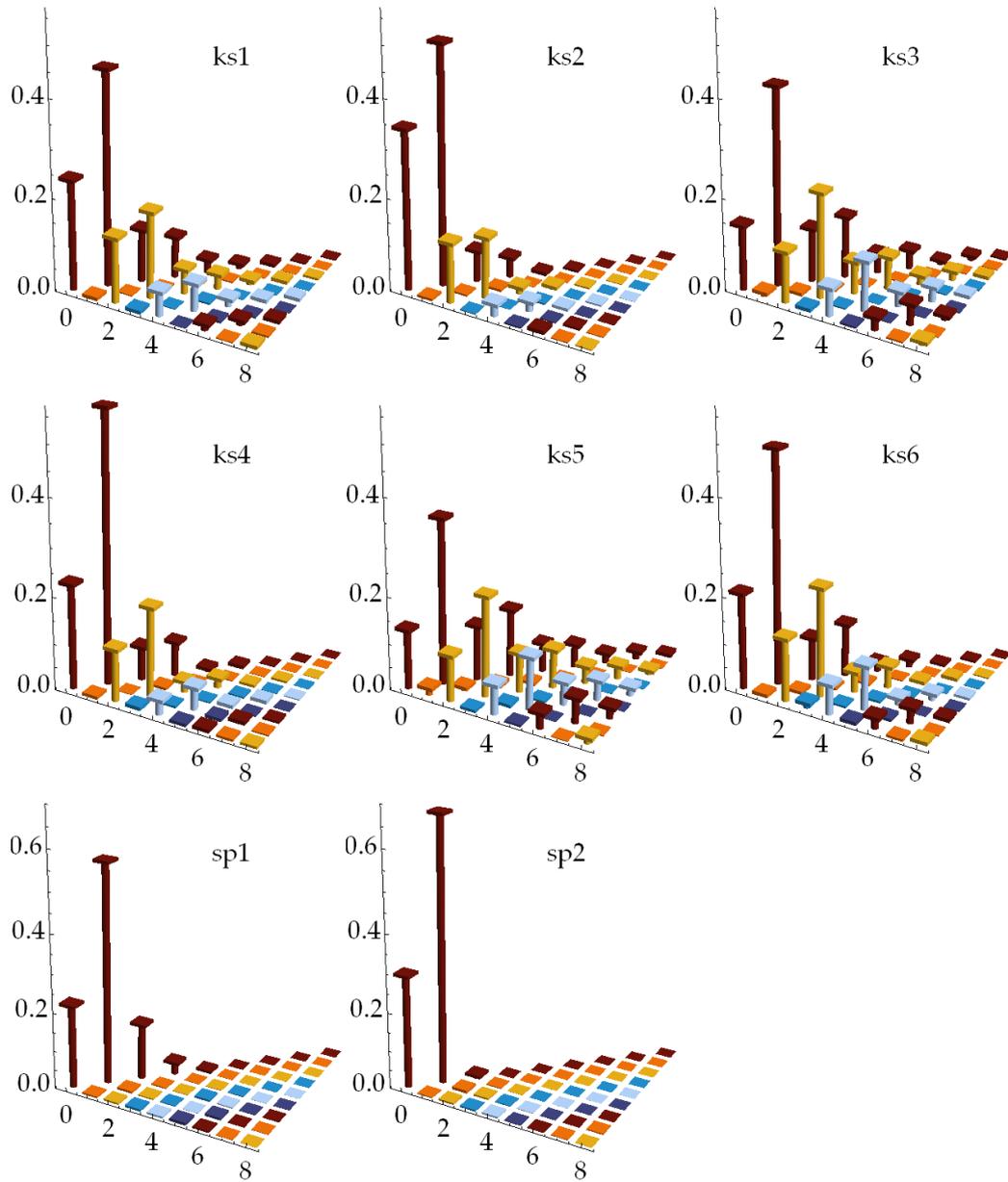


FIGURE 5.14: Density matrices in the number state basis (ρ_{mn}) for the eight states reconstructed with detector compensation. All bars show the absolute values of the density matrix elements, but since the states are rotated with the anti-squeezing direction along the x-axis, the imaginary parts of all the elements are close to zero, and there are almost no negative values. The numbers along one of the horizontal axes denote the photon number, i.e. the index m in ρ_{mn} . n is along the other axis.

5.3 Additional analysis

5.3.1 Error estimation

The estimation of the statistical uncertainties in the reconstructed quantum states can be done in several different ways:

- Since all of the phase-scanned data series (all but ks5-6) consist of a number of self-contained tomographic measurements with 20,000 samples each, these can be considered as separate measurements of the given state if a reconstruction is performed on each of them individually. Based on these measurements, a mean value and a standard deviation of the mean can be calculated for any parameter of the state, such as density matrix elements or Wigner function values. The mean value should be the same as the value obtained from the full reconstruction, while the standard deviation of the mean will be an estimate of the uncertainty in the full reconstruction including all samples.
- Based on the reconstructed state, a large number of data sets that approximately follow the same distribution as the original data can be Monte Carlo simulated. The statistical error in these simulated data will then give a good estimate of the uncertainty in the real data [28, 89].
- In principle, for a sufficiently large number of samples N , the variance of a maximum likelihood estimator \mathbf{s}_0 should reach the so-called Cramér-Rao lower bound on the variance of an unbiased estimator, $\text{Var } \mathbf{s}_0 \geq 1/NF(\mathbf{s}_0)$ [42, 61]. $F(\mathbf{s})$ is the Fisher information, defined as $F(\mathbf{s}) = \langle (\partial_{\mathbf{s}} \ln L(\mathbf{s}))^2 \rangle$. The averaging is over all possible outcomes of the specific observation (e.g. all possible distributions of quadrature values within the different bins). Because of this, $F(\mathbf{s})$ depends only on the parameters \mathbf{s} and the particular way in which the observation is carried out, *not* on the given observation/data set itself. We have had quite limited success with the calculation of this Fisher information based uncertainty. The numbers obtained are usually about a factor two smaller than those obtained by the methods mentioned above. This in itself does not seem unreasonable, but other factors make us doubt the validity of these error estimates – at least when it comes to homodyne measurements. A similar concern was stated by Lvovsky and Raymer [92].

We will now estimate the statistical errors of the reconstructed states – in particular their Wigner function values in the origin – based on the first two methods mentioned. In general, the standard deviations of inferred parameters scale by $1/\sqrt{N}$, where N is the total number of acquired samples.

Directly from data

For each of the series ks1-4, sp1-2, a full state tomography and reconstruction has been performed N_f times, this number referring to the number of 20,000-point

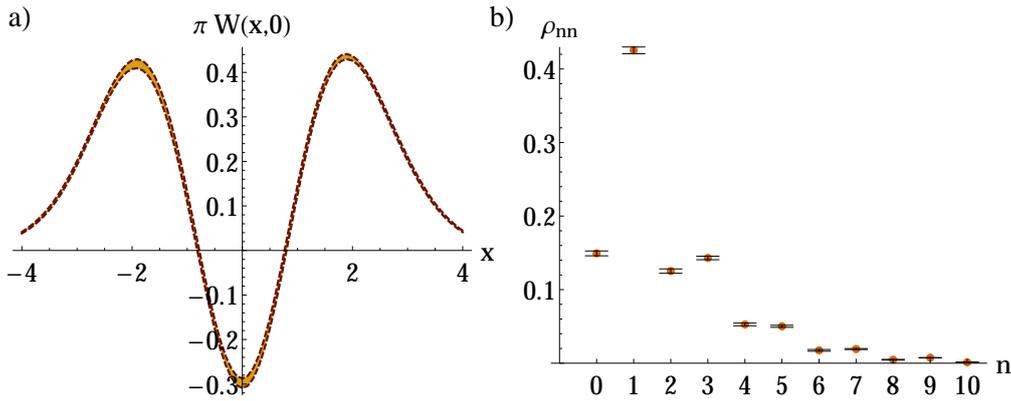


FIGURE 5.15: The estimated 68% confidence intervals of **a)** a slice through the Wigner function and **b)** the density matrix diagonal elements of the reconstructed ks3 state. The uncertainties are those obtained directly from the data, but corrected for the total number of data points.

files acquired (the first number in the last column of Table 5.1). Thus, N_f values for $\pi W(0,0)$ will be obtained. These will be scattered with a certain standard deviation

$$\sigma_{\pi W} = \sqrt{\langle (\pi W(0,0) - \langle \pi W(0,0) \rangle)^2 \rangle},$$

and the uncertainty of the mean is then

$$\bar{\sigma}_{\pi W} = \frac{\sigma_{\pi W}}{\sqrt{N_f}}.$$

The data of the phase-locked series ks5-6 were acquired in a way that they can not as easily be considered as a range of separate measurements. We can, however, separate the 50,000 points acquired at each phase settings into 10 groups of 5,000 points, and then perform 10 reconstructions, each based on 5,000 points from each phase setting.

The uncertainties in the Wigner function obtained in this way are then:

	ks1	ks2	ks3	ks4	ks5	ks6	sp1	sp2
$\sigma_{\pi W}$.028	.025	.034	.017	.010	.010	.023	.017
N_f	49	50	10	19	(10)	(10)	9	9
$\bar{\sigma}_{\pi W}$.004	.004	.011	.004	.003	.003	.008	.006

As they are obtained from the data, these numbers include not only statistical errors but also any fluctuations in the experimental conditions that might have happened within the time of acquisition of each series – although this argument does not hold for ks5-6. That can perhaps explain why $\sigma_{\pi W}$ for some series is twice as large as for others. The $\sigma_{\pi W}$ for ks5-6 are smaller than the others, simply because they are based on reconstructions of 60,000 points ($12 \times 5,000$), while the others are based on 20,000 points data sets.

For the data series with the largest uncertainty, ks3, the errors are shown in more detail in Figure 5.15. In a) a slice of the Wigner function through $p = 0$ is shown with the 68% confidence interval at all values (the yellow band), calculated in the same way as above, that is, with ± 0.011 in $x = 0$. Figure b) shows the density matrix diagonal elements with error bars, also obtained in the same way. We see that the errors are really quite small everywhere, so the 200,000 points used in ks3 is more than enough to give a really good estimate of the measured state.

From simulated data

Based on the reconstructed density matrices of series ks3 and ks6, a number of artificial data sets of different lengths were simulated (using pseudo-random numbers and inversion of the cumulative distributions at 24 or 12 different phase angles, as calculated from the density matrices). The uncertainties are then given in the same way as before, but this time they are purely statistical, since all data are simulated under the same conditions. For ks3, 40 sets of data were generated, while for ks6 20 sets of data were generated. Each set has a variable number of points, N_s , distributed equally over the 24 or 12 phase angles. The following table lists the uncertainties in the reconstructed $\pi W(0,0)$ values for different N_s , as well as the “per-point uncertainty”, $\sigma_{\pi W} \sqrt{N_s}$, which can then be used to estimate the uncertainty in the real measurements as $\sigma_{\pi W} \sqrt{N_s} / \bar{N}$ (N is again the total number of samples, e.g. 200,000 for ks3 and 600,000 for ks6):

	N_s	$\sigma_{\pi W}$	$\sigma_{\pi W} \sqrt{N_s}$
ks3	20,000	.019	2.66
	10,000	.030	2.97
	5,000	.039	2.76
	2,000	.049	2.18
ks6	20,000	.014	2.04
	10,000	.019	1.91
	5,000	.032	2.23

The per-point uncertainties are quite similar, irrespectively of the size of the data sets, as the should be. Taking the largest of the numbers for ks3 and ks6, the estimated uncertainties in $\pi W(0,0)$ for the measured states are then $2.97 / \sqrt{200,000} = .007$ for ks3, and $2.23 / \sqrt{600,000} = .003$ for ks6. For ks6, this is the same value as that obtained directly from the data before, while for ks3 it is a bit lower. As discussed before, this probably reflects changing experimental conditions during the ks3 measurements, giving an error larger than the purely statistical error. In any case, these simulations lend support to the error values from before being fair upper estimates.

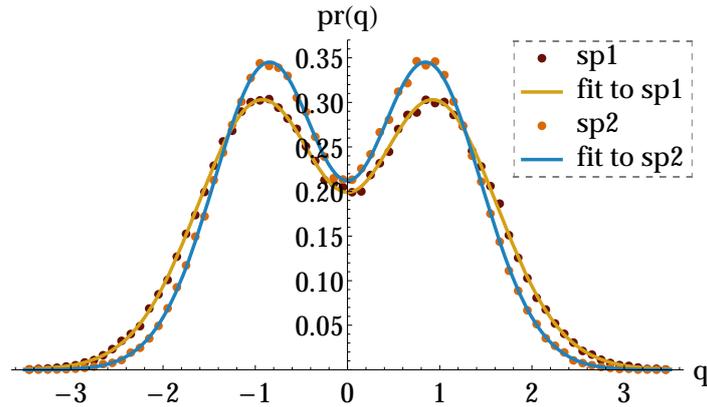


FIGURE 5.16: The full quadrature distributions (normalized histograms) of the sp1 and sp2 series together with the best fits to a mixture of 0- to 5-photon states.

5.3.2 Simple curve fitting to single photon data

The photon number distribution of the phase-invariant single photon states can also be estimated from a simple least-squares fit to the overall experimental distribution of quadrature values. The function to be fitted should be just a sum of the quadrature distributions for different number states:

$$\text{pr}_{\text{fit}}(q) = \sum_n w_n \text{pr}_n(q) = \sum_n w_n |\langle q|n\rangle|^2,$$

with the $\langle q|n\rangle$ given in (2.6). The weights w_n correspond to the diagonal entries ρ_{nn} of the density matrix, and are the free parameters to be fitted. Figure 5.16 shows the full 180,000 points quadrature distributions of the sp1 and sp2 data series together with the best least-squares fits to the above function, including photon numbers up to $n = 5$. The parameter estimates (up to $n = 4$) of the fits are listed in Table 5.4 together with the diagonal entries of the maximum likelihood density matrices, obtained from a reconstruction without detector efficiency compensation. The two estimation methods give basically identical results. It is worth noting, though, that the simple curve fitting does not have the built-in physical constraints that the maximum likelihood method has. Therefore, the resulting photon number weights need not add up to one, and they might even become negative as is the case here for w_3 of the sp2 set. The standard curve fitting is much faster than the iterative maximum likelihood method, but it is only applicable when the state is rotationally invariant in phase space. If no such assumption can be made, maximum likelihood (or another tomographic reconstruction method) must be employed.

5.3.3 Production rates, spectral brightness

One of the big advantages of using cavity-enhanced down-conversion is the hugely increased spectral brightness relative to pulsed sources. In this context, we consider the spectral brightness to be the state production rate per MHz

	n				
	0	1	2	3	4
sp1 w_n	.2827±.0016	.543±.002	.138±.002	.030±.003	.004±.003
sp1 ρ_{nn}	.2828	.543	.136	.032	.005
sp2 w_n	.3678±.0016	.616±.002	.017±.002	-.001±.003	.001±.003
sp2 ρ_{nn}	.3684	.617	.014	.001	.000

TABLE 5.4: Estimates of the photon number distribution of the sp1 and sp2 series. w_n are the parameters obtained from a simple curve fitting to the full quadrature distributions, while ρ_{nn} are from the maximum likelihood reconstruction.

bandwidth of the state. This figure becomes relevant for applications that require narrow bandwidths of the optical fields, such as atomic quantum memory (see §6.1.1). Pulsed SPDC sources often have production rates comparable with our system, but their bandwidths are several nanometers, so heavy spectral filtering would be needed to make them compatible with the atoms, thus compromising the production rate. As an example, the photon-subtracted squeezed vacuum states of Ourjoumtsev et al. [121] are produced at a rate of $4,500 \text{ s}^{-1}$, and with a Fourier-limited bandwidth of approximately 2 THz, giving a spectral brightness of about $0.002 \text{ s}^{-1}/\text{MHz}$. Slightly lower figures were reported for the single photon sources of U'Ren et al. [157] and Pittman et al. [127], while Fulconis et al. [44] achieved $\sim 0.15 \text{ s}^{-1}/\text{MHz}$ with a photonic crystal fiber based SPDC source. Higher spectral brightness single photon sources have been realized in atom-based systems, which are in general much more experimentally involved than SPDC sources, but offer inherently narrow bandwidths compatible with atoms. Systems such as single molecules [82], single atoms trapped in an optical tweezer [29] or cavity [96], or trapped ions [68] all achieved brightness levels of a few thousand photons per second per MHz. With an ensemble of cold atoms, Thompson et al. [155] obtained as many as $50,000 \text{ s}^{-1}/\text{MHz}$.

The average production rates for our 8 measurement series are stated in Table 5.5 together with the corresponding spectral brightness, based on a FWHM bandwidth of the fields of 9 MHz. Furthermore, the count rates expected from the expression in (3.20) are listed. These values take into account the T , η_{esc} , and x values from Table 5.1, together with an estimated trigger arm propagation efficiency of 14% (which is not necessarily correct for all series), and the APD quantum efficiency of 44%. We see that our spectral brightness values are 3–7 orders of magnitude larger than pulsed SPDC sources, but a bit lower than atoms based single photon sources. The experimentally observed rates are comparable to the expected values, considering the crudeness of these estimates. The ks5 and ks6 series are exceptions though. The expected rates based on (3.20) assumes a continuous production, but for the sequenced locking system employed for ks5 and ks6, the APD conditioning is only carried out in a fraction of the total run time. For some reason, the ks5 rate is even lower than expected, even taking this into account. I do not have a reasonable explanation for this large discrepancy. Among the single photon series, sp1 has a quite high production rate, but this is

	average production rate [states/sec]	spectral brightness [states/sec/MHz BW]	expected rate [states/sec]
ks1	6200 ± 400	690	6000
ks2	6800 ± 300	760	7400
ks3	9400 ± 400	1040	15000
ks4	7100 ± 500	790	8800
ks5	1500 ± 200	170	20500
ks6	1590 ± 40	180	7200
sp1	91600 ± 1500	10200	145000
sp2	12900 ± 1400	1400	23700

TABLE 5.5: The average rates of conditional state generation in the 8 measurement series, the corresponding spectral brightness (using a 9 MHz FWHM bandwidth), and the theoretically expected production rates.

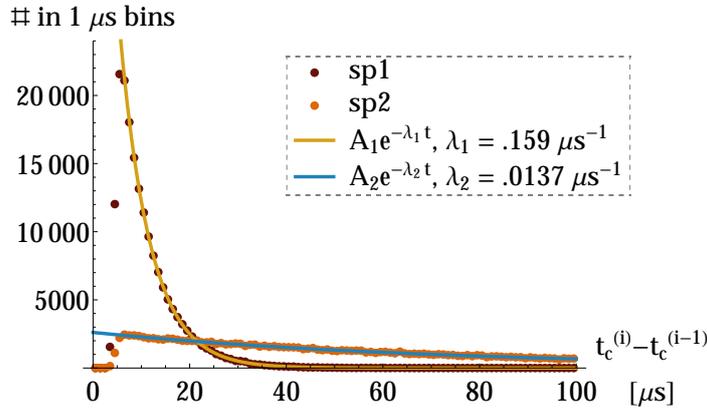


FIGURE 5.17: Histogram with $1 \mu\text{s}$ bin widths over the waiting times between two consecutive detector clicks, that is, $t_c^{(i)} - t_c^{(i-1)}$ with the superscript indicating the index of the given click time t_c in the recorded data set. The two data sets, sp1 and sp2, have quite different rates, but they both follow an exponential curve nicely (the down-conversion is a Poisson process), apart from the region at low interval lengths where the scope inhibits trigger events.

of course at the expense of the higher 2-photon contribution.

Figure 5.17 shows two examples of the distribution of the intervals between two consecutive trigger clicks. The graphs show histograms of these intervals for the sp1 series with a very high rate, and the sp2 series with a somewhat lower rate. For the most part, the curves follow nicely exponential decays as a Poisson process, such as SPDC, should. However, it is apparent that no clicks at all appear within $3 \mu\text{s}$ after the preceding click, and only at $6-7 \mu\text{s}$ intervals does the histogram reach its natural level, according to the exponential function. The APD dead time is not the cause for these lacking clicks, since it is only about 35 ns . Rather, the cause must be the scope – the histogram is not based on the actual clicks, but rather on those clicks that were recorded as trigger events by the

scope. Due to our settings with $2 \mu\text{s}$ long segments around each trigger time, we impose at least a $2 \mu\text{s}$ delay between subsequent trigger events. The remaining $1\text{-}5 \mu\text{s}$ minimum delay time must be due to processing time internally in the scope. In any case, the unrecorded clicks are a reason for why some of the observed production rates in Table 5.5 are lower than expected. The decay constants of the exponential fits in Figure 5.17 give directly the true trigger photon rates, i.e. $159,000 \text{ s}^{-1}$ for sp1 and $13,700 \text{ s}^{-1}$ for sp2. Especially for sp1, this brings the rate much closer to the expected value.

Outlook

6.1 Time gated state generation

6.1.1 Towards storage in atomic memory

An important goal for quantum information processing is to establish efficient interfaces between light and atoms. Light is a perfect means of long distance transportation of quantum states, while atomic systems – either individual atoms or collective ensembles – in controlled environments are well suited for long-lived storage of the states. The ability to efficiently exchange quantum states between light and atoms will therefore allow implementation of quantum networks with atomic systems as nodes, for example in a quantum repeater [32]. Memories for light based on atomic ensembles have been demonstrated for example with the EIT (electromagnetically induced transparency) interaction [81, 125], with off-resonant QND (quantum non-demolition) measurement and feedback [66, 140], and other systems – see the review by Hammerer et al. [54]. Mostly classical coherent pulses have been stored, but recently also the storage and retrieval of squeezed light in EIT was demonstrated [2, 59]. Probabilistic storage of single photons in atomic ensembles has also been carried out in several groups [24, 35], albeit with low efficiency. Common for these different systems is that the interaction is driven by a strong classical field. In particular for schemes like that of Julsgaard et al. [66], the quantum field to be stored is mixed with the strong driving field (local oscillator) in orthogonal polarizations, and the shape of the latter determines the temporal mode of the quantum field to be stored. In order to achieve a successful storage, the quantum field should match this mode well. The properties of the atomic systems set some limits on the possible bandwidth of the field – usually it should be on the order of 10 MHz at most [54].

With atomic storage in mind, the approach to nonclassical state generation using continuous cavity-enhanced pumping of the nonlinear crystal has some attractive properties compared with related pulsed sources. Most importantly the narrow bandwidth/high spectral brightness, which makes the states immediately

applicable for interactions with atomic media without need for strong spectral filtering. Another convenient property is the relative flexibility in the choice of mode function, as explored in §5.2.4, which, together with the long duration of the mode functions, makes timing and mode-matching issues fairly relaxed. There is one disadvantage, though, that comes with the continuous pumping, and that is the randomness of the arrival time of the heralded states. All other SPDC-based sources are probabilistic by nature as well, but with pulsed pumping, the arrival time of the heralded states – in case they are generated – is well defined by the timing of the pulsed laser source. The pulsed local oscillator can thus be prepared in advance and meet the generated state on the beamsplitter at the right time. Most of the pulses will not contain a non-Gaussian state, however, since the heralding trigger detection appears for only a small fraction of the pulses, so post-selection is necessary. For our continuous source the situation is similar, but there are some added difficulties, since the trigger click times are continuously distributed. It means that we do not know in advance at what time the local oscillator pulses should be prepared. The pulse shaping could be initiated at the instant of the trigger click, but even then it is problematic to make the signal field and LO match, since roughly half of the signal field is in the past relative to the click time. There are at least two conceivable ways to resolve this problem:

- 1. Delaying the signal field.** In the case that the local oscillator pulse is only produced whenever a state generation is heralded, it is essential that the signal state is optically delayed until the LO is ready. Apart from the fact that half of the temporal extent of the state is ahead of the trigger click (more than half, in fact, due to the trigger filtering), there will be unavoidable electronic delays in the APD module, in the optical pulse shaper for the LO (EOM or similar), in the shaper's amplifier etc. Figure 4.17b) shows the shortest delay we have had between the arrival to the scope of the homodyne signal and the trigger; about 25 ns. The arrival of the signal state on the LO mixing beamsplitter is of course even earlier. To get a decent mode-matching between the two fields, the LO pulse should have its front edge at a time $\sim \gamma^{-1}$ before the peak of the correlations, adding another ~ 40 ns (in our case) to the interval during which the generated state must be kept idle, waiting for the LO to be ready for the overlapping. That corresponds to a free-space delay line length of at least 20 m, most likely a lot longer. In order to preserve the high purity of our non-Gaussian, negative Wigner function states, the in- and out-coupling of the light to/from the delay line, as well as the propagation inside, must have low losses, which is difficult to achieve for such a long delay. For a long optical delay, fibers are an obvious choice. A polarization-maintaining fiber is probably needed, since most applications of the photon or kitten would require a well-defined polarization. The losses from the in-coupling to the fiber might be a limitation, though, so free-space delay lines could be an alternative. Pittman et al. [126] constructed a switchable free-space delay line with a short loop length and a switchable coupling based on a polarization-rotating Pockels cell in the loop and a PBS for the coupling. Their losses were quite large, though; 26% per round-trip of 4 meters length. This length is furthermore way too short to contain our

narrowband states which have an intensity distribution extent of about 15-30 meters. Some of the losses could be eliminated by using a passive delay line, such as a folded path between two large mirrors with in-coupling through a hole in one mirror or a small coupling mirror or prism somewhere in the path [58]. A group of bachelor students doing a project with our experiment worked on the implementation of such a Herriott cell. They achieved a delay of 70 ns, but with quite high losses of 90% [156]. These high losses were mainly due to the poor mirror quality, especially of the coupling mirror, but their results show that a delay of the necessary length is feasible. It requires high quality optics, though, since light is lost on each of the many reflections off the mirrors of the cell.

2. Time gating the state generation. Although delaying the signal field can leave enough time for the LO pulse to be generated conditioned on a trigger click, it does not remove the inherent randomness of the time of arrival of the combined LO and non-Gaussian state pulse. We have carried out some preliminary experiments to try to remedy this and make the arrival time somewhat more predictable. The idea is to make the OPO pump beam “semi-pulsed” such that photons can only be down-converted within predefined time windows. That way, the temporal gating of the pulsed sources is gained, while the cavity-limited bandwidth is maintained, at the expense of a lower rate. The LO pulses can then be prepared independently of whether or not a state is heralded, and therefore be ready to meet the generated signal state whenever one appears, without the need to delay the signal state. This, however, requires that the pump pulse is fairly short, such that the overlap between LO and signal state is good whether the state has been produced in the beginning or the end of the pump pulse – the temporal mode function defined by the LO shape cannot be optimal for each case, but can be optimized to give the best performance on average. The pump should not be too short either, since that will limit the production rate. A short pulse will be Fourier broadened, and for components of the pump field lying away from the central frequency, the down-conversion will be inhibited by the cavity’s spectral structure. Furthermore, if it is very short – on the order of the round-trip time – part of the blue light will be at frequencies in between two non-symmetric cavity resonances. Then, photon pairs will be down-converted into these modes, which will cause false clicks (“dark counts”) in the trigger mode.

The point of this proposal is thus to prepare the local oscillator pulses independently of the outcome of the state generation, but synchronized with the OPO pump pulses. Trigger clicks should then be subjected to a gating condition, such that they will only be accepted if they appear within a certain time window around the point of the highest down-conversion intensity. The exact arrival time of the trigger photons has no influence on the shape or timing of the LO pulse. The same thing could be done without pulsing the OPO pump; although photons were continuously down-converted, the trigger gating could be enforced such that only trigger photons arriving at times close to the LO pulse were counted. However, that would give a larger spread in the temporal distribution of the heralded states, compared to those obtained with the con-

finer pump interval, and hence the mode overlap and state quality would be degraded. Nielsen and Mølmer [108] made a detailed modeling of the pulsed OPO pump scenario described here and applied it to the case of kitten state generation, with the $W(0,0)$ value as the benchmark for state quality. They did indeed find that for time-gated triggering, short pulses result in deeper Wigner function dips than cw pumping. The shorter the pulse lengths T_p , the better, due to tighter confinement and thus better mode overlap (they did not consider the appearance of false clicks for very short pulses). A shorter gating interval T_g also improves the state quality, but of course a lower T_g , as well as T_p , will decrease the trigger probability per pulse/gating window and hence the overall production rate. It should be noted that Adamyan et al. [1] recently did some detailed calculations regarding the output of NDOPOs with modulated or pulsed pumping, including results for subsequent photon subtraction.

An important point of the time-gated state generation is that the shape of the optimal mode function no longer will be symmetrical around the trigger time (or at least roughly symmetrical). The symmetrical shape, in its basic form the double-sided exponential decay, stems from the symmetry between the trigger photon and its twin in the signal mode. It is equally probable that the trigger photon escapes the OPO before its twin or that it does so afterwards. On the other hand, the gated generation is independent of the trigger photon timing, and the optimal mode function will depend on the timing and duration of the pump pulse and gating window. As stated in [108], for long gating intervals T_g it is reasonable to assume that the optimal mode function will just be the square root of the intensity distribution of the down-converted light, which in turn is the pump field shape $z_p(t)$ convoluted with the OPO cavity response:

$$f_{s,\text{gated}}(t - t_c) \approx \int_{-\infty}^{t-t_c} z_p(t') e^{-\gamma(t-t_c-t')} dt'. \quad (6.1)$$

They numerically calculate the mode functions optimized for minimal $W(0,0)$ values, and they are indeed very close to the above convolution. The shape is quite asymmetric: for short pump pulses, the leading edge of the intensity distribution is roughly determined by the pump pulse shape, while the trailing edge is the exponential decay of the cavity.

6.1.2 Results of the pumped pulse experiment

So far, we have only carried out a proof-of-principle experiment, showing that single photon states with negative Wigner functions can be generated with a pulsed OPO pump and time-gated trigger. We have not made detailed studies with varying parameters, such as T_p and T_g or the gain x , in order to test the predictions of ref. [108] regarding kitten states. We also did not use a pulsed local oscillator, but used the usual digital mode selection.

We chop the blue beam from the SHG into short pulses with a pulse shaper from BME Bergmann (PCD dpp3.4c). It contains a KD*P double Pockels cell from Leysop and a circuit for fast switching of the half-wave voltage of about 1.5 kV for 426 nm. With a PBS after the device, the beam can be switched completely on or off in 6 ns. The PBS we used was not of the best quality, or perhaps

the polarization of the blue light was not completely pure, so the extinction ratio was only 1-2%. The maximum repetition rate is 200 kHz. The switching of the Pockels cell is triggered by sending TTL pulses to start and stop ports on the pulse shaper. We control the pulse length by using the same TTL signal for both the start and stop, but with a delaying coaxial cable of varying length in between. Since the blue light is off most of the time, it is not possible to lock the OPO and filtering cavities on the down-converted light. We therefore used an early implementation of the sequenced locking, where the lock beam and cavity locking PI-controllers were switched on for 0.8 s, followed by 0.2 s off-time where trigger events were accepted. There are more difficulties with the phase information. Since the production rate naturally is a lot lower when the OPO is only pumped a fraction of the time, it is not so feasible to use the calibrated phase scanning. The phase locking based on the intermediate gain lock is not possible at all, as long as the pump is pulsed, because the gain lock requires steady probe and pump beams. It is possible to make the phase locking work, by making the pulse shaper alternate between pulsing and being continuously on, synchronized with the lock beam switching. For the preliminary measurements, though, we just measured the phase-independent single photon state, not paying attention to the phase. We did not have the mechanical shutters to block the lock/probe beam at the time, but switched instead the AOMs used for the frequency shifting. Since the zeroth order of the AOM transmission is blocked in the double-pass configuration, the beam is completely shut when the AOM is switched off. The signal for the lock program is then obtained from the APD count rate.

Pulse testing

At first we studied the intensity distribution of the down-converted light with different pump pulse lengths. The three pulse shapes tested are well approximated by the functions in Figure 6.1 (with Gaussian error functions on the edges). We measured on the scope the delay between the TTL pulse sent to the pulse shaper to initiate the pulse shaping and the click from the APD, and delay histograms were slowly built up. The maximum blue power was low enough that there were very rarely two detection events per pulse – for the longest pulse there were about 10 clicks per second; the pulse repetition rate was 50 kHz.¹ Because of the not so good suppression of the pulse shaper + APD, there is a constant background level, on top of which lie the photons originating from the shaped part of the pump, as intended. The existence of the background counts is undesirable for the gated photon generation, but for the pulse shape analysis it is helpful since it can be used to normalize the three histograms to each other – the background level depends only on the total time of acquisition, not on the pulse length (and of course on the blue power, which was the same for the three). The background level includes dark counts from the APD, but only by about a tenth of the total level; the dark count rate was 75 c/s, while the total corresponded to roughly 800 c/s. In Figure 6.2, the delay histograms for the three pulse lengths

¹Two clicks from one pulse would not necessarily be recorded due to the APD dead time of 35 ns.

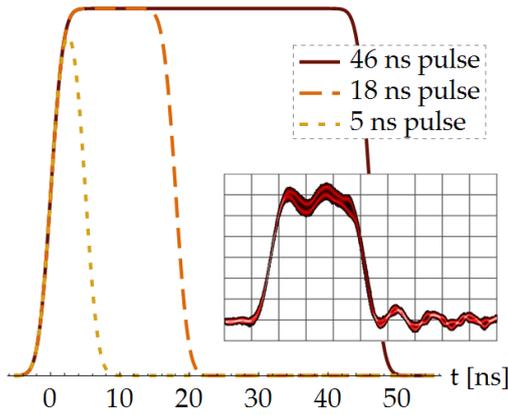


FIGURE 6.1: The shapes $|z_p(t)|^2$ of the three different pulses used for pumping the OPO. They are simple functions approximating the real pulse shapes – the inset shows a scope trace of the 18 ns pulse in a grid with 5 ns per horizontal box. It is measured with the fast homodyne detector (on a single diode), and the oscillations are due to the resonance peak around 100 MHz seen in Figure 4.13b).

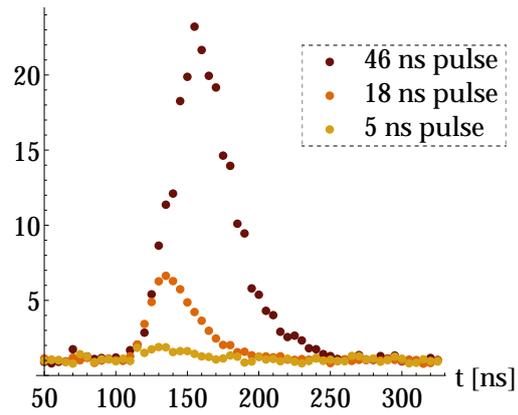


FIGURE 6.2: Histograms of the delay between the blue pulse and the APD click time (the horizontal offset is somewhat arbitrary, but consistent among the three histograms). All three histograms have been normalized to a level of 1 of the background level. Each time bin (point) covers 5 ns. The 18 ns pulse has the smoothest histogram since it was recorded for the longest time.

are plotted, normalized to a background level of 1. For longer pulses, the count histogram for the down-converted light becomes not only wider but also higher. That is because all three pulses are shorter than the build-up time of the OPO cavity, so the steady state level is not reached. It can also be understood in frequency space, as discussed above, as the filtering (or suppression) by the OPO of the part of the down-converted field with frequencies away from the cavity resonance – a filtering that becomes more severe for shorter and hence spectrally broader pulses. The acquisition time was limited, so especially the 5 ns pulse histogram (recorded over half an hour) is quite noisy. The same three histograms are repeated in Figure 6.3, now with the background level subtracted. Together with them are plotted theoretically calculated functions for the down-converted intensity distributions in the trigger mode, which fit well with the data. These functions, $|z_{dc,t}(t)|^2$, are given by the convolution of the pump field with the response of the OPO and trigger filter cavities:

$$z_{dc,t}(t) = \int_{-\infty}^t z_p(t')h(t-t')dt',$$

$$h(t-t') = \frac{1}{\gamma}e^{-\gamma|t-t'|} - \frac{1}{\kappa}e^{-\kappa|t-t'|}.$$

The cavity response function $h(t)$ is again assuming just a single Lorentzian trigger filtering with overall bandwidth κ . The parameters used for the OPO and filter bandwidths here are those that seem reasonable from the analysis in §5.2.4: $\gamma/2\pi = 4.6$ MHz, $\kappa/2\pi = 18$ MHz. The curves have been scaled vertically to fit the scaled histograms, and offset horizontally to get the right delays, but all three curves have been scaled and offset by the same amount, and they fit quite

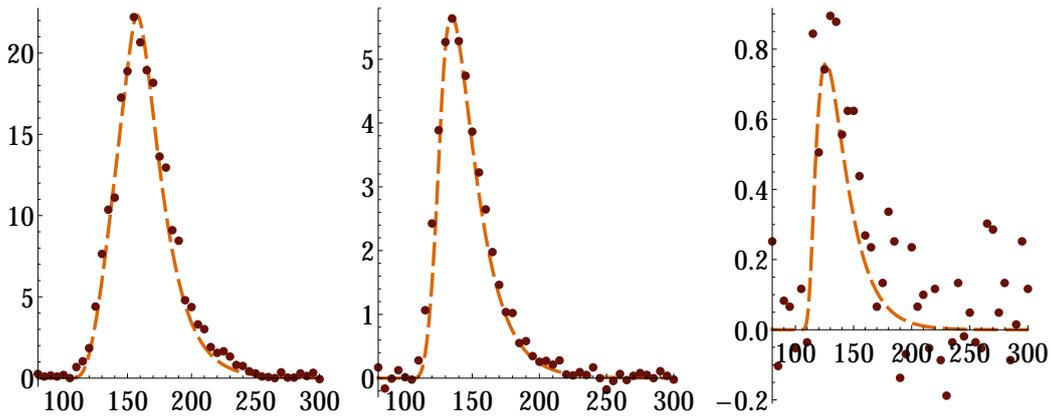


FIGURE 6.3: The histograms from Figure 6.2 with the background level subtracted. From left to right: 46 ns pulse, 18 ns pulse, 5 ns pulse. The orange curves are theoretical models for the intensity distribution. They are given by the pulse shapes from Figure 6.1 convoluted with the combined response of the OPO and filter cavities, with $\gamma/2\pi = 4.6$ MHz, $\kappa/2\pi = 18$ MHz. The vertical scaling factor and the horizontal offset are the same in all three plots.

well to the histograms. We see that the leading edges have durations roughly equal to the pulse lengths, while the trailing edges are primarily defined by the OPO lifetime and are basically the same for the three pulse lengths.

Photon state measurement with gating

We carried out the photon generation with homodyne characterization for both the 46 ns and the 5 ns pulse, but only got reasonable results for the long pulse. With the short pulse we got slightly non-Gaussian statistics, but far from a negativity in the Wigner function. A major reason must have been the comparatively high dark count rate due to the low down-conversion intensity, as seen above. For the long pulse, we did the measurement in the following way.

With the delay generator that makes the TTL pulses for the pulse generator we made a $T_g = 40$ ns long square pulse and sent it to a channel on the scope as a gating signal for the triggering. For each pump pulse signal there was a corresponding gating signal. We adjusted the offset such that it covered the time interval within which we observed the highest trigger counts in Figure 6.3 (left-most). The scope was then set to only accept APD click events occurring when this gating signal was in the high state. The actual trigger for the homodyne signal acquisition on the scope was not the APD clicks, though, but rather the gating signal. Therefore, the acquisition was always at a fixed time relative to the pump pulse and independent of the APD click timing, thus simulating a LO pulse synchronized with the pump.

We recorded 20,000 segments of homodyne data. The segment noise is extremely noisy, apparently due to some high frequency noise radiated by the pulse shaper during the fast switching of the high voltage – a noise that somehow got picked up by the homodyne detector. The segment noise could therefore not give any

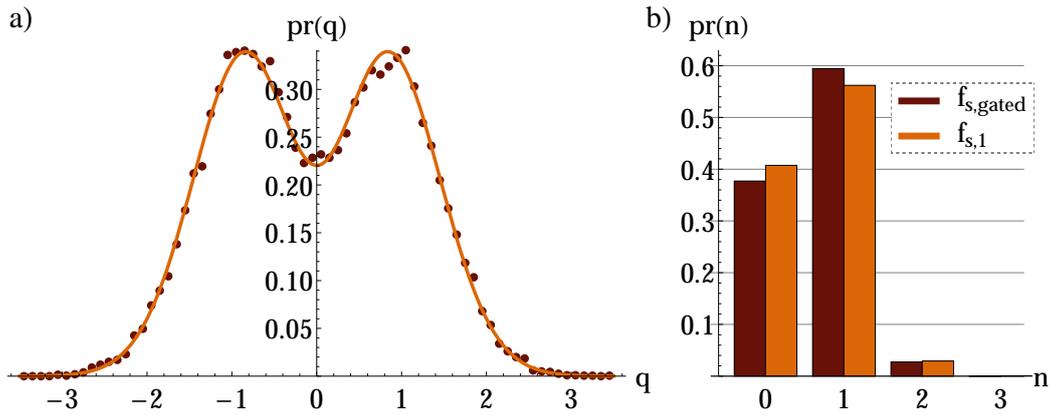


FIGURE 6.4: **a)** The quadrature distribution for the pulsed pump experiment using the mode function $f_{s,\text{gated}}(t - t_c)$. The orange curve is the best fit to a sum of the $n = 0 - 3$ number states. The photon number distribution is charted in **b)**, together with that obtained from using the symmetric exponential mode function $f_{s,1}(t - t_c)$.

hints for the mode function shape, apart from the delay. We assume, however, that the optimal mode function is close to that of (6.1). In this measurement, the gating interval $T_g = 40$ ns and the pulse length $T_p = 46$ ns have similar lengths, 15-18 times the OPO round trip time τ . Referring to the calculations in ref. [108], the optimal mode function will in this case have some small oscillations with period τ on top of the overall shape (6.1), but these are small corrections which will have little effect on the outcome. The trigger filtering should have little influence on the optimal mode function, since it no longer depends on the precise timing of the trigger click. In a first approximation, the trigger filtering just serves to delay the click. Thus, using the mode function $f_{s,\text{gated}}(t - t_c)$ of (6.1) with the 46 ns pump pulse $z_p(t)$ and an optimized horizontal offset t_c , we extract the quadrature data from the recorded segments. The results are plotted in Figure 6.4a) together with the best fit to a sum of number states, yielding a 1-photon fidelity of 59.4% with 37.7% vacuum and 2.7% 2-photon content. The photon number distribution is charted in Figure 6.4 together with that obtained by using the simple symmetric exponential mode function ($f_{s,1}(t - t_c)$ in (5.1)). The difference between the results with the two mode functions is not large – the 1-photon fidelity for $f_{s,1}$ is 56.2% – but it is still clear that the asymmetric $f_{s,\text{gated}}$ works best. For shorter pump pulses the leading edge is sharper, and the difference from choice of mode function should be more pronounced.

We have thus demonstrated the possibility of generating non-Gaussian states with a pulsed OPO pump, a gated trigger condition and a click-time independent signal mode function. Further studies should investigate more closely variations of the pump pulse length and the gating interval length (and perhaps offset), and how much is gained by the pulsed versus continuous pump when using the gated trigger. Optimally, the local oscillator should also be temporally shaped to fit the optimal mode function well. For an atomic memory experiment, there might be restrictions on the shape of the LO. A compromise should then be found between these restrictions and those imposed by the need for a high purity of the mode-matched non-Gaussian state. Luckily, the dependence

on the exact signal mode function shape is not very strong, as we have seen, and the shape can of course be adjusted by changing the OPO bandwidth, pump pulse length etc.

6.2 Other paths forward

The advances just mentioned for generation of narrow bandwidth, LO-matched single photons (and just as well kitten states) increase the feasibility of the storage of these states in an atomic quantum memory. The storage and retrieval of an externally prepared complex quantum state of light would be a strong demonstration of the capabilities of the memory. The single photon state could provide the first test, as its retrieval could be confirmed with just a photon detector. More challenging would be the full tomography of a stored non-Gaussian state, since, as we have seen, the dip in the Wigner function vanishes rapidly with increasing losses.

Apart from these prospects of interaction between the non-Gaussian light states and atomic quantum memories, other paths for future research within the purely optical domain lie ahead. These include e.g. generation of more complex non-Gaussian states, implementation of the states within optical quantum networks (like teleportation), and application of the developed methods to quantum information processing protocols such as distillation of entanglement [21] or to fundamental tests of quantum mechanics [46, 106]. Some progress has already been made within these areas. Especially in the Grangier group at Orsay, several interesting results have been achieved recently. They generated a two-photon state by conditioning on simultaneous clicks of two APDs in one mode of a two-mode squeezed state, which of course happens at a very low rate [121]. They subsequently used this two-photon state to prepare a special squeezed kitten state, employing conditional homodyne detection on a part of the two-photon state [120]. Finally, they were also able to observe an increase in the entanglement of an originally Gaussian two-mode squeezed state after a coherent subtraction of one photon from one beam or the other – a process reminiscent of entanglement distillation, although the final state was highly non-Gaussian [119]. These demonstrations were all done with a pulsed laser source. For a continuously pumped system like ours, there are some additional parameters to play with in non-Gaussian state generation schemes. In particular for subtraction of two or more photons from the initial Gaussian state, the interval between the two clicks become an important variable of the experiment, which can be optimized for e.g. large Wigner function negativity, high production rate, or large amplitude of the superposition state [109, 134, 154]. Recently, Takahashi et al. [148] could generate even kitten states with varying amplitudes based on variations of this two-click interval. Surely other sorts of non-Gaussian states will be possible to produce using the basic tools of squeezers, photon detectors and homodyne detectors. In particular high purity large amplitude cat states will be desirable, since the availability of these states could lead the way to all-optical quantum computation with continuous variables [131].

Quantum teleportation is an essential ingredient in many quantum information processing protocols. The usefulness of non-Gaussian states such as the kitten (or cat) states will be greatly increased if a teleporter can use them as input states for a high-fidelity state teleportation. During my stay at the Furusawa lab in Tokyo, we did some preliminary tests towards a demonstration of kitten state teleportation. So far, all cv teleportation experiments have been carried out at narrow frequency sidebands [16, 45, 51, 63, 150, 170, 178], so they are not compatible with the time localized/broadband states generated by conditional photon detection, such as the kittens. To be able to teleport such a state, two essential requirements must be met by the teleporter [160]: A high level of EPR-entanglement should be available at all frequencies occupied by the input state, and the classical channel (used to communicate the results of the Bell state measurement on the input and one of the EPR beams) should work with the correct electronic phase shift and gain settings throughout the same bandwidth. In ref. [149], we demonstrated the first of these two points, by observing more than 3 dB of EPR entanglement within an integrated top-hat temporal mode function, 0.2 μ s long, that is, within an integrated spectral range from DC to 5 MHz (weighted by a sinc-function). These “time-gated” correlations are clearly seen in Figure 6.5, where a sample of 50 quadrature measurements within consecutive top-hat mode functions are shown. Both x and p quadratures were measured for each of the two EPR beams. While the two x -traces follow each other neatly, the p -traces are more or less fluctuating oppositely to each other. Thus, the x -difference and the p -sum of the two modes are squeezed (-3.3 dB and -3.7 dB, respectively). This entanglement can then be applied to teleportation of broadband states of less than 5 MHz HWHM bandwidth – although the degree of entanglement should preferably be increased for high fidelity teleportation. As for the second requirement, the wideband classical channel, we build a very preliminary channel, consisting of homodyne detector, amplifier, a coherent displacement beam modulated by an EOM, and a 99/1 beamsplitter. The main issue for broadband operation is to obtain the correct phase of the EOM modulation at all frequencies. In standard sideband measurements, it is easy to just insert a variable phase shifter, but it is very hard to achieve a correct phase shift at different frequencies all the way down to DC. An alternative is therefore to make sure that there is basically no phase shift at all throughout the classical channel, which can only be achieved if all components have a wide bandwidth. We were somewhat limited by a slow detector, but could still show the ability of the channel to cancel up to 14 dB from DC to 500 kHz of a classical signal imprinted on two EPR beams, via homodyne detection on one beam and subsequent displacement (through the channel) of the second beam. As of this writing, the work on the classical channel is still ongoing, but may be close to a point where teleportation of a kitten or some other time-localized state could be carried out.

Finally, we can mention a few other relevant theoretical proposals. Regarding state generation, single photon states can be prepared at a higher rate if the trigger conditioning includes a prior interval of no detections [110]; kitten states can be purified with the help of conditional homodyne detection [146]; and arbitrary single-mode quantum states can be generated by continuous subtraction

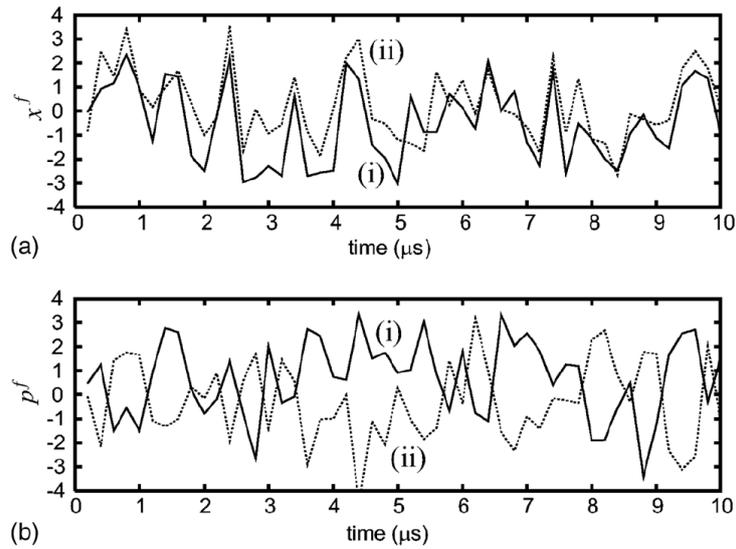


FIGURE 6.5: Time-domain EPR correlations. **a)** x -quadrature and **b)** p -quadrature measurements of the two EPR beams. Each point is obtained from an integration over a period of $0.2 \mu\text{s}$ of the oscilloscope-recorded quadrature data. From [149].

of photons [40]. Photon subtraction can also increase the performance of a teleportation protocol [115], and, as a more distantly related protocol, conditional photon detection also allows for universal quantum computation in a Gaussian cluster state setting [97]. These were just a few examples of the possible applications of conditional photon detection in continuous variable systems. Although several of them are rather infeasible with current technology, the number and variety of these theoretical proposals – together with the recent experimental advances – promise further progress in the near future within this area of applied quantum optics.

Various calculations

A.1 Gaussian integrals

The vector and matrix notation for multimode states as introduced on page 11 is both elegant and quite convenient for calculations. For example, it allows for easy evaluation of the characteristic function for a Gaussian state given the Wigner function (2.15)

$$W(\xi) = \frac{1}{\pi^N \sqrt{\det \Gamma}} e^{-\xi^T \Gamma^{-1} \xi}.$$

From (2.14), the characteristic function is

$$\begin{aligned} \chi(\zeta) &= \int e^{-i\zeta^T \sigma \zeta} W(\xi) d^{2N} \xi \\ &= \frac{1}{\pi^N \sqrt{\det \Gamma}} \int e^{-i\zeta^T \sigma \zeta - \xi^T \Gamma^{-1} \xi} d^{2N} \xi. \end{aligned} \tag{A.1}$$

Since Γ is positive, it can be factorized into square roots (and similarly for its inverse),

$$\Gamma = (\Gamma^{\frac{1}{2}})^T \Gamma^{\frac{1}{2}}, \quad \Gamma^{-1} = (\Gamma^{-\frac{1}{2}})^T \Gamma^{-\frac{1}{2}},$$

e.g. by the Cholesky decomposition, where $\Gamma^{\frac{1}{2}}$ is the an upper triangular matrix [124]. Define now the $2N$ -dimensional vector

$$z = \Gamma^{-\frac{1}{2}} (\zeta + \frac{i}{2} \Gamma \sigma \zeta).$$

We use this to “complete the square” in the exponent above. Using the symmetry of Γ ,

$$\begin{aligned} z^T z &= (\zeta^T + \frac{i}{2} \zeta^T \sigma^T \Gamma^T) (\Gamma^{-\frac{1}{2}})^T \Gamma^{-\frac{1}{2}} (\zeta + \frac{i}{2} \Gamma \sigma \zeta) \\ &= \zeta^T \Gamma^{-1} \zeta + i \zeta^T \sigma \zeta - \frac{1}{4} \zeta^T \sigma^T \Gamma \sigma \zeta. \end{aligned}$$

From the general substitution rule for multiple variables,

$$d^{2N} z = \left| \det \frac{\partial z}{\partial \zeta} \right| d^{2N} \zeta = \left| \det \Gamma^{-\frac{1}{2}} \right| = \frac{1}{\sqrt{\det \Gamma}}.$$

The final equality comes from standard rules for determinants. Substituting the new variable z into (A.1) gives

$$\chi(\zeta) = \frac{1}{\pi^N \sqrt{\det \Gamma}} e^{-\frac{1}{4} \zeta^T \sigma^T \Gamma \sigma \zeta} \int e^{-z^T z} \sqrt{\det \Gamma} d^{2N} z.$$

The plain $2N$ -dimensional Gaussian integral is $\int e^{-z^T z} d^{2N} z = \sqrt{\pi^{2N}}$, leaving the expression (2.16) for the characteristic function:

$$\chi(\zeta) = e^{-\frac{1}{4} \zeta^T \sigma^T \Gamma \sigma \zeta}.$$

A.2 Fidelity between 1-PSSqV and odd cat state

For $m' = 1$, the fidelity (3.2) becomes

$$\mathcal{F}(|\psi_1\rangle, |\tilde{\mathcal{O}}_-\rangle) = \frac{|v|^2}{\mathcal{N}_1 \sinh|\alpha|^2} \left| \sum_{n=0}^{\infty} \delta_{1n}^{[2]} \frac{(n+1)!}{\left(\frac{n+1}{2}\right)! n!} (\alpha v^*)^n \right|^2. \quad (\text{A.2})$$

\mathcal{N}_1 is simply

$$\begin{aligned} \mathcal{N}_1 &= \tanh^2 s {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, \frac{3}{2}, \tanh^2 s\right) \\ &= \frac{\tanh^2 s}{(1 - \tanh^2 s)^{3/2}} \\ &= \tanh^2 s \cosh^3 s. \end{aligned}$$

The factorials in (A.2) can be rewritten

$$\frac{(n+1)!}{\left(\frac{n+1}{2}\right)! n!} = \frac{(n+1)n!}{\frac{n+1}{2} \left(\frac{n-1}{2}\right)! n!} = \frac{2}{\left(\frac{n-1}{2}\right)!}.$$

Since only odd n contribute to the sum (represented by the $\delta_{1n}^{[2]}$), the summation index can be changed to k with $n = 2k + 1$. Using $|v|^2 = \tanh s/2$, (A.2) becomes

$$\begin{aligned} \mathcal{F}(|\psi_1\rangle, |\tilde{\mathcal{O}}_-\rangle) &= \frac{1}{2 \tanh s \cosh^3 s \sinh|\alpha|^2} \left| \sum_{k=0}^{\infty} \frac{2}{k!} (\alpha v^*)^{2k+1} \right|^2 \\ &= \frac{2|\alpha|^2 |v|^2}{\tanh s \cosh^3 s \sinh|\alpha|^2} \left| \sum_{k=0}^{\infty} \frac{(\alpha v^*)^{2k}}{k!} \right|^2 \\ &= \frac{|\alpha|^2}{\cosh^3 s \sinh|\alpha|^2} |e^{(\alpha v^*)^2}|^2. \end{aligned}$$

Writing explicitly $\alpha = |\alpha| e^{i \arg \alpha}$ and $v = \sqrt{\tanh s/2} e^{i\phi/2}$, the squared exponential is expanded as

$$\begin{aligned} |e^{(\alpha v^*)^2}|^2 &= \exp\left(|\alpha|^2 \frac{\tanh s}{2} (e^{i(2 \arg \alpha - \phi)} + e^{i(-2 \arg \alpha + \phi)})\right) \\ &= \exp\left(|\alpha|^2 \tanh s \cos(2 \arg \alpha - \phi)\right), \end{aligned}$$

such that the final expression is

$$\mathcal{F}(|\psi_1\rangle, |\tilde{\mathcal{O}}_-\rangle) = \frac{|\alpha|^2}{\cosh^3 s \sinh|\alpha|^2} \exp\left(|\alpha|^2 \tanh s \cos(2 \arg \alpha - \phi)\right). \quad (\text{A.3})$$

A.3 Equivalency of two detection models in the low efficiency limit

We show that the two models (3.4) and (3.11) for the signal state outcome after a detection event in the trigger mode of the two-mode state $\hat{\rho}$ are equivalent in the limit of low trigger detection efficiency η_t . In (3.4), the detection is modeled with the on-POVM $\hat{E}_t^{\text{on}}(\eta_t) = \hat{I}_t - \sum_m (1 - \eta_t)^m |m\rangle_t \langle m|$, with the signal state outcome

$$\hat{\rho}_{\text{on},s} = \frac{\text{tr}_t [\hat{E}_t^{\text{on}}(\eta_t) \hat{\rho}]}{\text{tr}_{s,t} [\hat{E}_t^{\text{on}}(\eta_t) \hat{\rho}]}.$$
 (A.4)

In (3.11), the detection is simply described by the annihilation operator:

$$\hat{\rho}_{\text{on},s} = \frac{\text{tr}_t [\hat{a}_t \hat{\rho} \hat{a}_t^\dagger]}{\text{tr}_{s,t} [\hat{a}_t \hat{\rho} \hat{a}_t^\dagger]}.$$
 (A.5)

The POVM model is more general, but in the limit $\eta_t \rightarrow 0$, it reduces to the annihilation operator model. In that limit, $(1 - \eta_t)^m \rightarrow 1 - \eta_t m$, so the on-POVM becomes

$$\hat{E}_t^{\text{on}}(\eta_t \rightarrow 0) = \hat{I}_t - \sum_m (|m\rangle_t \langle m| - \eta_t m |m\rangle_t \langle m|) = \eta_t \sum_m m |m\rangle_t \langle m|.$$

The numerator of (A.4) thus becomes

$$\text{tr}_t [\hat{E}_t^{\text{on}}(\eta_t \rightarrow 0) \hat{\rho}] = \text{tr}_t [\eta_t \sum_m m |m\rangle_t \langle m| \hat{\rho}] = \eta_t \sum_m m \langle m | \hat{\rho} | m \rangle_t.$$
 (A.6)

On the other hand, the numerator of (A.5) is

$$\text{tr}_t [\hat{a}_t \hat{\rho} \hat{a}_t^\dagger] = \sum_m \langle m | \hat{a}_t \hat{\rho} \hat{a}_t^\dagger | m \rangle_t = \sum_m (m+1) \langle m+1 | \hat{\rho} | m+1 \rangle_t,$$

which – apart from the factor η_t – is equal to (A.6). The η_t is taken account of, though, by the normalization factors in the denominator.

A.4 DOPO temporal correlation functions after filtering through an optical cavity

Here, we carry out the calculation of (3.30) and (3.31) for the DOPO correlation function $\langle : \hat{X}(t) \hat{X}(t') : \rangle$ with either one or both of the operators filtered through an optical cavity with the transformation given in (3.29),

$$\hat{X}(t) \rightarrow \hat{X}'(t) = \kappa \int_{-\infty}^t \hat{X}(y) e^{-\kappa(t-y)} dy + v(t) \hat{X}_{\text{vac}}(t).$$

The normal-ordered correlation function, as given by (3.21) and (3.22) is

$$\langle : \hat{X}(t) \hat{X}(t') : \rangle = \frac{\gamma_1 \epsilon}{\gamma - \epsilon} e^{-(\gamma - \epsilon)|t - t'|}.$$

Single filtering

With just one of the operators filtered through a cavity, such as it appears in the covariances between signal and trigger mode, the X -quadrature correlation function becomes

$$\langle : \hat{X}(t) \hat{X}'(t') : \rangle = \langle : \hat{X}(t) \left(\kappa \int_{-\infty}^{t'} \hat{X}(y) e^{-\kappa(t'-y)} dy + v(t') \hat{X}_{\text{vac}}(t') \right) : \rangle.$$

There are no correlations between the OPO output operator \hat{X} and the vacuum \hat{X}_{vac} , so the last term disappears. Furthermore, the integration can be taken outside the normal ordering and expectation value operations:

$$\begin{aligned} \langle : \hat{X}(t) \hat{X}'(t') : \rangle &= \kappa \int_{-\infty}^{t'} \langle : \hat{X}(t) \hat{X}(y) : \rangle e^{-\kappa(t'-y)} dy \\ &= \frac{\kappa \gamma_1 \epsilon}{\gamma - \epsilon} e^{-\kappa t'} \int_{-\infty}^{t'} e^{-(\gamma - \epsilon)|t-y| + \kappa y} dy. \end{aligned} \quad (\text{A.7})$$

Calculation of the integral must be separated into the two cases $t \geq t'$ and $t < t'$. In the first case, $t \geq y$, always, so $|t - y| = t - y$.

for $t \geq t'$:

$$\begin{aligned} \int_{-\infty}^{t'} e^{-(\gamma - \epsilon)|t-y| + \kappa y} dy &= e^{-(\gamma - \epsilon)t} \int_{-\infty}^{t'} e^{(\gamma - \epsilon + \kappa)y} dy \\ &= e^{-(\gamma - \epsilon)t} \frac{e^{(\gamma - \epsilon + \kappa)t'}}{\gamma - \epsilon + \kappa} \\ &= \frac{e^{-(\gamma - \epsilon)(t-t')} e^{\kappa t'}}{\gamma - \epsilon + \kappa}. \end{aligned}$$

The integral converges since $\gamma - \epsilon > 0$ in the sub-threshold regime.

In the case of $t < t'$, the integral can be split in two:

for $t < t'$:

$$\begin{aligned} \int_{-\infty}^{t'} e^{-(\gamma - \epsilon)|t-y| + \kappa y} dy &= e^{-(\gamma - \epsilon)t} \int_{-\infty}^t e^{(\gamma - \epsilon + \kappa)y} dy + e^{(\gamma - \epsilon)t} \int_t^{t'} e^{-(\gamma - \epsilon - \kappa)y} dy \\ &= e^{-(\gamma - \epsilon)t} \frac{e^{(\gamma - \epsilon + \kappa)t}}{\gamma - \epsilon + \kappa} - e^{(\gamma - \epsilon)t} \frac{e^{-(\gamma - \epsilon - \kappa)t'} - e^{-(\gamma - \epsilon - \kappa)t}}{\gamma - \epsilon - \kappa} \\ &= \frac{2(\gamma - \epsilon)e^{\kappa t}}{(\gamma - \epsilon + \kappa)(\gamma - \epsilon - \kappa)} - \frac{e^{-(\gamma - \epsilon)(t'-t)} e^{\kappa t'}}{\gamma - \epsilon - \kappa}. \end{aligned}$$

Inserting in (A.7),

$$\langle : \hat{X}(t) \hat{X}'(t') : \rangle = \begin{cases} \frac{\kappa \gamma_1 \epsilon}{(\gamma - \epsilon)(\gamma - \epsilon + \kappa)} e^{-(\gamma - \epsilon)(t-t')} & , t \geq t' \\ \frac{\kappa \gamma_1 \epsilon}{\gamma - \epsilon - \kappa} \left(\frac{2e^{-\kappa(t'-t)}}{\gamma - \epsilon + \kappa} - \frac{e^{-(\gamma - \epsilon)(t'-t)}}{\gamma - \epsilon} \right) & , t < t'. \end{cases}$$

Double filtering

In the variances of the trigger mode, both quadrature operators are filtered in the same way. The normal-ordered, double filtered correlation function is

$$\begin{aligned} \langle : \hat{X}'(t) \hat{X}'(t') : \rangle &= \kappa^2 \int_{-\infty}^t \int_{-\infty}^{t'} \langle : \hat{X}(y) \hat{X}(y') : \rangle e^{-\kappa(t-y)} e^{-\kappa(t'-y')} dy' dy \\ &= \frac{\kappa^2 \gamma_1 \epsilon}{\gamma - \epsilon} e^{-\kappa(t+t')} \int_{-\infty}^t \int_{-\infty}^{t'} e^{-(\gamma-\epsilon)|y-y'|} e^{-\kappa(y+y')} dy' dy. \end{aligned}$$

Because of the absolute value in the exponent, it is again useful to begin with the case $t \geq t'$. Then, $t \geq y'$ as well, and the integral over y can be separated in two at $y = y'$:

for $t \geq t'$:

$$\begin{aligned} &\int_{-\infty}^t \int_{-\infty}^{t'} e^{-(\gamma-\epsilon)|y-y'|} e^{-\kappa(y+y')} dy' dy \\ &= \int_{-\infty}^{t'} \left[e^{-(\gamma-\epsilon-\kappa)y'} \int_{-\infty}^{y'} e^{(\gamma-\epsilon+\kappa)y} dy + e^{(\gamma-\epsilon+\kappa)y'} \int_{y'}^t e^{-(\gamma-\epsilon-\kappa)y} dy \right] dy' \\ &= \int_{-\infty}^{t'} \left[\frac{e^{2\kappa y'}}{\gamma - \epsilon + \kappa} - \frac{e^{-(\gamma-\epsilon-\kappa)t} e^{(\gamma-\epsilon+\kappa)y'}}{\gamma - \epsilon - \kappa} + \frac{e^{2\kappa y'}}{\gamma - \epsilon - \kappa} \right] dy' \\ &= \frac{\gamma - \epsilon}{\kappa} \frac{e^{2\kappa t'}}{(\gamma - \epsilon)^2 - \kappa^2} - \frac{e^{-(\gamma-\epsilon)(t-t')} e^{\kappa(t+t')}}{(\gamma - \epsilon)^2 - \kappa^2}. \end{aligned}$$

For $t < t'$, the expression is the same, but with t and t' exchanged, such that in the end, all exponents are negative. The full expression for the correlation function thus becomes

$$\langle : \hat{X}'(t) \hat{X}'(t') : \rangle = \frac{\kappa^2 \gamma_1 \epsilon}{(\gamma - \epsilon)^2 - \kappa^2} \left[\frac{e^{-\kappa|t-t'|}}{\kappa} - \frac{e^{-(\gamma-\epsilon)|t-t'|}}{\gamma - \epsilon} \right].$$

A.5 Full expressions for the OPO state generation model

The kitten state generation scheme based on the OPO output that was summarized in §3.4.3 has the following parameters:

γ	total OPO decay rate
ϵ	pump/gain parameter
T	tapping beamsplitter transmittance
κ	decay rate of trigger filter cavity
η_s	overall efficiency in signal arm
η_t	overall efficiency in trigger arm
Ξ	modal purity (percentage of true trigger clicks)

The output coupler decay rate γ_1 is rewritten as $\gamma_1 = \eta_{\text{esc}}\gamma$, and η_{esc} is implicitly incorporated into η_s and η_t .

With the trigger mode function $f_t(t) = \delta(t - t_c)$, and the signal mode function $f_s(t) = \sqrt{\gamma_s}e^{-\gamma_s|t-t_c|}$, there is the additional parameter γ_s (we set $t_c = 0$). Using Mathematica, full analytical expressions for the entries of the final covariance matrix Γ^{final} can be calculated:

$$\begin{aligned}\Gamma_{11}^{\text{final}} &= 1 + T\eta_s \frac{4\gamma\epsilon(\gamma - \epsilon + 2\gamma_s)}{(\gamma - \epsilon)(\gamma - \epsilon + \gamma_s)^2} \\ \Gamma_{22}^{\text{final}} &= 1 - T\eta_s \frac{4\gamma\epsilon(\gamma + \epsilon + 2\gamma_s)}{(\gamma + \epsilon)(\gamma + \epsilon + \gamma_s)^2} \\ \Gamma_{33}^{\text{final}} &= 1 + (1 - T)\eta_t \frac{2\gamma\epsilon\kappa}{(\gamma - \epsilon)(\gamma - \epsilon + \kappa)} \\ \Gamma_{44}^{\text{final}} &= 1 - (1 - T)\eta_t \frac{2\gamma\epsilon\kappa}{(\gamma + \epsilon)(\gamma + \epsilon + \kappa)} \\ \Gamma_{13}^{\text{final}} &= -\sqrt{T(1 - T)}\sqrt{\eta_s\eta_t} \frac{4\gamma\epsilon\kappa\sqrt{\gamma_s}(\gamma - \epsilon + \kappa + \gamma_s)}{(\gamma - \epsilon)(\gamma - \epsilon + \kappa)(\gamma - \epsilon + \gamma_s)(\kappa + \gamma_s)} \\ \Gamma_{24}^{\text{final}} &= \sqrt{T(1 - T)}\sqrt{\eta_s\eta_t} \frac{4\gamma\epsilon\kappa\sqrt{\gamma_s}(\gamma + \epsilon + \kappa + \gamma_s)}{(\gamma + \epsilon)(\gamma + \epsilon + \kappa)(\gamma + \epsilon + \gamma_s)(\kappa + \gamma_s)}.\end{aligned}$$

For the single photon generation using the NDOPO, the entries of the covariance matrix $\Gamma^{\text{final,nd}}$ are simple combinations of the ones for the DOPO above. This can be seen by studying the correlations (3.21) and (3.27)-(3.28). Since there is no beamsplitter in the single photon scheme, the factors including T must be eliminated from the expressions for Γ^{final} . Apart from that, the NDOPO entries are basically averages of the X and P covariances for the DOPO:

$$\begin{aligned}\Gamma_{11}^{\text{final,nd}} = \Gamma_{22}^{\text{final,nd}} &= \frac{1}{2} \left(\frac{\Gamma_{11}^{\text{final}} - 1}{T} + \frac{\Gamma_{22}^{\text{final}} - 1}{T} \right) + 1, \\ \Gamma_{33}^{\text{final,nd}} = \Gamma_{44}^{\text{final,nd}} &= \frac{1}{2} \left(\frac{\Gamma_{33}^{\text{final}} - 1}{1 - T} + \frac{\Gamma_{44}^{\text{final}} - 1}{1 - T} \right) + 1, \\ \Gamma_{13}^{\text{final,nd}} = -\Gamma_{24}^{\text{final,nd}} &= -\frac{1}{2} \left(\frac{\Gamma_{13}^{\text{final}}}{\sqrt{T(1 - T)}} - \frac{\Gamma_{24}^{\text{final}}}{\sqrt{T(1 - T)}} \right).\end{aligned}$$

Bibliography

- [1] N. H. Adamyan, H. H. Adamyan, and G. Y. Kryuchkian, "Time-domain squeezing and quantum distributions in the pulsed regime," *Physical Review A* **77**, 023820 (2008).
- [2] J. Appel, E. Figueroa, D. Korystov, M. Lobino, and A. I. Lvovsky, "Quantum Memory for Squeezed Light," *Physical Review Letters* **100**, 093602 (2008).
- [3] J. Appel, D. Hoffman, E. Figueroa, and A. I. Lvovsky, "Electronic noise in optical homodyne tomography," *Physical Review A* **75**, 035802 (2007).
- [4] A. Arie, G. Rosenman, V. Mahal, A. Skliar, M. Oron, M. Katz, and D. Eger, "Green and ultraviolet quasi-phase-matched second harmonic generation in bulk periodically-poled KTiOPO₄," *Optics Communications* **142**, 265 (1997).
- [5] A. Aspect, P. Grangier, and G. Roger, "Experimental tests of realistic local theories via Bell's theorem," *Physical Review Letters* **47**, 460 (1981).
- [6] S. A. Babichev, J. Appel, and A. I. Lvovsky, "Homodyne tomography characterization and nonlocality of a dual-mode optical qubit," *Physical Review Letters* **92**, 193601 (2004).
- [7] H. A. Bachor and T. C. Ralph, *A Guide to Experiments in Quantum Optics* (New York: Wiley-VCH Verlag GmbH & Co. KgaA, 2004).
- [8] K. Banaszek, "Maximum-likelihood estimation of photon-number distribution from homodyne statistics," *Physical Review A* **57**, 5013 (1998).
- [9] K. Banaszek, "Quantum homodyne tomography with a priori constraints," *Physical Review A* **59**, 4797 (1999).
- [10] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, "Maximum-likelihood estimation of the density matrix," *Physical Review A* **61**, 010304 (1999).
- [11] J. A. Bergou and B. G. Englert, "Heisenbergs dog and quantum computing," *Journal of Modern Optics* **45**, 701 (1998).
- [12] P. Bertet, A. Auffeves, P. Maioli, S. Osnaghi, T. Meunier, M. Brune, J. M. Raimond, and S. Haroche, "Direct measurement of the Wigner function

- of a one-photon Fock state in a cavity," *Physical Review Letters* **89**, 200402 (2002).
- [13] J. Bertrand and P. Bertrand, "A tomographic approach to Wigner's function," *Foundations of Physics* **17**, 397 (1987).
- [14] B. Boulanger, I. Rousseau, J. P. Feve, M. Maglione, B. Menaert, and G. Marnier, "Optical studies of laser-induced gray-tracking in KTP," *IEEE Journal of Quantum Electronics* **35**, 281 (1999).
- [15] D. Bouwmeester, J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature* **390**, 11 (1997).
- [16] W. P. Bowen, N. Treps, B. C. Buchler, R. Schnabel, T. C. Ralph, H. A. Bachor, T. Symul, and P. K. Lam, "Experimental investigation of continuous-variable quantum teleportation," *Physical Review A* **67**, 032302 (2003).
- [17] W. P. Bowen, N. Treps, R. Schnabel, and P. K. Lam, "Experimental Demonstration of Continuous Variable Polarization Entanglement," *Physical Review Letters* **89**, 253601 (2002).
- [18] G. D. Boyd and D. A. Kleinman, "Parametric interaction of focused Gaussian beams," *Journal of Applied Physics* **39**, 3597 (1968).
- [19] S. L. Braunstein, C. A. Fuchs, and H. J. Kimble, "Criteria for continuous-variable quantum teleportation," *Journal of Modern Optics* **47**, 267 (2000).
- [20] S. L. Braunstein and P. van Loock, "Quantum information with continuous variables," *Reviews of Modern Physics* **77**, 513 (2005).
- [21] D. E. Browne, J. Eisert, S. Scheel, and M. B. Plenio, "Driving non-Gaussian to Gaussian states with linear optics," *Physical Review A* **67**, 062320 (2003).
- [22] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche, "Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement," *Physical Review Letters* **77**, 4887 (1996).
- [23] N. J. Cerf, G. Leuchs, and E. S. Polzik (eds.), *Quantum Information with Continuous Variables of Atoms and Light* (Imperial College Press, 2007).
- [24] T. Chanelière, D. N. Matsukevich, S. D. Jenkins, S. Y. Lan, T. A. B. Kennedy, and A. Kuzmich, "Storage and retrieval of single photons transmitted between remote quantum memories," *Nature* **438**, 833 (2005).
- [25] M. J. Collett and C. W. Gardiner, "Squeezing of intracavity and traveling-wave light fields produced in parametric amplification," *Physical Review A* **30**, 1386 (1984).
- [26] M. Dakna, J. Clausen, L. Knöll, and D.-G. Welsch, "Generation of arbitrary quantum states of traveling fields," *Phys. Rev. A* **59**, 1658 (1999).

-
- [27] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris, "Detection of the density matrix through optical homodyne tomography without filtered back projection," *Physical Review A* **50**, 4298 (1994).
- [28] G. M. D'Ariano, L. Maccone, and M. F. Sacchi, "Homodyne tomography and the reconstruction of quantum states of light," in N. J. Cerf, G. Leuchs, and E. S. Polzik (eds.), "Quantum Information with Continuous Variables of Atoms and Light," (Imperial College Press, 2007), chap. 8.
- [29] B. Darquie, M. P. A. Jones, J. Dingjan, J. Beugnon, S. Bergamini, Y. Sortais, G. Messin, A. Browaeys, and P. Grangier, "Controlled single-photon emission from a single trapped two-level atom," *Science* **309**, 454 (2005).
- [30] D. Dragoman, "Phase space correspondence between classical optics and quantum mechanics," *Progress in Optics* **43**, 433 (2002).
- [31] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," *Applied Physics B: Lasers and Optics* **31**, 97 (1983).
- [32] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, "Long-distance quantum communication with atomic ensembles and linear optics," *Nature* **414**, 413 (2001).
- [33] A. E. Dunlop, E. H. Huntington, C. C. Harb, and T. C. Ralph, "Generation of a frequency comb of squeezing in an optical parametric oscillator," *Physical Review A* **73**, 013817 (2006).
- [34] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Physical Review* **47**, 777 (1935).
- [35] M. D. Eisaman, A. Andre, F. Massou, M. Fleischhauer, A. S. Zibrov, and M. D. Lukin, "Electromagnetically induced transparency with tunable single-photon pulses," *Nature* **438**, 837 (2005).
- [36] J. Eisert and M. B. Plenio, "Introduction to the basics of entanglement theory in continuous-variable systems," *International Journal of Quantum Information* **1**, 479 (2003).
- [37] J. Eisert, S. Scheel, and M. B. Plenio, "Distilling Gaussian states with Gaussian operations is impossible," *Physical Review Letters* **89**, 137903 (2002).
- [38] J. Fiurášek, "Maximum-likelihood estimation of quantum measurement," *Physical Review A* **64**, 024102 (2001).
- [39] J. Fiurášek, "Gaussian transformations and distillation of entangled Gaussian states," *Physical Review Letters* **89**, 137904 (2002).
- [40] J. Fiurášek, R. García-Patrón, and N. J. Cerf, "Conditional generation of arbitrary single-mode quantum states of light by repeated photon subtractions," *Physical Review A* **72**, 033822 (2005).

- [41] G. T. Foster, L. A. Orozco, H. M. Castro-Beltran, and H. J. Carmichael, "Quantum state reduction and conditional time evolution of wave-particle correlations in cavity QED," *Physical Review Letters* **85**, 3149 (2000).
- [42] H.A. Freeman, *Introduction to Statistical Inference* (Addison-Wesley Educational Publishers Inc., US, 1963).
- [43] J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, and J. E. Lukens, "Quantum superposition of distinct macroscopic states," *Nature* **406**, 43 (2000).
- [44] J. Fulconis, O. Alibart, W. Wadsworth, P. Russell, and J. Rarity, "High brightness single mode source of correlated photon pairs using a photonic crystal fiber," *Optics Express* **13**, 7572 (2005).
- [45] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, "Unconditional quantum teleportation," *Science* **282**, 706 (1998).
- [46] R. García-Patrón, J. Fiurášek, N. J. Cerf, J. Wenger, R. Tualle-Brouiri, and P. Grangier, "Proposal for a loophole-free Bell test using homodyne detection," *Physical Review Letters* **93**, 130409 (2004).
- [47] C. W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods With Applications to Quantum Optics* (Springer Verlag, 2000).
- [48] C. C. Gerry and P. L. Knight, "Quantum superpositions and Schrödinger cat states in quantum optics," *American Journal of Physics* **65**, 964 (1997).
- [49] G. Giedke and J. I. Cirac, "Characterization of Gaussian operations and distillation of Gaussian states," *Physical Review A* **66**, 032316 (2002).
- [50] O. Glöckl, J. Heersink, N. Korolkova, G. Leuchs, and S. Lorenz, "A pulsed source of continuous variable polarization entanglement," *Journal of Optics B: Quantum and Semiclassical Optics* **5**, S492 (2003).
- [51] O. Glöckl, S. Lorenz, C. Marquardt, J. Heersink, M. Brownnutt, C. Silberhorn, Q. Pan, P. van Loock, N. Korolkova, and G. Leuchs, "Experiment towards continuous-variable entanglement swapping: Highly correlated four-partite quantum state," *Physical Review A* **68**, 012319 (2003).
- [52] H. Goto, Y. Yanagihara, H. Wang, T. Horikiri, and T. Kobayashi, "Observation of an oscillatory correlation function of multimode two-photon pairs," *Physical Review A* **68**, 015803 (2003).
- [53] F. Grosshans and P. Grangier, "Continuous variable quantum cryptography using coherent states," *Physical Review Letters* **88**, 057902 (2002).
- [54] K Hammerer, E. S. Polzik, and A. S. Sørensen, "Quantum interface between light and atomic ensembles," *to appear in Rev. Mod. Phys.* (2008).
- [55] R. Hanbury Brown and R. Q. Twiss, "Correlation between photons in two coherent beams of light," *Nature* **177**, 27 (1956).

-
- [56] G. Hansson, H. Karlsson, S. Wang, and F. Laurell, "Transmission measurements in KTP and isomorphic compounds," *Applied Optics* **39**, 5058 (2000).
- [57] J. Heersink, C. Marquardt, R. Dong, R. Filip, S. Lorenz, G. Leuchs, and U. L. Andersen, "Distillation of squeezing from non-Gaussian quantum states," *Physical Review Letters* **96**, 253601 (2006).
- [58] D. R. Herriott and H. J. Schulte, "Folded optical delay lines," *Applied Optics* **4**, 883 (1965).
- [59] K. Honda, D. Akamatsu, M. Arikawa, Y. Yokoi, K. Akiba, S. Nagatsuka, T. Tanimura, A. Furusawa, and M. Kozuma, "Storage and retrieval of a squeezed vacuum," *Physical Review Letters* **100**, 093601 (2008).
- [60] Z. Hradil, "Quantum-state estimation," *Physical Review A* **55**, 1561 (1997).
- [61] Z. Hradil, D. Mogilevtsev, and J. Řeháček, "Biased tomography schemes: An objective approach," *Physical Review Letters* **96**, 230401 (2006).
- [62] H. Jeong, A. P. Lund, and T. C. Ralph, "Production of superpositions of coherent states in traveling optical fields with inefficient photon detection," *Physical Review A* **72**, 013801 (2005).
- [63] X. Jia, X. Su, Q. Pan, J. Gao, C. Xie, and K. Peng, "Experimental demonstration of unconditional entanglement swapping for continuous variables," *Physical Review Letters* **93**, 250503 (2004).
- [64] R. Jozsa, "Fidelity for mixed quantum states," *Journal of Modern Optics* **41**, 2315 (1994).
- [65] B. Julsgaard, A. Kozhekin, and E. S. Polzik, "Experimental long-lived entanglement of two macroscopic objects." *Nature* **413**, 400 (2001).
- [66] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurasek, and E. S. Polzik, "Experimental demonstration of quantum memory for light," *Nature* **432**, 482 (2004).
- [67] K. Kato and E. Takaoka, "Sellmeier and thermo-optic dispersion formulas for KTP," *Applied Optics* **41**, 76 (2002).
- [68] M. Keller, B. Lange, K. Hayasaka, W. Lange, and H. Walther, "Continuous generation of single photons with controlled waveform in an ion-trap cavity system," *Nature* **431**, 1075 (2004).
- [69] M. S. Kim, E. Park, P. L. Knight, and H. Jeong, "Nonclassicality of a photon-subtracted Gaussian field," *Physical Review A* **71**, 043805 (2005).
- [70] H. J. Kimble, M. Dagenais, and L. Mandel, "Photon antibunching in resonance fluorescence," *Physical Review Letters* **39**, 691 (1977).
- [71] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature* **409**, 46 (2001).

- [72] C. E. Kuklewicz, F. N. C. Wong, and J. H. Shapiro, "Time-bin-modulated biphotons from cavity-enhanced down-conversion," *Physical Review Letters* **97**, 223601 (2006).
- [73] J. Laurat, T. Coudreau, N. Treps, A. Maître, and C. Fabre, "Conditional preparation of a quantum state in the continuous variable regime: Generation of a sub-Poissonian state from twin beams," *Physical Review Letters* **91**, 213601 (2003).
- [74] R. Le Targat, J. J. Zondy, and P. Lemonde, "75%-Efficiency blue generation from an intracavity PPKTP frequency doubler," *Optics Communications* **247**, 471 (2005).
- [75] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, R. Reichle, and D. J. Wineland, "Creation of a six-atom 'Schroedinger cat' state," *Nature* **438**, 639 (2005).
- [76] D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland, "Experimental determination of the motional quantum state of a trapped atom," *Physical Review Letters* **77**, 4281 (1996).
- [77] Lasse Leick, *The construction of a TiSapph laser and the use hereof in measurements on a Cs MOT*, Master's thesis, University of Aarhus (1998).
- [78] U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, 1997).
- [79] U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. G. Raymer, "Sampling of photon statistics and density matrix using homodyne detection," *Optics Communications* **127**, 144 (1996).
- [80] Z. M. Liao, S. A. Payne, J. Dawson, A. Drobshoff, C. Ebberts, D. Pennington, and L. Taylor, "Thermally induced dephasing in periodically poled KTP frequency-doubling crystals," *Journal of the Optical Society of America B* **21**, 2191 (2004).
- [81] C. Liu, Z. Dutton, C. H. Behroozi, and L. V. Hau, "Observation of coherent optical information storage in an atomic medium using halted light pulses," *Nature* **409**, 490 (2001).
- [82] B. Lounis and W. E. Moerner, "Single photons on demand from a single molecule at room temperature," *Nature* **407**, 491 (2000).
- [83] Y. J. Lu, R. L. Campbell, and Z. Y. Ou, "Mode-locked two-photon states," *Physical Review Letters* **91**, 163602 (2003).
- [84] YJ Lu and ZY Ou, "Optical parametric oscillator far below threshold: Experiment versus theory," *Physical Review A* **62**, 033804 (2000).
- [85] A. P. Lund, H. Jeong, T. C. Ralph, and M. S. Kim, "Conditional production of superpositions of coherent states with inefficient photon detection," *Physical Review A* **70**, 020101 (2004).

-
- [86] A. P. Lund and T. C. Ralph, "Coherent-state linear optical quantum computing gates using simplified diagonal superposition resource states," *Physical Review A* **71**, 032305 (2005).
- [87] J. H. Lundeman, O. B. Jensen, P. E. Andersen, S. Andersson-Engels, B. Sumpf, G. Erbert, and P. M. Petersen, "High power 404 nm source based on second harmonic generation in PPKTP of a tapered external feedback diode laser," *Optics Express* **16**, 2486 (2008).
- [88] N. Lütkenhaus and S. M. Barnett, "Nonclassical effects in phase space," *Physical Review A* **51**, 3340 (1995).
- [89] A. I. Lvovsky, "Iterative maximum-likelihood reconstruction in quantum homodyne tomography," *Journal of Optics B: Quantum and Semiclassical Optics* **6**, 556 (2004).
- [90] A. I. Lvovsky and S. A. Babichev, "Synthesis and tomographic characterization of the displaced Fock state of light," *Physical Review A* **66**, 011801 (2002).
- [91] A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller, "Quantum state reconstruction of the single-photon Fock state," *Physical Review Letters* **87**, 050402 (2001).
- [92] A. I. Lvovsky and M. G. Raymer, "Continuous variable quantum-state tomography of optical fields and photons," in N. J. Cerf, G. Leuchs, and E. S. Polzik (eds.), "Quantum Information with Continuous Variables of Atoms and Light," (Imperial College Press, 2007), chap. 21.
- [93] H. Mabuchi, E. S. Polzik, and H. J. Kimble, "Blue-light-induced infrared absorption in KNbO₃," *Journal of the Optical Society of America B* **11**, 2023 (1994).
- [94] L. P. Maguire, S. Szilagy, and R. E. Scholten, "High performance laser shutter using a hard disk drive voice-coil actuator," *Review of Scientific Instruments* **75**, 3077 (2004).
- [95] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).
- [96] J. McKeever, A. Boca, A. D. Boozer, R. Miller, J. R. Buck, A. Kuzmich, and H. J. Kimble, "Deterministic generation of single photons from one atom trapped in a cavity," *Science* **303**, 1992 (2004).
- [97] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, "Universal quantum computation with continuous-variable cluster states," *Physical Review Letters* **97**, 110501 (2006).
- [98] D. Mogilevtsev, Z. Hradil, and J. Perina, "Homodyne reconstruction of density matrix in Fock-state basis: deterministic versus maximum likelihood approach," *Journal of Modern Optics* **44**, 2261 (1997).

- [99] K. Mølmer, "Non-Gaussian states from continuous-wave Gaussian light sources," *Physical Review A* **73**, 063804 (2006).
- [100] C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland, "A "Schrödinger cat" superposition state of an atom," *Science* **272**, 1131 (1996).
- [101] M. Munroe, D. Boggavarapu, M. E. Anderson, and M. G. Raymer, "Photon-number statistics from the phase-averaged quadrature-field distribution: Theory and ultrafast measurement," *Physical Review A* **52**, 924 (1995).
- [102] V. Murk, V. Denks, A. Dudelzak, P. P. Proulx, and V. Vassiltsenko, "Gray tracks in KTiOPO₄: Mechanism of creation and bleaching," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **141**, 472 (1998).
- [103] J. S. Neergaard-Nielsen, B. M. Nielsen, C. Hettich, K. Mølmer, and E. S. Polzik, "Generation of a superposition of odd photon number states for quantum information networks," *Physical Review Letters* **97**, 083604 (2006).
- [104] J. S. Neergaard-Nielsen, B. M. Nielsen, H. Takahashi, A. I. Vistnes, and E. S. Polzik, "High purity bright single photon source," *Optics Express* **15**, 7940 (2007).
- [105] Jonas S. Neergaard-Nielsen, *Nonclassical light source for quantum networks*, Master's thesis, University of Copenhagen (2005), URL http://neer.dk/phd/jsnn_masterthesis.pdf.
- [106] H. Nha and H. J. Carmichael, "Proposed test of quantum nonlocality for continuous variables," *Physical Review Letters* **93**, 020401 (2004).
- [107] A. E. B. Nielsen and K. Mølmer, "Conditional generation of path-entangled optical $|N,0\rangle + |0,N\rangle$ states," *Physical Review A* **75**, 063803 (2007).
- [108] A. E. B. Nielsen and K. Mølmer, "Multimode analysis of the light emitted from a pulsed optical parametric oscillator," *Physical Review A* **76**, 033832 (2007).
- [109] A. E. B. Nielsen and K. Mølmer, "Photon number states generated from a continuous-wave light source," *Physical Review A* **75**, 043801 (2007).
- [110] A. E. B. Nielsen and K. Mølmer, "Single-photon-state generation from a continuous-wave nondegenerate optical parametric oscillator," *Physical Review A* **75**, 023806 (2007).
- [111] A. E. B. Nielsen and K. Mølmer, "Transforming squeezed light into a large-amplitude coherent-state superposition," *Physical Review A* **76**, 043840 (2007).
- [112] Bo Melholt Nielsen, *Generation of a Schrödinger kitten state of light*, Master's thesis, University of Copenhagen (2006).
- [113] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

-
- [114] T. Opatrny, N. Korolkova, and G. Leuchs, "Mode structure and photon number correlations in squeezed quantum pulses," *Physical Review A* **66**, 053813 (2002).
- [115] T. Opatrny, G. Kurizki, and D. G. Welsch, "Improvement on teleportation of continuous variables by photon subtraction via conditional measurement," *Physical Review A* **61**, 032302 (2000).
- [116] Z. Y. Ou and H. J. Kimble, "Probability distribution of photoelectric currents in photodetection processes and its connection to the measurement of a quantum state," *Physical Review A* **52**, 3126 (1995).
- [117] Z. Y. Ou and L. Mandel, "Violation of Bell's inequality and classical probability in a two-photon correlation experiment," *Physical Review Letters* **61**, 50 (1988).
- [118] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng, "Realization of the Einstein-Podolsky-Rosen paradox for continuous variables," *Physical Review Letters* **68**, 3663 (1992).
- [119] A. Ourjoumtsev, A. Dantan, R. Tualle-Brouri, and P. Grangier, "Increasing entanglement between Gaussian states by coherent photon subtraction," *Physical Review Letters* **98**, 030502 (2007).
- [120] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier, "Generation of optical "Schrödinger cats" from photon number states," *Nature* **448**, 784 (2007).
- [121] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, "Quantum homodyne tomography of a two-photon Fock state," *Physical Review Letters* **96**, 213601 (2006).
- [122] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier, "Generating optical Schrödinger kittens for quantum information processing," *Science* **312**, 83 (2006).
- [123] R. Paschotta, "Quasi-phase matching," *Encyclopedia of Laser Physics and Technology* (2008), URL http://www.rp-photonics.com/quasi_phase_matching.html.
- [124] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," (2007), URL <http://matrixcookbook.com/>.
- [125] D. F. Phillips, A. Fleischhauer, A. Mair, R. L. Walsworth, and M. D. Lukin, "Storage of light in atomic vapor," *Physical Review Letters* **86**, 783 (2001).
- [126] T. B. Pittman, B. C. Jacobs, and J. D. Franson, "Single photons on pseudo-demand from stored parametric down-conversion," *Physical Review A* **66**, 042303 (2002).
- [127] T. B. Pittman, B. C. Jacobs, and J. D. Franson, "Heralding single photons from pulsed parametric down-conversion," *Optics Communications* **246**, 545 (2005).

- [128] E. S. Polzik, J. Carri, and H. J. Kimble, "Atomic spectroscopy with squeezed light for sensitivity beyond the vacuum-state limit," *Applied Physics B: Lasers and Optics* **55**, 279 (1992).
- [129] E. S. Polzik and H. J. Kimble, "Frequency doubling with KNbO₃ in an external cavity," *Optics Letters* **16**, 1400 (1991).
- [130] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77* (Cambridge University Press, 1992).
- [131] T. C. Ralph, A. Gilchrist, G. J. Milburn, W. J. Munro, and S. Glancy, "Quantum computation with optical coherent states," *Physical Review A* **68**, 042319 (2003).
- [132] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, "Diluted maximum-likelihood algorithm for quantum tomography," *Physical Review A* **75**, 042108 (2007).
- [133] M. Sasaki and S. Suzuki, "Multimode theory of measurement-induced non-Gaussian operation on wideband squeezed light: Analytical formula," *Physical Review A* **73**, 043807 (2006).
- [134] M. Sasaki, M. Takeoka, and H. Takahashi, "Temporally multiplexed superposition states of continuous variables," *preprint arXiv:0804.0354* (2008).
- [135] W. Schleich, *Quantum optics in phase space* (Wiley-VCH New York, 2001).
- [136] C. Schori, J. L. Sørensen, and E. S. Polzik, "Narrow-band frequency tunable light source of continuous quadrature entanglement," *Physical Review A* **66**, 033802 (2002).
- [137] Christian Schori, *Frequency tunable light source of continuous quadrature entanglement*, Ph.D. thesis, University of Aarhus (2002).
- [138] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," *Naturwissenschaften* **23**, 807 (1935).
- [139] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, 1997).
- [140] J. F. Sherson, H. Krauter, R. K. Olsson, B. Julsgaard, K. Hammerer, I. Cirac, and E. S. Polzik, "Quantum teleportation between light and matter." *Nature* **443**, 557 (2006).
- [141] C. Silberhorn, P. K. Lam, O. Weiß, F. König, N. Korolkova, and G. Leuchs, "Generation of continuous variable Einstein-Podolsky-Rosen entanglement via the Kerr nonlinearity in an optical fiber," *Physical Review Letters* **86**, 4267 (2001).
- [142] R. Simon, "Peres-Horodecki separability criterion for continuous variable systems," *Physical Review Letters* **84**, 2726 (2000).

-
- [143] R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz, and J. F. Valley, "Observation of squeezed states generated by four-wave mixing in an optical cavity," *Physical Review Letters* **55**, 2409 (1985).
- [144] 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," *Physical Review Letters* **70**, 1244 (1993).
- [145] D. Stoler, "Equivalence classes of minimum uncertainty packets," *Physical Review D* **1**, 3217 (1970).
- [146] S. Suzuki, M. Takeoka, M. Sasaki, U. L. Andersen, and F. Kannari, "Practical purification scheme for decohered coherent-state superpositions via partial homodyne detection," *Physical Review A* **73**, 042304 (2006).
- [147] S. Suzuki, K. Tsujino, F. Kannari, and M. Sasaki, "Analysis on generation schemes of Schrödinger cat-like states under experimental imperfections," *Optics Communications* **259**, 758 (2006).
- [148] H. Takahashi, K. Wakui, S. Suzuki, M. Takeoka, K. Hayasaka, A. Furusawa, and M. Sasaki, "Generation of large-amplitude coherent-state superposition via ancilla-assisted photon-subtraction," *preprint arXiv:0806.2965* (2008).
- [149] N. Takei, N. Lee, D. Moriyama, J. S. Neergaard-Nielsen, and A. Furusawa, "Time-gated Einstein-Podolsky-Rosen correlation," *Physical Review A* **74**, 060101 (2006).
- [150] N. Takei, H. Yonezawa, T. Aoki, and A. Furusawa, "High-fidelity teleportation beyond the no-cloning limit and entanglement swapping for continuous variables," *Physical Review Letters* **94**, 220502 (2005).
- [151] Y. Takeno, M. Yukawa, H. Yonezawa, and A. Furusawa, "Observation of-9 dB quadrature squeezing with improvement of phase stability in homodyne measurement," *Optics Express* **15**, 4321 (2007).
- [152] Yuishi Takeno, private communication (2007).
- [153] M. Takeoka and M. Sasaki, "Conditional generation of an arbitrary superposition of coherent states," *Physical Review A* **75**, 064302 (2007).
- [154] M. Takeoka, H. Takahashi, and M. Sasaki, "Large amplitude coherent state superposition generated by a time-separated two-photon subtraction from a continuous wave squeezed vacuum," *preprint arXiv:0804.0464* (2008).
- [155] J. K. Thompson, J. Simon, H. Loh, and V. Vuletic, "A high-brightness source of narrowband, identical-photon pairs," *Science* **313**, 74 (2006).
- [156] Torsten Trantum-Jensen, Lars Skovgaard Madsen, and Jonas Jørgensen, "The construction of a folded optical delay line," Bachelor thesis (June 2006).

- [157] 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," *Physical Review Letters* **93**, 093601 (2004).
- [158] K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto, and K. Nakamura, "Accuracy of quantum-state estimation utilizing Akaike's information criterion," *Physical Review A* **68**, 022314 (2003).
- [159] H. Vahlbruch, M. Mehmet, S. Chelkowski, B. Hage, A. Franzen, N. Lastzka, S. Goßler, K. Danzmann, and R. Schnabel, "Observation of squeezed light with 10-dB quantum-noise reduction," *Physical Review Letters* **100**, 033602 (2008).
- [160] P. van Loock, S. L. Braunstein, and H. J. Kimble, "Broadband teleportation," *Physical Review A* **62**, 022309 (2000).
- [161] Y. Vardi and D. Lee, "From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems," *Journal of the Royal Statistical Society. Series B (Methodological)* **55**, 569 (1993).
- [162] F. Villa, A. Chiummo, E. Giacobino, and A. Bramati, "High-efficiency blue-light generation with a ring cavity with periodically poled KTP," *Journal of the Optical Society of America B* **24**, 576 (2007).
- [163] K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," *Physical Review A* **40**, 2847 (1989).
- [164] K. Wakui, H. Takahashi, A. Furusawa, and M. Sasaki, "Photon subtracted squeezed states generated with periodically poled KTiOPO₄," *Optics Express* **15**, 3568 (2007).
- [165] Eric W. Weisstein, "Double factorial," MathWorld—A Wolfram Web Resource (2008), URL <http://mathworld.wolfram.com/DoubleFactorial.html>.
- [166] J. Wenger, R. Tualle-Brouiri, and P. Grangier, "Non-Gaussian statistics from individual pulses of squeezed light," *Physical Review Letters* **92**, 153601 (2004).
- [167] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physical Review* **40**, 749 (1932).
- [168] Patrick Windpassinger, private communication (2007).
- [169] L. A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, "Generation of squeezed states by parametric down conversion," *Physical Review Letters* **57**, 2520 (1986).
- [170] H. Yonezawa, T. Aoki, and A. Furusawa, "Demonstration of a quantum teleportation network for continuous variables," *Nature* **431**, 430 (2004).

-
- [171] H. P. Yuen, "Two-photon coherent states of the radiation field," *Physical Review A* **13**, 2226 (1976).
- [172] B. Yurke and D. Stoler, "Generating quantum mechanical superpositions of macroscopically distinguishable states via amplitude dispersion," *Physical Review Letters* **57**, 13 (1986).
- [173] A. Zavatta, V. Parigi, and M. Bellini, "Experimental nonclassicality of single-photon-added thermal light states," *Physical Review A* **75**, 052106 (2007).
- [174] A. Zavatta, V. Parigi, and M. Bellini, "Cavity-enhanced generation of pulsed nonclassical multi-photon states," *preprint arXiv:0803.1712* (2008).
- [175] A. Zavatta, S. Viciani, and M. Bellini, "Quantum-to-classical transition with single-photon-added coherent states of light," *Science* **306**, 660 (2004).
- [176] A. Zavatta, S. Viciani, and M. Bellini, "Tomographic reconstruction of the single-photon Fock state by high-frequency homodyne detection," *Physical Review A* **70**, 053821 (2004).
- [177] A. Zavatta, S. Viciani, and M. Bellini, "Single-photon excitation of a coherent state: Catching the elementary step of stimulated light emission," *Physical Review A* **72**, 023820 (2005).
- [178] T. C. Zhang, K. W. Goh, C. W. Chou, P. Lodahl, and H. J. Kimble, "Quantum teleportation of light beams," *Physical Review A* **67**, 033802 (2003).