Quantum Optical Measurements for Practical Estimation and Information Thermodynamics

Mihai D. Vidrighin

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Physics

September 2016

Supervised by
Ian A. Walmsley and Myungshik Kim

Department of Physics

Imperial College London

Declaration of originality

I hereby declare that this thesis and the work presented in it are my own. This work was done wholly while in candidature for a research degree at this University.

I have done my best to clearly state which results have, to my knowledge, not been derived before and to appropriately reference the sources of all other results. The experimental data presented here has been obtained by myself, with support from research colleagues.

Oxford, September 2016

Mihai-Dorian Vidrighin

Copyright declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Quantum Optical Measurements for Practical Estimation and Information Thermodynamics

Mihai D. Vidrighin

Submitted for the degree of Doctor of Philosophy
September 2016

Abstract

The implementation of optical quantum technologies requires precise and complete characterisation tools for multi-mode nonclassical states of light. In this thesis, we propose and we implement experimentally three applications of optical quantum measurements, which combine aspects of light that are described by both discrete and continuous variables.

In the first part of this thesis, we present a new scheme for interferometric phase estimation, which uses double homodyne measurements as an alternative to photon counting. We show that, without requiring a phase reference, this scheme can achieve the optimum phase estimation precision for all path-symmetric probe states with a defined number of photons. Furthermore, the estimation procedure is robust against state preparation imperfections, and naturally yields the same precision for all values of the estimated phase. We implement the proposed scheme experimentally, using a polarisation interferometer probed by a single-photon signal, and temporally multiplexing a single homodyne detector. We repeat the experiment with classical probe states. For both cases, we demonstrate the accuracy and precision of the proposed method, our implementation deviating by 5% from the fundamental precision bound. We demonstrate that, in our scheme, single-photon probe states can provide better precision than weak coherent ones. These results indicate that hybrid quantum resources, which combine tools from the discrete and continuous-variable paradigms, can play a significant role in practical sensing scenarios.

In the second part of this thesis, we present an application in the field of information thermodynamics, reporting an experimental realisation, in a photonic setup, of Maxwell's demon. We show that a measurement at the few-photons level, followed by a feed-forward operation, allows the extraction of work from intense thermal light into an electric circuit. The interpretation of the experiment leads us to the derivation of an equality relating work extraction to information acquired by measurement. We prove a bound using this relation, and show that it is in agreement with the experimental results. Our work puts forward photonic systems as a platform for experiments in the field of information thermodynamics.

In the third part of this thesis, we present a new method for the characterisation of broadband parametric down-conversion, based on stimulated frequency generation. In particular, we analyse the information contained in the frequency generation produced in the same mode as a seeding beam, for a type-II down-conversion process. We derive a model for this signal and argue that its detection can provide a useful characterisation tool in the high-gain regime, allowing for a self-referenced estimation of squeezing gain. We propose a method for measuring this signal and present an experimental realisation, based on a wave-guided source, at telecommunication wavelengths. Our experimental results demonstrate that the proposed measurement is an effective experimental handle for discerning the intricate structure of broadband parametric down-conversion.

To my teachers, for their trust and support.

Acknowledgements

I dedicate this thesis to all the teachers who supported and guided me on my way here, in particular, Ms Mariana Brad, Mr Ion Nicula, and the professors of the Physics Department at the Babes-Bolyai University in Cluj-Napoca.

Danny Segal, Myungshik Kim and Terry Rudolph, the directors of the Controlled Quantum Dynamics Doctoral Training Centre, made this PhD possible. I will always be grateful to them for welcoming me to the DTC family. It is hard to imagine a richer learning experience than that of my first year at Imperial College. I hope that everyone involved in making the DTC what it is knows that their effort is appreciated. Within this group, I would also like to thank David Jennings and Marco Genoni, who very generously shared their knowledge with me throughout the PhD programme. It has been an honour to work alongside my colleagues from the wündercohort. Each of their personalities is a great source of inspiration. André and Mark, especially have helped me along, with their friendship and scientific discussions.

I am sincerely grateful to my supervisors, Ian Walmsley, from Oxford, and Myungshik Kim, from Imperial. Throughout my PhD candidacy, they have offered me their trust and support in pursuing my reasearch ideas. It has been a privilege to work with a supervising team of such caliber.

Vlatko Vedral and Oscar Dahlsten have welcomed me as an unofficial visitor in their research group. I hope that we can repeat the experience of working together. Brian Smith and Michal Karpinski also deserve my sincere gratitude for offering me the opportunity to leave a mark on the impressive setup designed by Merlin Cooper during his PhD.

As a member of the Ultrafast Quantum Optics group for the past 4 years, I have witnessed the turnover of a few generations of researchers. Marco Barbieri got me standing on my own two feet during my first year in the lab. Steven Kolthammer has always been supportive of my work, offering valuable scientific and strategic advice. Andreas Eckstein has been relentless in automating the lab instruments, and one of the chapters of this thesis would not have been possible without his help. I'd like to take this opportunity to thank the colleagues who have read my thesis, providing valuable input and encouragement: Helen Chrzanowski, Gil Triginer, Steven Kolthammer, and Magda Szczykulska.

Of course, the group ticks because of its excellent students. In particular, Merritt Moore and Magda Szczykulska have been very productive lab mates, contributing to the experimental efforts on which two chapters of this thesis are based. Gil Triginer has arrived when he was needed the most, and has proven to be a great colleague and friend. I am grateful to him for working alongside me on experiments on which a part of this thesis is based. I would also like to thank Joelle Boutari, with whom I have spent countless hours building outreach experiments.

I owe my sanity to Cora and to our friends. These people are making my life about as easy as I can imagine life being. A special acknowledgment goes to Cosmina, who has really helped a friend in need, giving the text of this thesis its most thorough reading.

The downside of doing this PhD has been living away from my family. Luckily, our support for each other comes so naturally, that I can easily feel it even when we are apart. I'd like to take this opportunity to thank my grandparents for nurturing my independent thinking, my elder brother for influencing basically every big decision I have made, and my parents for their unconditional support, which makes even the most daunting task seem achievable.

Author's publications

1. M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley. "**Photonic Maxwell's demon**". *Physical Review Letters*, **116**:050401 (2016).

   • *Work presented in Chapter 3 of this Thesis.*

2. M. Altorio, M. G. Genoni, M. D. Vidrighin, F. Somma, and M. Barbieri. "**Weak measurements and the joint estimation of phase and phase diffusion**". *Physical Review A*, **92**(3):032114 (2015).

3. G. Donati, T. J. Bartley, X. M. Jin, M. D. Vidrighin, A. Datta, M. Barbieri, and I. A. Walmsley. "**Observing optical coherence across Fock layers with weak-field homodyne detectors**". *Nature communications*, **5**:5584 (2014).

4. M. D. Vidrighin, G. Donati, M. G. Genoni, X. M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley. "**Joint estimation of phase and phase diffusion for quantum metrology**". *Nature communications*, **5**:3532 (2014).

   • *Discussed in Chapter 2 of this Thesis.*

5. M. D. Vidrighin, T. J. Bartley, G. Donati, X. M. Jin, M. Barbieri, W. S. Kolthammer, A. Datta, and I. A. Walmsley. "**Requirements for two-source entanglement concentration**". *Quantum Measurements and Quantum Metrology*, **1**:5 (2013).

In preparation

6. M. D. Vidrighin, M. Szczykulska, M. Karpinski, M. S. Kim, I. A. Walmsley, and B. J. Smith. "**Global estimation of phase with homodyne measurements and single photons**".

   • *Completed work, presented in Chapter 2 of this Thesis.*

7. M. D. Vidrighin, G. Triginer, M. Moore, A. Eckstein, W. S. Kolthammer, M. S. Kim and I. A. Walmsley. "**Precise measurement of two-mode squeezing by seeding a parametric down-conversion source**".

   • *Work in progress, presented in Chapter 4 of this Thesis.*

Contents

Declaration of originality **2**

Copyright declaration **3**

Acknowledgements **6**

Author's publications **8**

List of Figures **12**

List of Tables **14**

List of abbreviations **15**

List of symbols **17**

1 Introduction **18**
1.1 Thesis outline . . . 19
1.2 Basic concepts . . . 21
1.2.1 Quantum optics . . . 21
1.2.2 Nonlinear optics . . . 28
1.2.3 Information theory . . . 30
1.2.4 Error analysis . . . 34

2 Global estimation of phase with homodyne measurements and single photons **35**
2.1 Limits on precision in optical phase estimation . . . 36
2.1.1 The classical limit . . . 37
2.1.2 Quantum-enhanced precision . . . 40
2.1.3 Theoretical framework and fundamental limits . . . 43
2.1.4 The effect of decoherence . . . 51
2.1.5 Global estimation of phase . . . 54
2.1.6 Multi-parameter estimation . . . 56
2.2 A double homodyne measurement for quantum-enhanced phase estimation without prior information . . . 58

2.2.1 Optical homodyne detection . . . 58
2.2.2 Optimal global estimation of phase for a class of quantum probe states 59
2.2.3 Realistic precision limits . . . 63
2.3 Experimental implementation . . . 64
2.3.1 Pulsed optical homodyne detection . . . 64
2.3.2 A heralded source of single photons in pure quantum states . . . 66
2.3.3 Experimental setup . . . 69
2.3.4 Data analysis . . . 71
2.4 Experimental results: demonstrating phase estimation accuracy and precision 75
2.4.1 Single-photon probe state . . . 76
2.4.2 Classical probe state . . . 78
2.5 Summary and outlook . . . 81

3 Realisation of a photonic Maxwell's demon 83
3.1 Maxwell's demon – linking thermodynamics and information theory . . . 84
3.1.1 A brief history . . . 84
3.1.2 Maxwell's demon and quantum information . . . 88
3.1.3 Physical realisations of Maxwell's Demon . . . 89
3.2 Recent progress in thermodynamics . . . 90
3.2.1 Quantum thermal machines . . . 91
3.2.2 Thermodynamics as a resource theory . . . 94
3.2.3 Non-equilibrium work relations . . . 96
3.3 Light and thermal equilibrium . . . 98
3.3.1 Blackbody radiation . . . 99
3.3.2 Thermal light . . . 99
3.4 A photonic Maxwell's demon . . . 102
3.5 Experimental implementation . . . 105
3.5.1 Experimental methods . . . 105
3.5.2 The thermal light source . . . 106
3.5.3 Non-equilibrium preparation by measurement . . . 107
3.6 Information thermodynamics interpretation . . . 110
3.6.1 A non-equilibrium work-information equality . . . 110
3.6.2 Bound relating work extraction to information . . . 113
3.6.3 Experimental verification of the bound . . . 115
3.7 Summary and outlook . . . 116

4 Precise measurement of two-mode squeezing gain by seeding a parametric down-conversion source 118
4.1 Broadband parametric down-conversion in optical wave-guides . . . 119
4.1.1 Development of broadband parametric down-conversion sources . . . 119
4.1.2 General mode transformations for type-II down-conversion in optical wave-guides . . . 120
4.1.3 Perturbation analysis . . . 127
4.1.4 Schmidt mode decomposition . . . 132
4.2 Characterisation of a parametric down-conversion source using a seeding beam 133
4.2.1 Seeded measurements . . . 134
4.2.2 Application to self-referenced estimation of squeezing parameters . . . 136

4.3 Preliminary experimental results 138
4.3.1 Experimental details 138
4.3.2 Seeded measurements 139
4.4 Summary and outlook 141

5 Conclusion **144**

Bibliography **146**

A Fisher information for double homodyne detection **162**

B Predictions for the effect of the demon's measurements **164**
B.0.1 Effect of a lossy measurement 164
B.0.2 Energy imbalance between the two light modes produced by feed-forward 165
B.0.3 Non-disturbance condition applied to the APD measurement 165
B.0.4 Computation of the expected mutual information 166
B.0.5 Effect of multi-mode thermal statistics 167

List of Figures

2.1 Mach-Zehnder interferometer . . . 37
2.2 SU(2) picture of a prototypical interferometer . . . 39
2.3 SU(2) picture of squeezed vacuum interferometry . . . 42
2.4 Typical phase estimation precision in the presence of decoherence . . . 52
2.5 Double homodyne phase estimation scheme . . . 60
2.6 Measurement output distributions with Fock state inputs . . . 62
2.7 Fisher information with homodyne measurements in the presence of loss . . . 63
2.8 Homodyne detector electronic circuit . . . 65
2.9 KDP single-photon source . . . 67
2.10 Spectrum of the photon pairs produced by the KDP source . . . 68
2.11 Temporally multiplexed detection of two polarisation modes . . . 70
2.12 Homodyne signal . . . 72
2.13 Quadrature extraction . . . 73
2.14 Electronic noise in the homodyne detector . . . 75
2.15 Measured quadrature distribution for vacuum and a single photon state . . . 76
2.16 Phase modulation for sampling phase averaged coherent states . . . 76
2.17 Quadrature distribution for a single-photon probe state . . . 77
2.18 Phase estimation accuracy with single-photons . . . 78
2.19 Phase estimation precision with single photons . . . 79
2.20 Measured quadrature distribution with coherent light . . . 79
2.21 Phase estimation accuracy and precision with coherent light . . . 80
2.22 Comparison of single photon and coherent light estimation precision . . . 80
2.23 Phase tracking with single photons . . . 81

3.1 Szilard's engine . . . 86
3.2 Continuous quantum heat engine . . . 93
3.3 Laser speckle . . . 101
3.4 Photonic Maxwell's demon setup . . . 103
3.5 Intensity distribution of thermal light source . . . 107
3.6 Imbalance of conditional states . . . 108
3.7 Imbalance produced by measurement and feedback . . . 109
3.8 Extracted work bounded by information . . . 116

4.1 Transfer function for a typical broadband PDC process, coupling signal to idler 130

4.2 Transfer function for a typical broadband PDC process, coupling signal to signal . . . 131
4.3 Experimental setup for seeded PDC measurements . . . 139
4.4 Example of a PDC spectrum measured in the seeded mode . . . 140
4.5 Measured transfer functions . . . 141
4.6 Measured phase-matching profile . . . 142
4.7 Principal eigenvectors of measured transfer functions . . . 142

List of Tables

2.1 Precision limits in the presence of decoherence 53
2.2 Homodyne detector parameters 66
2.3 Single-photon source parameters 69
2.4 Homodyne detection efficiency for single-photons 71

List of abbreviations

APD	avalanche photodiode
BBO	barium borate
BS	beam splitters
BW	bandwidth
CMR	common-mode rejection
CR	Cramér-Rao
DFG	difference frequency generation
EPR	Einstein-Podolsky-Rosen
FBG	fiber-Bragg gratings
FI	Fisher information
GEP	global estimation of phase
GWD	gravitational-wave detectors
HB	Holland-Burnett
HD	homodyne detection
HWP	half-wave plate
KDP	potassium dihydrogen phosphate
LO	local oscillator
MBC	mean Bayesian cost
MSE	mean square error
OSA	optical spectrum analyser
PBS	polarising beam splitter
PD	probability distribution
PDC	parametric down-conversion

POVM quantum measurement

QFI quantum Fisher information

QPM quasi-phase-matching

SHG second harmonic generation

SLD symmetric logarithmic derivative

SNL shot noise limit

SNR signal-to-noise ratio

SPDC spontaneous parametric down-conversion

SQL standard quantum limit

TMSV two mode squeezed vacuum

List of symbols

$[\bullet, \bullet]$	commutator
$\{\ldots\}$	vector elements
$\dagger$	Hermite conjugation
$\hat{\bullet}$	linear operator
β	inverse temperature
c	speed of light
$\delta_{i,j}$	Kronecker symbol
$\delta(\bullet)$	Dirac delta
$\hbar$	Planck's constant
i	imaginary unit
e	base of the natural logarithm
$\hat{I}$	identity operator
ε_0	vacuum permittivity
∇	nabla differential operator
$\boldsymbol{r}$	position vector
t	time
$x.y(z)$	where x, y, z are digits, denotes x.y with standard deviation $0.z$

Chapter 1

Introduction

CURRENT research makes a concerted theoretical and experimental effort to refine our control over quantum dynamics. Due, in particular, to the development of laser systems, light serves as the main tool for the measurement and control of processes at the quantum level.

In fact, the nature of light itself has been closely linked to progress in the theory of quantum mechanics since its very beginning. Albert Einstein's arguments concerning the particle nature of light, presented in 1905 [1], marked the inception of wave-particle duality, a concept that lies at the heart of quantum mechanics. Since the 1970s, experimental quantum optics has been a key platform for testing the theory's fundamental concepts, with simple table-top experiments resulting in significant paradigm shifts. A good example is that of experimental tests for Bell nonlocality [2,3], which, with current technology, can be readily implemented in many quantum optics laboratories (albeit leaving open some well-known loopholes). The first realisation of such an experiment, reported by John F. Clauser in 1969 [4], was a landmark achievement. Clauser presented his work, titled "Experimental limitations to the validity of semiclassical radiation theories", at the Third Rochester Conference on Coherence and Quantum Optics. In the proceedings of the same conference, Edward T. Jaynes concluded his article, "Survey of the present status of neoclassical radiation theory" [5], with the following:

> "We have not commented on the beautiful experiment reported here by Clauser which opens up an entirely new area of fundamental importance to the issues

> facing us. [...] What it seems to boil down to, is this: a perfectly harmless looking experimental fact [...] can, at a single stroke, throw out a whole infinite class of alternative theories of electrodynamics [...] If it survives that scruitiny, and if the experimental result is confirmed by others, then this will surely go down as one of the most incredible intellectual achievements in the history of science, and my own work will lie in ruins. I wish John von Neumann were here to see it."

The potential impact of experimental quantum optics is as high now as it was then. Researchers are aiming to demonstrate controlled quantum evolution of nonclassical light at a scale at which the simulation of experimental results using classical computers becomes intractable. Although a daunting task, the scaling-up of current experiments in terms of the number of photons and the number of interfering modes is being actively pursued.

In addition, current research is also making an effort towards developing applications that use the laws of quantum mechanics in new ways (for instance, by exploiting quantum entanglement), in fields such as communications and metrology. In order for quantum technologies to prove successful, it is likely that a hierarchical approach to their design will be needed, as with most modern technologies. In the context of quantum optics, this requires the ability to abstract away the continuous modes of the electro-magnetic field and manipulate the quantum state of an effective system with reduced dimensionality. In order to achieve this, researchers need tools for the precise and complete characterisation of multi-mode nonclassical light.

In this thesis, we focus on the design of quantum optical measurements for a number of applications. Our work, consisting of three main projects, takes advantage of the versatility of photonic systems and highlights the potential of combining aspects of light described by both discrete and continuous variables.

1.1 Thesis outline

In Section 1.2, we define the basic concepts used throughout this work. In order to give the reader the precise context for each project presented in the thesis, we defer a detailed review of relevant topics to the beginning of the three main chapters.

In Chapter 2, we present a new measurement scheme for precise interferometric phase estimation, proposing double homodyne measurement as an alternative to photon counting. We show that, in an arrangement distinct from heterodyne detection, and which does not require an external phase reference, homodyne detection can achieve the optimum phase estimation precision for all path-symmetric interferometric probe states with definite number of photons. Moreover, the phase estimation procedure is robust against state preparation imperfections, and it performs global estimation, which does not require any prior information on the value of the estimated parameter. We begin the chapter with a review of quantum precision limits for phase estimation, with an emphasis on the issues that our scheme addresses. We then analyse the performance of the double homodyne measurement using the tools of quantum parameter estimation theory. Finally, we present the results of an experimental implementation, using a single-photon signal to probe the phase of a polarisation interferometer. We demonstrate that, for our particular scheme, the single-photon signal can provide an advantage in terms of precision compared to classical light.

In Chapter 3, we present an application of experimental quantum optics techniques in the field of information thermodynamics, translating a well-known thought experiment, that of Maxwell's demon, into a photonic setup. In particular, we demonstrate a measurement at the few-photons level that conditionally prepares an out-of-thermal-equilibrium state, starting with intense thermal light. The chapter begins with a review of recent progress in the fields of quantum and information thermodynamics, highlighting the concepts that help the interpretation of the experiment and which are, in turn, illustrated by our implementation. We then present our experimental results. Finally, we derive a new type of non-equilibrium equality, which allows us to interpret the experimental results in the context of information thermodynamics.

In Chapter 4, we present a new experimental tool for the characterisation of parametric down-conversion sources in the high-gain regime. Considering seeded down-conversion, we propose the measurement of frequency generation in the seeded mode, which, together with a measurement of the signal stimulated in the conjugate mode, allows the precise estimation of the gain that characterises the process. We begin the chapter with a general treatment of broadband parametric down-conversion, including a second order perturbation approach, which allows us to model the field generated in the seeded mode. We then present

an experimental implementation of the proposed measurements. Although we are able to report a full set of experimental results, this work is currently in progress. In this thesis, we include a discussion of the experimental data and an overview of future developments.

We present concluding remarks in Chapter 5.

1.2 Basic concepts

This section defines some of the key theoretical tools used throughout the thesis. We assume that the reader is familiar with the general formalism of quantum mechanics, including mixed quantum states, generalised quantum evolution and positive-operator valued measurements. Here, we begin with a presentation of quantum optics, which is relevant to all subsequent chapters. We then provide a brief introduction to nonlinear optics, focusing on parametric down-conversion (PDC), which is relevant for Chapter 2 and central to Chapter 4. We then introduce some concepts from information theory, which will prove useful in Chapter 3. Finally, we discuss the methods employed for the analysis of experimental data.

1.2.1 Quantum optics

Quantum optics is concerned with the semi-classical and quantum mechanical descriptions of light and its interaction with matter, bringing together the wave and particle pictures of the electromagnetic field. The standard quantisation procedure[6] starts with writing the free electric and magnetic fields in terms of the transverse vector potential, using the Coulomb gauge[7],

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{r},t) &= -\frac{\partial}{\partial t}\boldsymbol{A}(\boldsymbol{r},t); \\ \boldsymbol{B}(\boldsymbol{r},t) &= \boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r},t). \end{aligned} \tag{1.1}$$

Here, $\boldsymbol{r} = \{x, y, z\}$ is the position vector, and $\boldsymbol{\nabla} = \{\partial/\partial x, \partial/\partial y, \partial/\partial z\}$. Maxwell's equations describe the evolution of the electric and magnetic fields. Separating the space and time variables, we can write the vector potential as a superposition of spatial distributions, called modes, with complex time-dependent amplitudes. The spatial modes are solutions to the Helmholtz equation, depending on the set boundary conditions. For the electromagnetic field in a cubic box of volume $V = L^3$, with periodic boundary conditions, the mode

decomposition of the vector potential reads

$$\boldsymbol{A}(\boldsymbol{r},t) = \frac{1}{\sqrt{\varepsilon_0 V}} \sum_{\boldsymbol{k},s} \mathcal{A}_{\boldsymbol{k},s}(t)\, \boldsymbol{e}_{\boldsymbol{k},s}\, e^{\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{1.2}$$

with ε_0 the vacuum permittivity. The components of the wave-vector $\boldsymbol{k}$, which labels the electromagnetic modes, can take the values

$$\begin{aligned} k_x &= 2\pi n_x/L\ ,\ n_x = 0, \pm 1, \pm 2, \dots\ ; \\ k_y &= 2\pi n_y/L\ ,\ n_y = 0, \pm 1, \pm 2, \dots\ ; \\ k_z &= 2\pi n_z/L\ ,\ n_z = 0, \pm 1, \pm 2, \dots\ . \end{aligned} \tag{1.3}$$

We use this in Chapter 3, to calculate the density of states of the electromagnetic field. The unit vectors $\boldsymbol{e}_{\boldsymbol{k},s}$, with $s \in \{1, 2\}$, indicate the direction of the transverse vector potential, called polarisation, and they obey the conditions

$$\begin{aligned} \boldsymbol{k} \cdot \boldsymbol{e}_{\boldsymbol{k},s} &= 0; \\ \boldsymbol{e}^*_{\boldsymbol{k},s} \cdot \boldsymbol{e}_{\boldsymbol{k},s'} &= \delta_{s,s'}; \\ \boldsymbol{e}_{\boldsymbol{k},1} \times \boldsymbol{e}_{\boldsymbol{k},2} &= \boldsymbol{k}/k. \end{aligned} \tag{1.4}$$

For example, the two unit-vectors could indicate horizontal (H) and vertical (V) polarisations. Solving Maxwell's equations for the complex amplitudes introduced in Equation 1.2 yields

$$\mathcal{A}_{\boldsymbol{k},s} = c^+_{\boldsymbol{k},s} e^{-\mathrm{i}\,\omega t} + c^-_{\boldsymbol{k},s} e^{+\mathrm{i}\,\omega t}, \tag{1.5}$$

where we have introduced the mode frequency $\omega = c\,k$ (c is the speed of light). We associate the two terms of the solution to forward and backward propagating plane waves. As $\boldsymbol{A}(\boldsymbol{r},t)$ is real, we must have $c^-_{\boldsymbol{k},s} = \left(c^+_{-\boldsymbol{k},s}\right)^*$ and $\boldsymbol{e}_{\boldsymbol{k},s} = \boldsymbol{e}^*_{-\boldsymbol{k},s}$.

The Hamiltonian of the electromagnetic field is obtained by integrating the energy density over the volume V,

$$H = \frac{\varepsilon_0}{2} \int_V d^3\boldsymbol{r} \left[\boldsymbol{E}^2(\boldsymbol{r},t) + c^2 \boldsymbol{B}^2(\boldsymbol{r},t)\right] = \sum_{\boldsymbol{k},s} \omega^2 \left[\mathcal{A}^*_{\boldsymbol{k},s}(t)\mathcal{A}_{\boldsymbol{k},s}(t) + \mathcal{A}_{\boldsymbol{k},s}(t)\mathcal{A}^*_{\boldsymbol{k},s}(t)\right]. \tag{1.6}$$

According to the postulates of quantum mechanics, quantisation is performed by replacing

the complex amplitudes $\mathcal{A}_{\boldsymbol{k},s}(t)$ by the Hilbert space operators $\hat{a}_s(\boldsymbol{k},t)$, using the quantisation rules

$$\mathcal{A}_{\boldsymbol{k},s}(t) \to \sqrt{\frac{\hbar}{2\omega}}\hat{a}_s(\boldsymbol{k},t), \tag{1.7}$$

where $\hbar$ is the reduced Plank constant. We then impose the commutation relations

$$\begin{aligned}\left[\hat{a}_s(\boldsymbol{k},t),\hat{a}^\dagger_{s'}(\boldsymbol{k}',t)\right] &= \hat{a}_s(\boldsymbol{k},t)\,\hat{a}^\dagger_{s'}(\boldsymbol{k}',t) - \hat{a}_{s'}(\boldsymbol{k}',t)^\dagger\,\hat{a}_s(\boldsymbol{k},t) = \delta_{\boldsymbol{k},\boldsymbol{k}'}\delta_{s,s'};\\ \left[\hat{a}_s(\boldsymbol{k},t),\hat{a}_{s'}(\boldsymbol{k}',t)\right] &= \left[\hat{a}^\dagger_s(\boldsymbol{k},t),\hat{a}^\dagger_{s'}(\boldsymbol{k}',t)\right] = 0.\end{aligned} \tag{1.8}$$

We may allow the wave-vector components to take continuous values, in which case, the Kronecker symbol $\delta_{\boldsymbol{k},\boldsymbol{k}'}$ is replaced by the Dirac delta distribution, $\delta(\boldsymbol{k}-\boldsymbol{k}')$.

Transforming Equation 1.6 according to the rules given in Equation 1.7, and using the commutation relations, we find the Hamiltonian operator

$$\hat{H} = \sum_{\boldsymbol{k},s}\hbar\omega\left(\hat{a}_s(\boldsymbol{k},t)^\dagger\,\hat{a}_s(\boldsymbol{k},t) + \frac{1}{2}\right). \tag{1.9}$$

For continuous wave-vectors, the sum is replaced by an integral. Note that this representation of operators corresponds to the Heisenberg picture of quantum mechanics. In the Shrödinger representation, the time dependence is taken over by the quantum state. The free Hamiltonian generates the transformation

$$\hat{a}_s(\boldsymbol{k},t) = \hat{a}_s(\boldsymbol{k},0)\,e^{-\mathrm{i}\,\omega t}, \tag{1.10}$$

showing that, under free evolution, the field acquires a phase proportional with the propagation time.

It is useful to define modes that are superpositions of plane waves. For instance, all of the experiments presented in this thesis use mode-locked pulsed lasers. An isolated, polarised light pulse, propagating in the z direction ($k_x = k_y = 0$), is modeled by a broadband mode, to which we associate the quantised amplitude

$$\hat{A}(t) = \int dz\, f(k_z)\,\hat{a}_1(k_z,0)\,e^{-\mathrm{i}\,c\,k_z t}, \tag{1.11}$$

where $\int dk_z\,|f(k_z)|^2 = 1$. If the phase of $f(k_z)$ is constant, the pulse is called Fourier

transform limited, and it is as tightly localised in space as allowed by the spread in the wave-vector distribution. A phase that depends quadratically on the wave-vector describes a chirped pulse, which, in space, has high frequencies at one end and lower frequencies at the opposite end.

Let us now concentrate on a single mode, at a given moment in time, dropping the dependence on $\boldsymbol{k}$, s and t. The operator $\hat{n} = \hat{a}^\dagger \hat{a}$ is the photon number operator. Fock states, defined by $\hat{n}\,|n\rangle = n\,|n\rangle$, have well-defined energy, $E_n = \hbar\omega\,(n + 1/2)$, corresponding to precisely n excitations of the electromagnetic field, called photons. The energy of a photon is $\hbar\omega$, and $\hbar\omega/2$ is termed the vacuum energy. The state $|0\rangle$ is called the vacuum state.

The operators $\hat{a}^\dagger$ and $\hat{a}$, are called, respectively, the creation and annihilation operator, due to the way in which they act on Fock states:

$$\begin{aligned} \hat{a}^\dagger|n\rangle &= \sqrt{n+1}\,|n+1\rangle\,; \\ \hat{a}\,|n\rangle &= \sqrt{n}\,|n-1\rangle\,. \end{aligned} \tag{1.12}$$

Combining them, we can construct two Hermitian operators, $\hat{q} = \left(\hat{a}^\dagger + \hat{a}\right)/\sqrt{2}$ and $\hat{p} = \mathrm{i}\left(\hat{a}^\dagger - \hat{a}\right)/\sqrt{2}$, called field quadratures. In Chapter 2, we use the rotated quadrature, which we denote $\hat{X}_\theta$, according to a convention sometimes used in optical phase space, $\hat{X}_\theta \equiv \left[e^{\mathrm{i}\,\theta}\hat{a}_s^\dagger + e^{-\mathrm{i}\,\theta}\hat{a}_s\right]/\sqrt{2}$. The use of a capital letter in this case is simply meant to improve readability.

The quadratures $\hat{q}$ and $\hat{p}$ are physical observables orresponding to the sine and cosine components of the electromagnetic field amplitude. They are canonical conjugate operators, analogous to the position and momentum of the quantum harmonic oscillator. The measurement of a field quadrature is implemented using optical homodyne detection, which we introduce in Chapter 2. We can calculate the probability distribution corresponding to a quadrature measurement of the vacuum state, using the representation of the canonical operator $\hat{p}$ in the eigenbasis of $\hat{q}$, $\hat{p} = \int dq\,(-\mathrm{i}\,\partial/\partial q)\,|q\rangle\langle q|$:

$$\langle q|\,\hat{a}\,|0\rangle = 0 = \frac{1}{\sqrt{2}}\,\langle q|\,\hat{q} + \mathrm{i}\,\hat{p}\,|0\rangle = \frac{1}{\sqrt{2}}(q + \partial/\partial q)\,\langle q|0\rangle\,. \tag{1.13}$$

Solving this differential equation and normalising the solution, we obtain

$$|\langle q|0\rangle|^2 = \pi^{-\frac{1}{2}}\, e^{-x^2}. \tag{1.14}$$

The fact that this distribution has a non-zero width is a direct manifestation of vacuum fluctuations. In a similar manner, we can derive the quadrature distribution of a single-photon Fock state,

$$|\langle q|1\rangle|^2 = 2\,\pi^{-\frac{1}{2}}\, e^{-x^2} x^2. \tag{1.15}$$

We present the experimental results for a measurement of this distribution in Chapter 2. In general, the Fock basis representation of quadrature eigenstates can be computed by solving the Schrödinger equation for the stationary states of the quantum harmonic oscillator. We note that the eigenstates of the quadrature operators are not normalisable.

Using the commutation relations given in Equation 1.8, we can calculate the effect of some simple Hamiltonians. It is natural to use the Heisenberg picture of quantum mechanics, as it turns out that many relevant interactions result in linear transformations of the creation and annihilation operators. For convenience, we set $\hbar = 1$ throughout the rest of this thesis. Therefore, all Hamiltonians should be understood as scaled by $\hbar$.

Let us consider the simple Hamiltonian $\hat{H}_{\text{disp.}} = \mathrm{i}\,\xi\,\hat{a}^\dagger + c.c.$, where *c.c.* stands for the complex conjugate term (in this case, $-\mathrm{i}\,\xi^*\hat{a}$), which ensures that the operator is Hermitian. The transformation generated by this Hamiltonian is given by Heisenberg's equation (here, we reintroduce the time dependence of operators),

$$\begin{aligned} \frac{\partial}{\partial t}\hat{a}(t) &= \mathrm{i}\left[\hat{H}_{\text{disp.}}, \hat{a}(t)\right] = \xi; \\ \hat{a}(t) &= \hat{a}(0) + \xi\, t. \end{aligned} \tag{1.16}$$

This operation is called a displacement (anticipated in the notation). The unitary $\hat{D}(\alpha) = \exp(\alpha\,\hat{a}^\dagger - c.c.)$ is called the displacement operator, as it implements the displacement operation, $\hat{D}(\alpha)^\dagger \hat{a}\, \hat{D}(\alpha) = \hat{a} + \alpha$. Acting on the vacuum, this yields a coherent state, $|\alpha\rangle = \hat{D}(\alpha)\,|0\rangle$. For most applications, we can model the light generated by a laser using a coherent state [8], although, in principle, the phase of the laser emission is not well-defined [9]. It can be shown that the photon number distribution of a coherent state is Poissonian.

Optical fields can be superposed using semi-transparent optical elements, known as

beam-splitters. The transformation implemented by a beam-splitter for two perfectly overlapping input modes with annihilation operators $\hat{a}_1$ and $\hat{a}_2$ is modeled (up to free propagation of both modes) by

$$\begin{aligned} \hat{a}_1 &\rightarrow \sqrt{T}\,\hat{a}_1 + \sqrt{R}\,e^{\mathrm{i}\,\phi_1}\,\hat{a}_2; \\ \hat{a}_2 &\rightarrow \sqrt{R}\,e^{\mathrm{i}\,\phi_2}\,\hat{a}_1 - \sqrt{T}\,e^{\mathrm{i}\,(\phi_1+\phi_2)}\,\hat{a}_2, \end{aligned} \tag{1.17}$$

where T is the beam-splitter transmittance, and $R = 1 - T$.

In many instances described in this thesis, we implement the beam-splitter operation between two orthogonal polarisations of a spatial mode. We use wave-plates, which are thin slabs of birefringent material, to introduce a phase shift between two polarisation modes. By adjusting the orientation of the birefringent axiAs, half wave plates (HWP) can rotate the direction of oscillation of the electric field between the H and V directions. Quarter wave plates (QWP) can convert between linear and circular polarisations. A HWP and a QWP used together can implement any transformation described by Equations 1.17, where $\hat{a}_1$ and $\hat{a}_2$ correspond to two orthogonal polarisations of a spatial mode.

Another important operation is that generated by the squeezing Hamiltonian, $\hat{H}_{\mathrm{sq.}} = -\mathrm{i}\,\xi/2\,\hat{a}^\dagger\hat{a}^\dagger + c.c.$ The corresponding evolution of the mode operators is given by

$$\begin{aligned} \frac{\partial}{\partial t}\hat{a}(t) &= \mathrm{i}\left[\hat{H}_{\mathrm{sq.}}, \hat{a}(t)\right] = -\xi\,\hat{a}^\dagger(t); \\ \frac{\partial}{\partial t}\hat{a}^\dagger(t) &= \mathrm{i}\left[\hat{H}_{\mathrm{sq.}}, \hat{a}^\dagger(t)\right] = -\xi^*\,\hat{a}(t); \\ \implies \hat{a}(t) &= \cosh(r)\,\hat{a}(0) - e^{\mathrm{i}\,\theta}\sinh(r)\,\hat{a}^\dagger(0), \end{aligned} \tag{1.18}$$

where $\xi\,t = r\,e^{\mathrm{i}\,\theta}$, and the real value r is called the squeezing parameter. This is known as a squeezing transformation due to its effect on quadrature operators, $\hat{q}(t) = e^{-r}\hat{q}(0)$, $\hat{p}(t) = e^{+r}\hat{p}(0)$, assuming here that ξ is real, which means that the quadrature distributions corresponding to the input state become narrower for q and broader for p.

We employ, in Chapter 4 in particular, the two-mode squeezing operation, generated by

the Hamiltonian $\hat{H}_{2\text{sq.}} = -\mathrm{i}\,\xi\,\hat{a}_1^\dagger\hat{a}_2^\dagger + c.c.$ The corresponding mode transformations are

$$\begin{aligned}
\hat{a}_1(t) &= \cosh(r)\,\hat{a}_1(0) - e^{\mathrm{i}\,\theta}\sinh(r)\,\hat{a}_2(0)^\dagger;\\
\hat{a}_2^\dagger(t) &= \cosh(r)\,\hat{a}_2^\dagger(0) - e^{-\mathrm{i}\,\theta}\sinh(r)\,\hat{a}_1(0);\\
\hat{q}_1(t) + \hat{q}_2(t) &= e^{-r}\left[\hat{q}_1(0) + \hat{q}_2(0)\right];\\
\hat{p}_1(t) - \hat{p}_2(t) &= e^{-r}\left[\hat{p}_1(0) - \hat{p}_2(0)\right],
\end{aligned} \tag{1.19}$$

where $\xi\,t = r\,e^{\mathrm{i}\,\theta}$ and we assumed ξ to be real for the quadrature transformations. This shows that two-mode squeezing can generate correlations between the quadratures of two fields. For this reason, an interaction modeled by this operation can be used to generate light exhibiting experimentally accessible entanglement. In particular, the two-mode squeezed vacuum (TMSV) exhibits perfect photon number correlations between two modes [8],

$$\exp(-\mathrm{i}\,\hat{H}_{2\text{sq.}}t)\,|0,0\rangle = \cosh^{-1}(r)\sum_{n=0}^{\infty}\left[-e^{\mathrm{i}\,\theta}\tanh(r)\right]^n|n,n\rangle\,. \tag{1.20}$$

In Chapter 2, we use a result related to the transformation of quadrature eigenstates. It is well known that a TMSV can be created by combining two squeezed vacuum states on a balanced beam-splitter [10]. This equivalence can be extended to quadrature eigenstates of eigenvalue 0, which are analagous to infinitely squeezed vacuum states (although, as stated above, they are not normalisable). We take the balanced beam-splitter transformation

$$\begin{aligned}
\hat{a}_1 &\to \frac{1}{\sqrt{2}}\,\hat{a}_1 - \frac{1}{\sqrt{2}}\,\hat{a}_2;\\
\hat{a}_2 &\to \frac{1}{\sqrt{2}}\,\hat{a}_1 + \frac{1}{\sqrt{2}}\,\hat{a}_2,
\end{aligned} \tag{1.21}$$

implemented by the operator $\hat{B}^\dagger(50:50)$. Here, 50 : 50 stands for 50% transmittance. If we denote the quadrature eigenstates corresponding to the 0 eigenvalue by $|0\rangle_q$ and $|0\rangle_p$ for $\hat{q}$ and $\hat{p}$ respectively, then by applying $\hat{B}^\dagger(50:50)\,|0\rangle_q \otimes |0\rangle_p$, we obtain the Einstein-Podolsky-Rosen (EPR) state. This state has perfect number correlations, like the TMSV, but with equal weights for all photon numbers:

$$\hat{B}^\dagger(50:50)\,|0\rangle_q \otimes |0\rangle_p = \frac{1}{\sqrt{\pi}}\sum_{n=0}^{\infty}|n,n\rangle\,. \tag{1.22}$$

The name EPR is due to the fact that the state introduced above has perfect quadrature correlations, similar to those described in Equation 1.19. These are analogous to the position and momentum correlations discussed in the EPR argument, in 1935 [11].

All of the transformations described above, generated by Hamiltonians that are quadratic in the field amplitudes, maintain the Gaussian form of quadrature distributions, which we derived for vacuum in Equation 1.14. Together, and taking vacuum to be the input state, they form the tool-set of Gaussian quantum optics. For various tasks, such as universal quantum computation [12], and, as we will see in Chapter 4, work extraction in quantum thermal machines [13], non-Gaussian operations are necessary. These can be implemented, for instance, by conditional preparation techniques, using single-photon detection.

1.2.2 Nonlinear optics

The squeezing transformation, introduced in the previous subsection, can be implemented using parametric nonlinear optics [14]. In parametric processes, light interacts with matter without changing its state: there is no transfer of energy, momentum or angular momentum between light and the matter system. In effect, the material mediates the interaction between the modes of the electromagnetic field, without becoming correlated with them.

In conventional (linear) transparent media, the only effect of the light-matter interaction is the modification of the speed of wave propagation. The induced material polarisation $P(t) = \varepsilon_0 \chi^{(1)} E(t)$ describes the response of a linear medium to the electric field. The polarisation acts as a source term in Maxwell's equations. In nonlinear optics, the material polarisation is extended to a power series in $E(t)$,

$$P(t) = \varepsilon_0 \left(\chi^{(1)} E(t) + \chi^{(2)} E^2(t) + \chi^{(3)} E^3(t) + \dots \right), \tag{1.23}$$

where $\chi^{(n)}$ is the nth order nonlinear optical susceptibility. For clarity of the exposition, we describe here the simplified case of a scalar field. In general, $\chi^{(n)}$ is a rank-n tensor, coupling all vectorial components of the electric field.

The second order nonlinear susceptibility is non-zero only for crystals that do not display inversion symmetry. Second order nonlinearity mediates interactions between three fields. Let us consider two plane waves with frequencies ω_1 and ω_2, in a medium with non-zero

$\chi^{(2)}$. The electric field is given by

$$E(t) = E_1\, e^{-\mathrm{i}\,\omega_1 t} + E_2\, e^{-\mathrm{i}\,\omega_2 t} + c.c. \tag{1.24}$$

The induced nonlinear polarisation is then

$$\begin{aligned}\varepsilon_0\,\chi^{(2)}E^2(t) = \varepsilon_0\,\chi^{(2)}\,\big[&(2|E_1| + 2|E_2|) + \big(E_1^2\, e^{-\mathrm{i}\,2\omega_1} + E_2^2\, e^{-\mathrm{i}\,2\omega_2} + c.c.\big)\\ &+ \big(2E_1E_2\, e^{-\mathrm{i}\,(\omega_1+\omega_2)} + c.c.\big) + \big(2E_1E_2^*\, e^{-\mathrm{i}\,(\omega_1-\omega_2)} + c.c.\big)\big].\end{aligned} \tag{1.25}$$

The optical nonlinearity thus generates new frequencies, different from ω_1 and ω_2. The first term on the right-hand side in Equation 1.25 describes the generation of a 0 frequency (DC) field, called optical rectification. The second term is responsible for second-harmonic generation (SHG). This phenomenon was first observed in 1961 [15], just one year after the development of the ruby laser [16]. Optical rectification was observed in 1962 [17]. The third and fourth terms result in sum and difference frequency generation (DFG), respectively. In this thesis, we are mostly interested in the latter.

In the particle picture description of DFG, a photon with energy $\hbar\omega_1$ is converted into a pair of photons, with energies $\hbar\omega_2$ and $\hbar(\omega_1 - \omega_2)$ (assuming that $\omega_1 > \omega_2$). This obeys energy conservation, as expected for a parametric process. The generation of photon pairs is interesting for the study of quantum optics, as it can lead to non-classical statistics, i.e. the preparation of states that cannot be represented as a mixture of coherent states. In particular, if we assume that the field at frequency ω_1 is much stronger than at frequency ω_2, and that the optical nonlinearity has a negligible effect on the strong field, then we can treat it as an external source of Hamiltonian evolution for the two fields with frequencies ω_2 and $\omega_3 \equiv \omega_1 - \omega_2$. In the quantum optical formulation, the Hamiltonian corresponding to this evolution describes the creation of photon pairs in the two modes, and is equivalent to the two-mode squeezing Hamiltonian introduced in Section 1.2.1.

In order to observe efficient optical frequency conversion, the nonlinear polarisation, acting as a source at different points in a medium, must produce waves that interfere constructively. In the presence of dispersion, which is due to the linear polarisation of a transparent medium, the waves emitted at different points become out of phase over a given distance. Considering the spatial mode function of plane waves in transparent media, $e^{\mathrm{i}\,k(\omega)z}$ (assum-

ing propagation on the z axis), we find that the DFG signals produced at different positions become out of phase over a distance $\Delta z_{\text{deph.}} = \pi/\Delta k$, where $\Delta k = k(\omega_1) - k(\omega_2) - k(\omega_3)$ is called the phase mismatch. We note that, in the particle picture, phase-matching is equivalent to momentum conservation, as the momentum of a photon is $\hbar\, k(\omega)$. Efficient nonlinear phenomena are observed by phase-matching an interaction, making $\Delta k = 0$. This can be achieved by compensating for the linear dispersion using the birefringence of the nonlinear medium. The three interacting fields must then have different polarisations. We call type-I phase matching the arrangement in which the ω_2 and ω_3 fields have the same polarisation and type-II phase matching the arrangement where their polarisations are orthogonal.

In the context of optical frequency generation, the field with frequency ω_1 is called the pump field, and the process is referred to as PDC. A field at frequency ω_2 is called a seed. PDC also occurs spontaneously, or stimulated by vacuum fluctuations. The spontaneous process is used to generate individual photon pairs, which are the basis of many schemes in experimental quantum optics. For instance, using a single-photon detector to signal one of the two photons, we learn that its pair has been created. This is a common way of constructing a heralded single-photon source. We discuss spontaneous parametric down-conversion (SPDC) in more detail in Chapter 4.

If the pump is a broadband field, $E_1(t) = \int_0^\infty d\omega_1 E(\omega_1) e^{-\mathrm{i}\,\omega_1 t} + c.c.$, the Fourier component of the nonlinear polarisation at frequency ω_3 is given by

$$P^{(NL)}(\omega_3) = 2\varepsilon_0 \chi^{(2)} \int d\omega_1 E(\omega_1) E^*(\omega_1 - \omega_3) + c.c.. \tag{1.26}$$

Throughout this thesis, the integration limits for frequency integrals should be understood as 0 to ∞, although we may extend them to $-\infty$ when integrating over localised functions.

As $\chi^{(2)}$ is in general a tensor, an effective nonlinear coefficient can be introduced, which depends on the geometry of the interacting fields[14]. If we denote the unit polarisation vectors of the three interacting fields by $\boldsymbol{e}_i$, the second order effective nonlinear coefficient is defined as the tensor contraction $d_{\text{eff}} = 2\,\boldsymbol{e}_3 \cdot \boldsymbol{\chi}^{(2)} : \boldsymbol{e}_1 \boldsymbol{e}_2^*$.

1.2.3 Information theory

Information theory studies the quantification and the fundamental limits of the extraction, storage, processing and transmission of information. Claude E. Shannon (1916-2001) was a

pioneer of the theory, with a seminal paper published in 1948 [18].

There are many familiar examples of useful information stored in physical structures, such as DNA strands or solid state memory devices. A prototypical memory consists of a number of stable systems that each occupy one of two possible states. To communicate the state of such a memory, we can use a signal consisting of binary events, which are commonly labeled by 0 and 1. If these events are, for observers unfamiliar with the state of the memory, independent and randomly distributed, with equal probabilities for 0 and 1, then we call the amount of information communicated per event a "bit". In 1924, Harry Nyquist (1889-1976) found that the line speed at which "intelligence" can be transmitted via a telegraph line with m different voltage levels is $K \log m$ per time step, where K is a constant. If we replace K by $\log_2(e)$ bits, then this result is not surprising, as we know that integer numbers between 0 and $m-1$, which we can use to label the voltages, can be represented exactly using $\log_2(m)$ 0's and 1's.

The theory of probability is at the basis of information theory. [19]. The amount of information contained in the value of a random variable is a property of the probability distribution associated with it. Let us consider a string of independent realisations of the random variable X, which can take n values, each denoted x_i, to which we associate the probabilities $p(x_i)$. It can be shown that, for a string of length m, in the limit of $m \to \infty$, any message that can be encoded in the values of X can be compressed in a number of bits equal to m times [20]

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2(p(x_i)), \tag{1.27}$$

which is called Shannon entropy. For this thesis, we use a unit of information called "nat", which corresponds to replacing the base-2 logarithm in Equation 1.27 with a natural logarithm. Shannon entropy is the asymptotic measure of information, quantifying the information contained in a large ensemble of independent realisations of a random variable.

The amount of information shared (e.g. through previous communication) between two parties is called mutual information. Let us take two systems with states given by the random variables X and Y. Their mutual information is

$$I(X;Y) = H(X) + H(Y) - H(X,Y), \tag{1.28}$$

where $H(X,Y)$ is the Shannon entropy associated to the joint probability distribution of X and Y. Expanding this expression, using the natural logarithm and denoting the values taken by Y as y_j , we have

$$\begin{aligned} I(X;Y) &= -\sum_{i,j} p(x_i,y_j)\left[\log(p(x_i)) + \log(p(y_j)) - \log(p(x_i,y_j))\right] \\ &= \sum_{i,j} p(x_i,y_j)[\log(p(y_j|x_i)) - \log(p(y_j))]. \end{aligned} \tag{1.29}$$

Here, $p(y_j|x_i)$ denotes a conditional probability, corresponding to y_j, given the occurrence of x_i.

Due to the non-orthogonality of quantum states, the amount of information contained in the mixed quantum state $\hat{\rho}$ depends on the basis in which we choose to represent it. The maximum value is given by the von Neumann entropy [12,21],

$$S(\hat{\rho}) = -\mathrm{Tr}[\hat{\rho}\log(\hat{\rho})], \tag{1.30}$$

which gives the Shannon entropy corresponding to the diagonal representation of $\hat{\rho}$ i.e. the "letters" of a signal encoded in realisations of $\hat{\rho}$ are taken to be the eigenvectors of the state.

The function introduced in Equation 1.27 had already been used before Shannon's work, in the context of thermodynamics. It is said that the name "entropy", taken from thermodynamics, was suggested by John von Neumann (1903-1957) [22]. Entropy may be seen as a measure of the surprise created by learning the true state of a system described probabilistically. If the physical system occupies its ground state with high probability, the corresponding entropy is close to 0. Alternatively, if no state is more probable than any other (assuming a bounded system), the entropy is maximised. A system with high energy and low entropy (e.g. a laser field) is a resource, which can be used to perform mechanical work. As it becomes correlated with the systems on which work is performed, its entropy increases and it becomes less useful. For a fixed value of the internal energy of a system, entropy is maximised by Boltzman's distribution,

$$p(x_i) = e^{-\beta E(x_i)}\left[\sum_i e^{-\beta E(x_i)}\right]^{-1}, \tag{1.31}$$

where $E(x_i)$ is the energy of the state x_i. We can motivate the use of the Boltzman (or canonical) distribution in describing a system at thermal equilibrium by stating that our ignorance with respect to the actual configuration of the system must be maximal [23].

The free energy, introduced to quantify extractable work, is the difference between a system's internal energy and its entropy, scaled by temperature. We discuss in more detail the role of free energy in thermodynamics in Chapter 3. We take the view that thermodynamic and information entropy are equivalent, and specify in the text when we are referring to equilibrium quantities, which correspond to the thermal state of a system. Using the inverse temperature, $\beta = (k_B T)^{-1}$, where k_B is Boltzmann's constant, we define free energy as

$$F(X) = \sum_i p(x_i)E(x_i) - \beta^{-1}H(X), \tag{1.32}$$

for classical systems. For a quantum system with Hamiltonian operator $\hat{H}$, in the quantum state $\hat{\rho}$, we define

$$F(\hat{\rho}) = \mathrm{Tr}(\hat{H}\hat{\rho}) - \beta^{-1}S(\hat{\rho}). \tag{1.33}$$

In Chapter 3, we use the von Neumann relative entropy, which is a measure of distinguishability between two quantum states [12], defined as

$$S(\hat{\rho}|\hat{\sigma}) = -\mathrm{Tr}\left[\hat{\rho}\left(\log(\hat{\rho}) - \log(\hat{\sigma})\right)\right]. \tag{1.34}$$

For strings of finite length containing the realisations of a random variable, Shannon entropy is not the unique measure of information. In the single-shot regime (concerning a single realisation), only a partial order on probability distributions can be defined, by means of majorisation, which finds an operational interpretation in the theory of entanglement [12]. Generalised entropies are monotones of this partial order. We elaborate more on this in Chapter 3.

1.2.4 Error analysis

For a simple estimation procedure, where $f(x)$ is determined based on a measurement of x, with precision Δx, error propagation yields

$$(\Delta f(x))^2 \approx \left[\frac{\partial}{\partial x} f(x)\, \Delta x\right]^2 \tag{1.35}$$

(assuming that x is not too close to a stationary point of f).

Throughout this thesis, we report experimental values that are, in most cases, limited in precision by statistical errors i.e. errors that arise when we estimate a quantity based on a set of random variables (measurement outputs), using a finite set of realisations. We can define the covariance matrix of the set of measurement outputs $\boldsymbol{x} \equiv \{x_1, x_2, \dots x_n\}$ as $\sigma_{i,j} \equiv \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$, where $\langle \bullet \rangle$ denotes an expectation value. Then, error propagation yields

$$[\Delta f(\boldsymbol{x})]^2 \approx \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i} f(\boldsymbol{x}) \frac{\partial}{\partial x_j} f(\boldsymbol{x})\, \sigma_{i,j}. \tag{1.36}$$

In practice, as great computational power is readily available, it is feasible and often a robust option to estimate errors using re-sampling methods, instead of deriving analytical expressions. Re-sampling methods are also known as bootstrapping algorithms [24]. In order to describe how bootstrapping works, we consider the example of estimating the error that corresponds to the variance of a sampled normally-distributed variable. We can generate a set of m values, randomly sampled from the normal distribution (a Gaussian distribution of width 1). Bootstrapping works by re-sampling m values from the data set (allowing for repeated sampling of the same value), iterating this operation a large number of times (several thousands). For each re-sampling, we calculate the variance. We then compute the standard deviation of the resulting variances. We find that this method, applied to a sample size of $m = 200$, yields the estimate 0.10(1) for the error on the variance of the sampled variable, which agrees with the analytical solution, $\sqrt{\frac{2}{200-1}}$. The advantage of this method is that it can be straight-forwardly applied to arbitrary estimators and probability distributions (with some known caveats [24]). For this reason, we use it to determine error estimates in most instances of error analysis occurring in this thesis. A disadvantage of bootstrapping is that it sometimes deviates from the correct answer for small sample sizes ($\lesssim 100$). However, our data sets are always considerably larger.

Chapter 2

Global estimation of phase with homodyne measurements and single photons

Optical phase estimation is at the foundation of interferometry [25], which has a plethora of applications, ranging from basic length measurements to spectroscopy, astronomical observations and the celebrated detection of gravitational waves [26,27]. The fundamental limits on the precision of phase estimation are imposed by quantum fluctuations. In 1981, Carlton Caves proposed the use of non-classical light as a way to improve this precision [28], aiming to achieve the highest sensitivity in interferometers for gravitational-wave detection. The field of quantum-enhanced sensing has since expanded [29] and applications have been proposed to various quantum metrology problems, such as frequency estimation for atomic clocks [30–32], weak magnetic field sensing [33,34] or biological measurements [35].

In practical interferometry, attention has focused mostly on measurements of intensity and photon statistics. Typically, these schemes require adaptive phase control in order to achieve the highest sensitivity [36,37]. Here, we propose double-homodyne detection as an alternative to photon-counting, in an arrangement that is distinct from the well-known heterodyne technique. We show that, without requiring a phase reference, our measurement scheme achieves the same precision for all phase values, with no adaptive phase control. Moreover, using an intuitive estimation strategy that is robust to state-preparation imperfections, it yields the highest possible precision for all path-symmetric pure probe states with a fixed number of photons.

In the presence of decoherence, it becomes difficult to demonstrate quantum-enhanced sensitivity[38]. However, we find that while our proposed measurement scheme (with the already-mentioned advantages) is sub-optimal for classical light, it recovers the optimal precision for single-photon probe states. Therefore, it constitutes a scenario where non-classical resources can out-perform classical ones, when constrained to the same measurements. We demonstrate the precision and accuracy of our method experimentally, implementing a polarisation interferometer with a single-photon input and a double-homodyne measurement. Our results indicate that hybrid quantum resources, which combine tools from the discrete and continuous-variable paradigms (favouring descriptions in the photon number basis and in terms of quadrature distributions, respectively) can play an important role in achieving the highest sensitivity in practical sensing scenarios.

We begin this chapter by introducing the fundamental limits on precision in optical phase estimation, and the different frameworks in which these limits are derived. We then discuss global phase estimation and the role of prior information in interferometry, as well as some relevant aspects of multi-parameter estimation. In Section 2.2, we present the double-homodyne measurement scheme and discuss its advantages. We then introduce the experimental tools required to implement this measurement with a single-photon probe signal in Section 2.3. In the same section, we also describe a high repetition-rate pulsed optical homodyne detector[39] and a heralded source of single photons in pure quantum states[40], both of which have been previously developed in our laboratory. In Section 2.4, we present experimental results, showing the precision and accuracy achieved in our implementation. Finally, we summarise our work and discuss its implications in Section 2.5.

2.1 Limits on precision in optical phase estimation

An optical interferometer is designed to determine the phase accumulated by light waves, as they travel from a common source though different, overlapping routes. This constitutes a precise way of measuring small changes in the paths explored by the light waves, and it has numerous applications. Figure 2.1 shows a prototypical interferometer. In this figure, we set the phase accumulated in mode A as a reference and denote the additional phase acquired in mode B as $\phi = k\,\delta$. Here, k is the wave-vector of the light waves and δ is the optical path difference between modes B and A.

The spatial and temporal coherence properties of the light waves probing the interferometer determine the visibility of the interference observed at the output. In classical optics, there are no limitations on the precision with which ϕ can be determined using intensity measurements, as the signal-to-noise ratio (SNR) of these measurements can in principle be made arbitrarily high. A fundamental limit on precision exists, however, in the quantum mechanical description.

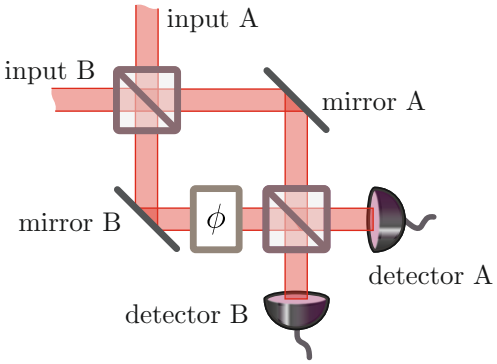


Figure 2.1: The prototype of a Mach-Zehnder interferometer with intensity detection. The gray squares depict optical beam-splitters.

2.1.1 The classical limit

The development of modern interferometric gravitational-wave detectors (GWD), which started in the 1970s, prompted the study of the precision limits introduced by optical quantum fluctuations. These detectors, arguably the most advanced interferometers realised to date[41], perform extremely sensitive measurements of spacetime distortions, by monitoring the optical path length between two test masses separated by distances on the order of kilometers. Their design is based on the Michelson interferometer, a folded version of the Mach-Zehnder layout depicted in Figure 2.1. Laser light enters through one of the two inputs and small phase changes are observed by monitoring the intensity at one of the outputs, set to an interference minimum. The test masses, two in each arm, double as mirrors, forming an optical cavity, so that light travels the path between them over several reflections. Power recycling is another standard feature. In the most recent runs, the displacement of test masses was measured with precision on the order of $10^{-19}\,\mathrm{m\,Hz^{-1/2}}$, which enables the detection of spacetime distortions with amplitudes as small as $10^{-22}\,\mathrm{Hz^{-1/2}}$ [41].

Quantum fluctuations are the main source of noise in interferometric GWD [42–45]. They

are manifested through (i) photon-counting noise at the monitored output, or optical shot noise, and (ii) phase noise caused by radiation pressure on the test masses. Both can be seen as consequences of vacuum fluctuations at the free input port of the interferometer. The first type of noise can be dealt with by increasing laser power, as the SNR of intensity measurements is inversely proportional to the square-root of the intensity. However, fluctuations of the intensity difference between the interferometer arms, which introduce phase noise through radiation pressure, are directly proportional to the square-root of the intensity. Adding the two effects, it turns out that the lowest noise and highest sensitivity is achieved for an optimum value of the power of circulating light.

The precision of path length measurements attained at the optimal power level is known as the standard quantum limit (SQL). This limit can be seen as arising from perturbations produced by the interaction between the test masses and the measurement light beams, like in Heisenberg's microscope [45]. We note, however, that test mass quantisation has no influence on GWD sensitivity [46,47]. Only the latest version of gravitational wave detectors [41] has reached the SQL over a limited signal frequency range. In many practical applications, laser power is much below the value necessary to achieve the SQL, and fluctuations due to radiation pressure are in effect negligible. Therefore, the main limitation of quantum origin to the sensitivity of interferometers is, in the majority of situations, that introduced by shot noise.

In order to discuss the effect of shot noise on phase estimation, we introduce a useful isomorphism between passive optical interferometers and spin systems, formulated by Yurke, McCall and Klauder in 1985 [48]. Let us denote by $\hat{a}$ and $\hat{b}$ the annihilation operators corresponding to interferometer arms A and B. The total number of photons is $\hat{n} = \hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}$. We further introduce the set of operators:

$$\hat{J}_x \equiv \frac{1}{2}\left(\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a}\right); \quad \hat{J}_y \equiv -\frac{\mathrm{i}}{2}\left(\hat{a}^\dagger\hat{b} - \hat{b}^\dagger\hat{a}\right); \quad \hat{J}_z \equiv \frac{1}{2}\left(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}\right). \tag{2.1}$$

Using the harmonic oscillator commutation relations, we can show that these operators satisfy the Lie algebra of the special unitary group SU(2):

$$\left[\hat{J}_x, \hat{J}_y\right] = \mathrm{i}\hat{J}_z; \quad \left[\hat{J}_y, \hat{J}_z\right] = \mathrm{i}\hat{J}_x; \quad \left[\hat{J}_z, \hat{J}_x\right] = \mathrm{i}\hat{J}_y. \tag{2.2}$$

Since we also have $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \frac{\hat{n}}{2}\left(\frac{\hat{n}}{2}+1\right)$, the set $\{\hat{J}_x, \hat{J}_y, \hat{J}_z\}$ can be seen as the angular momentum components of a system with spin $n/2$. Passive optical elements conserve photon number, and thus also the total angular momentum $\hat{J}$. Therefore, the evolution of operators through a passive interferometer (in the Heisenberg picture), corresponds to a 3-D rotation of the vector $\{\hat{J}_x, \hat{J}_y, \hat{J}_z\}$. We define the beam-splitter transformation as the unitary operation generated by $\hat{J}_x$ (in general, this generator is a linear combination of $\hat{J}_x$ and $\hat{J}_y$). Then, in the spin picture, a 50:50 beam-splitter corresponds to a $\frac{\pi}{2}$ rotation around the x axis. Relative phase-shifts between modes A and B (up to a global phase) are generated by the operator $\hat{J}_z$ and correspond to a rotation around the z axis. As depicted in Figure 2.2, the overall transformation implemented by an interferometer with relative phase ϕ amounts to a ϕ rotation around the y axis.

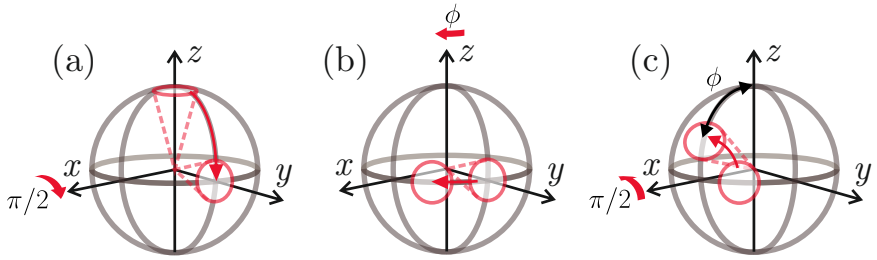


Figure 2.2: SU(2) picture of a prototypical interferometer. Using the isomorphism with a spin system, we depict the evolution of the quantum state of light through the interferometer (in the Scrödinger picture), representing the expectation value and uncertainty of angular momentum operators on a sphere of radius $\sqrt{\frac{n}{2}\left(\frac{n}{2}+1\right)}$. (a) Initially all excitations are in mode A, so the corresponding spin system occupies the eigenstate of $\hat{J}_z$ with the highest eigenvalue. The first beam-splitter corresponds to a rotation by $-\pi/2$ around the x axis. (b) A relative phase-shift ϕ corresponds to a rotation around the z axis. (c) The final beam-splitter corresponds to a rotation by $\pi/2$ around the x axis.

An intensity measurement at the output of the interferometer corresponds to determining the total number of photons n as well as the average intensity difference between outputs A and B, proportional to $\langle \hat{J}_z \rangle_{\text{out}}$. The phase ϕ can be estimated from the value of $\langle \hat{J}_z \rangle_{\text{out}}$. Using error propagation, we obtain the mean square error (MSE) of the estimation:

$$(\Delta\phi)^2 = \frac{(\Delta J_z)^2\Big|_{\text{out}}}{\left(\partial_\phi \langle \hat{J}_z \rangle_{\text{out}}\right)^2} = \frac{\langle \hat{J}_z^2 \rangle_{\text{out}} - \langle \hat{J}_z \rangle_{\text{out}}^2}{\left(\partial_\phi \langle \hat{J}_z \rangle_{\text{out}}\right)^2}. \tag{2.3}$$

As depicted in Figure 2.2, we have

$$\left(\hat{J}_z\right)_{\text{out}} = \cos(\phi)\left(\hat{J}_z\right)_{\text{in}} - \sin(\phi)\left(\hat{J}_x\right)_{\text{in}}. \tag{2.4}$$

For an interferometer where light enters at input A and the total number of photons is measured to be n, using Definition 2.1 we can calculate the following expectation values:

$$\begin{aligned} &\langle \hat{J}_z \rangle_{\text{in}} = n/2; \quad \langle \hat{J}_x \rangle_{\text{in}} = 0; \\ &\langle \hat{J}_z^2 \rangle_{\text{in}} = n^2/4; \quad \langle \hat{J}_z \hat{J}_x \rangle_{\text{in}} = 0; \quad \langle \hat{J}_x^2 \rangle_{\text{in}} = n/4. \end{aligned} \tag{2.5}$$

Plugging these into Equation 2.3 and using Equation 2.4, we get

$$(\Delta\phi)_0^2 = \frac{n/4\sin(\phi)^2}{n^2/4\sin(\phi)^2} = \frac{1}{n}. \tag{2.6}$$

If the input is a coherent state and we only consider $\langle \hat{J}_z \rangle_{\text{out}}$, without the total photon-number measurement, then performing the calculation above yields $(\Delta\phi)_\alpha^2 = 1/\left(\langle \hat{n} \rangle_\alpha \sin(\phi)^2\right)$ [38], indicating a dependence of the estimation precision on ϕ.

2.1.2 Quantum-enhanced precision

We have seen that when only one input port is used, the MSE characterising phase estimation is limited by $1/n$. In order to improve phase estimation precision beyond this value, which we will refer to as the classical, or shot noise limit (SNL), more general input states must be considered.

The particle picture of light is particularly useful in approaching this problem [29]. Let us formally consider that each input photon occupies one of the two modes of the interferometer, in a symmetrised quantum state. The multi-particle quantum state that describes coherent light in both input ports is separable in this picture. Moreover, passive optical transformations act independently on each particle. Therefore, in an interferometer probed by classical (coherent) light, photons effectively evolve independently. The phase estimator MSE in this case is limited by the shot noise, as discussed in Section 2.1.1. Any quantum state of light that offers an improvement on the classical limit must therefore contain correlations between photons i.e. particle entanglement [49]. A beam-splitter can convert particle entanglement into mode entanglement [50]. However, the two types of correlations are inde-

pendent and it is the former that constitutes a resource for quantum-enhanced sensing.

Particle entanglement can be produced by the squeezing interaction, which is discussed in several instances throughout this thesis. In 1981, Carlton Caves [28] proposed the injection of squeezed vacuum in the free input port of GWD interferometers in order to reduce the effect of quantum noise, taking into consideration the limited power available from laser sources. Let us denote a squeezed vacuum state $|r\rangle$ and choose the squeezing parameter r to be real, meaning that the squeezed quadrature is $\frac{1}{2}(\hat{a} + \hat{a}^\dagger)$. The expected number of photons in this state is $\langle r|\,\hat{n}\,|r\rangle = \sinh^2(r)$.

For inputs of coherent light and squeezed vacuum, $|\alpha\rangle_{\mathrm{A}} \otimes |r\rangle_{\mathrm{B}}$, we can compute the following expectation values for the operators introduced in the previous section:

$$\begin{aligned}
\langle\hat{n}\rangle = |\alpha|^2 + \sinh^2(r);\ \langle\hat{J}_z\rangle_{\mathrm{in}} = \frac{1}{2}\left[|\alpha|^2 - \sinh^2(r)\right];\ \langle\hat{J}_x\rangle_{\mathrm{in}} = 0;\ \langle\hat{J}_z\hat{J}_x\rangle_{\mathrm{in}} = 0; \\
(\Delta J_z)^2\Big|_{\mathrm{in}} = \langle\hat{J}_z^2\rangle_{\mathrm{in}} - \langle\hat{J}_z\rangle_{\mathrm{in}}^2 = \frac{1}{4}\left[|\alpha|^2 + \frac{1}{2}\sinh^2(2r)\right]; \\
(\Delta J_x)^2\Big|_{\mathrm{in}} = \frac{1}{4}\left[|\alpha|^2\cosh 2r - \mathrm{Re}(\alpha^2)\sinh(2r) + \sinh^2(r)\right]; \\
(\Delta J_y)^2\Big|_{\mathrm{in}} = \frac{1}{4}\left[|\alpha|^2\cosh 2r + \mathrm{Re}(\alpha^2)\sinh(2r) + \sinh^2(r)\right].
\end{aligned} \tag{2.7}$$

The squeezed noise effect is manifested through the interference term $\mathrm{Re}(\alpha^2)\sinh(2r)$, which lowers the variance of $\hat{J}_x$ while increasing that of $\hat{J}_y$. The effect on $\hat{J}_x$ is maximised by choosing $\mathrm{Re}(\alpha) = |\alpha|$. We illustrate the basic principle of squeezed vacuum quantum-enhanced phase estimation in Figure 2.3. Using the expressions above and Equation 2.4, we obtain the MSE

$$(\Delta\phi)^2_{\{\alpha,r\}} = \frac{(\Delta J_z)^2\Big|_{\mathrm{out}}}{\left(\partial_\phi\langle\hat{J}_z\rangle_{\mathrm{out}}\right)^2} = \frac{\cot^2(\phi)\left[|\alpha|^2 + \frac{1}{2}\sinh^2(2r)\right] + |\alpha|^2\,e^{-2r} + \sinh^2(r)}{\left[|\alpha|^2 - \sinh^2(r)\right]^2}. \tag{2.8}$$

Note that this expression is minimised by setting $\phi = \pi/2$. We can further optimise with respect to r, under the constraint of a fixed average number of photons. This optimisation can only be performed analytically in the limit $\langle\hat{n}\rangle \gg 1$, in which case the minimum MSE is

$$(\Delta\phi)^2_{\mathrm{coh.+sq.}} \approx \frac{1}{\langle\hat{n}\rangle^{3/2}}. \tag{2.9}$$

The optimal number of photons in the squeezed vacuum input is $\sinh(r)^2 = \langle\hat{n}\rangle^{\frac{1}{2}}\Big/2$, which

means that most of the energy is in the coherent light input.

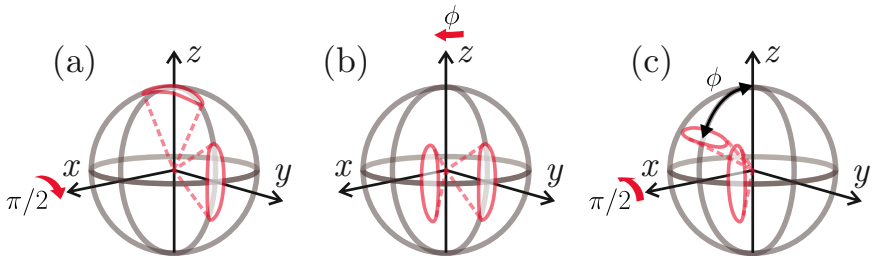


Figure 2.3: SU(2) picture of squeezed vacuum interferometry. The variance of the angular momentum operators is reduced in one direction at the expense of increased variance in the others. The operation of the interferometer ((a) input 50:50 beam-splitter, (b) phase-shift, (c) output 50:50 beam-splitter, overall a rotation around the y axis) is estimated with better precision than using uncorrelated particles.

Sub-shot-noise precision (by a factor of ≈ 2) using squeezed vacuum was demonstrated experimentally in 1987, detecting phase modulation signals over a limited bandwidth around 400 kHz [51] and 1.6 MHz [52]. In 2011, a squeezed vacuum source was added to the GEO 600 GWD observatory, yielding a factor of ≈ 2 noise reduction for signals at a few hundred Hz [53,54]. For gravitational wave detectors approaching the SQL, further theoretical analysis has lead to the proposal of frequency dependent squeezing, as well as frequency dependent homodyne detection [55–57].

Better-still scaling of the estimation precision with photon number can be obtained using a different input quantum state [48]:

$$|\psi\rangle_{\text{Fock}} = \frac{1}{\sqrt{2}}\left(|n/2\rangle_{\text{A}} \otimes |n/2\rangle_{\text{B}} + |n/2-1\rangle_{\text{A}} \otimes |n/2+1\rangle_{\text{B}}\right), \tag{2.10}$$

where $|n/2\rangle$ is the Fock state with $n/2$ photons. For simplicity, we assume that n is even. In this case, we have

$$\begin{gathered} \langle \hat{J}_z \rangle_{\text{in}} = \frac{1}{2};\ \langle \hat{J}_x \rangle_{\text{in}} = \frac{1}{2}\sqrt{\frac{n}{2}\left(\frac{n}{2}+1\right)}; \\ (\Delta J_z)^2\Big|_{\text{in}} = \frac{1}{2};\ (\Delta J_x)^2\Big|_{\text{in}} = \frac{n}{4}\left(\frac{n}{2}+1\right) - \frac{1}{4};\ \text{cov}(J_x, J_z)|_{\text{in}} = 0 \end{gathered} \tag{2.11}$$

which results in the phase estimation MSE

$$(\Delta\phi)^2_{\text{Fock}} = \frac{1}{\frac{n}{2}\left(\frac{n}{2}+1\right)}. \tag{2.12}$$

The quadratic scaling of the MSE with respect to photon number is known as Heisenberg scaling. Holland and Burnett (HB) have shown that if photon statistics at the output are considered, instead of just the average intensity, the simpler state $|\psi\rangle_{\mathrm{HB}} = |n/2\rangle_{\mathrm{A}} \otimes |n/2\rangle_{\mathrm{B}}$ yields the same precision scaling [58].

2.1.3 Theoretical framework and fundamental limits

We now introduce a general framework for the systematic analysis of estimation precision limits, which we use to derive some of our results in subsequent sections.

A general setting for parameter estimation consists of three main steps: probe preparation, parametrised evolution, and measurement. The analysis of measurement results can be considered a fourth step, and more generally, allowing for ancillary systems and controlled operations, we can describe adaptive schemes for preparation and measurement [59,60]. The task of detecting small deviations of parameter values can be reformulated as that of distinguishing neighboring states of a system, on a manifold defined by the parametrisation. If both preparation and measurement are fixed, then this amounts to statistically distinguishing the parametrised probability distributions (PDs) that define the output of the estimation experiment. This is the setting for classical parameter estimation theory [61]. If, on the other hand, the measurement is undecided, then this amounts to distinguishing parametrised quantum states, which is the setting for the theory of quantum parameter estimation [62].

Let us first consider the classical setting, where we require a statistical distinguishability metric for PDs. The Fisher Information (FI), named after the statistician Ronald A. Fisher (1890-1962), provides such a metric. To introduce this construct, we consider a single-parameter estimation scenario. We define a family of PDs, $p_\phi(\xi)$, and we want to estimate ϕ from a measurement register $\{\xi_1, \ldots, \xi_m\}$. The estimation is performed via a function $\phi_{\mathrm{est}}(\xi_1, \ldots, \xi_m)$. The estimator's deviation is $\Delta\phi_{\mathrm{est}} \equiv \phi_{\mathrm{est}}(\xi_1, \ldots, \xi_m) - \langle\phi_{\mathrm{est}}\rangle_\phi$, where $\langle\phi_{\mathrm{est}}\rangle_\phi$ is the estimator's expected value. By simply expanding this expression, we get the following equality:

$$\int d\xi_1 \ldots \xi_m \, p_\phi(\xi_1) \ldots p_\phi(\xi_m) \Delta\phi_{\mathrm{est}} = 0 \tag{2.13}$$

and taking the derivative with respect to ϕ,

$$\int d\xi_1 \dots \xi_m \, p_\phi(\xi_1) \dots p_\phi(\xi_m) \left(\sum_{i=1}^{m} \frac{\partial_\phi p_\phi(\xi_i)}{p_\phi(\xi_i)} \right) \Delta\phi_{\text{est}} = \frac{d\langle \phi_{\text{est}} \rangle_\phi}{d\phi}. \tag{2.14}$$

By applying the Cauchy-Schwartz inequality to the left-hand term, we obtain

$$\left[\int d\xi_1 \dots \xi_m \, p_\phi(\xi_1) \dots p_\phi(\xi_m) \left(\sum_{i=1}^{m} \frac{\partial_\phi p_\phi(\xi_i)}{p_\phi(\xi_i)} \right)^2 \right] \left\langle (\Delta\phi_{\text{est}})^2 \right\rangle_\phi \geq \left(\frac{d\langle \phi_{\text{est}} \rangle_\phi}{d\phi} \right)^2. \tag{2.15}$$

We can now use error propagation, in the same manner as in the previous subsections, to obtain

$$(\Delta\phi)^2 \equiv \left\langle (\Delta\phi_{\text{est}})^2 \right\rangle_\phi \Bigg/ \left(\frac{d\langle \phi_{\text{est}} \rangle_\phi}{d\phi} \right)^2 \geq \left[m \int d\xi \, p_\phi(\xi) \left(\frac{\partial_\phi p_\phi(\xi)}{p_\phi(\xi)} \right)^2 \right]^{-1}. \tag{2.16}$$

Finally, FI is defined as

$$F(p_\phi) \equiv \int d\xi \, \frac{(\partial_\phi p_\phi(\xi))^2}{p_\phi(\xi)} \tag{2.17}$$

and the inequality above is known as the Cramér-Rao (CR) bound [63],

$$(\Delta\phi)^2 \geq \frac{1}{mF(p_\phi)}. \tag{2.18}$$

An estimator that achieves this bound only exists if $p_\phi(\xi)$ belongs to the family of exponential PD functions. However, for $m \gg 1$ and any well-behaved $p_\phi(\xi)$, the CR bound is achieved by the maximum likelihood estimator [64,65],

$$\phi_{\text{max.-lik.}}(\xi_1, \dots, \xi_m) \equiv \arg\max_\phi \left[\sum_{i=1}^{m} \log\left(p_\phi(\xi_i)\right) \right]. \tag{2.19}$$

We sketch below a brief proof of this result. Using Bayes' theorem and ignoring any prior information, the PD of ϕ, given the measurement register $\{\xi_1, \xi_2 \dots, \xi_m\}$, is the likelihood function

$$\begin{aligned} \mathcal{L}\left(\phi | \xi_1, \xi_2 \dots, \xi_m\right) =& \mathcal{N} \exp\left[\sum_{i}^{m} \log(p_\phi(\xi_i)) \right] \\ =& \mathcal{N} \exp\left(c_0 + c_1(\phi - \phi_0) + c_2(\phi - \phi_0)^2 + c_3(\phi - \phi_0)^3 + \dots \right). \end{aligned} \tag{2.20}$$

In this expansion, c_0 contributes a constant factor, which can be included in the normalisation constant $\mathcal{N}$, while $c_1 = 0$ if we choose $\phi_0 = \phi_{\text{max.-lik.}}$. The second order term in the exponent is

$$c_2 = \frac{1}{2}\left(\frac{1}{m}\sum_i^m \partial_\phi^2 \log(p_\phi(\xi_m))\right)\left[\sqrt{m}(\phi - \phi_{\text{max.-lik.}})\right]^2. \tag{2.21}$$

The expected value of the sum in Equation 2.21 is $-F(p_\phi)$, and the deviations around the expected value scale as $1/\sqrt{m}$. If we rewrite the expansion in Equation 2.20 in terms of the variable $\sqrt{m}(\phi - \phi_{\text{max.-lik.}})$, the constant coefficients corresponding to terms of order k will scale as $m^{1-k/2}$. Therefore, for $m \gg 1$, we have

$$\mathcal{L}\left(\phi|\xi_1, \xi_2 \ldots, \xi_m\right) \approx \mathcal{N}' \exp(-mF(p_\phi)\ (\phi - \phi_{\text{max.-lik.}})^2/2). \tag{2.22}$$

The width of this approximation to the likelihood function is $1/\sqrt{mF(p_\phi)}$. This shows that the CR bound is achieved asymptotically by the maximum likelihood estimator.

FI defines a natural (Riemannian) metric on PDs, with the infinitesimal distance between two neighboring points given by

$$(ds_{\text{PD}})^2 \equiv F(\phi)(d\phi)^2. \tag{2.23}$$

In quantum parameter estimation, ϕ describes a quantum state $\hat{\rho}_\phi$ and the PD $p_\phi(\xi)$ is given by a positive-operator valued measure (POVM) with elements $\hat{\Pi}_\xi \geq 0$ ($\int d\xi\, \hat{\Pi}_\xi = \hat{I}$), such that $p_\phi(\xi) = \text{Tr}\left(\hat{\rho}_\phi \hat{\Pi}_\xi\right)$. The optimal estimation strategy is then found by maximising $F(\phi)$ over all possible quantum measurements. The maximal value is referred to as quantum Fisher information (QFI) and we denote it $F_{\text{Q}}(\hat{\rho}_\phi)$:

$$F_{\text{Q}}(\hat{\rho}_\phi) \equiv \max_{\text{meas.}}\left(F(p_\phi)\right) = \max_{\hat{\Pi}_\xi}\left\{\int d\xi\ \left[\text{Tr}\left(\partial_\phi \hat{\rho}_\phi \hat{\Pi}_\xi\right)\right]^2 \Big/ \text{Tr}\left(\hat{\rho}_\phi \hat{\Pi}_\xi\right)\right\}. \tag{2.24}$$

The optimisation procedure yields a valid distinguishability metric on quantum states, $(ds_{QS})^2 \equiv \max_{\text{meas.}}\left[(ds_{PD}(\phi))^2\right]$. In order to find an expression for $F_{\text{Q}}(\hat{\rho}_\phi)$, it is useful to define the symmetric logarithmic derivative (SLD), $\hat{L}(\hat{\rho}_\phi)$, such that

$$\frac{1}{2}\left[\hat{\rho}_\phi \hat{L}(\hat{\rho}_\phi) + \hat{L}(\hat{\rho}_\phi)\hat{\rho}_\phi\right] = \partial_\phi \hat{\rho}_\phi. \tag{2.25}$$

Since $\hat{\rho}_\phi$, $\hat{\Pi}_\xi$ and $\hat{L}(\hat{\rho}_\phi)$ are all Hermitian, we have

$$\begin{aligned}\mathrm{Tr}\left(\partial_\phi\hat{\rho}_\phi\hat{\Pi}_\xi\right) &= \frac{1}{2}\left[\mathrm{Tr}\left(\hat{L}(\hat{\rho}_\phi)\hat{\rho}_\phi\hat{\Pi}_\xi\right) + \mathrm{Tr}\left(\hat{\Pi}_\xi\hat{\rho}_\phi\hat{L}(\hat{\rho}_\phi)\right)\right] \\ &= \mathrm{Re}\left[\mathrm{Tr}\left(\hat{L}(\hat{\rho}_\phi)\hat{\rho}_\phi\hat{\Pi}_\xi\right)\right] \leq \left|\mathrm{Tr}\left(\hat{L}(\hat{\rho}_\phi)\hat{\rho}_\phi\hat{\Pi}_\xi\right)\right|.\end{aligned} \tag{2.26}$$

Plugging this into Definition 2.24,

$$\begin{aligned}F_\mathrm{Q}(\hat{\rho}_\phi) &\leq \max_{\hat{\Pi}_\xi}\left[\int d\xi \left|\mathrm{Tr}\left(\hat{L}(\hat{\rho}_\phi)\hat{\rho}_\phi\hat{\Pi}_\xi\right)\right|^2 \Big/ \mathrm{Tr}\left(\hat{\rho}_\phi\hat{\Pi}_\xi\right)\right] \\ &= \max_{\hat{\Pi}_\xi}\left\{\int d\xi \left|\mathrm{Tr}\left[\hat{\Pi}_\xi^{\frac{1}{2}}\,\hat{L}(\hat{\rho}_\phi)\,\hat{\rho}_\phi^{\frac{1}{2}}\,\hat{\rho}_\phi^{\frac{1}{2}}\,\hat{\Pi}_\xi^{\frac{1}{2}}\right] \Big/ \left(\mathrm{Tr}\left(\hat{\rho}_\phi\hat{\Pi}_\xi\right)\right)^{\frac{1}{2}}\right]\right|^2\right\}.\end{aligned} \tag{2.27}$$

Applying the Cauchy-Schwartz inequality, in the form $\left|\mathrm{Tr}\left(\hat{O}^\dagger\hat{P}\right)\right|^2 \leq \mathrm{Tr}\left(\hat{O}^\dagger\hat{O}\right)\mathrm{Tr}\left(\hat{P}^\dagger\hat{P}\right)$, for $\hat{O} = \hat{\rho}_\phi^{\frac{1}{2}}\,\hat{L}(\hat{\rho}_\phi)\,\hat{\Pi}_\xi^{\frac{1}{2}}$ and $\hat{P} = \hat{\rho}_\phi^{\frac{1}{2}}\,\hat{\Pi}_\xi^{\frac{1}{2}} \Big/ \mathrm{Tr}\left(\hat{\rho}_\phi\hat{\Pi}_\xi\right)$, we get

$$F_\mathrm{Q}(\hat{\rho}_\phi) \leq \int d\xi\,\mathrm{Tr}\left[\hat{\Pi}_\xi\left(\hat{L}(\hat{\rho}_\phi)\right)^2\hat{\rho}_\phi\right] = \mathrm{Tr}\left[\left(\hat{L}(\hat{\rho}_\phi)\right)^2\hat{\rho}_\phi\right]. \tag{2.28}$$

This result was first derived by Helstrom[66] and Holevo[67], yet it was established as central to quantum metrology by Braunstein and Caves[68], who explicitly separated the classical and quantum optimisation steps and showed that the limit obtained in Equation 2.28 can be saturated. Indeed, a measurement that projects $\hat{\rho}_\phi$ on the eigenstates of $\hat{L}(\hat{\rho}_\phi)$[69] saturates both inequalities used above. We now have a distinguishability metric for quantum states, which can be used to derive limits for quantum parameter estimation:

$$(ds_{QS})^2 = F_\mathrm{Q}(\hat{\rho}_\phi)(d\phi)^2 = \mathrm{Tr}\left[\left(\hat{L}(\hat{\rho}_\phi)\right)^2\hat{\rho}_\phi\right](d\phi)^2. \tag{2.29}$$

In order to derive precision bounds, one can optimise the QFI for probe states that obey some relevant constraint. It is useful to note that the QFI is convex: $F_\mathrm{Q}\left(\sum_i p_i\rho_\phi^{(i)}\right) \leq \sum_i p_i F_\mathrm{Q}\left(\rho_\phi^{(i)}\right)$.

Let us consider states with a fixed number of photons, n. It is easy to show, using the convexity of the QFI, that in this case the optimal state must be pure. For pure states, the QFI reduces to $F_\mathrm{Q}(|\psi\rangle_\phi) = 4(\langle\dot{\psi}|\dot{\psi}\rangle_\phi - |\langle\dot{\psi}|\psi\rangle_\phi|^2)$, where $|\dot{\psi}\rangle_\phi \equiv \partial_\phi|\psi\rangle_\phi$. In the particular case of interferometric phase estimation, we have $|\psi\rangle_\phi = e^{-\mathrm{i}\hat{J}_z\phi}|\psi\rangle_0$ where $|\psi\rangle_0$ is the state

after the input beam-splitter, so the QFI is

$$F_{\mathrm{Q}}(|\psi\rangle_\phi) = 4(\langle\psi|\,\hat{J}_z^2\,|\psi\rangle_0 - \langle\psi|\,\hat{J}_z\,|\psi\rangle_0^2) = 4\,\Delta\hat{J}_z\Big|_0 . \tag{2.30}$$

The variance of $\hat{J}_z = \frac{1}{2}(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b})$ is maximised by the N00N state[29], $|\mathrm{N00N}\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle_\mathrm{A} \otimes |n\rangle_\mathrm{B} + |n\rangle_\mathrm{A} \otimes |0\rangle_\mathrm{B})$, which sets the ultimate precision limit for all inputs with n photons,

$$(\Delta\phi)^2_{\mathrm{N00N}} = \frac{1}{n^2}. \tag{2.31}$$

As we mentioned in the previous subsection, this is referred to as the Heisenberg limit. The name is due to the similarity between the Heisenberg uncertainty principle and Inequality 2.18, with FI given by Equation 2.30.

In practice, N00N states are notoriously difficult to realise. A comparatively simple way to achieve the same precision scaling is to pass the probing light multiple times through the sampled optical path[70], which is implemented in GWD. After n passes, a single photon picks up the phase $n\phi$, which is equivalent to the particle-entangled N00N state, corresponding to the same number of photons passing through the sampled path. The disadvantage of multi-pass interferometry is a decrease in temporal resolution, as well as the accumulation of losses. N00N states are also strongly affected by losses, with the same scaling for both strategies[71,72]. We note, however, that the effect of detection loss is less significant in multi-pass than in N00N state interferometry. Moreover, if the relevant resource is the amount of light absorbed by a sample rather than light passing through it (like in the case of photo-sensitive samples), the multi-pass scheme is more effective, as at most one photon per realisation can be absorbed by the sample. As we detail in Section 2.1.4, the main practical motivation for investigating particle-entangled states for phase estimation is the possibility of achieving improved tolerance to loss, and other decoherence effects, compared to multi-pass interferometry.

We note that the QFI is not invariant to phase-shifts generated by $\hat{n}$. This issue can lead to misleading results for states that do not have a fixed number of photons. While coherences between components of the state with different total photon-number contribute to the QFI, in practice they can only be accessed in the presence of a phase reference[9]. In the optical domain, such a reference is implemented by a strong coherent field, which is not

part of our model of an interferometer and, unless extra assumptions about phase stability are made, it must be accounted for as a resource[73]. The issue is resolved by averaging states with respect to the phase generated by $\hat{n}$, before computing the QFI[73].

Now, taking a new look at phase estimation with coherent light and squeezed vacuum with the help of QFI, we find[73]

$$F_{\mathrm{Q}}(\rho_{\phi}^{\mathrm{coh.+sq.}}) = |\alpha|^2 e^{2r} + \sinh(r)^2, \tag{2.32}$$

with the notation already defined in the previous subsection, and after total-phase averaging, as discussed above. Fixing the average photon-number, $\langle\hat{n}\rangle = |\alpha|^2 + \sinh(r)^2$ and taking the limit $\langle\hat{n}\rangle \gg 1$, this expression is maximised when $\sinh(r)^2 \approx |\alpha|^2$, in contrast with the scheme discussed in Section 2.1.2. The Cramér-Rao bound then yields

$$(\Delta\phi_{\mathrm{coh.+sq.}})^2 \approx \frac{1}{\langle\hat{n}\rangle^2}. \tag{2.33}$$

It has been shown that, in principle, this limit can be achieved with photon-counting[74], considering the full information available in the photon statistics, rather than just the mean output intensity.

While these results provide useful ultimate bounds, there are several issues characterising the QFI approach to quantum parameter estimation. First, FI only depends on local quantities: the output probabilities and their derivatives with respect to ϕ. This can give misleading results. For example, using N00N states, all phases $\phi_i = \phi + 2i\,\pi/n$ with integer i give an equivalent output, as the state picks up n times the phase ϕ. Therefore, in practice, the prior distribution of ϕ must be very sharp in order for a N00N state measurement to be useful. Furthermore, for estimation schemes where the output signal is at an extremal point, the sign of a small phase-shift cannot be straight-forwardly determined. This issue may not be captured by the FI approach.

Secondly, QFI is not appropriate for deriving limits with a constraint on the average number of photons. Consider, for example, the state $\frac{1}{\sqrt{n}}\,|\mathrm{N00N}\rangle\langle\mathrm{N00N}| + \frac{\sqrt{n-1}}{\sqrt{n}}\,|00\rangle\langle 00|$, where $|\mathrm{N00N}\rangle$ has n photons. This has 1 photon on average, yet its QFI is n, which can be made arbitrarily high.

The Bayesian approach to quantum phase estimation effectively tackles the issues stated

above[38,75–78]. In this framework, the phase ϕ is treated as a random variable in the interval $[0, 2\pi)$, distributed according to the prior $p(\phi)$. Registering the output ξ produced by an interferometric measurement leads to a posterior distribution, according to Bayes' theorem, $p(\phi|\xi) = p(\xi|\phi)\frac{p(\phi)}{p(\xi)}$. In general, an estimator's MSE can be calculated using the posterior PD, and the minimum is achieved by the Bayesian mean estimator, $\phi_{\mathrm{BME}}(\xi) = \langle\phi\rangle_{p(\phi|\xi)}$. We note that, in the case of phase estimation, the squared error is not an appropriate cost (error) function; a well defined cost function should have the same value for ϕ and $\phi + 2\pi$. A natural choice is $C(\phi_{\mathrm{est}}(\xi), \phi) = 4\sin^2\left(\frac{\phi_{\mathrm{est}}(\xi)-\phi}{2}\right)$, which for sharp posterior distributions is approximately the squared error. Phase estimation precision is naturally quantified by the mean Bayesian cost (MBC),

$$\langle C\rangle \equiv \int d\phi\, d\xi\, p(\phi, \xi)\, C(\phi_{\mathrm{est}}(\xi), \phi), \tag{2.34}$$

which, for our choice of cost function, is minimised by the estimator $\phi_C(\xi) = \arg\left[\langle e^{\phi}\rangle_{p(\phi|\xi)}\right]$.

It is important to note that in the Bayesian framework, we consider the performance of a single experiment. The full estimation strategy, including repeated runs, (adaptive) measurements and data analysis, is effectively encapsulated in the measurement operators[70]. Every possible outcome corresponds to a phase estimate, so we can identify the POVM elements by this phase: $\hat{\Pi}_\varphi \equiv \hat{\Pi}_\xi\Big|_{\phi_{\mathrm{est}}(\xi)=\varphi}$. The measured state is $\rho_\phi = \hat{U}_\phi\rho_0\hat{U}_\phi^\dagger$, with $\hat{U}_\phi \equiv e^{-i\hat{J}_z}$. We choose the uniform prior, $p(\phi) = \frac{1}{2\pi}$, denoting complete lack of initial knowledge. Using these definitions, the MBC is

$$\langle C\rangle = \frac{1}{2\pi}\int d\phi\, d\varphi\, \mathrm{Tr}\left(\hat{U}_\phi\hat{\rho}_0\hat{U}_\phi^\dagger\hat{\Pi}_\varphi\right) C(\varphi, \phi). \tag{2.35}$$

We can always define a new POVM,

$$\hat{\Pi}_\varphi^{(c)} \equiv \hat{U}_\varphi\hat{\Pi}\,\hat{U}_\varphi^\dagger, \text{ with } \hat{\Pi} \equiv \frac{1}{2\pi}\int d\theta\, \hat{U}_\theta^\dagger\hat{\Pi}_\theta\hat{U}_\theta, \tag{2.36}$$

which describes a covariant measurement. Using Equations 2.35 and 2.36, and the symmetry of the cost function, it can be shown that this new POVM yields the same MBC as the

original one. We then have

$$\langle C\rangle = \frac{1}{2\pi}\int d\phi\, d\varphi\, \mathrm{Tr}\left[\hat{U}_\phi \rho_0 \hat{U}_\phi^\dagger \hat{\Pi}_\varphi^{(c)}\right] C(\varphi,\phi) = 4\,\mathrm{Tr}\left[\hat{\Pi}\int_0^{2\pi} d\phi\, \hat{\rho}_\phi \sin^2(\phi)\right]. \tag{2.37}$$

This leaves only the task of optimising with respect to $\hat{\Pi}$, with the constraints

$$\int_0^{2\pi} d\varphi\, \hat{U}_\varphi \hat{\Pi}\, \hat{U}_\varphi^\dagger = \hat{I} \text{ and } \hat{\Pi} \geq 0. \tag{2.38}$$

The minimum MBC is a concave function of the probe state: $\left\langle C\left(\sum_i p_i \hat{\rho}_\phi^{(i)}\right)\right\rangle_{\min} \geq \sum_i p_i \left\langle C\left(\hat{\rho}_\phi^{(i)}\right)\right\rangle_{\min}$. Therefore, the n-photon state that optimises the MBC must be a pure state of the form $|\psi\rangle = \sum_{i=0}^n \psi_i |i, n-i\rangle$. We can compute the integral in Equation 2.37:

$$\begin{aligned}\left[\int_0^{2\pi} d\phi\, \hat{U}_\phi |\psi\rangle\langle\psi| {U_\phi}^\dagger \sin^2(\phi)\right]_{i,j} &= \psi_i \psi_j^* \int_0^{2\pi} d\phi\, \sin^2(\phi) e^{\mathrm{i}\phi(j-i)} \\ &= \psi_i \psi_j^* \left[\pi\delta_{i,j} + \frac{\pi}{2}(\delta_{i,j+1} + \delta_{i,j-1})\right],\end{aligned} \tag{2.39}$$

where $\delta_{i,j}$ is the Kronecker delta. Using this, we find $\langle C\rangle = 4\pi \sum_{i=0}^n |\psi_i|^2 \Pi_{i,i} - 4\pi \mathrm{Re}\left(\sum_{i=1}^n \psi_i^* \psi_{i-1} \Pi_{i,i-1}\right)$. From Equation 2.38, we get $\Pi_{i,i} = \frac{1}{2\pi}$ and $\Pi_{i,i-1} \leq \frac{1}{2\pi}$. Together with $\langle\psi|\psi\rangle = 1$, this yields a lower limit for the MBC,

$$\langle C\rangle \geq 2 - 2\sum_{i=1}^n |\psi_i \psi_{i-1}| = 2 - \langle\psi|\hat{A}|\psi\rangle, \tag{2.40}$$

with $A_{i,i-1} = A_{i-1,i} = 1$. The limit can be achieved by setting $\hat{\Pi} = |\chi\rangle\langle\chi|$ with $|\chi\rangle = \frac{1}{\sqrt{2\pi}} \sum_{i=0}^n \psi_i/|\psi_i|$. Setting all ψ_i to be real, we find the optimal POVM:

$$\Pi_\varphi^{(\mathrm{n})} = \frac{1}{2\pi}\sum_{i,j} e^{\mathrm{i}\varphi(i-j)} |i, n-i\rangle\langle j, n-j|, \tag{2.41}$$

which is referred to as the canonical phase measurement [79–81]. The expression in Equation 2.40 is minimised by the eigenstate of $\hat{A}$ with the highest eigenvalue, which is known as the sine state:

$$|\psi\rangle_{\text{sine}} = \sum_{i=0}^n \sqrt{\frac{2}{n+2}} \sin\left(\pi\frac{i+1}{n+2}\right) |i\rangle_{\mathrm{A}} |n-i\rangle_{\mathrm{B}}, \tag{2.42}$$

with $\langle C\rangle_{\text{sine}} = 2\left[1-\cos\left(\frac{\pi}{n+2}\right)\right]$. In the limit of $n \gg 1$, we get the lowest possible error, in the Bayesian sense, for states with n photons:

$$(\Delta\phi_{\text{sine}})^2 \approx \frac{\pi^2}{n^2}. \tag{2.43}$$

We have thus found the optimal probe state for quantum phase estimation in the absence of prior information. This is quite different from the N00N state, yet it yields similar estimation precision.

For states with an indefinite number of photons, in the absence of a phase reference, we can use the concavity of the MBC to show that $(\Delta\phi)^2 \geq \frac{\pi^2}{\langle\hat{n}\rangle^2}$.

2.1.4 The effect of decoherence

The practical applicability of quantum-enhanced sensing is severely limited by decoherence. In optical interferometry, three decoherence mechanisms must be taken into account: (i) optical loss, (ii) phase diffusion, and (iii) imperfect visibility.

Arguably the main impediment towards the implementation of quantum-enhanced interferometry is optical loss[82,83]. It occurs at all stages of an experimental setup, due to absorption and scattering in optical elements: in the preparation through imperfect mode-selection, and in the measurement, due to sub-unit detector efficiency and detection noise. Balanced optical loss in the arms of an interferometer can be modeled all in one stage (preparation, evolution or detection).

Phase diffusion, also known as collective dephasing, denotes fluctuations in the value of the estimated parameter. It can arise due to thermal fluctuations, seismic noise, air movement, temperature changes, radiation pressure, etc. The effect of phase diffusion depends on the time scale and frequency range of the phase modulation to be detected.

Imperfect visibility, also known as local dephasing, arises because the modes combined on the output beam-splitter are not perfectly matched. Equivalently, it can be seen as phase diffusion acting individually on each photon, or as optical loss where the scattered light is incoherently fed back into the detection stage.

In the presence of decoherence, quantum FI can be optimised numerically for states with a low number of photons, up to ~ 100[54,84–86]. In addition, several methods have been proposed for deriving asymptotic limits. In general, these methods use a representa-

tion of the quantum channel modeling the interferometer (consisting of a phase-shift plus decoherence processes) from which an upper bound on QFI can be derived. This bound can be optimised over a set of possible channel representations. For example, one general method is that of channel purification, which produces the limit $F_{\mathrm{Q}}(\rho_\phi) = \min_{|\psi\rangle_\phi} \left[F_{\mathrm{Q}}(|\psi\rangle_\phi)\right]$ with $\rho_\phi = \mathrm{Tr}_{\mathrm{E}}(|\psi\rangle_\phi\langle\psi|)$. Here, the state $|\psi\rangle_\phi$ extends over the system and additional environmental degrees of freedom, denoted by E[87,88]. Other methods are classical simulation, where the channel is approximated locally by a classical distribution over quantum operations, with parametrised classical probabilities, and quantum simulation or channel extension[89]. In the presence of loss or imperfect visibility, all of these methods produce bounds of the form $\langle\Delta\phi^2\rangle \geq \text{const.}/n$, which limit any improvement over the SNL to a constant factor. This is due to the fact that probe states achieving Heisenberg scaling in the absence of decoherence are more fragile in a practical setting than uncorrelated states.

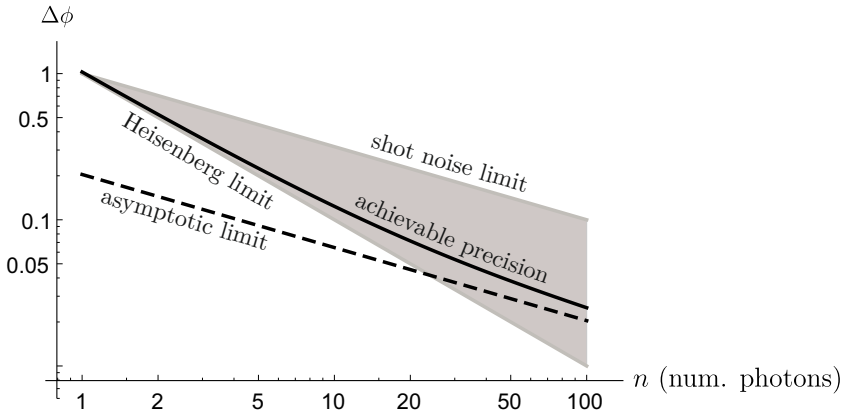


Figure 2.4: Typical scaling of phase estimation precision, in the presence of decoherence. The optimum performance is bounded from above by the classical limit, $\Delta\phi = 1/\sqrt{n}$ and from below by the Heisenberg limit $\Delta\phi = 1/n$ and the asymptotic bound, $\Delta\phi = \text{const.}/\sqrt{n}$, where the constant depends on the decoherence parameters, as shown in Table 2.1.

Intuitively, as we increase the number of available photons n, it becomes preferable to limit the size of N00N-like states to $k < n$ photons, and use these individually n/k times. For large n, the improvement on precision therefore scales like the SNL. We illustrate this in Figure 2.4. For example, in the case of N00N state (or multi-pass) interferometry with transmittance η in one arm and 1 in the other, the optimal choice is $k \approx 1.478/|\log(\eta)|$, yielding the precision bound $\langle(\Delta\phi)^2\rangle \geq C|\log(\eta)|/n$, with $C \approx 1.618$[72,84].

The general limits produced by channel extension methods[87–90], and confirmed by

Table 2.1: Lower bounds on the phase estimation MSE $\langle(\Delta\phi)^2\rangle$ and the MBC $\langle C\rangle$, respectively (the two are equivalent for high photon numbers), in the presence of decoherence processes.

Dephasing process	Classical limit	Quantum limit (QFI)	Quantum limit (Bayesian)
Phase diffusion (variance Γ)	$\Gamma + \frac{1}{n}$	$\Gamma + \frac{\pi^2}{n^2}$	$2\left(1 - e^{-\frac{\Gamma}{2}}\right) + \frac{\pi^2}{n^2} e^{-\frac{\Gamma}{2}}$
Loss (efficiencies $\eta_A,\ \eta_B$)	$\frac{1}{4}\left(\sqrt{\frac{1}{\eta_A}} + \sqrt{\frac{1}{\eta_B}}\right)^2 \frac{1}{n}$	$\frac{1}{4}\left(\sqrt{\frac{1-\eta_A}{\eta_A}} + \sqrt{\frac{1-\eta_B}{\eta_B}}\right)^2 \frac{1}{n}$	
Imperfect visibility (mode overlap η)	$\frac{1}{n\,\eta^2}$	$\frac{1-\eta^2}{n\,\eta^2}$	

other techniques[91,92], are presented in Table 2.1. The Bayesian approach produces similar bounds[90,93], through the optimisation presented in the previous subsection for states affected by decoherence. We include these, too, in Table 2.1. Note that the QFI and Bayesian approaches agree in the case of uncorrelated noise (loss and local dephasing), despite the fact that in the absence of decoherence they yield different limits. For phase diffusion, the two approaches only agree in the limit of small fluctuations. We note that various strategies employing ancillary systems have been proposed, which may achieve better precision than the limits presented in Table 2.1[94–97].

The limits derived for uncorrelated decoherence processes are saturated asymptotically by the coherent light and squeezed vacuum scheme discussed in Section 2.1.2. In fact, the sensitivity of the GEO-600 detector, where the squeezed vacuum technique has been implemented, is limited by loss, rather than difficulties in the engineering of quantum correlations[54]. More exotic approaches aiming to achieve Heisenberg scaling can only significantly surpass squeezed vacuum interferometry as described in Section 2.1.2 if losses are very low ($\eta > 90\%$) and only a small number of photons ($n \lesssim 100$) is used. So far, experimental realisations of NOON or HB states with a few photons have been shown to achieve high enough interference visibility, yet they have been limited by high preparation and detection losses[36,82,83,98,99].

However, there is still scope in considering phase estimation with exotic quantum states, such as in a situation where the phase must be estimated in the absence of any prior knowledge, in a single-shot manner, as we detail in the following section. These requirements

are relevant in applications like imaging or communication using phase-shift keying [100], which motivate the work that we present in this chapter.

2.1.5 Global estimation of phase

In the already familiar example of GWD interferometry, the phase is set in order to obtain an interference minimum at the monitored output. At this setting, the rate of change of the measured intensity as a function of phase shifts vanishes. However, the phase sensitivity of the measurement is optimal. The reason for this is captured by the FI expression given by Definition 2.17: the sensitivity remains high if the average signal vanishes. This scenario is possible in GWD because we are only interested in phase perturbation signals at a non-zero frequency, where there is no systematic noise, and the output is carefully filtered.

In general phase estimation, however, there are several reasons why it is optimal to set the measurement on a steep slope of the detected fringe. In fact, several effects lead to a divergent variance of the phase estimator around phase values that correspond to fringe minima or maxima. First, due to decoherence, the measured signal does not vanish at fringe extrema, leading to low phase sensitivity close to these points [99]. Recently, it has been shown that the situation can be improved, by accessing additional degrees of freedom of the probing light [101]. Secondly, as indicated in Section 2.1.1, the phase estimator variance diverges for states with an undefined total number of photons, an issue that can be addressed by keeping track of this quantity, if the measurement setup allows it [78]. Thirdly, there is an ambiguity in choosing the phase estimator in the proximity of a fringe extremum.

We denote the task of determining a phase-shift with no prior information and with sensitivity independent of its value as the global estimation of phase (GEP). One simple way of implementing GEP is to perform measurements with a uniformly sampled (known) phase offset for each measurement repetition [76,102,103]. More elaborate schemes must be considered in order to obtain sub-SNL precision, and adaptive measurements are often employed [70,77]. These schemes rely on a controlled phase offset that is constantly updated, taking into account all previous measurement outcomes.

Previous research has demonstrated that, in principle (ignoring, for the moment, decoherence), adaptive estimation algorithms can achieve GEP with Heisenberg-like scaling, using N00N-state or multi-pass interferometry [36,104,105]. The protocols are, in effect, op-

tical implementations of the quantum phase estimation algorithm, a basic building-block of quantum computation [12]. They rely on subsequent photon-counting measurements, employing probe states that pick up the phase $2^k\phi$: either N00N states with 2^k photons or single-photons making 2^k passes. These measurements are performed $m(K,k)$ times, with k taking the values $K, K-1, \ldots 0$, each time adjusting the interferometer phase. For constant $m(K,k) > 3$, it has been shown that this protocol asymptotically achieves the $1/n^2$ precision scaling (for $n \gtrsim 100$), departing from the Heisenberg limit by a constant factor. In fact, if a variable $m(K,k)$ is used (with higher values for smaller k), Heisenberg scaling can be achieved without adaptive measurements [105]. Experimental implementations have used N00N states up to 4 photons [36] and multi-pass interferometry [104,105], post-selecting on lossless realisations. Analysing their performance in the presence of decoherence is an interesting future avenue for research.

The sensitivity of GEP is naturally assessed by the MBC, introduced in Section 2.1.3. As we have seen, the measurement that minimises this figure-of-merit is the canonical phase POVM [81], with elements $\hat{\Pi}_\varphi^{(n)} = \frac{1}{2\pi}\sum_{i,j}^{n} e^{-\mathrm{i}\varphi(i-j)}\,|i, n-i\rangle\langle j, n-j|$ (written in the photon number basis). There is no feasible way to realise this measurement exactly. However, there are practical schemes that provide a good approximation [70,77,106,107]. The algorithm presented above is such an approximation, if we use a representation of multiple temporal modes as a single interferometric probe state [70].

There is an ample body of work focusing on GEP schemes that use optical homodyning. This type of detection, which we describe in more detail in Section 2.2, uses a strong classical field, referred to as a local oscillator (LO), to perform a measurement of a weak field's quadrature, $\hat{X}_\theta = \frac{1}{\sqrt{2}}(e^{\mathrm{i}\theta}\,\hat{a}^\dagger + e^{-\mathrm{i}\theta}\,\hat{a})$. The parameter θ is determined by the LO phase with respect to the measured mode. In optical heterodyne detection, θ is modulated, usually by offsetting the LO frequency with respect to that of the measured field. Before discussing phase estimation with these measurements, note that we must depart from the model developed throughout this Chapter: two-mode interferometers with no external phase reference. For the moment, we relax the requirement that LOs be accounted for as resources, in order to convey the relevant GEP strategies.

Heterodyne measurement readily achieves single-shot GEP, as it samples all quadratures of the measured mode with equal probability. This is a standard way of implementing

GEP [108], although the estimation variance achieved with classical probe states is two times larger than the SNL (adapted to the single-mode scenario). The intuitive reason for the factor of 2 is that heterodyne detection yields information not only about the phase, but also about the amplitude of the measured field, and this information remains unused. A phase uncertainty is introduced by the measurement scheme in addition to the intrinsic uncertainty of the probe state. While the heterodyne phase variance scales with photon number at best as the SNL, it has been shown that homodyne measurements with an adaptive phase can perform better [106,109–111], reaching $\sim \log(n)/n^2$, with the appropriate feedback and estimation strategy [107]. Adaptive homodyning schemes have been demonstrated experimentally using classical light [37,112], as well as squeezed states, beating the SNL [113–115].

In Section 2.2, we present our homodyne detection scheme, which combines the advantages of heterodyne detection with two-mode interferometry, enabling GEP with sub-SNL capability, without an external reference and without the need for adaptive measurements.

2.1.6 Multi-parameter estimation

The simultaneous quantum estimation of multiple parameters can yield better sensitivity than determining them in a sequential, individual fashion [?]. Szczykulska, Baumgratz and Datta have recently presented a review of multi-parameter estimation, including historical development and recent advances [116]. In an interferometric setting, multi-parameter estimation has been studied for the estimation of multiple phases [117], as well as for the estimation of phase and the strength of a decoherence process (subject to fluctuations or simply not known in advance) [103,118].

The theoretical framework of local parameter estimation, discussed in Section 2.1.3, can be extended to the multi-parameter regime. The FI metric corresponding to the vector of parameters $\boldsymbol{\theta}$ is given by $(ds_{\mathrm{PD}})^2 = d\boldsymbol{\theta}\boldsymbol{F}d\boldsymbol{\theta}$, where $\boldsymbol{F}$ is now the FI matrix, with elements:

$$[\boldsymbol{F}(p_{\boldsymbol{\theta}})]_{i,j} \equiv \int d\xi \, \frac{\partial_{\theta_i} p_{\boldsymbol{\theta}}(\xi) \, \partial_{\theta_j} p_{\boldsymbol{\theta}}(\xi)}{p_{\boldsymbol{\theta}}(\xi)}. \tag{2.44}$$

The multi-parameter CR bound reads:

$$\mathrm{cov}(\boldsymbol{\theta}_{\mathrm{est}}) \geq \frac{1}{N} \boldsymbol{F}(p_{\boldsymbol{\theta}})^{-1}, \tag{2.45}$$

where $\mathrm{cov}(\boldsymbol{\theta}_{\mathrm{est}})$ denotes the covariance matrix of the estimators. For two parameters, it is easy to see that the matrix inverse on the right-hand side yields a higher bound on the estimator variances of θ_1 and θ_2 than the corresponding single-parameter FI, if $[\boldsymbol{F}(p_{\boldsymbol{\theta}})]_{1,2}$ is non-zero. The interpretation is that not knowing one of the parameters increases the estimation error on the other. One implication is that interferometers must be precisely characterised, including decoherence effects, in order to achieve the highest phase estimation sensitivity.

The multi-parameter QFI provides a valid metric on density matrices [119], called the Bures metric. The QFI matrix elements are

$$[\boldsymbol{F}_{\mathrm{Q}}(\hat{\rho}_{\boldsymbol{\theta}})]_{i,j} = \frac{1}{2}\mathrm{Tr}\left[\left(\hat{L}_{\theta_i}(\hat{\rho}_{\boldsymbol{\theta}})\hat{L}_{\theta_j}(\hat{\rho}_{\boldsymbol{\theta}}) + \hat{L}_{\theta_i}(\hat{\rho}_{\boldsymbol{\theta}})\hat{L}_{\theta_j}(\hat{\rho}_{\boldsymbol{\theta}})\right)\hat{\rho}_{\phi}\right]. \tag{2.46}$$

The bound on FI in terms of the QFI is not sharp in the multi-parmeter regime (it can be achieved if $\hat{L}_{\theta_i}(\hat{\rho}_{\boldsymbol{\theta}})$ and $\hat{L}_{\theta_j}(\hat{\rho}_{\boldsymbol{\theta}})$ commute). For quantum state estimation, which is equivalent to the estimation of all d^2-1 parameters of a d-dimensional density matrix, Gill and Massar showed that, in the limit of many measurement repetitions ($N \gg 1$), we have [120]

$$\mathrm{Tr}\left[\boldsymbol{F}_{\mathrm{Q}}(\hat{\rho}_{\boldsymbol{\theta}})^{-1}\mathrm{cov}(\boldsymbol{\theta})^{-1}\right] \leq (d-1)N. \tag{2.47}$$

Note that the quantum Cramér-Rao bound yields the much weaker limit $(d^2-1)N$. An estimation strategy must trade-off the errors associated to different parameters, in order to satisfy this bound [121].

We have shown, in work performed during the author's PhD candidacy but which we do not fully develop here [103], that, for the joint interferometric estimation of phase and the strength of phase diffusion, we have

$$\frac{F_{\mathrm{Q}}(\hat{\rho}_{\phi})^{-1}}{(\Delta\phi)^2} + \frac{F_{\mathrm{Q}}(\hat{\rho}_{\Gamma})^{-1}}{(\Delta\Gamma)^2} \leq N, \tag{2.48}$$

which holds for classical interferometry and N00N states. We provided numerical evidence (for HB states with up to 6 photons) that the same bound holds more generally, in the limit of small phase diffusion variance Γ. Furthermore, we found that in this scenario, collective (entangled) measurements on multiple probe states that evolve sequentially through the interferometer can offer an advantage over separable measurements. This is not the

case for single-parameter estimation, where we have $F_{\mathrm{Q}}\left(\hat{\rho}_{\phi}^{\otimes N}\right) = NF_{\mathrm{Q}}(\hat{\rho}_{\phi})$. We found that, even though the tools used to derive the bound given by Inequality 2.48 are those of local parameter estimation, the bound is attained by measurements implementing GEP. In particular, the double homodyne scheme that we present in the following section was proposed as a means of achieving the multi-parameter bound on phase and phase diffusion estimation. This shows an advantage of the measurement, in the fact that it provides information about parameters other than phase. On the other hand, in view of the trade-off relations discussed above, or simply considering the finite capacity of quantum channels [122], the measurements that yield the most information about phase cannot provide information about other parameters.

2.2 A double homodyne measurement for quantum-enhanced phase estimation without prior information

In this section, we introduce a measurement scheme that uses homodyne detection for interferometric phase estimation. As this scheme does not make use of a phase reference external to the interferometer, it can be considered an alternative to photon counting. We describe the estimation procedure and quantify the attainable precision, first in the ideal case and then in the presence of losses.

2.2.1 Optical homodyne detection

Balanced homodyne detection (HD) is an important experimental tool for quantum optics, as it can be employed to measure the quadratures of electromagnetic modes. One of its applications is in the implementation of optical tomography [123–125], a method for reconstructing the quantum state of an optical mode based on measured quadrature distributions.

HD relies on the interference of the measured signal with an LO. The two modes are combined on a balanced beam-splitter and the outputs are monitored by linear photodiodes [126]. An electrical signal proportional to the quadrature of the weak field is obtained by subtracting the photodiode currents, and amplified using a transimpedance amplifier [127].

Photodiode detection can be modeled as a photon-number measurement. We denote by

$\hat{a}_+$ and $\hat{a}_-$ the annihilation operators for the two detected modes, and by $\hat{a}_{\mathrm{LO}}$ and $\hat{a}_{\mathrm{sig.}}$ the annihilation operators corresponding to the modes preceding the HD beam-splitter. Using the balanced beam-splitter transformation defined in Section 1.2.1, we can derive the photon number operator transformations

$$\hat{n}_\pm = \frac{1}{2}(\hat{n}_{\mathrm{LO}} + \hat{n}_{\mathrm{sig.}} \pm (\hat{a}^\dagger_{\mathrm{LO}}\hat{a}_{\mathrm{sig.}} + \hat{a}_{\mathrm{LO}}\hat{a}^\dagger_{\mathrm{sig.}})). \tag{2.49}$$

As the LO is a strong classical field, $\hat{a}_{\mathrm{LO}}$ and $\hat{a}^\dagger_{\mathrm{LO}}$ can be replaced by the coherent field amplitudes α_{LO} and α^*_{LO}, respectively. In the difference signal, the first two terms in Equation 2.49 are canceled. The interference terms survive and produce a measurement output proportional to $|\alpha_{\mathrm{LO}}|(e^{\mathrm{i}\theta}\,\hat{a}^\dagger_{\mathrm{sig.}} + e^{-\mathrm{i}\theta}\,\hat{a}_{\mathrm{sig.}})$, with $\theta = \arg(\alpha)$. This represents the quadrature of the weak field, $X_\theta = \frac{1}{\sqrt{2}}(e^{\mathrm{i}\theta}\,\hat{a}^\dagger_{\mathrm{sig.}} + e^{-\mathrm{i}\theta}\,\hat{a}_{\mathrm{sig.}})$, scaled by the LO amplitude, $\sqrt{2}|\alpha_{\mathrm{LO}}|$. The finite mode-overlap between the LO and signal, as well as the finite quantum efficiency of the photodiodes, play equivalent roles, contributing to the efficiency of the HD.

2.2.2 Optimal global estimation of phase for a class of quantum probe states

Several experimental groups have observed high-resolution interference fringes, produced by probing interferometers with higher-order Fock states [36,82,83,98,99]. These experiments show that, with increasing photon number, the fringe patterns become more complex and prone to experimental imperfections, which makes GEP difficult. Here, we propose a measurement that allows for a more intuitive phase estimation strategy, without requiring an adaptive phase offset.

Figure 2.5 depicts our measurement scheme, which consists of orthogonal (or conjugate) quadrature projections in the two outputs of the interferometer. We denote by $|x\rangle_\theta$ the eigenstate of the quadrature operator $\hat{X}_\theta$ with the eigenvalue x. For the moment, we set $\theta = 0$. Then, the measurement output $\{x_1, x_2\}$ corresponds to a projection on the state $|x_1\rangle_0 \otimes |x_2\rangle_{\pi/2}$. Using the displacement operator, defined in Section 1.2.1, we can write $|x_1\rangle_0 = \hat{D}\left(x_1/\sqrt{2}\right)|0\rangle_0$ and $|x_2\rangle_{\pi/2} = \hat{D}\left(\mathrm{i}\,x_2/\sqrt{2}\right)|0\rangle_{\pi/2}$. We now propagate the measurement backwards, in order to obtain its effect on the probe state before the output beam-splitter. The vectors that describe the projection inside the interferometer are obtained by applying

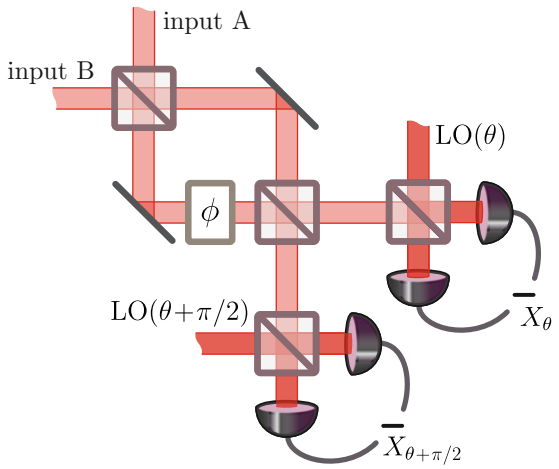


Figure 2.5: Double homodyne phase estimation scheme. Orthogonal quadrature measurements are performed in the two outputs of the interferometer, using homodyne detectors.

the inverse beam-splitter transformation, $\hat{B}^\dagger(50{:}50)$, defined in Section 1.2.1:

$$\begin{aligned}|\text{meas.}: x1, x2\rangle &= \hat{B}^\dagger(50{:}50)\,|x_1\rangle_0 \otimes |x_2\rangle_{\pi/2} = \hat{B}^\dagger(50{:}50)\,\hat{D}(x_1/\sqrt{2})|0\rangle_0 \otimes \hat{D}(\mathrm{i}\,x_2/\sqrt{2})|0\rangle_{\pi/2} \\ &= \hat{D}\left[(x_1+\mathrm{i}x_2)/2\right] \otimes \hat{D}\left[(-x_1+\mathrm{i}x_2)/2\right] \hat{B}^\dagger(50:50)\,|0\rangle_0 \otimes |0\rangle_{\pi/2}\,.\end{aligned} \tag{2.50}$$

As discussed in Section 1.2.1, the balanced beam-splitter transformation acting on $|0\rangle_0 \otimes |0\rangle_{\pi/2}$ gives an EPR state, which in the photon-number basis is written as $\frac{1}{\sqrt{\pi}}\sum_i |i\rangle \otimes |i\rangle$. Using this, we find

$$\begin{aligned}\langle j_1, j_2|\text{meas.}: x_1, x_2\rangle &= \frac{1}{\sqrt{\pi}}\sum_i \left\langle j_1 \left| \hat{D}\left((x_1+\mathrm{i}x_2)/2\right)\right| i\right\rangle \left\langle j_2 \left| \hat{D}\left((-x_1+\mathrm{i}x_2)/2\right)\right| i\right\rangle \\ &= \frac{1}{\sqrt{\pi}}\sum_i \left\langle j_1 \left| \hat{D}\left((x_1+\mathrm{i}x_2)/2\right)\right| i\right\rangle \left\langle i \left| \hat{D}\left((x_1+\mathrm{i}x_2)/2\right)\right| j_2\right\rangle \\ &= \frac{1}{\sqrt{\pi}}\left\langle j_1 \left| \hat{D}\left(x_1+\mathrm{i}x_2\right)\right| j_2\right\rangle = \frac{1}{\sqrt{\pi}} e^{-\mathrm{i}(j_1-j_2)\varphi}\left\langle j_1 \left| \hat{D}\left(r\right)\right| j_2\right\rangle,\end{aligned} \tag{2.51}$$

with $r \equiv \sqrt{x_1^2+x_2^2}$ and $\varphi \equiv -\arg(x_1+\mathrm{i}x_2)$. In the second line, we applied the resolution to the identity, formally taking the operators to act on the same space. For total photon-number n, we can now write the measurement operators for the double homodyne

measurement,

$$\begin{aligned}\hat{\Pi}^{(n)}_{r,\varphi} &= \frac{1}{2\pi}\sum_{i,j}\xi^{(n)}_i(r)\,\xi^{(n)}_j(r)\,e^{-\mathrm{i}(i-j)2\varphi}\,|i,n-i\rangle\langle j,n-j|\,,\ \text{where}\\ \xi^{(n)}_i(r) &= \sqrt{2}\left\langle i\left|\hat{D}\,(r)\right|n-i\right\rangle = \sqrt{2}\frac{\sqrt{i!}}{\sqrt{(n-i)!}}e^{-\frac{r^2}{2}}(-r)^{n-2i}L_i^{n-2i}(r^2),\end{aligned} \tag{2.52}$$

L denoting the generalised Laguerre polynomials[128].

Considering the phase of the LO, θ, we note that this translates to a relative phase between blocks with different total photon number n. Therefore, if we let θ be random (in the absence of a phase reference), the measurement effectively acts on the subspaces with definite total number of photons independently. Equation 2.52 is in the form of a covariant POVM, and it shows that a relative phase between the interferometer arms results in a rotation of the vector $\{x_1, x_2\}$.

For a single repetition of the measurement (a single-shot), the optimal estimator is $\varphi' = 2(\varphi \bmod \pi)$. In this case, we can easily associate a measurement operator to each estimate, integrating over r. This yields a covariant POVM: $\hat{\Pi}^{(n)}_{\varphi'} = \hat{U}_{\varphi'}\hat{\Pi}_{\text{s.-s.}}\hat{U}^{\dagger}_{\varphi'}$, where

$$\left\langle i\left|\hat{\Pi}^{(n)}_{\text{s.-s.}}\right|j\right\rangle = \frac{1}{2\pi}\int_0^{\infty} r\,dr\,\xi^{(n)}_i(r)\xi^{(n)}_j(r) \tag{2.53}$$

This integral can be evaluated analytically[129]. In particular, we have $\left\langle i\left|\hat{\Pi}^{(n)}_{\text{s.-s.}}\right|i\right\rangle = \frac{1}{2\pi}$ and $\left\langle i\left|\hat{\Pi}^{(n)}_{\text{s.-s.}}\right|n-i\right\rangle = \frac{(-1)^n}{2\pi}$. This shows that for $n = 1$, the POVM is equivalent with the canonical phase measurement. However, that is not the case for higher n, at least not considering the simple single-shot estimator, as in general $\left\langle i\left|\hat{\Pi}^{(n)}_{\text{s.-s.}}\right|j\right\rangle \leq \frac{1}{2\pi}$.

In order to study the asymptotic limit corresponding to many measurement repetitions, we want to calculate the FI for a pure probe state with n photons, $|\psi\rangle = \sum_{i=0}^{n}\psi_i\,|i,n-i\rangle$. The probability of obtaining the measurement outputs $\{r,\varphi\}$ reads

$$p_\phi(r,\varphi) = \frac{1}{2\pi}\left|\sum_{i=0}^{n}\psi_i\xi^{(n)}_i(r)e^{\mathrm{i}(n/2-i)(\phi-2\varphi)}\right|^2. \tag{2.54}$$

The corresponding FI is

$$F(p_\phi) = \int r\,dr\,d\varphi\ \left(\partial_\phi p_\phi(r,\varphi)\right)^2\Big/p_\phi(r,\varphi) = \int r\,dr\,d\varphi\ \left(2\,\partial_\phi\sqrt{p_\phi(r,\varphi)}\right)^2. \tag{2.55}$$

The FI expression is greatly simplified for probe states that are path-symmetric [130], fulfilling the condition $\psi_i = \psi^*_{n-i} e^{-2i\chi_0}$. This family of states is of great relevance, as it includes most examples studied so far, describing classical, squeezed light, N00N-state and HB-state interferometry. We show in Appendix A that, for path-symmetric states

$$F(p_\phi) = \sum_i^n |\psi_i|^2 (n-2i) = 4\langle \Delta \hat{J}_z \rangle^2 = F_Q(\hat{U}_\phi |\psi\rangle), \tag{2.56}$$

which means that these states achieve the quantum CR bound. It has been shown that precisely the same states achieve the quantum CR bound with photon counting measurements [130]. There is, however, a difference between the two detection schemes: while photon-counting naturally defines the total number of photons, n, homodyne detection does not, which means that the CR bound cannot be achieved in practice with homodyne measurements and probe states with indefinite photon-number (if we do not allow an external phase reference).

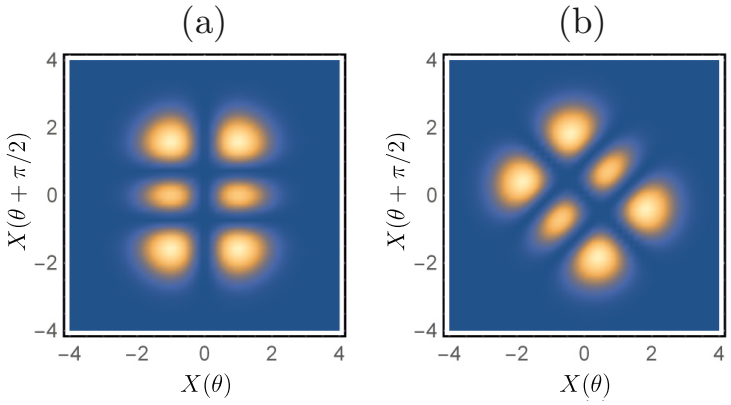


Figure 2.6: Measurement output distributions with Fock state inputs. The interferometer arms are prepared in states with with 1 and 2 photons, respectively. Orthogonal quadratures are measured in the two outputs. The phase is 0 in (a) and $\pi/2$ in (b).

To exemplify how the estimation works, let us consider Fock states in both interferometer inputs, $|i,j\rangle_\phi \equiv \hat{U}_\phi \hat{B}(50:50)|i\rangle \otimes |j\rangle$. The QFI is $F_Q(|i,j\rangle_\phi) = 2ij + i + j$. In Figure 2.6, we depict the PD of the measurement outputs $\{x_1, x_2\}$ corresponding to the input state $|1,2\rangle_\phi$. For $\phi = 0$, the PD is separable in the two outputs, corresponding to quadrature measurements of the two input modes. A phase shift results in a rotation of the PD, with correlations between the two homodyne signals. The details of the quadrature distribution can also provide information about other parameters, such as the photon

number distribution of the input states.

2.2.3 Realistic precision limits

In a practical setup, affected by optical losses, it is difficult to achieve a quantum enhancement over the SNL. In Figure 2.7, we show the FI per photon corresponding to the homodyne measurement, as a function of the total efficiency, for two-photon interferometry (with phase encoded in the state $|1,1\rangle_\phi$), as well as using a single-photon probe state. We also include the precision attained using coherent light, after phase averaging with respect to the LO. For strong coherent light, the FI is 1/2 of that corresponding to the SNL, just as it is the case for heterodyne phase estimation. In contrast, single photon probe states can achieve the SNL.

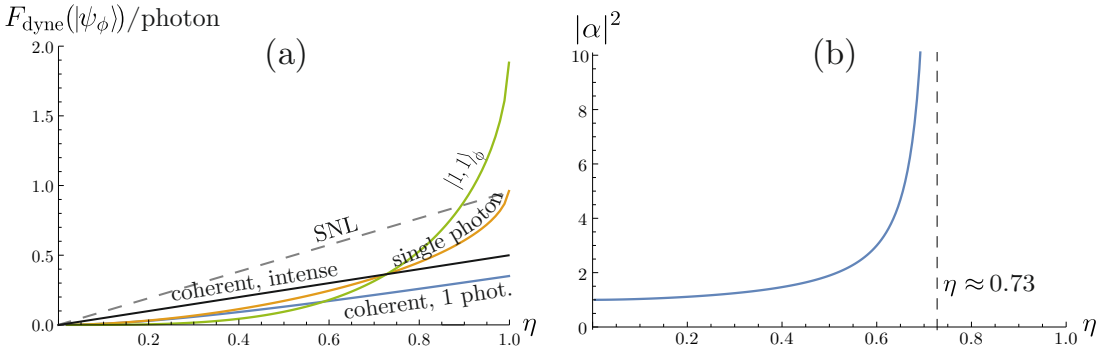


Figure 2.7: Fisher information per photon, with homodyne measurements, in the presence of loss. (a) FI achieved by strong coherent light (dashed black), coherent light with 1 photon on average (blue), single photon probe states (orange) and 2-photon N00N states (green), as a function of the total efficiency, η. (b) The largest number of photons in a coherent state that achieves the same FI per photon as a single photon input, as a function of the total efficiency, η. The plotted values were calculated numerically.

For weak coherent states, the FI per photon is less than the limit corresponding to intense fields. For a probe with 1 photon on average, it is always favourable to use a state with precisely one photon. The higher the efficiency, the larger the coherent state amplitude that is equivalent to single photons. As shown in Figure 2.7, for $\eta > 0.73$, a single-photon probe state provides an advantage over any classical state. In practice, the number of photons per experimental repetition is relevant in scenarios where a varying phase is monitored, or when implementing multi-pass interferometry.

For $\eta > 0.73$, the two-photon N00N state offers better precision than the single photon and for $\eta > 0.9$, this strategy beats the SNL (when considering the same efficiency for the

SNL measurement).

We note that 79% total efficiency for the homodyne detection of single photons has been reported in a continuous-wave experimental setup [131]. A practical advantage of homodyne detection over photon counting is the ability to implement fast and robust GEP, with high-bandwidth detectors and no need for a controlled phase offset.

2.3 Experimental implementation

In this section, we present the main experimental challenges and methods used to demonstrate the phase estimation scheme described above. We first introduce the HD setup and the single-photon source. We then sketch the experiment, and discuss in detail the data analysis process.

2.3.1 Pulsed optical homodyne detection

The key performance requirements of HD are high efficiency, SNR, bandwidth (BW), and stability. The development of fast, efficient HD has been stimulated by applications in quantum information processing.

Several research laboratories [39,132–134] have developed detectors that can measure pulsed optical modes at the full repetition rate of standard mode-locked Ti:Sapphire oscillators ($\sim 100\,\mathrm{MHz}$). Pulsed HD delivers fast data collection, as well as the possibility to sample, in real time, modes prepared by temporal filtering using click detectors [135]. Due to these capabilities, it is an essential tool for hybrid protocols that make combined use of discrete and continuous variables of the optical field [136].

Intensity measurements on a pulse train using linear photodiodes effectively integrate over the duration of ultrafast pulses. In order to sample each temporal mode, the photodiodes, combined with electronic amplification and read-out, must be fast i.e. have a constant response for frequencies ranging from 0 Hz to the pulse repetition rate. In our implementation of hybrid phase estimation, we use a mode-locked Ti:Sapphire laser (Spectra-Physics Tsunami) producing a 100 fs (300 fs filtered) pulse train, with a repetition rate of 80 MHz. Our detector has a bandwidth of 70 MHz, sufficiently high to sample each pulse.

We use Silicon PIN photodiodes (Hamamatsu S3883), with 90% quantum efficiency at 830 nm, 6 pF junction capacitance and 300 MHz bandwidth, reverse-biased at $\approx 20\,\mathrm{V}$.

These photodiodes have a low noise-equivalent power ($6.7\,\mathrm{W/Hz^{1/2}}$). Their active area has a diameter of 1.5 mm, which is large enough to ensure that beam-pointing drift does not have a strong effect on the stability of the detector, without needing to tightly focus the light modes. Moderate focusing maintains a highly linear response to signals with optical power below 10 mW. The difference photocurrent is amplified using a transimpedance amplifier, including a voltage amplification stage, with an overall gain of $36\,\mathrm{kV\,A^{-1}}$.

Figure 2.8 depicts the HD electronic circuit. The two photodiodes are mounted directly on a circuit board, and all electronic elements are in surface-mount form, minimising any delays and parasitic capacitances and inductances. The power supply lines and the signal are isolated by using top and bottom ground planes. Decoupling capacitors and voltage regulators filter any electronic noise that might affect the signal. The operational amplifiers (Texas Instruments OPA847) offer a high gain-bandwidth product (3.9 GHz) and low noise ($0.85\,\mathrm{nV/Hz^{1/2}}$). The feedback capacitances were optimised using SPICE[137], and are essential for eliminating amplification instabilities (smoothing the frequency response of the amplifier). The HD signal is monitored using an oscilloscope (Tektronix MSO5104) with a 10 GHz digital sampling rate and a 1 GHz bandwidth.

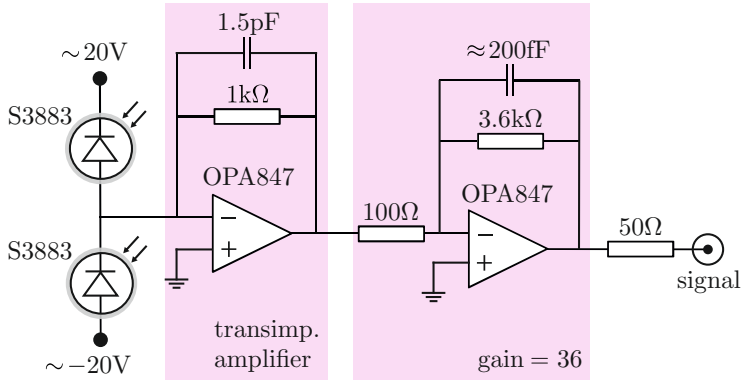


Figure 2.8: Homodyne detector electronic circuit. The difference photocurrent is amplified in two stages: a transimpedance amplifier with gain 1 kV/A and a voltage amplifier with gain = 36.

In our experiment, we use 2 mW of LO average power, which corresponds to 10^8 photons/pulse. Thus, the difference signals produced by quantum fluctuations are at the level of $\sim 10^4$ photons/pulse. The ratio between the average pulse intensity and the difference signal produced by quantum fluctuations gives an indication of how good the common-mode rejection (CMR) of the detector must be. The CMR depends on the overlap

between the pulse response functions of the two photodiodes, which must be matched in amplitude, timing and duration. The amplitudes are balanced by fine-tuning the HD beam-splitter ratio, implemented in polarisation, using a HWP and a polarising beam-splitter (PBS). The relative arrival time of pulses on the photodiodes is adjusted using an optical translation stage. The width of the response functions is tuned using individual bias voltages for the two photodiodes, as the junction capacitance, and thus the duration of the response function for each photodiode, depends on the bias voltage.

Table 2.2 summarises the performance of our pulsed homodyne detector. The detection setup has been used previously by Cooper, Smith et al., to measure the quadrature distribution of Fock states up to 3 photons[135], state[138] and process tomography[139]. For the work presented in this Chapter, we required two quadrature measurements. We implemented these by temporally multiplexing a single homodyne detector. As we describe in this section, using adequate data analysis, the detector can be used to perform two measuremets per pulse, at twice the repetition rate that it was originally designed for.

Table 2.2: Homodyne detector parameters.

PD efficiency	3 dB bandwidth	CMR (temp. multiplex.)	SNR (2 mW LO power)	Detection efficiency	Stability
90%	70 MHz	≈ 30 dB	11 dB	82%	$\gtrsim 2$ s [135]

2.3.2 A heralded source of single photons in pure quantum states

In our experimental implementation, we prepare a single-photon signal using SPDC. We use type-II phase matching, allowing us to separate SPDC photon pairs on a PBS, using one of the modes as a herald and one as a signal. Observing herald photons with a simple click detector, any spatio-temporal correlations between the modes in which the photon pairs are emitted results in decoherence for the signal photon. We note that the generation of a signal in a pure quantum state is necessary for the implementation of efficient HD, as this measurement effectively filters a single optical mode. Multi-mode emission results in non-unit interference visibility between the signal and the LO, which has an effect equivalent to optical loss. The efficiency of HD is at most equal to the purity of the measured signal, attained when the LO matches its marginal spatio-temporal distribution[140].

The single-photon source that we use was designed by Peter Moseley and reported in

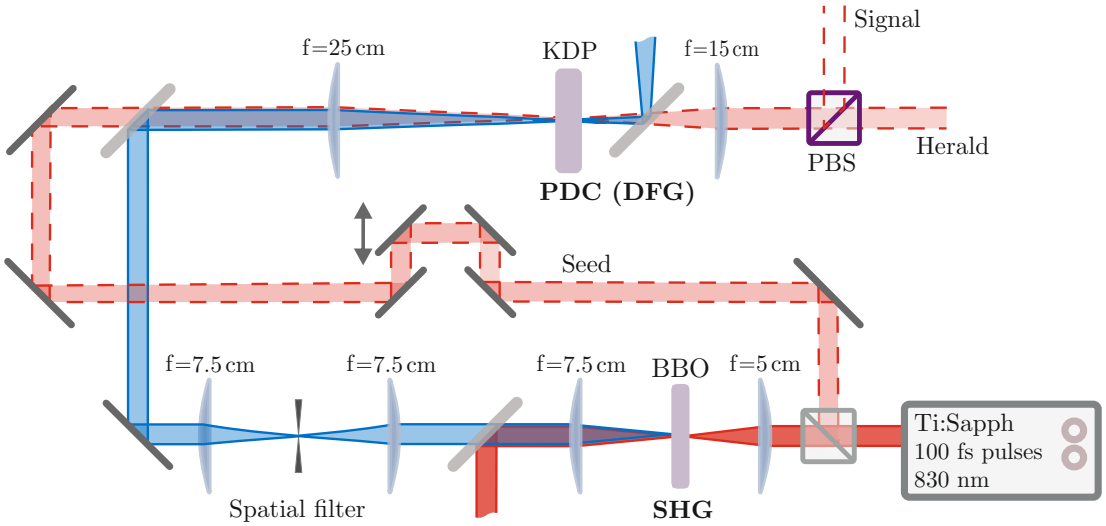


Figure 2.9: Sketch of the KDP type-II PDC single-photon source. This source generates photon pairs at 830 nm, in separable quantum states (uncorrelated in frequency or space). The alignment is performed with the help of a strong seeding field, which produces a DFG signal that simulates the spatio-temporal mode of the PDC.

2007[40,141]. It is the first PDC source engineered to produce a single-mode signal without requiring any filtering of the heralding mode. The PDC nonlinear medium is potassium dihydrogen phosphate (KDP). A KDP crystal is cut with the optical axis at an angle of 67.8° with respect to the direction of propagation, which allows the phase matching of a degenerate PDC process at 830 nm, in a collinear configuration. Moreover, the pump, an e-polarised pulse with 415 nm central wavelength, travels at the same group velocity as the o-polarised PDC signal, generated at 830 nm. This results in uncorrelated emission times for the PDC photon pairs, with the heralding photon well localised temporally. Together with the collinearity of the source, this ensures that the state of the signal and herald photons is separable.

The pump field is produced by second harmonic generation (SHG) in a barium borate (BBO) crystal. Special care is given to the spatial filtering of the SHG beam, which comes out elliptical, due to birefringent walk-off in the BBO. The Rayleigh length of the pump focusing in the KDP medium is matched to the length of the crystal, which ensures the optimal collection efficiency of the generated PDC[141]. Heralding photons are collected in an optical fiber and detected using an avalanche photodiode (APD). An optical fiber is also used as a mode reference for aligning the LO and the PDC signal. The LO is spatially cleaned using a short length of fiber, and coupled to the mode reference with 90% efficiency.

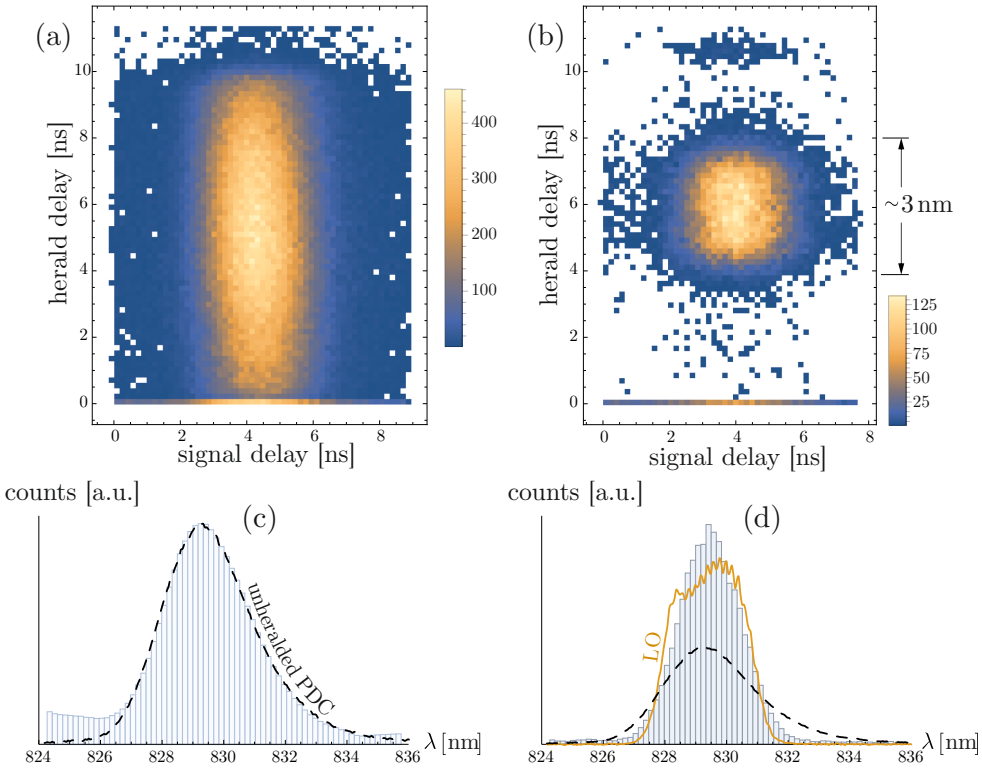


Figure 2.10: Spectrum of the photon-pairs produced by the KDP source. (a) Photon-pair joint spectrum, measured by converting wavelength into time-delay using linearly chirped fiber-Bragg gratings (FBG). This density histogram shows the binning of 3×10^5 coincidences. (b) The same type of joint spectrum, after inserting a 3 nm band-pass thin-film filter in the herald arm. This density histogram shows the binning of 3.5×10^4 coincidences. Measurement artifacts can be observed, such as a wavelength-independent reflection on the entrance-face of the FBG, as well as the finite wavelength range of the Bragg gratings. (c) We calibrate the wavelength-dependent delay by matching the (unheralded) arrival time distribution of signal photons (blue histogram, showing the distribution of 8.5×10^6 counts) with their marginal spectrum, measured on a grating spectrometer with a CCD sensor (black dashed line). (d) Using the calibration, we can compare the spectrum of heralded single photons, measured using FBGs (blue histogram) with the LO spectrum measured using the grating spectrometer (orange line).

Coupling the PDC signal to the same fiber, we employ an APD to detect coincidences with the heralding arm, which we use to optimise the heralding efficiency of the PDC source. For alignment, we also employ a classical seeding field that stimulates the PDC process, producing a difference frequency generation (DFG) signal that simulates the optical mode of the signal photons. This allows us to perform most of the spatial and temporal alignment using strong fields.

The separability of photon pairs in the time-frequency domain translates into the condition for having a factorable joint spectrum for the emmited photons. We measure the joint spectrum using a novel method based on chirped fiber Bragg gratings (FBG) [142]. Imprinted

in an optical fiber using UV writing, these gratings are reflective for different wavelengths at different positions. With the addition of an optical circulator, a chirped FBG can be used to introduce a large wavelength-dependent delay in an optical signal. We use two FBGs and APDs to detect coincidences in the two modes of the PDC source. We record the photon arrival times using a time-to-digital converter (qutools quTAU), and we use the same circuit to keep track of the laser pulse arrival times on a photodiode, reducing the repetition rate with a pulse delay generator (Stanford Instruments DG645). With this method, joint spectrum measurements, which would take a long time using gratings and scanning slits, can be performed at a rate of 10^4 coincidences/min, limited by the bandwidth of the time tagger output. We show the joint spectrum of our source in Figure 2.10. For calibration of the FBG delays, we measure the marginal spectrum of the signal photon both with the FBG method and using a grating spectrometer (Andor Shamrock 303) with a CCD camera (Andor iDus 401 series). This allows us to compare the spectrum of the heralded single photons with the spectrum of the LO, which is picked-off from the Ti:Sapph laser beam and filtered using a 3 nm band-pass thin-film filter, with angle-tuned central wavelength.

Using singular value decomposition of the joint spectrum, we estimate the purity of the source to be $> 96\%$ (greater than, because the higher-order modes that we find are mostly due to measurement noise, and smoothing the joint spectrum readily improves the estimated purity). For the work presented in this chapter, we improved the heralding efficiency of the source from $\approx 68\%$ to $\approx 75\%$, by reducing optical losses and adding a 3 nm bandwidth filter in the heralding arm. The performance of the single-photon source is summarised in Table 2.3.

Table 2.3: Single-photon source parameters.

Brightness (3 nm herald filtering)	Heralding efficiency – raw	Heralding efficiency – corrected for APD efficiency	Spectral purity
$\sim 10^5$ counts/s	34%	$\sim 75\%$	$> 96\%$

2.3.3 Experimental setup

We implemented the measurement scheme described in Figure 2.5 using a polarisation interferometer and a single homodyne detector, temporally multiplexed. This experimental setup has the advantages of excellent phase stability and precise phase calibration, sym-

metry between the two interferometer modes, which ensures balanced detection efficiency, as well as resource efficiency.

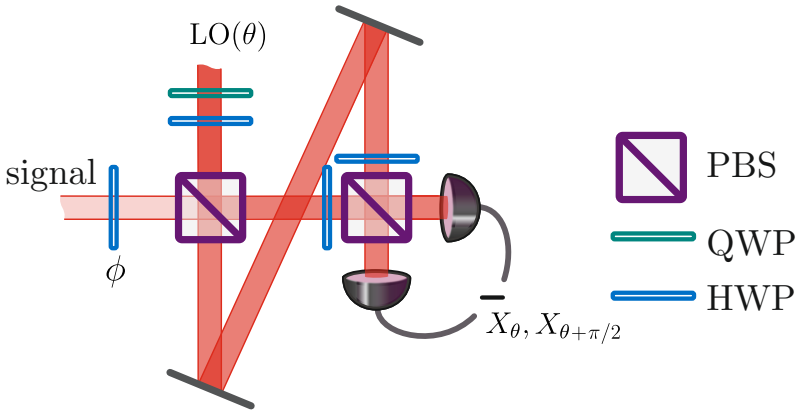


Figure 2.11: Temporally multiplexed detection of two polarisation modes. The H and V modes of a polarisation interferometer are split and overlapped spatially with the V and H polarisations of the LO on a PBS. Homodyne detection is implemented in both outputs, using a HWP and a second PBS, but with a delay of $\approx 1.87\,\mathrm{m}$ in one of the arms, which corresponds, in propagation time, to half of the repetition period of the laser.

The setup (not including signal preparation) is depicted in Figure 2.11. The two interferometer arms are the H and V polarisations of the signal spatial mode. Passively going between the H/V and the R/L polarisation bases corresponds to a 50 : 50 beam-splitter transformation. A half-wave plate (HWP) effectively introduces a phase between the R and L polarisations and is therefore equivalent to a Mach-Zehnder interferometer in the H/V basis. Setting the birefringent axis to an angle θ with respect to the vertical direction is equivalent to setting the interferometer phase to $\phi = 4\,\theta$. Using a PBS, the H- and V-polarisation modes of the interferometer are respectively overlapped with the V- and H-polarisation modes of the LO. A HWP, with its axis set to $\pi/8$ with respect to the vertical direction, then plays the role of the HD balanced beam-splitter and a PBS directs the two interference outputs to the HD photodiodes. We use the two input ports of the same PBS for both measurements, multiplexing a single homodyne detector. The two modes are distingushed by a temporal delay that corresponds to half of the laser repetition period. The delay is implemented using highly reflective dielectric mirrors ($R > 99.8\%$), ensuring that the added loss in the delayed arm is negligible.

In order to implement the measurement described in Section 2.2, we must appropriately set the relative phase between the H and V polarisations of the LO. The dielectric coating

on the first PBS introduces an unspecified phase-shift between the transmitted and the reflected modes, which can only be observed by interfering fields in both polarisations and both inputs. We use a D-polarised, phase-modulated, weak coherent field as input, in order to determine the phase-shift introduced by the PBS, by observing the phase-offset between the interference signals measured on the homodyne detector for the two PBS outputs. We compensate for this phase-shift using a HWP and QWP in the LO beam.

Table 2.4: Homodyne detection efficiency for single-photons.

Transmittance of opt. elem.	Spatial overlap	Spectral overlap	Spectral purity	HD efficiency	Observed total efficiency
87%	86%	96%	96%	82%	55%

We summarise the different parameters that contribute to the total efficiency of single-photon HD, in the temporally multiplexed scheme, in Table 2.4. The transmission of optical elements is determined by measuring the attenuation of a classical beam travelling the same path as the single photon signal. Due to polarisation dependence of the transmission, this value is valid to within 2%. The spatial overlap of signal and LO is estimated based on the parameters given in Table 2.3, and depends on a previous estimation of the APD's absolute efficiency. The spectral overlap assumes flat spectral phases for LO and single photons. The HD efficiency is estimated based on the parameters in Table 2.2. The different factors that contribute to the final efficiency match well the total efficiency reported here, which is determined directly from the data presented in this Chapter. During the data-taking process this efficiency drifted by at most 0.5% and we have, on other occasions, observed higher efficiencies, up to $\approx 57\%$.

For phase estimation with weak coherent light, we use a pick-off from the Ti-Sapph laser, attenuated and pulse-picked using an acousto-optical modulator (APE pulseSwitch).

2.3.4 Data analysis

For each measurement, we acquire an oscilloscope trace representing 100 ns of HD signal. The acquisition time window spans the two temporal modes probing the interferometer, as well as 7 pairs of unpopulated temporal modes. We use the latter to determine the variance corresponding to vacuum fluctuations, for detector calibration. A typical set of outputs is depicted in Figure 2.12.

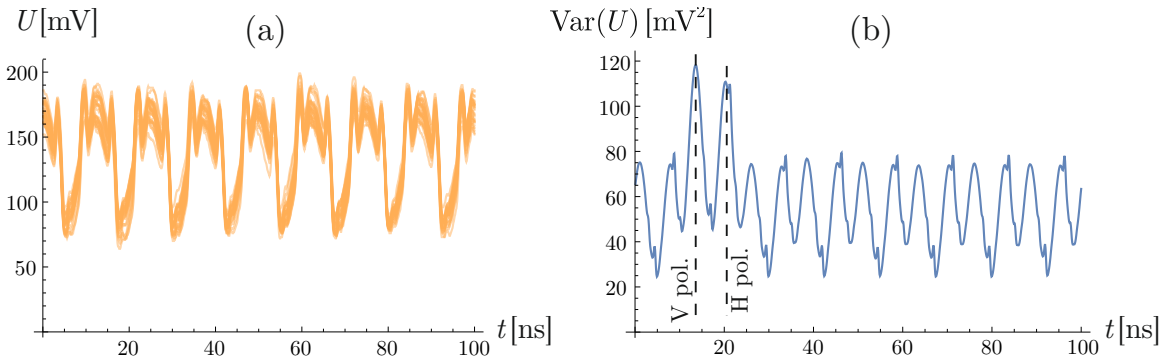


Figure 2.12: Homodyne signal. (a) 30 oscilloscope traces triggered by heralding APD events, showing typical HD signals. Quantum fluctuations are superposed with a modulation at 80 MHz, which is due to imperfect subtraction of the photodiode signals. (b) Variance of the HD signal as a function of the time passed from the triggering event, calculated using 4×10^5 traces. The peaks indicate quantum fluctuations that correspond to the different temporal modes of the pulsed optical signal, two per laser repetition period. The peaks with higher variance correspond to the modes in which we prepare a split single-photon. Note that the variance of a single-photon signal detected with unit efficiency is 3 times the variance corresponding to vacuum fluctuations.

The HD output is a linear combination of several signals. First, for each measured temporal mode, we have an output proportional to the impulse response function of the detector, weighed by the quadrature value that we wish to extract. Secondly, we have a modulation at 80 MHz due to imperfect rejection of the detector's common mode. Thirdly, the signal is affected by electronic noise, which is caused by photodiode dark currents, thermal noise in the amplification circuit, digitisation, and other noise sources. In the temporally multiplexed scheme, the CMR cannot be optimised to the same extent as with simple HD. In this arrangement, the two HD inputs have slightly different spot sizes, due to the longer propagation of the delayed mode, leading to slightly different response functions that cannot be adjusted independently.

Let us define a matrix A, which contains the data acquired for m measurements and where each oscilloscope trace, consisting of l points, forms a column. Assuming that the common mode signal is identical for all measurements, it can be easily removed by subtracting the mean value in each line of A. However, due to triggering jitter, the common mode signal is slightly displaced in time for each measurement. The jitter is greatly reduced by precisely triggering on a fast photodiode that monitors the laser pulses, preconditioned by a heralding or pulse-picking event. However, due to the persisting jitter, there remains a contribution to each measurement which is proportional to the time-derivative of the common mode signal, with randomly varying amplitude. This contribution can introduce artificial

correlations between the sampled quadratures. Another effect that introduces artificial correlations, if a simple integration is used to extract the quadrature values, is the overlap between the impulse response functions corresponding to subsequent temporal modes. This has a significant effect in the temporally multiplexed scheme, as in this case the temporal modes are spaced closer than the resolution of the detector. In order to eliminate unwanted correlations between the extracted quadrature values, we must deconvolve the HD signal. We do this by fitting the data to a linear combination of all the contributions described above, as illustrated in Figure 2.13.

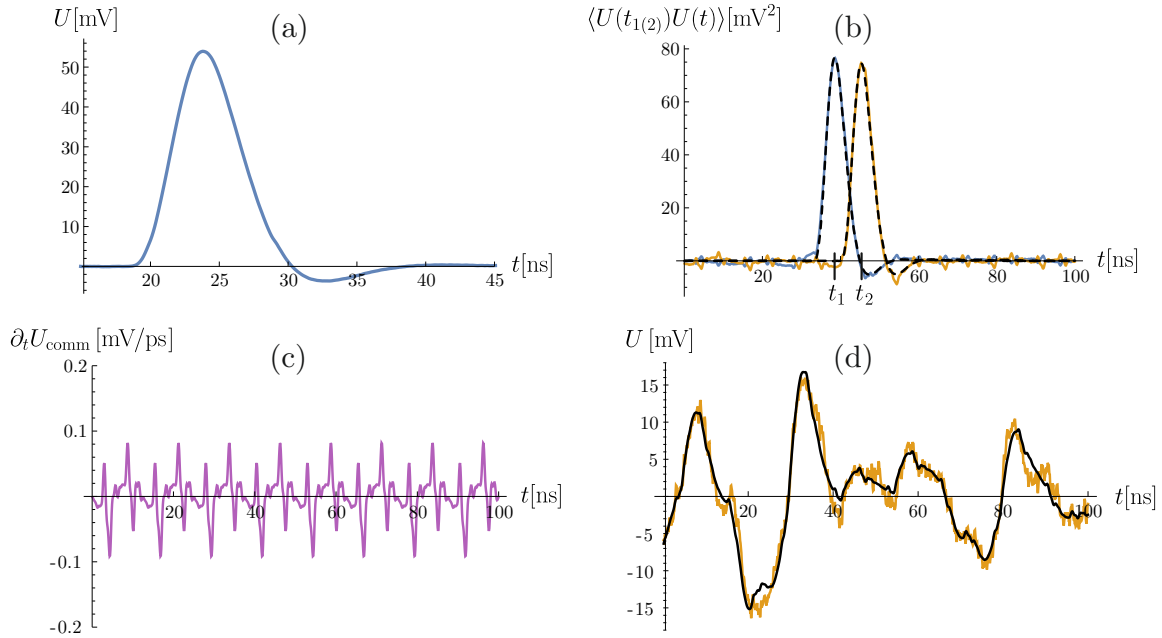


Figure 2.13: Quadrature extraction. (a) The impulse response function of the detector, measured by sending an isolated laser pulse onto one of the photodiodes. (b) Matching the temporal correlation of the HD signal (obtained using 4×10^5 traces) with respect to a point of maximum variance (for each of two neighboring temporal modes), with the measured response function. (c) Temporal derivative of the common mode signal (obtained by averaging 4×10^5 traces), which contributes to the HD output due to triggering jitter. (d) Example of a 100 ns HD signal (orange line) fitted as a linear combination of impulse response functions, centered on the arrival time of each pulse, and the common mode temporal derivative (the fit is given by the black line). The difference between the fit and the original signal is due to electronic noise.

We determined the impulse response function of the homodyne detector by sending a weak isolated laser pulse on one of the photodiodes (we call this an impulse response function because the femtosecond laser pulses are much faster than the correlation time of the detection signal). In order to discern the HD response to each temporal mode, we deconvolve the detector signal, as illustrated in Figure 2.13 (b). We calculate the correlation

between the signal values taken at different delays and the signal values taken at the delay that maximises the fluctuations corresponding to each temporal mode. We then fit the deconvolved signals to the measured inpulse response function, as illustrated in Figure 2.13 (b). We can then create a matrix, C, with columns consisting of impulse response functions, centered to each measured temporal mode, with one extra column for the time-derivative of the common mode. Let us arrange the quadrature values (which we wish to extract) as well as the amplitude of the common mode contribution, in the columns of a matrix Q. We then have $A \approx C \cdot Q$. We can easily find the matrix Q_{fit} that provides the best approximation to A, with respect to the Frobenius norm, using the Moore-Penrose pseudoinverse: $Q_{\text{fit}} = C^{+}A$. This reduces the quadrature extraction procedure to a simple matrix multiplication, followed by a rescaling by the standard deviation of vacuum fluctuations.

If we denote by $\mathcal{E}$ the electronic noise contribution, then the resulting error in the quadrature estimates is given by $C^{+}\mathcal{E}$. If the singular values of C^{+} are high, which is the case if there is a large overlap between response functions corresponding to subsequent temporal modes, the noise contribution will be amplified by the deconvolution. Another parameter that determines the effect of the electronic noise is the autocorrelation function: if this is sharp, corresponding to uncorrelated noise, then the error contribution to the quadrature estimates is reduced by effectively averaging the signal over the width of an optical response function, which is implicit in the deconvolution procedure. A detailed characterisation of the electronic noise in our homodyne detector is presented in Figure 2.14.

Applying the algorithm described above, we extract the quadrature distribution corresponding to a single photon signal, which we show in Figure 2.15. By calculating the variance of the quadrature distribution, from 4×10^5 samples, we estimate the overall single-photon detection efficiency to be $\eta = 0.553(2)$.

For phase estimation with weak coherent light, we aim to sample uniformly the phase of the probe state with respect to the LO. As the two beams come from the same laser, they have a well defined relative phase, which drifts due to temperature changes and air fluctuations. We introduce a modulation of this phase using a piezoelectric actuator attached to a mirror in the signal beam path. We drive the piezoelectric actuator with a saw-tooth wave at $\approx 30\,\text{Hz}$, with an amplitude that corresponds to a phase shift larger than 2π. We then select a window in the modulation signal, as illustrated in Figure 2.16. Vibrations

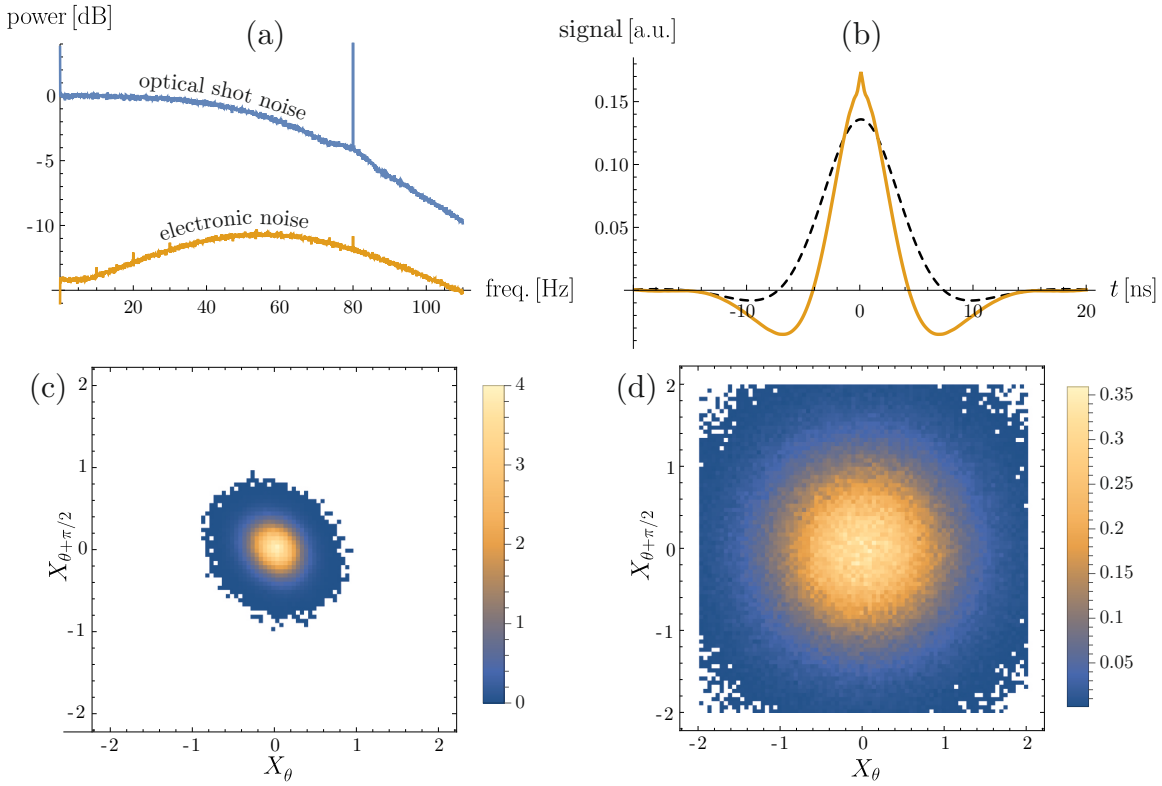


Figure 2.14: Electronic noise in the HD signal. (a) Spectrum of the HD signal, with the LO (blue line) and without the LO (orange line). The measurements were performed for the temporally multiplexed detector, with 2 mW LO power in each polarisation. The spectra were obtained from 40 ms of signal. We smoothed the Fourier transform with a Gaussian filter of 125 Hz width. (b) The autocorrelation function of the detector electronic noise (orange line) extracted from the spectrum shown in (a), compared to the autocorrelation of the optical response function (black dashed line). (c) Distribution of measurement outputs extracted from the HD signal without the LO (electronic noise) and (d) for comparison, the measurement output distribution corresponding to vacuum fluctuations (optical shot noise). The density histograms are scaled to represent sampled normalised PD functions.

introduced by the piezoelectric actuator introduce noise in the HD signal, which we remove with a digital low-pass filter.

2.4 Experimental results: demonstrating phase estimation accuracy and precision

In this section, we present our main experimental results. We have implemented the phase estimation scheme introduced in Section 2.2, using single photons, as well as coherent light. We use Bayesian estimation to deduce the interferometer phase from the experimental data, demonstrating both accuracy and precision over the full range of values that the phase can

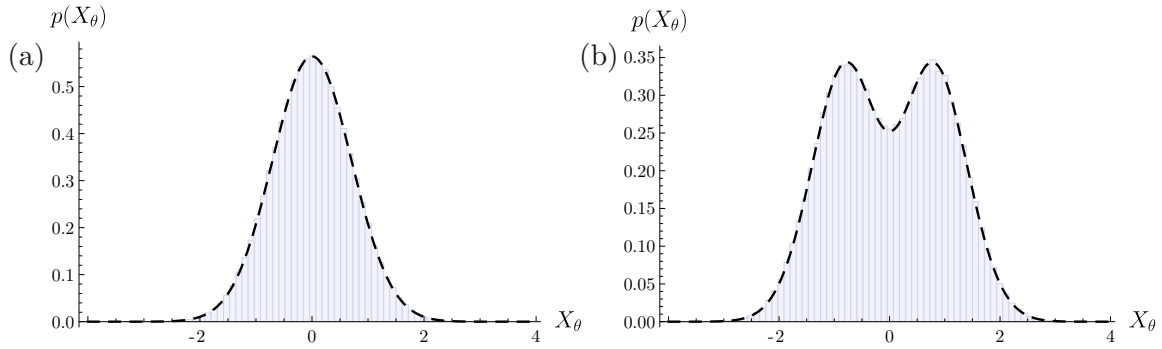


Figure 2.15: Measured quadrature distribution for (a) vacuum and (b) a single photon state. The blue histograms depict the distribution of 4×10^5 samples. The black dashed lines are theoretical fits, with 55.3% efficiency.

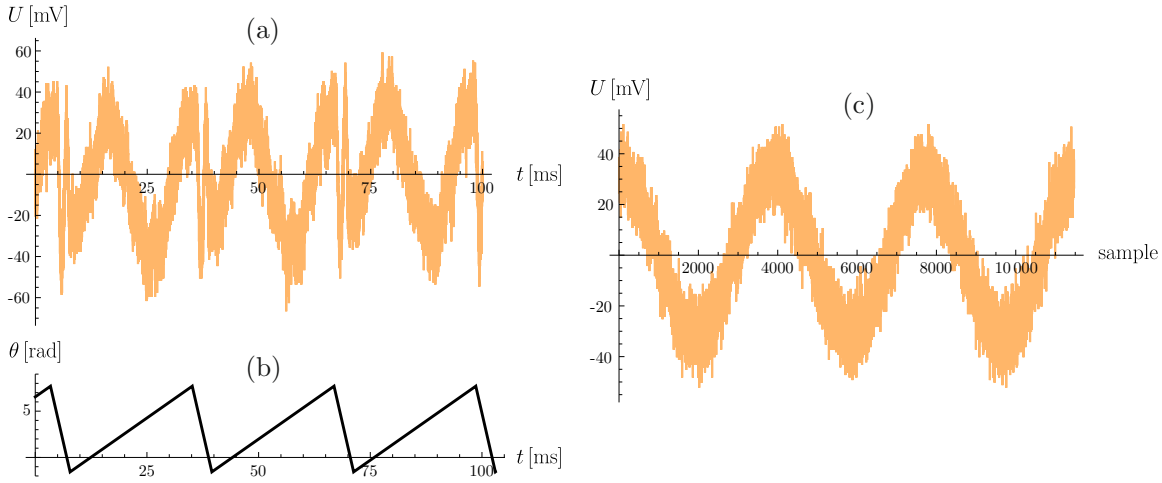


Figure 2.16: Phase modulation for sampling phase averaged coherent states. (a) Quadratures extracted for a phase modulated coherent state. (b) The phase modulation. (c) The processed signal, keeping only a 2π interval for each period of the modulation and discarding the rest of the data.

take.

2.4.1 Single-photon probe state

Here, we present the experimental results with single-photon probe states, i.e., the case when we send a single photon in one of the input ports of the interferometer (in our implementation, this is the V-polarised mode). For $\phi = 0$ or $\phi = \pi$, the single photon is localised at one of the outputs, and the two measured quadrature signals are uncorrelated. A much more interesting setting is $\phi = \pm\pi/2$, when the single photon occupies a superposition of the two output modes. In this case, we observe correlations between the two quadrature

measurements, produced by a single excitation. In Figure 2.17, we show the experimental distribution of quadrature values, which is in excellent agreement with the theoretical prediction derived in Section 2.2, where we showed that a phase-shift results in a rotation of the distribution of measurement outputs.

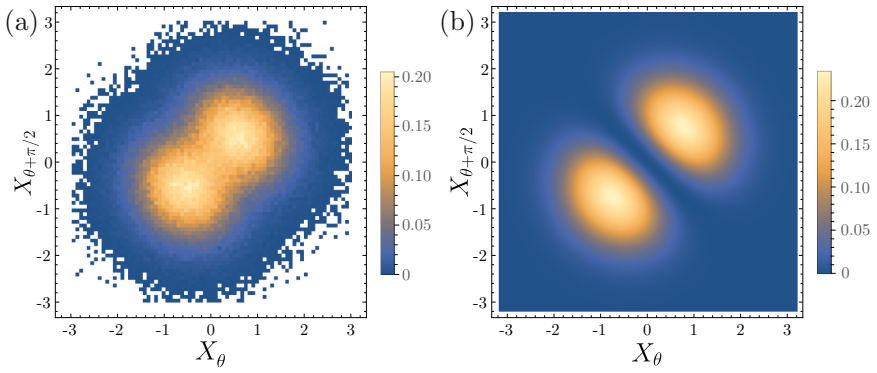


Figure 2.17: Quadrature distribution for a single-photon probe state, with $\phi = -\pi/2$. (a) Experimental distribution, reconstructed using 4×10^5 measurements. (b) For comparison, the expected distribution in the case of unit efficiency. Both figures represent normalised PD functions.

In order to demonstrate the accuracy of our phase estimation scheme, we rotate the HWP that implements polarisation interference between 0° and 90°, in steps of 4°. This corresponds to a scan of the phase ϕ between 0 and 2π. For each step, we record 10^5 measurements. Using the sampled quadratures, we calculate the posterior distribution for the interferometer phase. We model the phase-dependent probability distribution of the measurement outcomes including the effect of finite efficiency. Using the posterior distribution, we then calculate the estimator $\phi_{\text{est}} = \arg\left(\langle e^{i\phi}\rangle\right)$. This estimator minimises the Bayesian cost, defined in Section 2.1.3, as well as the Holevo variance[67], $\text{Var}_{\text{H}} \equiv \left|\langle e^{i\phi}\rangle\right|^{-2} - 1$. The results are shown in Figure 2.18. For the error bars, we use Var_{H}, which is equivalent to the standard quadratic variance for sharp distributions.

Some error sources that can limit the accuracy of our implementation are wave-plate imperfections, finite visibility of the polarisation analysis, artificial correlations between quadrature measurements, and unbalanced efficiency for the two detection arms. Despite these possible errors, we find that in our experimental implementation, without any corrections, the deviation of the phase estimate from the actual set value is at most on the order of 1°, which demonstrates the robustness of our method.

Let us now consider the precision of the phase estimation. In principle, for a large

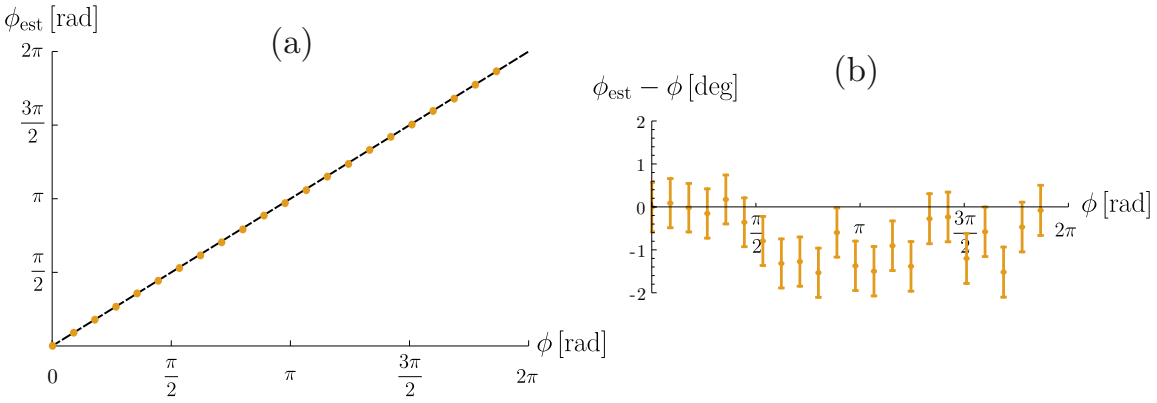


Figure 2.18: Phase estimation accuracy with single-photons. (a) The estimated phase (using 10^5 mesurements) versus the set phase, for 22 values in the interval $[0, 2\pi)$ (orange points). Errors are of the same order as the size of the points. The black dashed line gives the function $f(x) = x$. (b) The difference between the estimated phase and the set phase.

number of measurements, this should approach the Cramér-Rao bound. We divide the recorded data into blocks of m measurements, using each block to provide an estimate. We then calculate the variance of these estimates, which, for large m, is proportional to $1/m$. In Figure 2.19 we show that the proportionality constant is approximately equal to the FI, which we calculate taking the finite efficiency of the setup into account. The estimator variance determined from the experimental data is larger than the minimum variance given by the CR bound by only 5.3(10) %.

As discussed in Section 2.1.3, the estimator variance for a given realisation deviates from the expected value with a standard deviation proportional to $1/\sqrt{m}$. In Figure 2.19 (b), we show the Holevo variance of the posterior distribution, determined for different values of m, as well as the standard deviation of this variance, which matches the theoretical prediction very well. The Holevo variance is close to the CR bound for $100 < m < 2 \times 10^5$. For $m < 100$, Var_{H} increases, as this quantity diverges from the standard quadratic variance unless the posterior distribution is sharp. The data presented in Figure 2.19 corresponds to $\phi = -\pi/2$.

2.4.2 Classical probe state

We have performed the same phase estimation experiments with phase-averaged coherent probe states. Figure 2.20 shows the measured quadrature distribution for $\phi = \pi/2$, using a coherent probe state with average photon number $|\alpha|^2 \approx 3$. This distribution is in excellent

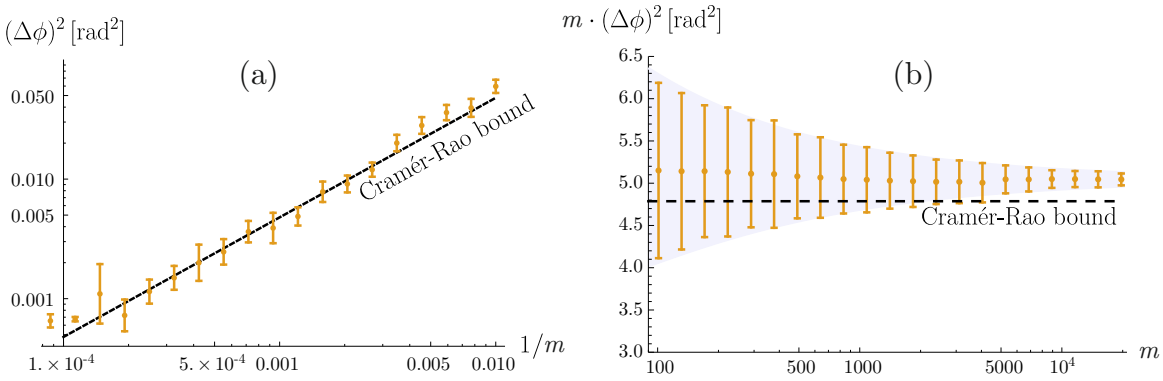


Figure 2.19: Phase estimation precision with single photons. (a) The variance of the phase estimates as a function of $1/m$ where m is the number of measurements used to calculate the estimates (orange points) and the minimum variance given by the Cramér-Rao bound (black line). (b) The Holevo variance of the phase probability distribution, scaled by the number of measurements m, as a function of m (orange points) and a theoretical prediction (light gray filling).

agreement with the theoretical prediction.

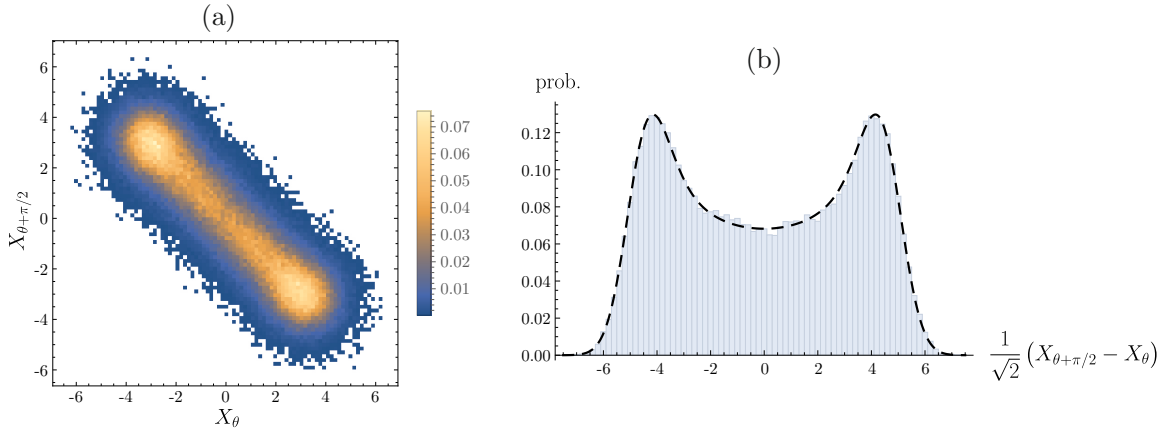


Figure 2.20: Measured quadrature distribution with coherent light. (a) Joint distribution of the two quadrature measurements for a probe state with ≈ 3 photons on average, and $\phi = \pi/2$. 1.8×10^5 samples were used. (b) The distribution of the rotated quadrature, corresponding to a projection of the data in (a) on the anti-diagonal (blue histogram) and the expected distribution, assuming a phase averaged coherent state (dashed black line).

Just like in the single-photon case, we determine the accuracy and precision of the phase estimation scheme with coherent light. We show the results for $|\alpha|^2 \approx 3$ in Figure 2.21. The accuracy is slightly worse than when using single photons, with a deviation of $\sim 2°$. The estimator variance is $5(1)\,\%$ higher than the Cramér-Rao bound, for $100 < m < 10^5$.

We find that single-photon probe states outperform weak coherent states. In Figure 2.22, we compare the precision attainable in our setup with the two strategies. The single-

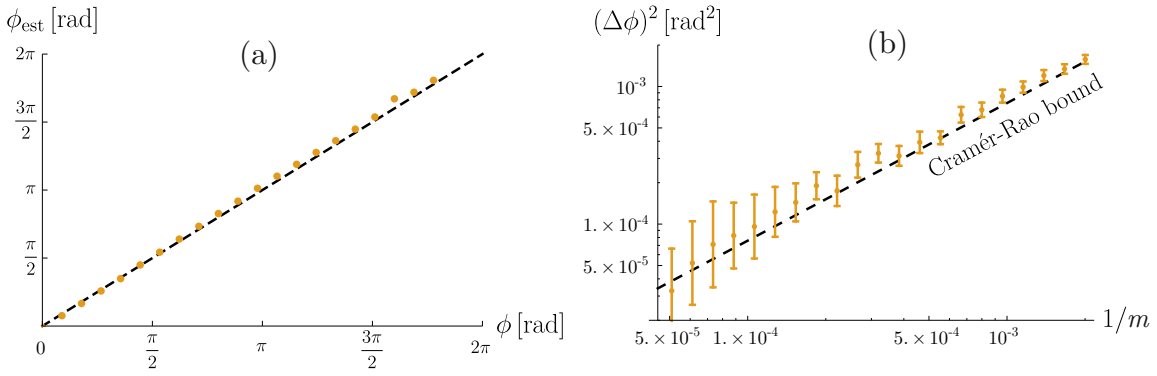


Figure 2.21: Phase estimation accuracy and precision with coherent light. (a) Phase estimate versus actual phase, for 21 values in the interval $[0, 2\pi)$, with ≈ 3 photons per probe state and $\sim 5 \times 10^4$ measurements. Errors are smaller than the point size. The black dashed line gives the function $f(x) = x$. (b) Variance of phase estimates as a function of $1/m$, where m is the number of measurements used to obtain the estimates (orange points), and the minimum variance allowed by the Cramér-Rao bound (black dashed line).

photon signal outperforms a coherent probe state with less than ≈ 2 photons on average, if we assume unit detection efficiency for the latter (this situation is depicted in Figure 2.22). If we impose the same detection efficiency (55.3%) to both scenarios, then we find that the single-photon signal outperforms coherent probe states with up to ≈ 4 photons on average.

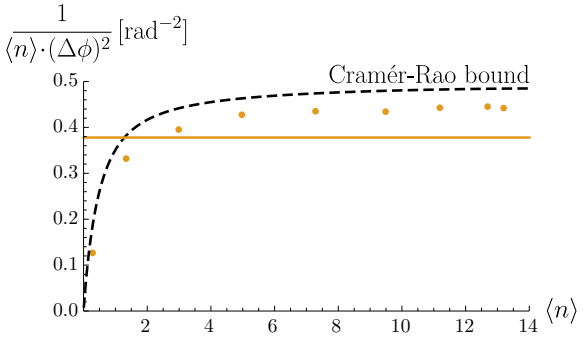


Figure 2.22: Comparison of single photon and coherent light estimation precision. Experimentally determined inverse variance of the phase estimate with coherent light, divided by the average number of photons per probe state, $\langle n\rangle$, as a function of $\langle n\rangle$ (orange points) and the maximum allowed values, given by the Cramér-Rao bound (black dashed line). The horizontal orange line gives the inverse variance, divided by the average number of detected photons per pulse, for a single-photon probe state, determined with our experimental setup.

An application of GEP is optical phase tracking. We demonstrate the tracking capability of our experimental setup, by following a phase that varies randomly, using 500 single-photon measurements per point. For continuously varying phases, smoothing strategies have been developed to improve phase estimation precision [37,113].

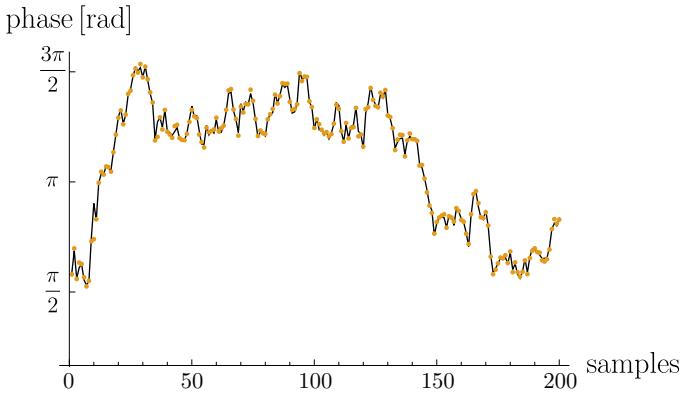


Figure 2.23: Phase tracking with single photons. A randomly varying phase (black line) and the estimates obtained using single-photons, with 500 measurements per point (orange points).

2.5 Summary and outlook

We have proposed and realised a phase estimation scheme that combines two-mode interferometry with double-homodyne measurement. The LO required to implement HD represents a resource in terms of optical energy, but not as a phase reference, as it does not need to have a well-defined phase with respect to the light that probes the interferometer.

Through our experimental implementation, we have demonstrated that two quadrature measurements can be performed using a single homodyne detector. To this end, we have developed optimal data analysis techniques that extend the functionality of our HD setup, using it at double the repetition-rate that it was designed for. Our experimental results are in excellent agreement with theoretical predictions, and we were able to demonstrate good estimation accuracy over a full 2π range, with no prior knowledge, no calibration and no corrections. We demonstrated a precision close to the theoretical limit, with the estimator variance only $\approx 5\%$ larger than that given by the Cramér-Rao bound. We showed that with our scheme, using a single-photon probe state can offer an advantage over weak coherent light.

Our HD achieved $\approx 55\%$ efficiency for the detection of single-photon signals. There are several contributions to this efficiency that can be readily improved: the loss on optical elements (by investing in better coatings), the quantum efficiency of the photodiodes (detectors with efficiency in excess of 98% are available), the electronic SNR (in the present implementation, the LO power is limited by self-phase modulation in the fiber used as a spatial-mode cleaner), and the spectral overlap between the LO and signal. With these im-

provements, we can foresee an efficiency in excess of 69%. For higher efficiency, continuous wave pumped sources may offer a better solution. Achieving the 90% efficiency required to beat the SNL would certainly be technologically challenging, but it is not out of the question. On a more exotic note, a non-destructive photon-number measurement in advance of the HD would greatly enhance the capability of the phase estimation scheme.

Possible applications of our method can be made in fast phase tracking with very low light levels, optical imaging, possibly with spatially multi-mode probe states, and optical communications using phase-shift keying. Estimation precision in these scenarios can be improved using multi-pass interferometry. Our experiment is a demonstration that going beyond the discrete variable paradigm in protocols such as phase estimation allows the design of advantageous quantum optical measurements, which can address practical problems.

Chapter 3

Realisation of a photonic Maxwell's demon

THIS chapter presents an application of experimental quantum optics techniques to a well-known scenario pertaining to the field of thermodynamics. Maxwell's demon has had a rich, thought-provoking history [143,144], highlighting the fact that information can be traded for thermodynamic work. In recent years, the demon is receiving renewed attention, this time embodied in a number of physical implementations [145–153] made possible by ongoing progress in nano-technology.

Here, we present an experimental implementation of Maxwell's demon in a photonics setup [154]. In particular, we show that a measurement at the few-photons level applied on intense thermal light and accompanied by feedback can allow extraction of work into an electric circuit. Our aims are to pinpoint some of the challenges that arise when putting the thought experiment into practice and to gain insights into the relation between thermodynamic work and information. Thermodynamics, a field of study with overarching implications on practical issues, is undergoing rapid progress, and we believe it is important that experiments guide this progress. To this end, our work contributes by putting forward photonic systems as a promising experimental platform.

We begin the Chapter with a brief history of Maxwell's demon and a discussion of its role in tackling questions of quantum information theory using thermodynamic arguments. We then review recently reported realisations of the demon in various physical systems. We introduce, in Section 3.2, some of the recent progress and current topics in thermodynamics, focusing on quantum thermal machines and resource theories, as both play an important role in clarifying the concept of work in small systems. We also introduce non-equilibrium

work equalities, which provide a means of linking work to information in our experiment. In Section 3.3, we briefly discuss the interface between optics and thermodynamics, presenting the properties of thermal light. Section 3.4 describes our experimental setup, which models a demon with light as the working medium. Section 3.5 present the details and results of our experimental implementation. In Section 3.6, we present theoretical results stemming from our analysis of the experiment, including a to-our-knowledge new non-equilibrium work–information relation. We use this relation to derive a bound linking work extraction and information in our setup, which we test with experimental data. We summarise and discuss the implications of our work in Section 3.7.

3.1 Maxwell's demon – linking thermodynamics and information theory

3.1.1 A brief history

The four laws of thermodynamics define the theory's fundamental physical quantities (energy, entropy and temperature) and impose limits on their dynamics. Maxwell's demon is a scenario that challenges the second law of thermodynamics, a statement about irreversibly in thermodynamic processes. Likely the best way to fathom the meaning of the second law is to gain familiarity with its many faces. Its formulations range from technical to formal: "The efficiency of a quasi-static or reversible Carnot cycle depends only on the temperatures of the two heat reservoirs, and is the same, whatever the working substance. A Carnot engine operated in this way is the most efficient possible heat engine using those two temperatures." stated by Carnot (1796-1832), or "In every neighborhood of any state S on an adiabatically enclosed system there are states inaccessible from S." stated by Caratheódory (1873-1950). Formulations by Clausius (1822-1888) and Kelvin (1824-1907) are the most fitting for understanding the challenge that Maxwell's demon lays out for the second law: "Heat can never pass from a colder to a warmer body without some other change, connected therewith, occurring at the same time" and "It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects". Finally, the second law of thermodynamics can be seen as a statement that thermodynamic systems evolve toward

macroscopic states characterised by higher entropy, or disorder.

James Clerk Maxwell (1831-1879) made an important contribution to the kinetic theory of gases in 1860, deriving the probability distribution of the velocities of particles in an ideal gas, now known as the Maxwell-Boltzmann distribution. While his result provides an underpinning for the laws of thermodynamics, Maxwell noted that it only does so in a statistical sense: "The 2nd law of thermodynamics has the same degree of truth as the statement that if you throw a tumblerful of water into the sea, you cannot get the same tumblerful of water out again".

In a letter from 1867, Maxwell introduced a thought experiment, questioning whether the second law could be hypothetically violated. William Thomson (Lord Kelvin) would later call this "Maxwell's demon". Maxwell describes the scenario in his thermodynamics textbook "Theory of Heat" [155]. The thought experiment is centered around "a being whose faculties are so sharpened that he can follow every molecule in its course, [...] whose attributes are as essentially finite as our own". The being (or demon) sits in a vessel full of air at uniform temperature, divided by a wall in two parts, A and B. Molecules of air in the vessel are moving with varying velocities. The demon can manipulate an opening in the wall, and makes sure that only the faster molecules in A are allowed to pass to B and only the slower molecules in B are allowed to pass to A. This selection process requires an arbitrarily small amount of work, since the movement of the door in the wall can be made arbitrarily slow, yet as time passes, part B of the vessel becomes hotter and part A colder. Furthermore, the wall can then be used as a piston and work can be extracted from the system initially at equilibrium, at the same time cooling it. The implication that Maxwell proposed is that the second law is true when dealing with large systems, but can be violated by controlling individual molecules.

Leo Szilard (1898-1964) made important progress towards reconciling Maxwell's demon with the second law of thermodynamics. He had several major contributions [156], the first being to replace Maxwell's intelligent being with a mechanical device. The second was to introduce the Szilard engine, which greatly simplifies the thought experiment, reducing the gas to a single molecule and the demon's measurement to a binary one, as illustrated in Figure 3.1. Szilard's model consists of a box containing a single particle, initially with an unknown position. The demon inserts a wall in the middle of the box and observes

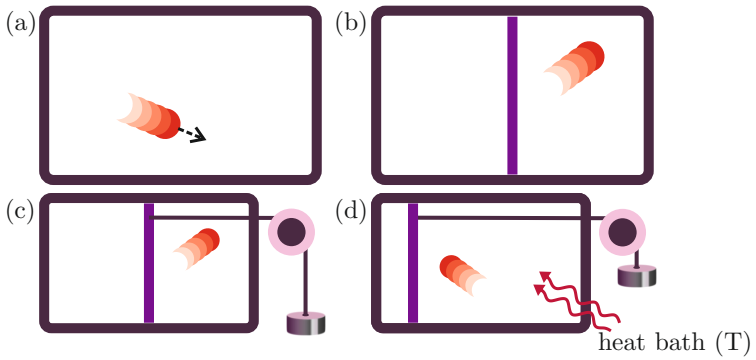


Figure 3.1: Szilard's engine. (a) A single molecule is enclosed in a box. (b) A demon introduces a wall separating the box in two halves. It then measures which side of the wall the molecule is on. (c) The demon sets up a work extraction mechanism according to the measurement outcome. (d) The demon extracts work by moving the wall quasistatically such as to increase the volume occupied by the molecule, while keeping it in equilibrium with a heat bath at temperature T. $k_B T \log 2$ work is extracted.

the particle's position, either to the left or to the right of the wall, thus reducing the volume it occupies. Assuming quasistatic evolution and that the box is in contact with a thermal bath at temperature T, the particle has on average $k_B T$ kinetic energy per degree of freedom. Szilard calculated that moving the wall in these conditions, such as to expand the volume occupied by the particle back to the volume of the initial box produces an average of $k_B T \log 2$ mechanical work, thus quantifying the result of the demon's operation. Another contribution by Szilard was to suggest that the demon's measurement and storage of its result cannot come for free. He proposed that the measurement must dissipate an amount of heat higher than $k_B T \log 2$, keeping the second law intact. Brillouin (1889-1969) suggested that the measurement converts "negentropy" (a term he introduced to signify negative entropy, a measure of non-equilibrium) into information [157], which is then converted back to negentropy through work extraction.

The work of Rolf Landauer (1927-1999) later clarified the connection to information theory. Aiming to derive physical limits for the speed and power consumption of computers, Landauer argued that the logical irreversibility of computer operations (like logical gates with two inputs and a single output) implies thermodynamic irreversibility [158]. He quantified the thermodynamic costs of irreversible logical operations, obtaining what is known as "Landauer's principle", which states that in order to erase a bit of information in an environment at temperature T, $k_B T \log 2$ of heat must be dissipated. Landauer's principle is presently widely accepted, with one caveat: it has been shown that information eras-

ure can take place in the presence of a bath corresponding to a conserved quantity other than energy, such as angular momentum[159]. Therefore, while the generation of entropy amounting to $k_B \log 2$ is unavoidable, energy expenditure can be replaced with expenditure of angular momentum.

Charles Bennett made the connection between Landauer's principle and Maxwell's demon[160,161], noting that the information referring to the particle position in a given cycle of Szilard's engine must occupy a memory bit. Erasure of this bit of information must generate enough entropy as to balance the reduction of entropy produced by measurement. Furthermore, the work extracted through each cycle of the engine, $k_B T \log 2$ is exactly the amount required for erasing (reseting) the memory, which is necessary if the engine is to be truly cyclic. This observation leads to a reconciliation of Maxwell's demon with the second law of thermodynamics.

Bennett also showed that measurements can be constructed in a thermodynamically reversible fashion, implying that it is not the measurement process itself that must generate entropy, like Szilard suggested, but really the necessary resetting of the Demon's memory. To understand the operation of Maxwell's demon, we can consider the dynamics of Shannon entropy throughout an operating cycle, as follows. Shannon and thermodynamic entropy are often used interchangeably, with $k_B \log 2$ as a scaling factor between the two. The demon and the working medium start out uncorrelated, with their initial states characterized by entropies $H_0(\mathrm{D})$ and $H_0(\mathrm{S})$, respectively. A measurement takes place when an interaction between the demon and the working medium creates correlations between the two systems. If the interaction is reversible, the total entropy, equal to $H_0(\mathrm{D}) + H_0(\mathrm{S})$, will be conserved. However, the entropy of each system will change, such that, after the measurement, we have $H(\mathrm{D}) + H(\mathrm{S}) = H_0(\mathrm{D}) + H_0(\mathrm{S}) + I(\mathrm{D};\mathrm{S})$. $I(\mathrm{D};\mathrm{S})$ is mutual information as defined in **??** and must be positive. Let us assume that the measurement is a reversible operation changing the state of the demon controlled on the state of the working medium, leaving the latter unchanged. This models a measurement in Szilard's engine and, with the demon's states degenerate, can be performed with no energy consumption. In this case, $H(S) = H_0(S)$ and we then get $H(\mathrm{D}) = H_0(\mathrm{D}) + I(\mathrm{D};\mathrm{S})$. An operation on the system controlled on the demon's state (we will call this feedback) then leaves the demon's state unaltered, $H(D)' = H(\mathrm{D}) = H_0(\mathrm{D}) + I(\mathrm{D};\mathrm{S})$. Making this operation is reversible, we can have

$H(\mathrm{S})' = H_0(\mathrm{S}) - I(\mathrm{D};\mathrm{S})$, with the two systems uncorrelated again in the end. The demon, therefore, uses mutual information created by measurement to decrease the entropy of the system it is acting upon, while its own entropy increases by the same amount. Hence, Maxwell's demon can be thought of as a device that moves entropy from a working medium into a memory, without an associated heat transfer. The low entropy working medium can then be used to extract work.

Maxwell's demon made a brief new attempt at evading the second law of thermodynamics in 1990. Wojciech Zurek had proposed a demon that compresses measurement information, and he identified algorithmic information as a measure of the memory space required to store measurement outcomes[162]. At William G. Unruh's suggestion, Carlton Caves analysed a scenario in which the demon monitors a large number of Szilard engines simultaneously and only records a measurement in the rare event in which the microstate of the system can be efficiently represented in its memory e.g. when every particle is on the left side of its box[163]. Then the demon can extract work while only storing a small amount of information. The second law survived once again[164], since the decision of whether to store a measurement result or not occupies memory, which must be reset, using up the work extracted by "Unruh demons".

3.1.2 Maxwell's demon and quantum information

More than a century after its debut, Maxwell's demon plays an important role in illuminating the relation between information and thermodynamics. Landauer's principle can be deduced from the second law of thermodynamics by applying it to the demon scenario (note however that the result has also been derived for physical models of memories, under classical and quantum evolution[165,166]). In particular, insights from thermodynamics have guided progress in the theory of quantum information. John von Neumann (1903-1957) used a thermodynamic consideration when discussing the entropy of a quantum state[21]. Asher Peres (1934-2005) used similar arguments in discussing the distinguishability of quantum states[167,168]. Important information theoretic results, such as Holevo's bound on the classical capacity of quantum channels[169] can be derived by invoking Landauer's cost of erasure[170,171]. Landauer's principle has also been used to derive measures of quantum and classical correlations by quantifying the minimum work cost that must be payed for des-

troying them[172]. The analysis of work extraction scenarios involving correlated Szilard engines has provided further insights into quantum correlations, yielding an operational interpretation of quantum discord[173] and a witness of entanglement[174].

3.1.3 Physical realisations of Maxwell's Demon

In 1912, Marian von Smoluchowski (1872-1917) discussed a machine similar to Maxwell's demon in effect, yet simpler in construction[175]. Instead of a demon performing measurements and feedback, he considered a trap door between the two sides of a box, acting like a valve. The trap door is held closed by a weak elastic mechanism and it can be opened by receiving a kick from a molecule coming from one side but not from the other. This type of device is a Brownian ratchet, and has been popularised by Richard Feynman (1918-1988)[176]. Its effect is not in contradiction with the second law of thermodynamics, as thermalisation with the working system introduces fluctuations in the trap door, hindering it from working as a valve unless it is cooled to a lower temperature. In a broad sense, the Brownian ratchet is not too far removed from practical issues. Consider, for example, a windmill designed to pivot such that it always faces the wind which blows from a randomly changing direction. The pivoting movement is naturally dampened by friction. It is easy to overlook the importance of this effect, but without it, the windmill would end up rotating randomly, so in effect, the dampening can be seen as analogous to the cooling required by Brownian ratchets.

When it comes to measuring and controlling microscopic degrees of freedom, the realisation of Maxwell's demon type experiments has had to wait until recently to become possible. Measurement and feedback techniques are likely to find practical applications not as work extraction mechanisms, but as methods for confining systems to a small volume of phase space, which is desired in many experimental situations. A molecular "information ratchet" has been shown to achieve this[145], driving a molecular ring to occupy with high probability one of two sites on a molecular axis. This effect was achieved by conditionally lowering and raising a barrier between the two sides, using a photochemical effect. A similar idea was demonstrated as a general method for increasing the density of cold atomic or molecular gases[146,147], confining magnetically trapped particles into an optical trap with a smaller volume. The method is reminiscent of Maxwell's demon, as it relies on a laser beam that

"measures" when a particle enters the optical trap and changes its state, such that it becomes insensitive to magnetic fields. This has been coined single photon cooling, as a single photon is scattered for each trapped particle.

Experiments with colloidal particles manipulated using optical tweezers have been an appropriate testbed for protocols related to information in thermodynamics like Szilard's engine and Landauer's erasure[148–150]. In these experiments, work is inferred by measuring particle trajectories in known external fields. Pekola et al. realised a single electron Szilard engine[151] using an electron box, a device that can be fabricated using small metal islands connected by tunneling junctions. A single electron charge can be detected in this setting using a single electron transistor based on the Coulomb blockade effect, while the potential energy of electrons can be changed by applying a control voltage. The same setup was used to validate a non-equilibrium work-information relation by Sagawa and Ueda (Equation 3.3)[152]. An autonomous demon working as an information powered refrigerator was realised by replacing the measurement device with a second electron box[153]. Huard et. al. realised a Szilard engine in a superconducting microwave cavity coupled to a superconducting qubit[177]. In their experiment, a microwave pulse is used to detect the state of the qubit. When it is in the excited state, another microwave pulse stimulates the emission of a microwave photon. The emission of this extra photon was confirmed experimentally.

These works have focused on demonstrating an effect in close analogy with Szilard's single particle engine. Our work described in Section 3.4 offers a slightly different perspective. The photonic demon uses tools of quantum optics to prepare light in a non-equilibrium state starting from equilibrium. Work extracted due to this non-equilibrium is stored in a passive element (a capacitor) which can be used as an intermediary between the small system and a practical work reservoir (a battery). We provide new insights into the relation between information and thermodynamics, even though our setup does not aim to extract the optimum amount of work per bit of information.

3.2 Recent progress in thermodynamics

In this section, we review some of the recent progress in thermodynamics, focusing on quantum thermal machines, the resource theory of thermodynamics and fluctuation theorems. Throughout this review, we discuss some of the models that have been proven effective

for thermodynamic concepts in small systems, the aim being to motivate the choices that we will make when interpreting our experiment in Section 3.6.

In recent years, the advancement of experimental techniques has enabled us to control physical systems down to the level of single particles. This progress has prompted significant interest in the analysis of thermodynamics in microscopic systems and into the quantum regime. Single molecule experiments[178], made possible by technologies like optical tweezers and atomic force microscopes, can probe biological processes at a scale where large thermal fluctuations are the rule[179]. A multitude of small devices driven by fluctuations have been recently developed, such as molecular motors[180], information ratchets[181] and nanowalkers[148]. The theoretical framework necessary for extracting useful information in these noisy conditions is provided by fluctuation theorems and non-equilibrium work relations[182–186]. In even smaller systems, quantum effects such as coherence and entanglement can play an important role. Transport of electrical charge or energy in molecular-size systems can be affected by quantum interference[187–189] and there is evidence of quantum effects playing a relevant role in biology[190]. Quantum thermodynamics aims to describe the fundamental constraints that govern these processes and to clarify the emergence of macroscopic thermodynamics from quantum mechanics. This is of great importance for the implementation of quantum information processing tasks, which require quantum systems to be purified and maintained in error–free subspaces by dissipative processes.

3.2.1 Quantum thermal machines

Quantum thermal machines are analogues of traditional heat engines and refrigerators, with quantum systems as the working medium. Reviewing this topic will aid in the interpretation of our experiment, introduced in Section 3.4, as we highlight results referring to the type of interactions that allow work extraction and to fluctuations in the work extraction process.

For a quantum system, the internal energy is $U = \left\langle \hat{H} \right\rangle = \sum_n p_n E_n$ where $\{p_n\}$ are the occupation probabilities of the Hamiltonian eigenstates and $\{E_n\}$ the corresponding eigenenergies. Its infinitesimal change is $dU = \sum_n dp_n E_n + p_n dE_n$. This can be interpreted as an expression of the first law of thermodynamics: the first term, which is due to the redistribution of occupation probabilities, corresponds to heat transfer to or from the system while the second term, due to shifts of energy levels, corresponds to work performed on the system.

Starting with this interpretation, we can define protocols analogous to traditional thermodynamic transformations, including thermalisation of the working medium by interaction with a heat bath, and expansion/compression implemented by time dependent Hamiltonians. Following this approach, quantum mechanical analogues of Carnot's ideal heat engine as well as other idealised thermodynamic cycles have been analysed [191]. Although these thermal machines, known as quantum reciprocating engines/refrigerators behave mostly in the same way as their classical counterparts, some dynamic effects of quantum nature have been identified, such as the enhancement of work extraction due to the preparation of a working substance with coherence between energy levels [192], the oscillation of occupation probabilities during thermalisation [193–195] or other coherent effects [196].

The time dependent, sequential approach to quantum thermal machines assumes a high degree of classical control, exerted on the working medium through zero-entropy baths, such as intense laser fields. An alternative model that is closely related to systems of practical interest is that of continuous quantum thermal machines [197], including autonomous systems [198]. Geusic, Schultz-DuBois and Scoville pointed out in 1959 [199] that the 3-level maser (the microwave analogue and predecessor of lasers) is an example of a quantum thermal machine and showed that the efficiency of masers is bounded by that of the Carnot cycle (actually the hydrogen maser works in a similar way to Maxwell's demon, as the population inversion is realised by a magnetic measurement).

We illustrate an idealised 3-level thermal machine in Figure 3.2. This model of great conceptual simplicity serves as prototype for many important processes that stand at the foundation of current experimental techniques, such as lasing, optical pumping and laser cooling. The 3 or more energy levels of the model act as frequency filters, selectively coupling to heat bath modes with a fixed energy splitting. A heat bath is a large quantum system at thermal equilibrium, with a continuum of energy levels. It can be a bosonic bath, consisting of photons or phonons, or a fermionic bath composed of electrons. The interaction driving the heat machines must be non-linear, coupling three systems with different energy splittings. Linear operations, generated by Hamiltonians quadratic in the creation and annihilation operators, as exemplified in Section 1.2.1, can only lead to heat transfer from a higher temperature to a lower one and are thus insufficient for realising non-trivial heat machines [13]. The work extraction dynamics are essentially captured by the Hamilto-

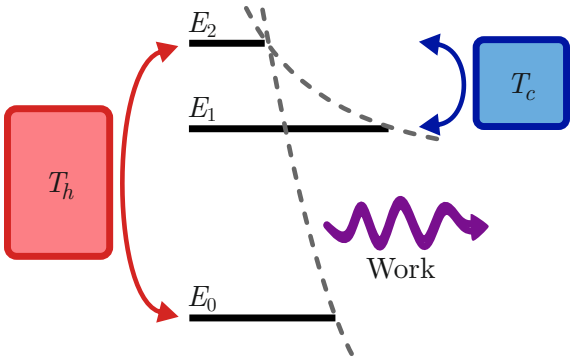


Figure 3.2: Continuous quantum heat engine: a 3-level system is in contact with two thermal baths. E_i denotes the energy of level i and $\{T_c, T_h\}$ are the bath temperatures. The lengths of the energy levels represent their respective populations when the machine is in thermal equilibrium with the heat baths (obeying the canonical probability distribution). For $T_h/T_c > (E_2 - E_0)/(E_2 - E_1)$, there is a population inversion between levels 0 and 1 and a field stimulating the transition from E_1 to E_0 will extract work from the system. The efficiency of converting heat from the bath at T_h into work is $(E_1 - E_0)/(E_2 - E_0)$ which in the limiting case is $1 - T_c/T_h$, the same as for the Carnot cycle. If $T_h/T_c > (E_2 - E_0)/(E_2 - E_1)$ the machine functions as a refrigerator, cooling the cold bath at T_c.

nian $\hat{H}_I = \hbar\epsilon(\hat{a}^\dagger\hat{b}^\dagger\hat{c} + \hat{a}\hat{b}\hat{c}^\dagger)$, where $\hat{a}^\dagger$ creates an excitation in one of the sub-systems that acts as frequency filter (corresponding to one of the possible transitions in Figure 3.2) and $\hat{b}^\dagger$ and $\hat{c}^\dagger$ belong, analogously, to the other two subsystems. This Hamiltonian commutes with the non-interacting one $\hat{H}_0 = \hbar\omega_a\hat{a}^\dagger\hat{a} + \hbar\omega_b\hat{b}^\dagger\hat{b} + \hbar\omega_c\hat{c}^\dagger\hat{c}$ if $\hbar\omega_a + \hbar\omega_b = \hbar\omega_c$ (as depicted in Figure 3.2 where $\hbar\omega_a = E_1 - E_0$, $\hbar\omega_b = E_2 - E_1$ and $\hbar\omega_c = E_2 - E_0$). The work extraction dynamics, therefore essentially consists of energy-conserving transitions between three systems.

In the framework of open quantum systems, a Markovian master equation in the Lindblad form [200] can be derived for the evolution of quantum thermal machines [197,198], treating the thermal baths implicitly. This approach has been employed for several example systems to calculate the trade-off between the power and efficiency of work extraction/cooling. Moreover, a study of the cooling rate in a bath as it approaches absolute 0 temperature has provided a quantitative, dynamical underpinning for the third law of thermodynamics [201].

The simplest thermal machines are composed of two qubits in contact with two different thermal baths. We denote the ground and excited states of the qubits by $|g_i\rangle$ and $|e_i\rangle$ respectively ($i \in \{1, 2\}$). Then the two levels $|g_1\rangle\,|e_2\rangle$ and $|e_1\rangle\,|g_2\rangle$ form a "virtual qubit" [202] which will tend to equilibrate at a temperature that depends on the bath temperatures and qubit energy splittings. The energy splittings can be chosen such as to make this

virtual temperature take any value, positive or negative. Autonomous thermal machines act by putting an external system in contact with a virtual qubit. In this framework, work extraction corresponds to creating a population inversion in the external system, which acts as a work reservoir. If the energy of the work reservoir has an upper bound, it will tend to equilibrate with the virtual qubit at negative temperature. For any temperature different from -0, the final energy of the reservoir will not be sharply defined. Autonomous quantum thermal machines thus introduce a new measure for classifying work extraction processes, apart from efficiency and power: the "temperature of work" [202]. For unbounded reservoirs, this corresponds to the amount of fluctuations produced per average energy increase of the reservoir. We use this concept (and the explicit model of a work reservoir) in Sections 3.5 and 3.6, to quantify the performance of our photonic Maxwell's demon.

3.2.2 Thermodynamics as a resource theory

The concept of work in finite systems can be grasped through models like the ones outlined in the previous subsection. The more general concept of thermal non-equilibrium is best captured by a recently formulated resource theory that provides a general framework for thermodynamic transformations. This framework, which we introduce in this subsection, has been particularly successful in elucidating the role of quantum coherence in thermal processes.

A resource theory is constructed around a restricted class of "free" operations. Specifying this class divides the state space of the theory into a set of free states which can be prepared using the free operations and a set of resource states, which cannot be prepared by these operations alone. The foremost example is the theory of entanglement [203,204], which is constructed by allowing free local operations and classical communication. Other examples of resource theories deal with purity [205] and asymmetry [206]. It turns out that the main questions of thermodynamics can be addressed rigorously within a hybrid resource theory, combining purity and asymmetry [207].

In the resource theory of thermal non-equilibrium, the free transformations are all quantum operations $\mathcal{E} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ of the form $\mathcal{E}(\hat{\rho}) = \mathrm{Tr}_{\mathcal{A}}\left[\hat{V}_{\mathcal{SA}}(\hat{\rho} \otimes \hat{\gamma})\hat{V}_{\mathcal{SA}}^{\dagger}\right]$. Here, $\hat{\gamma}$ is the thermal state of an ancillary system with Hamiltonian $\hat{H}_{\mathcal{A}}$ and inverse temperature β, $\hat{\gamma} = e^{-\beta\hat{H}_{\mathcal{A}}}/\mathrm{Tr}\left(e^{-\beta\hat{H}_{\mathcal{A}}}\right)$ and $\hat{V}_{\mathcal{SA}}$ is a unitary transformation that commutes with

the free Hamiltonian, $\left[\hat{V}_{\mathcal{SA}}, \hat{H}_{\mathcal{S}} \otimes \hat{I} + \hat{I} \otimes \hat{H}_A\right] = 0$. The latter condition ensures that the closed evolution of the system and ancilla is strictly energy conserving. Free use of thermal ancillary systems (in the state $\hat{\gamma}$) corresponds to free access to a single heat bath. It is a generalization of the free preparation of maximally mixed states in the resource theory of purity. The requirement of strict energy conservation ensures that all energy flows are well accounted for. As it is equivalent to enforcing symmetry with respect to time translations, this is related to the resource theory of asymmetry. We note that a unitary transformation that preserves energy only on average can always be constructed using an ancillary system and a strictly energy conserving global unitary. Such an operation will introduce fluctuations in the energy of the ancillary system, which must be accounted for in light of the discussion in Section 3.1. This is why strict energy conservation must be enforced.

Using only thermal transformations, the rate of asymptotic conversion between non-equilibrium states is given solely by their free energy. The state $\hat{\rho}^{\otimes n}$ with $n \to \infty$ can be reversibly converted to $\hat{\sigma}^{\otimes m}$ using thermal operations if $m/n = S\left(\hat{\rho}|\hat{\gamma}\right)/S\left(\hat{\sigma}|\hat{\gamma}\right)$[208]. $S\left(\hat{\rho}|\hat{\gamma}\right)$ is the von Neumann relative entropy, defined in Section 1.2.3. The conversion rate can be written in terms of free energies, $F(\hat{\rho}) = \mathrm{Tr}\left(\hat{H}_s\hat{\rho}\right) - \beta^{-1}S(\hat{\rho})$, using the equality $S\left(\hat{\rho}|\hat{\gamma}\right) = \beta\left[F(\hat{\rho}) - F(\hat{\gamma})\right]$. Extraction of work W in this context is defined as the conversion of a resource state into the fiducial state $\hat{\sigma}_W$, which has sharp energy distribution: $\mathrm{Tr}\left(\hat{H}_S\hat{\sigma}_W\right) = W$, $S(\hat{\sigma}_W) = 0$, and the free energy $F(\hat{\sigma}_W) = W$. Inserting this into the relation for the conversion rate and setting $n = m$, we get that the work extractable from $\hat{\rho}$ in the asymptotic regime is $W = F(\hat{\rho})$. In the non-asymptotic regime, when only one copy of the system is available, the same relation holds true only on average, $\langle W\rangle = F(\hat{\rho})$ and only for states that commute with the system Hamiltonian[209]. In this regime, states with coherence in the energy eigenbasis possess "locked" work, which can only be extracted in the presence of a reference frame.

In the single shot regime, conversion rates are given in terms of generalised free energies defined by replacing von Neumann entropy with single shot entropies. When a sharp amount of work is required, the relevant quantities are smooth entropies, and transformations become irreversible[210]. In the presence of a catalyst (a non-thermal ancillary system that is left unchanged by thermal operations), the conversion of resource states is governed by an infinity of free energies[211], $F_\alpha(\hat{\rho}) = \beta^{-1}\left(S_\alpha(\hat{\rho}|\hat{\gamma}) + F(\hat{\gamma})\right)$, $\alpha \geq 0$ where S_α are Rényi

relative entropies. These are not sufficient for describing allowed thermal transformations for states exhibiting coherence in the Hamiltonian basis. There is a second set of divergences that govern the transformation of these states[212]: $A_\alpha(\rho) = S_\alpha\left(\hat{\rho}|\mathcal{D}_H(\hat{\rho})\right)$ where $\mathcal{D}_H$ eliminates coherences in the Hamiltonian eigenbasis. All F_α and A_α must decrease during a thermal transformation.

Models in which reference systems are included explicitly have been studied recently[213–215], aiming to elucidate how microscopic work is turned into its macroscopic counterpart. In our experiment described in Section 3.4, we produce a macroscopic effect by a microscopic measurement, using bright thermal states of light. The measurement converts an equilibrium (thermal) state into a non-equilibrium one. As we have seen, any non-equilibrium state is a resource in the theory of thermodynamic transformations.

In Section 3.6 we model work extraction assuming a strictly energy conserving global interaction between the system, environment and work reservoir. Our assumption there relies on physical arguments, but is also in accordance with the resource theory formulation of thermodynamics.

3.2.3 Non-equilibrium work relations

In Section 3.6 we derive a non-equilibrium work relation that includes the effects of measurement and feedback, which allows us to quantify the performance of our experimental setup. We introduce this type of relation in this subsection.

The resource theoretic formulation of thermodynamics reveals the importance of free energy, which turns out to quantify the resource of thermal non-equilibrium. For canonical states, the expression of free energy at inverse temperature β given Hamiltonian operator $\hat{H}$ is $F_{\hat{H}}(\hat{\gamma}_{\hat{H}}) = \mathrm{Tr}\left(\hat{H}\hat{\gamma}_{\hat{H}}\right) - \beta^{-1}S(\hat{\gamma}_{\hat{H}}) = -\beta^{-1}\log\left[\mathrm{Tr}(e^{-\beta\hat{H}})\right]$, which we will denote equilibrium free energy. When a system kept in contact with a heat bath undergoes a quasistatic transformation such that it is always in the canonical state, the work extracted in the time interval $t = 0$ to τ is $W = -\int_0^\tau dt\, \partial_t \mathrm{Tr}\left[\hat{\gamma}_{\hat{H}}\hat{H}\right]$. Using this expression and that of the equilibrium free energy, and denoting the Hamiltonian at times 0 and τ by $\hat{H}_0$ and $\hat{H}_\tau$ respectively, yields $W = F_{\hat{H}_0}(\gamma_{\hat{H}_0}) - F_{\hat{H}_\tau}(\gamma_{\hat{H}_\tau}) \equiv -\Delta F$. The minus sign appears here because we are writing all equations in terms of extracted work, instead of the work cost. This choice ensures consistency within the current Chapter, although it is not the usual convention in

the literature.

In finite time transformations, the Hamiltonian is changed at a rate too high for the system to follow in the corresponding equilibrium state. In this case, the extracted work is always lower that the consumed free energy, $W < -\Delta F$. The offset $-\Delta F - W$ is called dissipated work, signifying free energy that is converted to heat as the system equilibrates with the thermal bath, generating entropy. Several relations go beyond inequalities in this regime, maintaining their validity arbitrarily far from equilibrium. One example is entropy production fluctuation theorems, which can be derived from statistical mechanics with different ranges of applicability[184]. Their general form is $P(+\chi)/P(-\chi) = e^{\chi\tau}$ where $P(\chi)$ is the probability distribution of the irreversible entropy production χ, meaning that the probability of negative entropy production decreases exponentially with the protocol time τ.

Another example is the non-equilibrium work equality introduced by Jarzynski in 1997[185]. This equality relates equilibrium free energy change to the work statistics in non-equilibrium processes:

$$\langle e^{\beta W} \rangle = e^{-\beta \Delta F}. \tag{3.1}$$

The assumption is that the working system starts in thermal equilibrium. However the control parameters can be changed arbitrarily fast. For a closed system, the derivation of the equality relies solely on the properties of Hamiltonian dynamics, in particular the conservation of phase space density (Liouville's theorem), which ensures that the initial thermal distribution of microstates can be extended to the trajectories that the system takes throughout the transformation. The extension to an open process in the presence of a heat bath can be achieved by simply interpreting part of the initial system as the heat bath. The closed system derivation then suffices, with the assumption that the interaction Hamiltonian between system and bath is negligibly small, such that correlations between the two can be ignored in the global canonical state. The equality can also be derived for stochastic evolution[216]. A different derivation can be obtained starting from a fluctuation theorem by Crooks[217]. In the limit of instantaneous Hamiltonian switching, Equation 3.1 had been used before as a standard method in Monte-Carlo simulations of free energy differences[218]. Using Jensen's inequality for the exponential function, we have $\langle e^x \rangle \geq e^{\langle x \rangle}$ and applying this to Equation 3.1 we obtain $W < -\Delta F$, as expected. For a Gaussian work

distribution, we obtain $-W-\Delta F = \beta \,\mathrm{Var}(W)/2$, which is a fluctuation-dissipation theorem. The left side of this equation is dissipated work. On the right side, $\mathrm{Var}(W)$ quantifies work fluctuations.

Sagawa and Ueda extended Jarzynski's equality to include preparation by measurement and feedback [219], in effect an intervention from Maxwell's demon. As discussed in Section 3.1, the resource in this scenario is the average of the mutual information between system and demon, generated by measurement. This is defined point-wise as

$$I = \log\left[p(m|i)\right] - \log\left[m\right] \tag{3.2}$$

where i denotes the state of the working system and m the measurement outcome. The equality by Sagawa and Ueda is

$$\langle e^{\beta W-(I-\beta\Delta F)}\rangle = 1, \tag{3.3}$$

which puts mutual information on an equal footing with free energy as a resource for work extraction.

The Jarzynski and Sagawa-Ueda relations can be straightforwardly extended to closed quantum systems, where unitarity plays the role of Liouville's theorem [220], with the caveat that a measurement of energy must be assumed at the end of a protocol (and implicitly at the beginning) [221,222]. In open quantum systems, analogues of the Jarzynski relation can be derived [223], but cannot be straightforwardly interpreted in terms of work.

Non-equilibrium work relations and fluctuation theorems can be tested in systems with few degrees of freedom and have been proven accurate in laboratory experiments [148,224–226].

3.3 Light and thermal equilibrium

The previous sections have emphasized the importance of thermal equilibrium. Here, we discuss this concept in the context of optical fields and introduce some of the properties of thermal light.

3.3.1 Blackbody radiation

Blackbody radiation is the name given to a set of electromagnetic modes that are in equilibrium with a heat bath. As with any physical system, the equilibrium state of the electromagnetic field is $e^{-\beta\hat{H}}/\operatorname{Tr}\left[e^{-\beta\hat{H}}\right]$. The Hamiltonian is $\hat{H} = \sum_{\mathbf{k},s} \hbar\omega(\mathbf{k})\hat{n}_{\mathbf{k},s}$ where $\omega(\mathbf{k}) = c\,|\mathbf{k}|$ is the angular frequency and $\hat{n}_{\mathbf{k},s}$ is the photon number operator belonging to the mode with wave-vector $\mathbf{k}$ and polarisation s. The average number of excitations in a mode at thermal equilibrium is $\langle n_{\mathbf{k},s}\rangle = 1/(e^{\beta\hbar\omega(\mathbf{k})} - 1)$ and the corresponding photon number distribution is

$$p(n_{\mathbf{k},s}) = \frac{1}{(1 + \langle n_{\mathbf{k},s}\rangle)(1 + 1/\langle n_{\mathbf{k},s}\rangle)^{n_{\mathbf{k},s}}}. \tag{3.4}$$

To calculate the spectrum of blackbody radiation, we need the density of states of the electromagnetic field. The density of modes in phase space [7] is equal to $(2\pi)^{-3}$ in the $\mathbf{k}$-vector parametrisation. Using this, we can calculate the density of states for modes with angular frequency in the vicinity of $\omega = c\,k$: $\phi(\omega) = 2(2\pi)^{-3}4\pi k^2 dk/d\omega = \omega^2/(\pi^2 c^3)$ (a factor of 2 is due to the polarisation degree of freedom). We then get the energy density as a function of angular frequency,

$$u(\omega) = \phi(\omega)\frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} = \frac{\hbar\omega^3}{\pi^2 c^3\left(e^{\beta\hbar\omega} - 1\right)}. \tag{3.5}$$

This is the blackbody radiation spectrum first found by Max Plank (1858-1947) in 1900. Satyendra Bose (1894-1974) re-derived this formula in 1924 using the particle picture of light. He assumed that photons are indistinguishable for the purpose of counting microstates [227], which we did here implicitly when defining the Hamiltonian. In 1905, Albert Einstein (1879-1955) had made the journey in reverse, starting from the blackbody radiation spectrum and using arguments from thermodynamics to arrive at the particle nature of light [1]. Equation 3.4 is known as the Bose-Einstein distribution, and it describes the equilibrium state for all indistinguishable particles with integer spin.

3.3.2 Thermal light

Many sources at thermal equilibrium emit light that is not blackbody radiation, as it is directional, polarised or with a different spectrum compared to the one given in Equation 3.5. In general, light emitted by these sources is equivalent to filtered blackbody radiation

and is known as thermal light. Specifically, if only a narrow spectrum is emitted like in the case of atomic gas lamps, only the modes within a certain bandwidth are in equilibrium with a heat bath and the rest are unpopulated. The photon number distribution of each mode of thermal light is still of the form given by Equation 3.4. Using the coherent state representation introduced in Section 1.2, the state of a thermal light mode can also be written as

$$\gamma_{\text{phot}} = \frac{1}{\pi\langle n\rangle}\int d\alpha\, e^{-|\alpha|^2/\langle n\rangle}\,|\alpha\rangle\,\langle\alpha| \tag{3.6}$$

with $\langle n\rangle = \text{Tr}\,(\gamma_{\text{phot}}\hat{n})$. This is a mixture of classical states with Gaussian distributed amplitudes. The probability distribution for the intensity $I = |\alpha|^2$ is then

$$\mathcal{P}(I) = \langle I\rangle^{-1} e^{-I/\langle I\rangle}. \tag{3.7}$$

The intensity of thermal modes fluctuates following this distribution, with a coherence time given by the inverse bandwidth of the signal. These fluctuations lead to intensity correlations, which are quantified by the second order auto-correlation function (introduced in Section 1.2), $g^{(2)}(\mathbf{r1}, t, \mathbf{r2}, t+\tau) = \langle I(\mathbf{r1}, t) I(\mathbf{r2}, t+\tau)\rangle / \left(\langle I(\mathbf{r1}, t)\rangle\langle I(\mathbf{r2}, t+\tau)\rangle\right)$. Using Equation 3.7 we can compute $g^{(2)}(\mathbf{r}, t, \mathbf{r}, t) = \langle I(\mathbf{r}, t)\rangle^2 / \langle I(\mathbf{r}, t)\rangle^2 = 2$. Thermal intensity correlations were measured in the laboratory in 1958 in a famous experiment by Hanbury Brown and Twiss[228]. In 1956 they were used to determine the spatial coherence properties of light emitted by a star[229] from which the star's angular size could be inferred. The correlations measured by Hanbury Brown and Twiss were of the form $\langle \Delta I(\mathbf{r1}, t)\Delta I(\mathbf{r2}, t+\tau)\rangle / \left(\langle \Delta I(\mathbf{r1}, t)\rangle\langle \Delta I(\mathbf{r2}, t+\tau)\rangle\right)$ where $\Delta I(\mathbf{r}, t) = I(\mathbf{r}, t) - \langle I(\mathbf{r}, t)\rangle$ is the deviation of the intensity at position $\mathbf{r}$ from its mean value.

Let us now consider the average number of photons per coherence volume in thermal light. For blackbody radiation, integrating the photon number density over the full spectrum yields $\int_0^\infty d\omega\, u(\omega)/(\hbar\omega) = 2\zeta(3)/\left(\pi^2(\beta hc)^3\right)$ where $\zeta(3) \approx 1.2$. The coherence length of blackbody radiation is of the order $\beta\hbar c$[7], making the coherence volume $\sim (\beta\hbar c)^3$. Thus, it turns out that the number of photons per coherence volume in blackbody radiation is always of order unity. For a thermal source that emits only at low frequency compared to the frequency of maximum intensity in blackbody radiation, the energy in a coherence volume is of the order kT. For practical thermal lamps, reaching temperatures on the order

of 3000 K and operating at optical frequencies, these conditions do not hold. The number of photons per coherence volume in the thermal radiation emitted by atomic gas lamps is only $10^{-2} \sim 10^{-3}$.

The photonic Maxwell's demon setup described in Section 3.4 requires intense thermal fluctuations that can be resolved using linear silicon photodiodes. A thermal atomic gas lamp would not be bright enough to achieve this. However, Equation 3.6 suggests that thermal light can be simulated by a random process that produces a classical field with amplitude sampled from a Gaussian distribution. A widely used method of achieving precisely this [230–232], which we adopt in our experiment, is to collect light from a cone of laser speckle [233]. Depicted in Figure 3.3 (a), this is the interference pattern created by a diffusing surface placed in a laser beam. As illustrated in Figure 3.3 (b) the interference pattern is obtained by adding the effects of secondary sources, with random phases introduced by the diffusing surface. The addition of these amplitudes resembles a two-dimensional random walk and therefore the total amplitude is sampled from a Gaussian distribution. Thermal radiation is simulated by collecting light around a point in the speckle pattern and moving the diffuser that creates it.

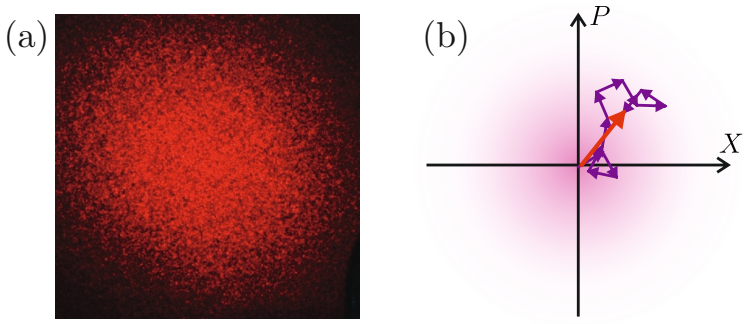


Figure 3.3: (a) A picture of a laser speckle pattern (image from the public domain). (b) Phase space diagram showing secondary sources with random phases (purple) summed up to create the total amplitude (red) in a point of the speckle pattern. Like in the case of a random walk, the total amplitude is distributed according to a Gaussian probability density (light purple).

An important property of the laser speckle pattern is its correlation length, the length over which the intensity stays relatively constant. This must be larger than the collection aperture to ensure that the collected signal has the same statistics as single mode thermal states. If this condition is not satisfied, the signal intensity is averaged over many coherence

areas, yielding the intensity distribution of multi-mode thermal light,

$$\mathcal{P}_m(I) = \frac{m\,(m\,I/\langle I\rangle)^{m-1}}{\langle I\rangle(m-1)!}e^{-mI/\langle I\rangle}, \tag{3.8}$$

for m modes. The corresponding second order auto-correlation is $g^{(2)}_m(0) = 1 + 1/m$. When m is large, Equation 3.8 gives a sharp distribution and the statistics are Poissonian.

The correlation length of the speckle pattern is determined by the geometry of the setup. For clarity, we model this in one dimension. Let the optical field at the scattering surface be $A_s(z) = A_0 \exp\left[-z^2/(2w^2)\right] h(z)$ where z is the spatial coordinate orthogonal to propagation in our model, w is the waist of the laser beam and $h(z)$ is a random phase pattern. The angular distribution of the amplitude in the far field, $A_f(\theta)$ is given by the Fourier transform:

$$A_f(\theta) = \int dz A_s(z) e^{\mathrm{i}\,k\,\theta z} \equiv \mathcal{F}_k[A_s(z)] \tag{3.9}$$

where k is the amplitude of the wave vector characterising the laser light. Using the convolution theorem, $A_f(\theta) \sim \exp\left[-(kw\theta)^2\right] * \mathcal{F}_k[h(z)]$. The scattering surface is itself characterised by a correlation length l_s, which means that $\mathcal{F}_k[h(z)]$ is a random distribution of width $(kl_s)^{-1}$. For $l_s < w$, this angle gives the divergence of the scattered cone of light. We can identify the correlation length of the speckle pattern with $d(kw)^{-1}$ where d is the distance from the scattering surface. This shows that the correlation length is inversely proportional to the waist of the laser beam at the scattering surface.

3.4 A photonic Maxwell's demon

In the previous sections, we have introduced Maxwell's demon, its role in physics and its recent physical implementations. We have reviewed the current context in thermodynamics and introduced the concept of thermal equilibrium for light.

We now propose a physical scenario analogous to Maxwell's demon, where light is the working medium, as depicted in Figure 3.4. Instead of the two boxes occupied by gas molecules, we take two optical spatio-temporal modes initially in thermal equilibrium, prepared in thermal states with the same average energy. We prepare these modes using the random scattering method described in the previous section. We use laser pulses in order to achieve high average energy, and collect light using optical fibers. The two modes are thus defined

spatially by the optical fibers and temporally by the laser pulse shape.

Like in the scenario of Maxwell's demon, we implement a measurement that distinguishes the microstates of the equilibrium system, creating thermal non-equilibrium. To achieve this, each light mode propagates through a high transmittance beam splitter from which the reflected light is coupled to a highly sensitive avalanche photo-diode (APD). A way to picture this measurement is with the demon, equipped with a sensitive APD, "hiding" behind a highly reflective mirror that directs the thermal light.

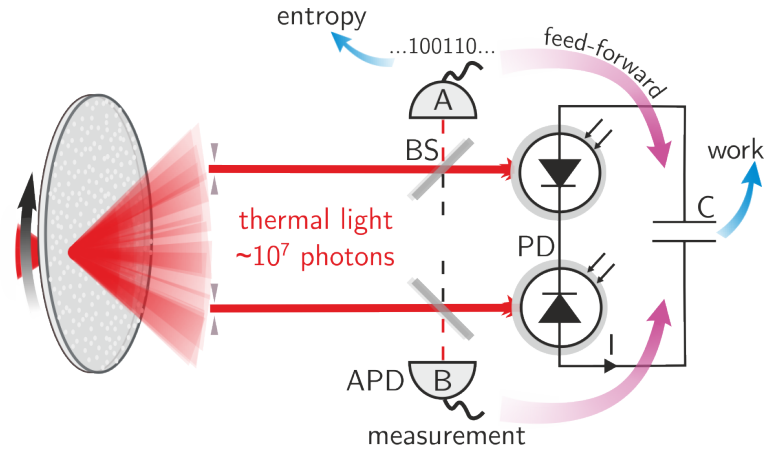


Figure 3.4: Photonic Maxwell's demon setup. A source of time-varying laser speckle simulates thermal light. Two modes of the same intensity ($\approx 4 \times 10^7$ photons/pulse) are selected using optical fibre. A microscopic measurement is implemented on each mode using beam splitters (BS) with high transmittance and single-photon sensitive avalanche photodiodes (APD). Two linear photodiodes (PD) form an electromotive source that charges a capacitor (C). Non-zero average voltage across C can be obtained by swapping its polarity according to the APD measurement outcomes, alowing work extraction. We label the two sides A and B.

The measurement described has a binary output for each mode, as an APD simply heralds the detection of one or more photons by producing a "click", or otherwise remains in its default state. We will treat the measurement as pertaining to the demon, without explicitly considering the energy cost of running the detectors. Knowledge of a measurement outcome updates the state of the corresponding light mode. In general for this measurement and any input light, when the detector clicks, the expected average number of photons is updated to a higher value if the input state is super-Poissonian ($g^{(2)}(0) > 1$, e.g. thermal light) and to a lower value if it is sub-Poissonian (see Section 1.2). As an example of how this works for super-Poissonian statistics, consider the extreme case in which an incident pulse contains either a large number of photons or no photons, with equal probabilities. In this example, detection of even a single photon from the incident pulse informs that the

transmitted mode has high energy. Therefore, the expected number of photons transmitted conditioned by a detection event is two times the expected number of photons in the input mode (minus the few photons that triggered the detector).

The limit of small photon detection probability is a relevant regime for the demon's measurement. In this limit, the operation conditioned by a click approximates "single photon subtraction", which has been extensively used for preparing non-classical, non-Gaussian states of light [234,235]. In this limit, for thermal light at the input, the average number of photons conditional on a click is double that in the initial state, an effect that has been observed experimentally [232]. In our setup we do not restrict the APD click probabilities. Changing these probabilities by choosing how much light gets sent to the APDs provides a way of tuning the effect of the measurement as well as the mutual information generated between light and measurement apparatus (or demon).

We have introduced a measurement that produces conditional states that are out of thermal equilibrium. The average state however, observed if measurement outcomes are ignored, is no different from the input. The average effect of the measurement amounts to negligibly small loss introduced by the high transmittance beam splitters. In order to create the non-equilibrium resource on average, an operation controlled on the state of the detectors is required. For example, the APD outputs can be fed-forward to a beam splitter with controllable reflectance on which the two light modes are incident. The beam splitter can be changed from totally reflective to totally transmitting, swapping the two incident modes in the event that detector A clicks and detector B does not. In this way, an asymmetric energy distribution can be created, with higher energy on average on side B than on side A.

For work extraction we let the light fall on two photodiodes, connected with opposing polarities such that on average they produce zero voltage when the two light modes are balanced. This photodiode circuit charges a capacitor according to the fluctuating energy difference between the two light modes. If an asymmetry is produced by forward-feeding the APD measurement outcomes, the capacitor has a non-zero average voltage. As discussed in Section 3.2.1, for true work extraction we must accumulate energy in a work reservoir. To achieve this, a battery can be connected to the capacitor for each pulse and if there is a non-zero average voltage across the capacitor, the battery can be charged. We propose this

setup for its conceptual and experimental simplicity, aiming to observe what effects may be achieved by measurement and feedback and without aiming to realise the optimal work extraction strategy. We use a capacitor as intermediary energy storage, as this makes the experimental setup symmetric with respect to the exchange of sides A and B. This allows us to measure all the quantities of interest (we simply measure the voltage across the capacitor instead of connecting a battery) without actually performing any fast feed-forward.

3.5 Experimental implementation

We now present the details and results of an experimental implementation of the scenario described above. We observe the effect of the single-photon-level measurement on bright thermal light and show that, accompanied by feedback, this measurement allows work extraction. The results are interpreted in the context of information thermodynamics in Section 3.6, where we present a bound on the work distribution in terms of mutual information and verify it using experimental data.

We make a series of simplifications for our proof-of-principle implementation. Firstly, the demon's measurement can be realised with imperfect photo-detectors. In fact, the effect of the measurement does not depend on detection efficiency, as we discuss in detail in Appendix B. This means that we may implement the measurement beam splitters with reflectance small enough to have a negligible effect on the transmitted light, and further use variable attenuators to fix the photon flux sent towards the two APDs, at the level of few photons per pulse. Secondly, the symmetry of the setup depicted in Figure 3.4 ensures that swapping the two light modes has an effect indistinguishable from switching the polarity of capacitor C. We actually replace feed-forward altogether with a logical operation on the measured voltage across C, switching sign as a function of the measured APD outputs.

3.5.1 Experimental methods

We simulate thermal light using the method outlined in Section 3.3.2. In our experiment, laser speckle is produced by a pulsed laser beam scattered on a spinning glass diffuser and collected into multi-mode optical fibers. The average energy of the collected pulses is $\approx 10\,\mathrm{pJ}$ (in each fiber), high enough to allow intensity measurements with good signal-to-noise ratio using standard Silicon photodiodes (Thorlabs Det10A). According to the analysis in Section

3.3.2, the laser beam waist at the scattering surface must be small in order to make the correlation length of the speckle pattern large, producing single-mode thermal statistics in the collected modes. We focus the laser beam down to $\approx 15\,\mu m$ and set the distance between the diffuser and the collection plane to $d \approx 15\,cm$. With central wavelength $\lambda \approx 830\,nm$, the correlation length of the collected speckle is $\approx 1\,mm$, much larger than the $25\,\mu m$ core diameter of the optical fibers. For the glass diffuser we use a fine grit (1500), as we found by trial and error that this is optimal in terms of the source's intensity fluctuations. The speckle is produced by scattering a train of $\approx 4\,\mu J$ laser pulses from a Ti:Sapphire regenerative amplifier (Rega9000). We use a pulse stretcher to chirp the femtosecond pulses, in order to avoid nonlinear damage mechanisms that smooth out the surface of the glass diffuser.

Two light modes prepared by this source impinge on photodiodes connected through a capacitor, as depicted in Figure 3.4. We use $C \approx 2\,pF$, which is of the same order as a photodiode's junction capacitance. The voltage across the capacitor is monitored by an oscilloscope with $50\,\Omega$ impedance. We observe simultaneously this voltage and the output of the two APDs performing the demon's measurement. Light is sent towards the APDs by two anti-reflection coated glass windows, acting as beam splitters with 99.5% transmittance. The amount of light falling on the APDs is determined by variable absorbers, which we use to tune the rate at which they click. This rate takes values between 0 (the detector never clicks) and 1 (the detector always clicks).

Throughout this Chapter, the uncertainties of estimated quantities are calculated using the method outlined in Section 1.2.4.

3.5.2 The thermal light source

The simulated thermal light source was found to produce a pulse intensity distribution in very good agreement with that of a single-mode thermal state. We measured the intensity distribution using a linear photodiode, associating to each light pulse an amplitude by averaging the photodiode signal around the peak voltage, as illustrated in Figure 3.5 (a). A histogram of pulse amplitudes measured in mode A is shown in Figure 3.5 (b). The data follows closely the model of a multi-mode thermal state, with the intensity distribution given by Equation 3.8. This model has two parameters, m and $\langle I \rangle$, which we set by making the model's mean and second order auto-correlation $g^2(0)$ equal to those of the measured pulse

amplitudes.

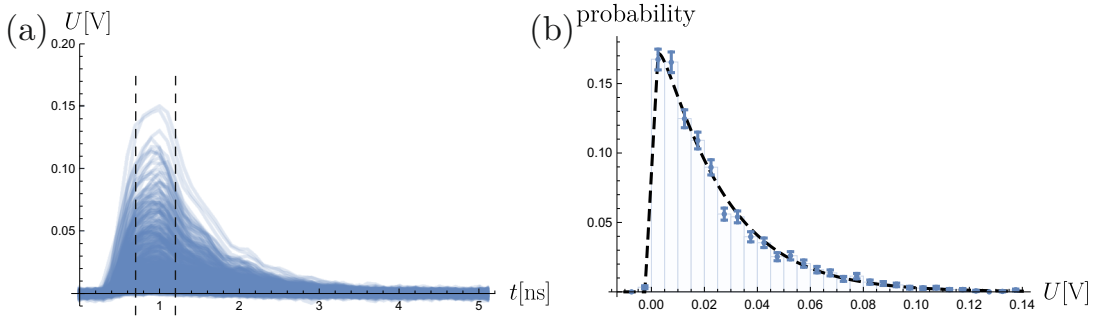


Figure 3.5: Intensity measurements characterising the thermal light source. (a) photodiode voltage response measured for 600 pulses of light. The plot is partially transparent, so that darker regions correspond to many overlapping traces. Dashed lines indicate the interval over which voltage is averaged in order to estimate each pulse amplitude, proportional to the light intensity. (b) Histogram of 3000 measured pulse amplitudes (blue points). Error bars correspond to counting statistics. The height of each solid bar indicates the frequency of events in the corresponding interval. The dashed line corresponds to multi-mode thermal statistics with the same mean and variance as the data. The model distribution was convolved with a top hat function of the same width as the histogram intervals, to make it directly comparable with the data points.

For the second order auto-correlation we use the estimate $g^{(2)} = \langle I^2\rangle/\langle I\rangle^2 = \left(\langle U^2\rangle - \Delta U_e^2\right)/\langle U\rangle^2$ where we denote by U the photodiode signal amplitude and the correction term ΔU_e^2 is the variance of the electronic noise, which artificially increases $\langle U^2\rangle$ without affecting $\langle U\rangle$. Recording the intensity of 3000 pulses for each source mode, we obtain the estimates $g_{\mathrm{A}}^{(2)} = 1.93(3)$ and $g_{\mathrm{B}}^{(2)} = 1.95(3)$, which indicate thermal mode numbers of 1.07(3) and 1.05(3) respectively.

3.5.3 Non-equilibrium preparation by measurement

The two thermal modes in our experiment are balanced so that the average voltage produced by the "work extraction" photodiodes is zero. In practice, this means that the average value of the voltage is much smaller than its standard deviation: $\langle U\rangle_0/\sqrt{\mathrm{Var}(U)_0} = 0.002(7)$. However, conditioned on an outcome of the demon's measurement, the average voltage can be significantly different from zero. We will use the displacement $\langle U\rangle/\sqrt{\mathrm{Var}(U)}$ to quantify the imbalance of conditional distributions. In the context of work extraction, this ratio is akin to the notion of "temperature of work" that was introduced in Section 3.2.1.

The demon's measurement is tuned by the amount of light sent towards APDs. In the case of very rare detection events, as discussed in Section 3.4, the average energy of a

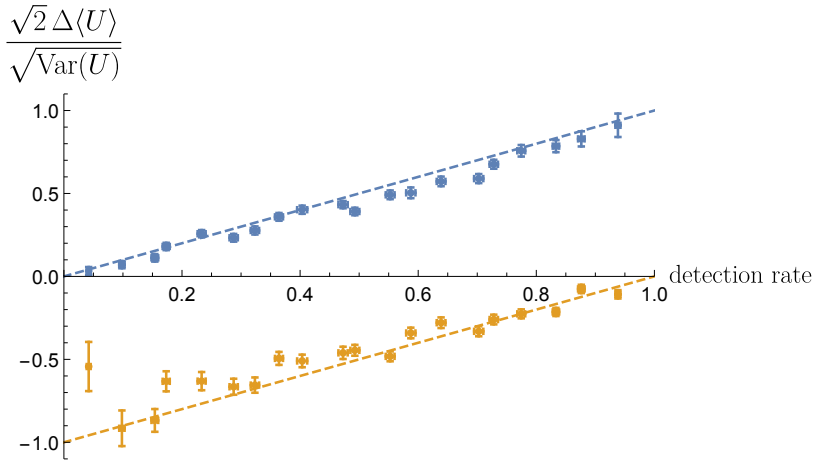


Figure 3.6: Relative displacement of the average voltage U produced by the "work extraction" photodiodes conditional on the output of APD A, as a function of the APD's detection rate. Orange points correspond to conditioning on click events and blue points to conditioning on trials when the APD does not click. The unconditioned average voltage is 0. The dashed lines plot the functions $1-x$ (orange) and x (blue).

transmitted pulse conditioned on a detection event is twice the average energy of the incident pulses. In contrast, when a detector clicks on every trial, there is no way of distinguishing pulses and therefore the conditional state must be the same as the initial one. The case of intermediate detection rates can also be easily calculated. For the sake of clarity, we only give the results here and present the calculation in Appendix B, where we also work out the case of multi-mode thermal states. For single mode thermal light, if the probability of a click is $p^{(1)}$, then the average energy conditioned on registering a click is different by a factor of $2-p^{(1)}$ from the initial average energy. For the state conditioned on the absence of a click, the change is by a factor of $1-p^{(1)}$. Remarkably, the effect of the measurement depends solely on the outcome probabilities. In Figure 3.6 we depict the displacement of the voltage distribution across C, conditioned on the outcomes registered at APD A. We scan the photon detection rates between 0 and 1, showing how the conditional states depend on the probabilities of measurement outcomes.

We note that the voltage produced by the photodiodes turns out to be linear in the energy difference between the two impinging light modes, but this is only true within the linear regime of the photodiodes and for the small value of C. We choose to plot $\sqrt{2}\langle U\rangle/\sqrt{\mathrm{Var}(U)}$ in Figure 3.6 with the extra factor $\sqrt{2}$ because the standard deviation of the energy difference between modes A and B is $\sqrt{2}$ times the mean energy in one mode, so we are really

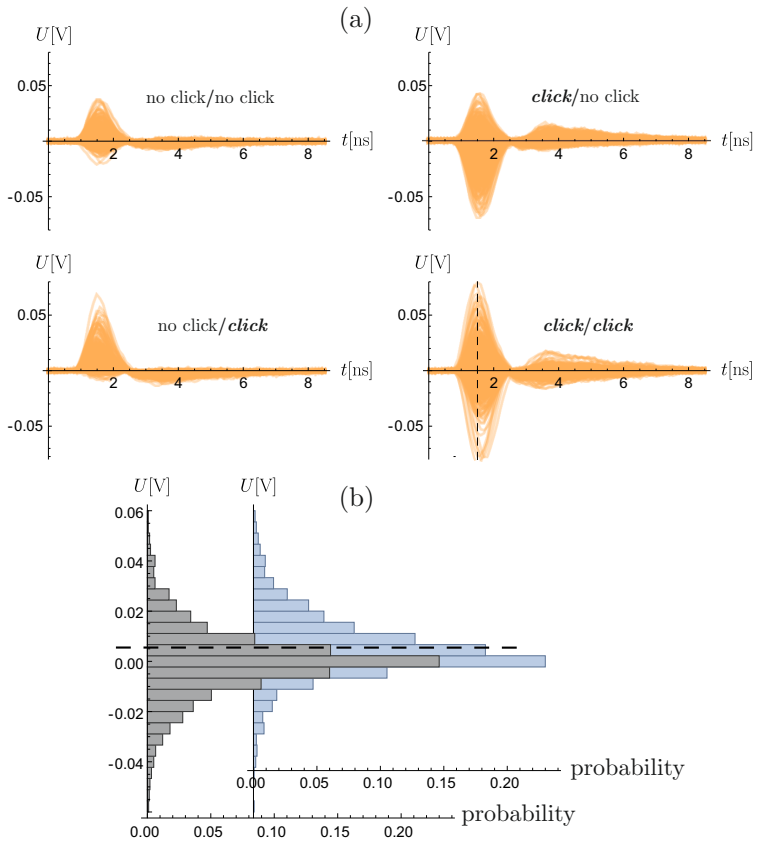


Figure 3.7: Observing the average effect of measurement and feedback. (a) Photodiode traces (4000 in total) sorted according to the demon's measurement outcomes. The dashed vertical line in the bottom-right quadrant indicates the time at which the maximum voltage is sampled. (b) Histogram of the maximum voltage. In gray, on top, the case when no feedback is applied, when the average deviation from 0 is very small. In blue, below and displaced to the right, the case when the sign of the voltage is flipped if the measurement outcome is "click/no click". The dashed horizontal line shows the average value for the second distribution, significantly displaced from 0. The detection rates for the two APDs were $p_A^{(1)} = 0.702(8)$ clicks/pulse and $p_B^{(1)} = 0.311(8)$ clicks/pulse.

normalising by the average voltage that would be created by one mode.

To convert the conditional effect depicted in Figure 3.7 into an average one, we must implement feedback. Here, we simulate this by setting the sign of the voltage across C according to the APD outputs. The efficiency of this measurement and feedback operation depends on the probabilities of the demon's measurement outcomes. One might expect that the greatest average imbalance can be created when the detection probability is 1/2, corresponding to the highest entropy (information content) of the measurement register. However, an optimisation of the average energy difference between the two modes (performed not experimentally but using the model detailed in Appendix B) yields an optimum for

$p_{\mathrm{A}}^{(1)} = 2/3$ and $p_{\mathrm{B}}^{(1)} = 1/3$. Figure 3.7 shows results for this setting.

3.6 Information thermodynamics interpretation

In the previous Sections we have presented the concept and experimental implementation of a photonic Maxwell's demon scenario, where work extraction is allowed by a measurement and feedback operation. The advantage of using light as the working medium is that we can easily implement a tunable microscopic measurement on thermal states that are intense enough to have an observable effect on regular room temperature electronics. While it is clear that the non-equilibrium resource is created by measurement and feedback, it is also easy to see that we could simply use the large temperature difference between simulated thermal light and electronics (photodiodes) to convert the energy of light pulses into work. Our work-extraction mechanism uses symmetry to make this non-equilibrium resource inaccessible, balancing the effect of two thermal light modes on the photodiodes. We note that the feedback that we envisage (swapping the two sides of the setup) is an energy conserving unitary operation.

Due to the existence of different temperatures in our setup, the scenario we have presented does not immediately fit within models of work extraction that assume contact with a single heat bath. Therefore we cannot apply previously derived bounds to our scenario. Furthermore, with our choice of parameters the work that can be extracted for each pulse is much smaller than the average energy of the light pulses, which gives the order of Landauer's limit for work extraction. Still, we are able to show that there is a clear connection between the information acquired by measurement and the work distribution achieved, in the form of a bound on the latter. Here we derive this bound and test it with experimental data.

3.6.1 A non-equilibrium work-information equality

The work-information theorem by Sagawa and Ueda (Equation 3.3) was derived within a model assuming interaction with a single thermal bath. As noted above, this is not the case for our system of interest. We now present a relation similar to that of Sagawa and Ueda that can be applied to our scenario despite this fact. The new work-information relation makes use of work extraction measurements performed when the demon's actions are turned off. We denote this case, when no measurement or feedback are preformed and

no work can be extracted by "0" and the case that includes measurement and feedback by "m+f". Modeling these two situations, we can derive the following result:

$$\langle e^{\beta W - I} \rangle_{\mathrm{m+f}} = \langle e^{\beta W} \rangle_0 \tag{3.10}$$

where I is the point-wise mutual information defined in Section 3.2.3 (Equation 3.2). The derivation of this theorem requires that the feedback operation be an energy conserving unitary transformation acting solely on the equilibrium system that provides the inverse temperature β (the two thermal light modes) and that a non-disturbance condition applies to the measurement, as we detail below.

We now introduce a general model that encompasses the photonic Maxwell's demon. We divide the model into three parts: (i) The first system $\mathcal{S}$ starts in thermal equilibrium, with inverse temperature β. This represents the working medium, the two thermal light modes in our setup. The initial and final energy eigenstates that this system occupies are labeled $|s\rangle$ and $|S\rangle$ respectively, and the Hamiltonian is $\hat{H}_{\mathcal{S}}$. Measurement and feedback operate only on this system, with feedback described by a unitary transformation with no energy cost. The measurement outcome is labeled m with the corresponding set of measurement operators $\hat{M}_j^{(m)}$. The index j allows for the possibility of several measurement operators associated with a given measurement outcome (e.g. several possible physical processes are associated with an APD click). The measurement operators are normalized according to $\sum_{m,j} \hat{M}_j^{(m)\dagger} \hat{M}_j^{(m)} = \mathbb{1}$. We denote the feed-forward operation $\hat{U}_m$. (ii) The work reservoir/ battery $\mathcal{B}$ occupies initial and final energy levels $|b\rangle$ and $|B\rangle$ respectively with energies E_b and E_B; its Hamiltonian is $\hat{H}_{\mathcal{B}}$. We define the extracted work as the energy increase of this system: $W = E_B - E_b$. (iii) All ancillary systems and the environment, we denote as the rest, $\mathcal{R}$. This system occupies initial and final energy levels r and R respectively, and has Hamiltonian $\hat{H}_{\mathcal{R}}$. We also find it convenient to label the final joint energy of $\mathcal{S}$ and $\mathcal{R}$ as $E_{SR} \equiv E_S + E_R$. We denote the set of random variables of interest $\xi \equiv \{s, b, r, m, SR\}$.

We assume that owing to decoherence the total initial state is diagonal in the product basis of the free Hamiltonians so that we can treat the initial energies of the three systems as classical variables. We also assume that the work reservoir undergoes decoherence in its energy eigenbasis at the end, so that its final energy is well defined. It is a priori not necessarily the case that the sum of the free-Hamiltonian energies of the systems is

conserved. However we can model light-matter interaction using the Jaynes-Cummings interaction Hamiltonian (with the rotating-wave approximation) plus decoherence in the energy eigenbasis, in which case we do have energy conservation. Furthermore, no energy is required in principle to connect or disconnect the capacitor from the battery. We shall accordingly assume that $E_s+E_b+E_r = E_S+E_B+E_R$. This implies that $W = E_s+E_r-E_{SR}$. The evolution of the entire closed system is given by the unitary operator $\hat{V}$.

Using these definitions, the left hand side of Equation 3.10 is

$$\langle e^{\beta W-I}\rangle_{\mathrm{m+f}} = \sum_{\xi} p(s,b,r,m,SR)\, e^{\beta W-I} = \sum_{\xi} p(m,SR|s,b,r)p(s,b,r)\, e^{\beta W-I}, \tag{3.11}$$

using Bayes' theorem in the second step. The probability distribution of the initial system energies is $p(s,b,r) = \frac{1}{Z}e^{-\beta E_s}p(b,r)$. Introducing the expressions for W and I we get

$$\begin{aligned}\langle e^{\beta W-I}\rangle_{\mathrm{m+f}} &= \sum_{\xi} p(m,SR|s,b,r)\frac{1}{Z}e^{-\beta E_s}p(b,r)\, e^{\beta(E_s+E_r-E_{SR})}\frac{p(m)}{p(m|s)} \\ &= \sum_{\xi} p(m,SR|s,b,r)\frac{p(m)}{p(m|s)}\frac{1}{Z}p(b,r)\, e^{\beta(E_r-E_{SR})}.\end{aligned} \tag{3.12}$$

Isolating the sum over s, we have

$$\begin{aligned}&\sum_s p(m,SR|s,b,r)/p(m|s) = \\ &\mathrm{Tr}\left[|SR\rangle\langle SR|\,\hat{V}\left(\hat{U}_m\left(\sum_s \frac{\sum_j \hat{M}_j^{(m)}|s\rangle\langle s|\,\hat{M}_j^{(m)\dagger}}{p(m|s)}\right)\hat{U}_m^\dagger \otimes |b\,r\rangle\langle b\,r|\right)\hat{V}^\dagger\right].\end{aligned} \tag{3.13}$$

The sum over s is over all normalized states produced by the demon's measurement given a sharp energy input. If the measurement is non-disturbing, it leaves the state $|s\rangle\langle s|$ unchanged and the sum reduces to the identity. We show in Appendix B that for our measurement, although it is not non-disturbing, we have

$$\sum_s \frac{\sum_j \hat{M}_j^{(m)}|s\rangle\langle s|\,\hat{M}_j^{(m)\dagger}}{p(m|s)} = \hat{I}. \tag{3.14}$$

Using this and $\hat{U}_m\hat{U}_m^\dagger = \hat{I}$, we get

$$\sum_s p(m, SR|s, b, r)/p(m|s) = \mathrm{Tr}\Big[|SR\rangle\langle SR|\ \hat{V}\Big(\hat{I}\otimes|b\,r\rangle\langle b\,r|\Big)\hat{V}^\dagger\Big] = \sum_s p_0(SR|s, b, r) \tag{3.15}$$

where p_0 denotes the probability distribution corresponding to the evolution of the systems when the measurement and conditional operations are not implemented. Plugging this into the expression for the left-hand side of Equation 3.10 we get

$$\langle e^{\beta W - I}\rangle_{\mathrm{m+f}} = \sum_\xi p_0(SR|s, b, r)p(s, b, r)\, e^{\beta W} = \langle e^{\beta W}\rangle_0. \tag{3.16}$$

We note that Equation 3.10, which we have derived here, is really an information-theoretic equality, applied to an energy transfer scenario. This fact exposes the information-theoretic underpinning of thermodynamics [236].

3.6.2 Bound relating work extraction to information

We can derive a bound on work extraction using Equation 3.10. Applying the concave log function to the two sides of the equation and using Jensen's inequality, we get

$$\beta\,\langle W\rangle_{\mathrm{m+f}} - \langle I\rangle < \log\Big\langle e^{\beta W}\Big\rangle_0. \tag{3.17}$$

Let us rewrite the argument of the log function $\big\langle e^{\beta W}\big\rangle_0 = e^{\beta\langle W\rangle_0}\Big\langle e^{\beta\left(W-\langle W\rangle_0\right)}\Big\rangle_0$. We then get

$$\langle W\rangle_{\mathrm{m+f}} - \langle W\rangle_0 < \beta^{-1}\langle I\rangle + \beta^{-1}\,\log\Big\langle e^{\beta\left(W-\langle W\rangle_0\right)}\Big\rangle_0. \tag{3.18}$$

Expanding up to second order in β we have

$$\log\Big\langle e^{\beta\left(W-\langle W\rangle_0\right)}\Big\rangle_0 = \beta^2/2\,\mathrm{Var}(W)_0 + \mathcal{O}(\beta^3). \tag{3.19}$$

The second order expansion is exact in the case of a Gaussian work distribution. Taking advantage of this intuition, we define the following quantity:

$$\sigma(\beta, W) \equiv \sqrt{2\,\beta^{-2}\log\Big\langle e^{\beta\left(W-\langle W\rangle_0\right)}\Big\rangle_0}. \tag{3.20}$$

For a Gaussian distribution of W, as well as up to first order in β for any distribution, $\sigma(\beta, W)$ is equal to the standard deviation of W. Inequality 3.18 can be rewritten:

$$\langle W \rangle_{\mathrm{m+f}} - \langle W \rangle_0 < \beta^{-1} \langle I \rangle + \beta \left[\sigma(\beta, W)\right]^2 / 2. \tag{3.21}$$

Let us now consider the work that can be extracted in N runs of our experiment. Let U_k denote the voltage created across the capacitor C in the k^{th} iteration. Work extraction occurs when energy stored in the capacitor during each iteration is accumulated in a work reservoir. Connecting the capacitor to a battery with voltage U_0 transfers the charge $C(U_k - U_0)$ to the battery, increasing its energy by $CU_0(U_k - U_0)$ (we assume that the voltage U_0 is constant). The capacitor then equilibrates to voltage 0 again. Over N cycles the work extracted is $W^{(N)} = CU_0 \sum_k^N (U_k - U_0)$ and the mutual information gain is $I^{(N)} \equiv \sum_k^N I_k$. We define the sum of voltages across the capacitor over N runs $U^{(N)} \equiv \sum_k^N U_k$. Let us also define $\xi \equiv \beta C U_0$. Replacing the expression for work in Equation 3.10 and canceling the constant factor $e^{-\beta C N U_0^2}$, we can apply Jensen's inequality in the same way as above, obtaining

$$\left\langle U^{(N)} \right\rangle_{\mathrm{m+f}} - \left\langle U^{(N)} \right\rangle_0 < \xi^{-1} \langle I^{(N)} \rangle + \xi \left[\sigma(\xi, U^{(N)})\right]^2 / 2. \tag{3.22}$$

Using only the fact that U_k are independently and identically distributed, we have the following equality:

$$\left\langle e^{x\, U^{(N)}} \right\rangle_0 = \left\langle e^{x\, U} \right\rangle_0^n \tag{3.23}$$

which is a general property of the moment generating function. We can then show using Definition 3.20 that $[\sigma(\xi, U^{(N)})]^2 = N[\sigma(\xi, U)]^2$. We also have $\left\langle U^{(N)} \right\rangle = N \langle U \rangle$ and $\left\langle I^{(N)} \right\rangle = N \langle I \rangle$. Equation 3.22 then becomes

$$\langle U \rangle_{\mathrm{m+f}} - \langle U \rangle_0 < \frac{1}{\xi} \langle I \rangle + \xi \left[\sigma(\xi, U)\right]^2 / 2. \tag{3.24}$$

Since this is valid for any value of U_0, we can minimize the right-hand term of the inequality with respect to the variable ξ, in order to obtain the tightest bound:

$$\langle U \rangle_{\mathrm{m+f}} - \langle U \rangle_0 < \min_\xi \left\{ \frac{1}{\xi} \langle I \rangle + \xi \left[\sigma(\xi, U)\right]^2 / 2 \right\}. \tag{3.25}$$

Note that for a Gaussian work distribution of U we have $[\sigma(\xi, U)] = \sqrt{\mathrm{Var}(U)_0}$ and $\min_\xi \left\{ \frac{1}{\xi}\langle I\rangle + \xi \left[\sigma(\xi, U)\right]^2/2 \right\} = \sqrt{2I}\sqrt{\mathrm{Var}(U)_0}$ which yields

$$\frac{|\langle U\rangle_{\mathrm{m+f}} - \langle U\rangle_0|}{\sqrt{\mathrm{Var}(U)_0}} < \sqrt{2\langle I\rangle} \tag{3.26}$$

and, using the definition of work in our setup

$$\frac{|\langle W\rangle_{\mathrm{m+f}} - \langle W\rangle_0|}{\sqrt{\mathrm{Var}(W)_0}} < \sqrt{2\langle I\rangle}. \tag{3.27}$$

The bound given by Inequality 3.26 is an approximation of that given by Inequality 3.25, which works best for small values of $\langle I\rangle$. The left hand term of Inequality 3.26 can be interpreted as a measure of distinguishability between work distributions with and without measurement and feedback. This expression relates the work produced by the setup to its fluctuations and can be linked to the concept of temperature of work, introduced in Section 3.2.1. We have shown that this quantity is limited by information acquired by measurement. Our relation approaches the regime of Landauer's bound [158] when the work fluctuations are of the same order as the characteristic thermal energy of the equilibrium system on which measurement and feedback are applied. In the following section we show how experimentally measured quantities relate to the bound derived here.

3.6.3 Experimental verification of the bound

The extractable work being proportional to the voltage across capacitor C, here we focus on measuring this voltage in relation to APD outputs. We measure the peak voltage for each pulse, like in Figure 3.6. We relate the observed distribution to mutual information using Inequality 3.25. Note that the entropy of the measurement register provides an upper bound for the average mutual information $\langle I\rangle$. However we can calculate $\langle I\rangle$ using a model of the measurement, as detailed in Appendix B and we use this result in order to provide a tighter bound.

In order to maximise the observable effect in our setup, as discussed in Section 3.5.3, we set the detection rate of APD B to approximately 1/3. We then scan the detection rate corresponding to the measurement on mode A. Our results are reported in Figure 3.8. There are two optimum strategies, corresponding to a change of the measured voltage sign every

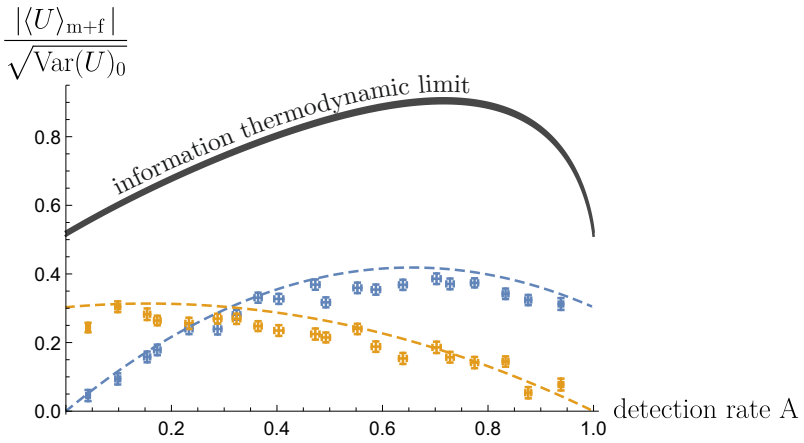


Figure 3.8: Extracted work bounded by information. We depict the displacement of the measured distribution of voltage U on capacitor C due to measurement and feedback. This is related to the distribution of extractable work: $|\langle U\rangle_{\mathrm{m+f}}|/\sqrt{\mathrm{Var}(U)_0} = (\langle W\rangle_{\mathrm{m+f}} - \langle W\rangle_0)/\sqrt{\mathrm{Var}(W)_0}$. The horizontal axis gives the detection rate for the demon's measurement on mode A, effectively tuning the effect of the measurement as well as the extracted information. The measurement on mode B yields 0.311(8) clicks/pulse. Blue and orange points correspond to two feedback strategies: changing the polarity of C when the measurement yields click/no-click (blue) or no-click/click (orange). The black line gives the bound established in terms of mutual information, Inequality 3.25, rescaled by the standard deviation of U. The width of the line represents the uncertainty with which the bound is determined using the measured voltage distribution (two times the standard deviation). Dashed lines give predictions based on the average photon number difference between the two light modes after feedback. The calculation of the photon number difference is detailed in Appendix B.

time APD A clicks and B does not, or B clicks and A does not, respectively. We find that for different measurement and feedback settings, the extracted work weighted by the work fluctuations is below the bound given in terms of mutual information, yet of the same order. As Jensen's inequality is only saturated for sharp voltage (work) distributions, it is to be expected that our bound is not saturated. We emphasise however that the bound demands that no work can be extracted when no information is acquired.

3.7 Summary and outlook

We have presented an experiment similar to Maxwell's demon, in which work extraction is ultimately bound by the amount of information acquired through measurement on a system in equilibrium. We derived a bound relating work extraction to mutual information characterizing the measurement setup by employing an information theoretic analysis, building on previous literature and considering less idealised levels of complexity, as dictated by the experimental arrangement. Differently from previous studies, the focus of our analysis is

not the absolute energy efficiency, but the ratio between the average extracted work and the fluctuations of the work distribution in the unperturbed setup, per each cycle. This measure is relevant in scenarios where fluctuations are important. The derivation presented here could be adapted to impose limits on the way in which the distribution of parameters other than work can be affected by measurement and feed-forward strategies.

The ability of addressing light in both the single-photon and intense-field regimes has been key in our experiment, and constitutes one of the main advantages of a photonic approach. Our demonstration has been carried out in the classical regime, exploiting the properties of second-order coherence. The setup can be adapted to include non-classical states and measurements, investigating the role of quantum coherence in thermodynamic processes. Furthermore, single particle measurement techniques similar to that presented here are pertinent to opto-mechanical [237] or spin-ensemble [238] systems, and can be used to investigate the effect of weak probing in the dynamics of open systems.

Chapter 4

Precise measurement of two-mode squeezing gain by seeding a parametric down-conversion source

The ability to prepare and manipulate non-classical states of light has enabled many important experimental achievements, such as the loophole-free violation of Bell's inequality [239,240] and quantum-enhanced interferometric measurements [53]. Most of these applications rely on spontaneous parametric down-conversion (SPDC) for squeezed light generation. In recent years, the refinement of dispersion engineering [40] and the development of wave-guided SPDC sources [241] have made possible the observation of high degrees of squeezing in well-defined spatio-temporal modes [242–244].

In this Chapter, we focus on the characterisation of pulsed SPDC sources in the high-gain regime. We propose a method for measuring squeezing gain that uses seeding and frequency-resolved intensity measurements and that does not require precise efficiency calibration. Our method is applicable to squeezing in multiple spectro-temporal modes, and can be a useful tool for analysing the intricate structure of broadband SPDC [245–247].

Using a seeding beam, we can determine the full broadband transformation describing an SPDC source by classical simulation of the spontaneous process [248–250]. In particular, for a two-mode squeezing process, the generation of light in the unseeded mode, known as difference frequency generation (DFG), can be used to simulate the joint-spectral distribu-

tion of photon-pairs produced by SPDC[250,251]. In this Chapter, we go one step further and consider the generation of light in the seeded mode. We argue that this signal has a dependence on the squeezing strength that differs from the DFG, and which can be used to provide a self-referenced measurement of the squeezing gain.

We begin the chapter with a brief look at the evolution of SPDC source engineering. In Section 4.1, we develop the theory of type-II broadband SPDC. We devote the remainder of 4.1 to a discussion of second-order perturbation, which, to our knowledge, has not received much attention to date. In Section 4.2, we present experimental methods that allow for the field generated in the same mode as a seeding beam to be observed. We also describe an application of these experimental methods towards the estimation of squeezing gain. In Section 4.3, we report an experimental implementation, presenting preliminary results and discussing the challenges that must be overcome in order to fully develop and apply our method. We summarise our work in Section 4.4.

4.1 Broadband parametric down-conversion in optical wave-guides

4.1.1 Development of broadband parametric down-conversion sources

David Klyshko (1929-2000) provided a theoretical description of SPDC (which he called parametric fluorescence) in the early 1970s[252]. Klyshko also proposed the use of stimulated and spontaneous PDC signals for self-referenced intensity measurements[253], a technique that we refine in this chapter. In 1970, Burnham and Weinberg realized the first experimental implementation of SPDC[254]. Later on, Leonard Mandel (1927-2001) and others developed applications related to coherence[7]. Over the years, SPDC sources based on $\chi^{(2)}$ and $\chi^{(3)}$ nonlinear crystals have become an established method for generating squeezed light[255,256], heralded single-photons[257–259] and polarisation-entangled photon-pairs[260–262].

Spontaneous frequency generation is a highly inefficient process due to typical values for the nonlinearity of optical crystals. The squeezing gain of SPDC sources can, however, be enhanced, either by using optical cavities or by pumping with ultrafast pulses. Of these two options, pulsed operation has the advantage of allowing conditional state preparation based

on multiple photon detections. As an application, homodyne tomography of Fock states with up to 3 photons has been demonstrated using pulsed SPDC[135,263,264].

In 2008, Peter Mosley developed a single-photon source based on KDP. This was the first pulsed source ever engineered to generate squeezing in a single spatio-temporal mode, and we use it for singe-photon generation in the experiment described in Chapter 2. For the experimental implementation presented in this chapter, we use an SPDC source based on wave-guided periodically-poled potassium titanyl phosphate (ppKTP), which was developed by Andreas Eckstein and others[241]. We present the experimental details of this source in Section 4.3.1.

Wave-guides sources have the advantage of allowing down-converted light to be constrained to propagate in a single spatial mode, which eliminates correlations between the generated signals in the spatial degree of freedom. Furthermore, the large interaction lengths lead to greatly enhanced brightness compared to bulk crystals. These advantages enabled recent research to report the generation of two-mode squeezed light with up to 80 photons on average in a single spatio-temporal mode, using the ppKTP source we employ for the work presented in this chapter[242].

4.1.2 General mode transformations for type-II down-conversion in optical wave-guides

First order perturbation theory is commonly used in the analysis of broadband SPDC[40,265,266]. This is a suitable approach in the limit of low squeezing gain, leading to the correct spectral distribution of SPDC-generted photon-pairs for single-photon sources. However, as the squeezing gain increases, the effects of time-ordering become significant. Several authors analysed these effects, using numerical methods to calculate the broadband mode transformations for some example processes[245,247]. In this section, we present a derivation of the general mode transformation for type-II down-conversion, starting from first principles. We obtain expressions that are suitable for efficient numerical evaluation, as well as for deriving perturbative approximations. Both of these characteristics are useful for the work presented in this chapter.

From Maxwell's equations, we can show that for a medium with no free charges, we have

$$\nabla \times \nabla \times \boldsymbol{E}(\boldsymbol{r},t) + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E}(\boldsymbol{r},t) = -\frac{1}{\varepsilon_0 c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{P}(\boldsymbol{r},t). \tag{4.1}$$

For cases of practical interest, we can approximate $\nabla \cdot \boldsymbol{E} = 0$[14], although this is not, in general, implied by Gauss's law ($\nabla \cdot \boldsymbol{D} = 0$). Using this, we obtain the simpler wave-equation

$$\nabla^2 \boldsymbol{E}(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{\varepsilon_0 c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{P}(\boldsymbol{r},t). \tag{4.2}$$

We can write the electric field in terms of its frequency components: $\boldsymbol{E}(\boldsymbol{r},t) = \int_0^\infty d\omega' \boldsymbol{E}(\boldsymbol{r},\omega') e^{-\mathrm{i}\,\omega' t} + c.c.$, and perform an analogous spectral decomposition of the polarisation vector. Inserting these into Equation 4.2, multiplying both sides by $e^{\mathrm{i}\,\omega t}$, and integrating over t (after taking the time derivatives), we obtain the wave-equation in the frequency representation:

$$\nabla^2 \boldsymbol{E}(\boldsymbol{r},\omega) + \frac{\omega^2}{c^2}\boldsymbol{E}(\boldsymbol{r},\omega) = -\frac{\omega^2}{\varepsilon_0 c^2}\boldsymbol{P}(\boldsymbol{r},\omega). \tag{4.3}$$

The polarisation of the propagation medium has a linear term,

$$\boldsymbol{P}^{\mathrm{L}}(\boldsymbol{r},\omega) = \varepsilon_0 \boldsymbol{\chi}^{(1)}(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega), \tag{4.4}$$

and a nonlinear one,

$$\boldsymbol{P}^{\mathrm{NL}}(\boldsymbol{r},t) = \varepsilon_0 \left[\boldsymbol{\chi}^{(2)}(\boldsymbol{r}) : \boldsymbol{E}(\boldsymbol{r},t)\boldsymbol{E}(\boldsymbol{r},t) + \boldsymbol{\chi}^{(3)}(\boldsymbol{r}) \vdots \boldsymbol{E}(\boldsymbol{r},t)\boldsymbol{E}(\boldsymbol{r},t)\boldsymbol{E}(\boldsymbol{r},t) + \ldots\right], \tag{4.5}$$

where we approximate the nonlinear interaction to be instantaneous. In the frequency domain, the electric field products in Equation 4.5 become convolutions, as described in Section 1.2.2. We use this later, when we introduce the nonlinear polarisation for a $\chi^{(2)}$ interaction in the frequency representation.

We consider waves traveling along a crystal axis, with a well-defined polarisation, indicated by the unit vector $\boldsymbol{e}$. We can therefore define the refractive index $n(\boldsymbol{r},\omega) \equiv \sqrt{1 + \boldsymbol{e}\,\boldsymbol{\chi}^{(1)}(\boldsymbol{r},\omega)\,\boldsymbol{e}}$, which yields

$$\nabla^2 \boldsymbol{E}(\boldsymbol{r},\omega) + \frac{\omega^2 n(\boldsymbol{r},\omega)^2}{c^2}\boldsymbol{E}(\boldsymbol{r},\omega) = -\frac{\omega^2}{\varepsilon_0 c^2}\boldsymbol{P}^{\mathrm{NL}}(\boldsymbol{r},\omega). \tag{4.6}$$

In an optical wave-guide, a finite number of guided modes with well-defined transverse profiles are supported. Light propagation in this case can be described by a 1D wave-equation, with the bulk medium refractive index $n(\boldsymbol{r},\omega)$ replaced by an effective value, $\mu(\omega)$. This effective value is a property of the guided mode. The dependence on x and y is now implicit:

$$\frac{\partial^2}{\partial z^2}\boldsymbol{E}(z,\omega)+\frac{\omega^2\mu(\omega)^2}{c^2}\boldsymbol{E}(z,\omega)=-\frac{\omega^2}{\varepsilon_0 c^2}\boldsymbol{P}^{\mathrm{NL}}(z,\omega). \tag{4.7}$$

For this chapter, we consider forward-propagating pulses that interact through a parametric process on length scales that are much longer than the oscillation period of the fields. To make this explicit, we write the forward-propagating electric field as the product of a slowly varying envelope and a planar carrier-wave, $\boldsymbol{E}^+(z,\omega)=\boldsymbol{E}_0(z,\omega)e^{\mathrm{i}k_0 z}$. Here, the wave-vector k_0 corresponds to the carrier frequency ω_0. We then assume that $\left|\partial^2/\partial z^2\ \boldsymbol{E}_0(z,\omega)\right| \ll |k_0\ \partial/\partial z\ \boldsymbol{E}_0(z,\omega)|$, which means that the second-order derivative of $\boldsymbol{E}_0(z,t)$ can be neglected in Equation 4.7. This yields

$$\begin{aligned}
&2\,\mathrm{i}\,k_0\frac{\partial}{\partial z}\boldsymbol{E}_0(z,\omega)-k_0^2\boldsymbol{E}_0(z,\omega)+\frac{\omega^2\mu(\omega)^2}{c^2}\boldsymbol{E}_0(z,\omega)=-\frac{\omega^2}{\varepsilon_0 c^2}e^{-\mathrm{i}k_0 z}\boldsymbol{P}^{\mathrm{NL}}(z,\omega)\\
\Longleftrightarrow\ &\frac{\partial}{\partial z}\boldsymbol{E}_0(z,\omega)-\mathrm{i}\,\frac{k(\omega)^2-k_0^2}{2k_0}\boldsymbol{E}_0(z,\omega)=\mathrm{i}\,\frac{\omega^2}{2\varepsilon_0 k_0 c^2}e^{-\mathrm{i}k_0 z}\boldsymbol{P}^{\mathrm{NL}}(z,\omega),
\end{aligned} \tag{4.8}$$

where we have introduced $k(\omega)=\omega\mu(\omega)/c$. Now, going back to the forward-propagating field $\boldsymbol{E}^+(z,\omega)$, we have

$$\frac{\partial}{\partial z}\boldsymbol{E}^+(z,\omega)-\mathrm{i}\left(\frac{k(\omega)^2-k_0^2}{2k_0}+k_0\right)\boldsymbol{E}^+(z,\omega)=\mathrm{i}\,\frac{\omega^2}{2\varepsilon_0 k_0 c^2}\boldsymbol{P}^{\mathrm{NL}}(z,\omega). \tag{4.9}$$

For long pulses (with bandwidth that is much smaller than the carrier-wave frequency), we have $k(\omega)=k_0+\delta k(\omega)$, with $\delta k(\omega)\ll k_0$. This gives

$$\frac{k(\omega)^2-k_0^2}{2k_0}+k_0=k_0+\delta k(\omega)+\frac{\delta k(\omega)^2}{2k_0}\approx k(\omega), \tag{4.10}$$

and allows us to obtain the simplified nonlinear wave-equation

$$\frac{\partial}{\partial z}\boldsymbol{E}^+(z,\omega)=\mathrm{i}\,k(\omega)\boldsymbol{E}^+(z,\omega)+\mathrm{i}\,\frac{\omega^2}{2\varepsilon_0 k_0 c^2}\boldsymbol{P}^{\mathrm{NL}}(z,\omega), \tag{4.11}$$

where $\mathrm{i}\,k(\omega)$ generates the free evolution, and the nonlinear polarisation plays the role of a

source-term.

We now turn to type-II PDC. In this case, the nonlinear polarisation mixes three fields: the two down-conversion modes referred to as signal and idler, and a pump field. Writing the nonlinear polarisation, as described in Section 1.2.2, we obtain, for the forward-propagating signal and idler,

$$\begin{aligned}\frac{\partial}{\partial z}E_s^+(z,\omega) &= \mathrm{i}\,k_s(\omega)E_s^+(z,\omega) + \mathrm{i}\,\frac{\omega^2}{2k_{0s}c^2}\,d_{\text{eff}}\int d\omega'\,E_p(\omega+\omega')\,E_i^{+*}(z,\omega');\\ \frac{\partial}{\partial z}E_i^{+*}(z,\omega) &= -\mathrm{i}\,k_i(\omega)E_i^{+*}(z,\omega) - \mathrm{i}\,\frac{\omega^2}{2k_{0i}c^2}\,d_{\text{eff}}\int d\omega'\,E_p^*(\omega+\omega')\,E_s^+(z,\omega'),\end{aligned} \tag{4.12}$$

where d_{eff} is the effective nonlinearity, combining $\boldsymbol{\chi}^{(2)}$ and the polarisation directions of the interacting fields. Analogous relations hold for $E_i(z,\omega)$ and $E_s^*(z,\omega)$. As the process in which we are interested is very weak, we can approximate that the pump field is not affected by the nonlinear interaction. Solving for the free propagation of the pump field (which we assume to be forward-propagating), we obtain $E_p(z,\omega) = A_p f(\omega) e^{\mathrm{i}\,k_p(\omega)z}$. Here, $f(\omega)$ defines the pump spectral density, and is normalised such that $\left|\int d\omega\, f(\omega)\right| = 1$. $A_p \equiv \int d\omega\, E(0,\omega)$ is a real value and represents the effective amplitude of the pump field. The propagation equations for the signal and idler then read

$$\begin{aligned}\frac{\partial}{\partial z}E_s^+(z,\omega) &= \mathrm{i}\,k_s(\omega)E_s^+(z,\omega) + \mathrm{i}\,\frac{\omega^2 d_{\text{eff}}}{2k_{0s}c^2}A_p\int d\omega'\,f(\omega+\omega')\,e^{\mathrm{i}\,k_p(\omega+\omega')z}E_i^{+*}(z,\omega');\\ \frac{\partial}{\partial z}E_i^{+*}(z,\omega) &= -\mathrm{i}\,k_i(\omega)E_i^{+*}(z,\omega) - \mathrm{i}\,\frac{\omega^2 d_{\text{eff}}}{2k_{0i}c^2}A_p\int d\omega'\,f^*(\omega+\omega')\,e^{-\mathrm{i}\,k_p(\omega+\omega')z}E_s^+(z,\omega').\end{aligned} \tag{4.13}$$

We are now ready to quantise the electric field. For wave-guided modes, the quantisation relation is[245,267]

$$E(\boldsymbol{r},\omega) = \mathrm{i}\,2\pi\zeta(x,y)\sqrt{\frac{\hbar\omega}{4\pi\varepsilon_0 cn}}\hat{b}(z,\omega), \tag{4.14}$$

where $\zeta(x,y)$ defines the transverse mode, and the bosonic operators $\hat{b}(z,\omega)$ satisfy the commutation relations[268–270]

$$\left[\hat{b}(z,\omega),\hat{b}(z,\omega')\right] = 0;\ \left[\hat{b}(z,\omega),\hat{b}^\dagger(z',\omega')\right] = \delta(\omega-\omega'). \tag{4.15}$$

The quantisation of Equations 4.13 leads to coupled differential equations describing a linear

mode-transformation:

$$\begin{aligned}\frac{\partial}{\partial z}\hat{b}_s(z,\omega) &= \mathrm{i}\,k_s(\omega)\hat{b}_s(z,\omega) + \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f(\omega+\omega')\, e^{\mathrm{i}\,k_p(\omega+\omega')z}\hat{b}_i^\dagger(z,\omega');\\ \frac{\partial}{\partial z}\hat{b}_i^\dagger(z,\omega) &= -\mathrm{i}\,k_i(\omega)\hat{b}_i^\dagger(z,\omega) - \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f^*(\omega+\omega')\, e^{-\mathrm{i}\,k_p(\omega+\omega')z}\hat{b}_s(z,\omega').\end{aligned} \tag{4.16}$$

Here, we introduced a nonlinear interaction length, $L_{\mathrm{NL}} \equiv \left(2c^2\sqrt{k_{0s}k_{0i}}\right)/(\omega_{0i}\omega_{0s}d_{\mathrm{eff}}A_p)$, and in order to simplify the equations, we replaced all frequencies and wave-vectors by the corresponding carrier-wave values. This operation assumes, again, a small bandwidth for all interacting signals.

We now make a change of variables, writing $\hat{b}_s(z,\omega) = \hat{c}_s(z,\omega)e^{\mathrm{i}\,k_s(\omega)z}$ and $\hat{b}_i^\dagger(z,\omega) = \hat{c}_i^\dagger(z,\omega)e^{-\mathrm{i}\,k_i(\omega)z}$. This yields

$$\begin{aligned}\frac{\partial}{\partial z}\hat{c}_s(z,\omega) &= \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f(\omega+\omega')\, e^{\mathrm{i}\,\Delta k(\omega,\omega')z}\hat{c}_i^\dagger(z,\omega');\\ \frac{\partial}{\partial z}\hat{c}_i^\dagger(z,\omega) &= -\mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f^*(\omega+\omega')\, e^{-\mathrm{i}\,\Delta k(\omega,\omega')z}\hat{c}_s(z,\omega'),\end{aligned} \tag{4.17}$$

where $\Delta k \equiv k_p(\omega+\omega') - k_s(\omega) - k_i(\omega')$ is the phase mismatch. We follow a common approach to linearise this function, using an expansion around the carrier-wave frequencies, $\Delta k(\omega,\omega') \approx \Delta k_0 + \left[(\partial k/\partial\omega)_p - (\partial k/\partial\omega)_s\right](\omega-\omega_{0s}) + \left[(\partial k/\partial\omega)_p - (\partial k/\partial\omega)_i\right](\omega'-\omega_{0i})$, with $\Delta k_0 = k_{0p} - k_{0s} - k_{0i}$. Let us define $\Delta k_s(\omega) \equiv \left[(\partial k/\partial\omega)_s - (\partial k/\partial\omega)_p\right](\omega-\omega_{0s})$ and $\Delta k_i(\omega) \equiv \left[(\partial k/\partial\omega)_i - (\partial k/\partial\omega)_p\right](\omega-\omega_{0i})$. Making a new change of variables, $\hat{c}_s(z,\omega) = \hat{a}_s(z,\omega)e^{-\mathrm{i}\,\Delta k_s(\omega)z}$ and $\hat{c}_i^\dagger(z,\omega) = \hat{a}_i^\dagger(z,\omega)e^{\mathrm{i}\,\Delta k_i(\omega)z}$, we obtain

$$\begin{aligned}\frac{\partial}{\partial z}\hat{a}_s(z,\omega) &= \mathrm{i}\,\Delta k_s(\omega)\,\hat{a}_s(z,\omega) + \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f(\omega+\omega')\, e^{\mathrm{i}\,\Delta k_0 z}\hat{a}_i^\dagger(z,\omega');\\ \frac{\partial}{\partial z}\hat{a}_i^\dagger(z,\omega) &= -\mathrm{i}\,\Delta k_i(\omega)\,\hat{a}_i^\dagger(z,\omega) - \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int d\omega'\, f^*(\omega+\omega')\, e^{-\mathrm{i}\,\Delta k_0 z}\hat{a}_s(z,\omega').\end{aligned} \tag{4.18}$$

The parametric process is phase-matched if $\Delta k_0 = 0$. This can be achieved in the desired frequency range either by taking advantage of the birefringence of the propagation medium, or with the help of quasi-phase-matching (QPM).

QPM has greatly enhanced the versatility of parametric processes in nonlinear crystals[271,272]. The simplest type of QPM is implemented by periodic poling: switching the sign of the $\chi^{(2)}$ nonlinearity of a crystal after each length $l_{\mathrm{p.p.}}/2$. This results in a modulation of the effective nonlinear coefficient by a sawtooth function with period $l_{\mathrm{p.p.}}$, which can

be written as the Fourier series

$$h_{\text{square}}(z) = \frac{2}{\pi}\sum_{j=1}^{\infty}\frac{\mathrm{i}}{2j-1}\left[\exp(\mathrm{i}\,2\pi(2j-1)z/l_{\text{p.p.}}) - \exp(-\mathrm{i}\,2\pi(2j-1)z/l_{\text{p.p.}})\right]. \tag{4.19}$$

Multiplying the nonlinear term in Equations 4.18 by h_{square} results in a sequence of terms, with different effective phase mismatch. For example, taking the first part of the term with $j = 1$, we see that the phase mismatch is changed to $\Delta k_0 + 2\pi/l_{\text{p.p.}}$. This can be made 0 by choosing the appropriate value for the poling period, $l_{\text{p.p.}} = -2\pi/\Delta k_0$. For this first-order QPM process, the effective nonlinear coefficient is $2/\pi\ d_{\text{eff}}$.

Throughout the rest of this section, we consider phase matched processes. Setting $\Delta k_0 = 0$ (meaning that we are expanding the wave-vectors around the values that correspond to a phase mismatch of 0), the right-hand side of Equation 4.18 becomes independent of z, and a solution can be easily computed. The general form of this solution is

$$\begin{aligned}
\hat{a}_s(z,\omega) &= \int d\omega'\, U_{s,s}(z,\omega,\omega')\hat{a}_s(0,\omega') + \int d\omega'\, U_{s,i}(z,\omega,\omega')\hat{a}_i^{\dagger}(0,\omega');\\
\hat{a}_i^{\dagger}(z,\omega) &= \int d\omega'\, U_{i,i}^{*}(z,\omega,\omega')\hat{a}_i^{\dagger}(0,\omega') + \int d\omega'\, U_{i,s}^{*}(z,\omega,\omega')\hat{a}_s(0,\omega'),
\end{aligned} \tag{4.20}$$

with analogous relations for $\hat{a}_i$ and $\hat{a}_s^{\dagger}$. The functions introduced here completely describe the propagation of frequency modes along the z axis. They are sometimes referred to as Green's functions[245,270], but in this work we will use the term "transfer functions". In order to preserve the bosonic commutation relations, the transfer functions must satisfy

$$\begin{aligned}
&\int d\omega''\, U_{s,s}(z,\omega,\omega'')\, U_{s,s}^{*}(z,\omega',\omega'') - U_{s,i}(z,\omega,\omega'')\, U_{s,i}^{*}(z,\omega',\omega'') = \delta(\omega-\omega');\\
&\int d\omega''\, U_{i,i}(z,\omega,\omega'')\, U_{i,i}^{*}(z,\omega',\omega'') - U_{i,s}(z,\omega,\omega'')\, U_{i,s}^{*}(z,\omega',\omega'') = \delta(\omega-\omega');\\
&\int d\omega''\, U_{s,s}(z,\omega,\omega'')\, U_{i,s}(z,\omega',\omega'') - U_{s,i}(z,\omega,\omega'')\, U_{i,i}(z,\omega',\omega'') = 0.
\end{aligned} \tag{4.21}$$

Before moving on to treat pulsed PDC, let us take a look at the CW-pumped processes[270]. In this case, the integrals in Equations 4.18 can be readily evaluated, after replacing the pump spectral density by the Dirac delta distribution $f(\omega) \to \delta(\omega - \omega_{0p})$, with

$\omega_{0p} = \omega_{0s} + \omega_{oi}$. This gives

$$\begin{aligned}\frac{\partial}{\partial z}\hat{a}_s(z,\omega_{0s}+\Delta\omega) &= \mathrm{i}\,\Delta k_s(\omega_{0s}+\Delta\omega)\,\hat{a}_s(z,\omega_{0s}+\Delta\omega) + \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\hat{a}_i^\dagger(z,\omega_{0i}-\Delta\omega);\\ \frac{\partial}{\partial z}\hat{a}_i^\dagger(z,\omega_{0i}-\Delta\omega) &= -\mathrm{i}\,\Delta k_i(\omega_{0i}-\Delta\omega)\,\hat{a}_i^\dagger(z,\omega_{0i}-\Delta\omega) - \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\hat{a}_s(z,\omega_{0s}+\Delta\omega).\end{aligned} \tag{4.22}$$

These differential equations have the solution

$$\begin{aligned}\hat{a}_s(z,\omega_{0s}+\Delta\omega) &= C_{s,s}(z,\Delta\omega)\,\hat{a}_s(0,\omega_{0s}+\Delta\omega) + C_{s,i}(z,\Delta\omega)\,\hat{a}_i^\dagger(0,\omega_{0i}-\Delta\omega);\\ \hat{a}_i^\dagger(z,\omega_{0i}-\Delta\omega) &= (C_{i,i}(z,\Delta\omega))^*\hat{a}_i^\dagger(0,\omega_{0i}-\Delta\omega) + (C_{i,s}(z,\Delta\omega))^*\hat{a}_s(0,\omega_{0s}+\Delta\omega),\end{aligned} \tag{4.23}$$

with

$$\begin{aligned}C_{s,i}(z,\Delta\omega) = -\left(C_{i,s}(z,\Delta\omega)\right)^* = \frac{\mathrm{i}\,z}{L_{\mathrm{NL}}}\exp\left(\mathrm{i}\,\left[\Delta k_s(\omega_{0s}+\Delta\omega) - \Delta k_i(\omega_{0i}-\Delta\omega)\right]\,z/2\right)\\ \times\,\mathrm{sinc}\left(\frac{z}{2}\sqrt{\left[\Delta k_s(\omega_{0s}+\Delta\omega)+\Delta k_i(\omega_{0i}-\Delta\omega)\right]^2 - 4/L_{\mathrm{NL}}^2}\right).\end{aligned} \tag{4.24}$$

The absolute-squared values of the functions $C_{s,s}(z,\Delta\omega)$ and $C_{i,i}(z,\Delta\omega)$ are given by $1 + |C_{s,i}(z,\Delta\omega)|^2$. In the interest of space, we do not include the full expressions for these functions here.

Note that, at 0 phase mismatch, we have $C_{s,i}(z,0) = \mathrm{i}\sinh(z/L_{\mathrm{NL}})$. Moreover, in the weak interaction limit, given by $z \ll L_{\mathrm{NL}}$, we have

$$\begin{aligned}C_{s,i}(z,\Delta\omega) \approx \mathrm{i}\sinh(z/L_{\mathrm{NL}})\,\mathrm{sinc}\left(\left[\Delta k_s(\omega_{0s}+\Delta\omega)+\Delta k_i(\omega_{0i}-\Delta\omega)\right]\,z/2\right)\\ \times\exp\left(\mathrm{i}\,\left[\Delta k_s(\omega_{0s}+\Delta\omega)-\Delta k_i(\omega_{0i}-\Delta\omega)\right]\,z/2\right).\end{aligned} \tag{4.25}$$

We will find a similar answer for pulsed PDC in the weak interaction limit.

For pulsed PDC in general, we resort to numerical evaluation. We discretise the frequency scale taking N points spaced by $\delta\omega$: $\omega_j^{(s)} \equiv \omega_{0s} + (j - N/2)\delta\omega$ and $\omega_j^{(i)} \equiv \omega_{0i} + (j - N/2)\delta\omega$. We construct the $2N$-dimensional vector $\boldsymbol{A}(z)$ with the components $A_j(z) = \hat{a}_s\left(z,\omega_j^{(s)}\right)$ and $A_{N+j}(z) = \hat{a}_i^\dagger\left(z,\omega_j^{(i)}\right)$. If $\delta\omega$ is small enough compared to the bandwidths of the interacting fields, solving the discretised differential equations provides a good approximation for the evolution of the continuous mode-operators. The finite system

of linear differential equations can be written in a compact form:

$$\begin{aligned}&\partial_z \boldsymbol{A}(z) = \mathrm{i}\,\boldsymbol{G}\boldsymbol{A}(z),\ \text{with}\ \boldsymbol{G} = \begin{pmatrix} \boldsymbol{D}^{(s)} & \boldsymbol{V} \\ -\boldsymbol{V}^* & -\boldsymbol{D}^{(i)} \end{pmatrix},\ \text{where}\\ &D^{(s)}_{j,j'} = \delta_{j,j'}\Delta k_s(\omega^{(s)}_j);\ D^{(i)}_{j,j'} = \delta_{j,j'}\Delta k_i(\omega^{(i)}_j); V_{j,j'} = \delta\omega/L_{\mathrm{NL}}\ f\left(\omega^{(s)}_j + \omega^{(i)}_{j'}\right).\end{aligned} \tag{4.26}$$

The solution is given by $\boldsymbol{A}(z) = \boldsymbol{T}(z)\,\boldsymbol{A}(0)$, with $\boldsymbol{T}(z) = \exp(\mathrm{i}\,\boldsymbol{G}\,z)$. Then, the transfer functions introduced in Equation 4.20 are given (approximately) by:

$$\begin{aligned}&U_{s,s}(z,\omega^{(s)}_j,\omega^{(s)}_{j'}) \approx \frac{1}{\delta\omega}\left[\boldsymbol{T}(z)\right]_{j,j'};\ U_{s,i}(z,\omega^{(s)}_j,\omega^{(i)}_{j'}) \approx \frac{1}{\delta\omega}\left[\boldsymbol{T}(z)\right]_{j,N+j'};\\ &U_{i,s}(z,\omega^{(i)}_j,\omega^{(s)}_{j'}) \approx \frac{1}{\delta\omega}\left[\boldsymbol{T}(z)\right]^*_{N+j,j'};\ U_{i,i}(z,\omega^{(i)}_j,\omega^{(i)}_{j'}) \approx \frac{1}{\delta\omega}\left[\boldsymbol{T}(z)\right]^*_{N+j,N+j'}.\end{aligned} \tag{4.27}$$

4.1.3 Perturbation analysis

In the previous sections, we have derived the differential equations describing mode-propagation for type-II PDC. The evolution of mode-operators described by Equations 4.18 is generated by two terms, both independent of z. The first term describes free propagation: the walk-off between pulses in the signal, idler, and pump modes, which arises due to different dispersion relations. The second term describes nonlinear frequency conversion. Since the nonlinear interaction is weak, we can treat the second term as a perturbation.

Let us first define the linear superoperators $\hat{O}$ and $\hat{V}$, such that

$$\begin{aligned}&\hat{O}\,\hat{a}_s(z,\omega) = \Delta k_s(\omega)\hat{a}_s(z,\omega);\ \hat{O}\,\hat{a}^\dagger_i(z,\omega) = -\Delta k_i(\omega)\hat{a}^\dagger_i(z,\omega);\\ &\hat{V}\,\hat{a}_s(z,\omega) = \int d\omega'\, f(\omega+\omega')\hat{a}^\dagger_i(z,\omega');\ \hat{V}\,\hat{a}^\dagger_i(z,\omega) = -\int d\omega'\, f^*(\omega+\omega')\hat{a}_s(z,\omega'),\end{aligned} \tag{4.28}$$

and introduce the propagator $\hat{U}(z,z_0)$ such that $\hat{a}_s(z,\omega) = \hat{U}(z,z_0)\,\hat{a}_s(z_0,\omega)$ and $\hat{a}^\dagger_i(z,\omega) = \hat{U}(z,z_0)\,\hat{a}^\dagger_i(z_0,\omega)$. Equations 4.18 can then be recast as

$$\frac{\partial}{\partial z}\hat{U}(z,z_0) = \mathrm{i}\left(\hat{O} + \frac{1}{L_{\mathrm{NL}}}\hat{V}\right)\hat{U}(z,z_0);\ \hat{U}(z,z) = \mathbb{1}, \tag{4.29}$$

with the solution $\hat{U}(z,z_0) = \exp\left(\mathrm{i}\left[\hat{O} + 1/L_{\mathrm{NL}}\,\hat{V}\right](z-z_0)\right)$. We now expand this expres-

sion using the Suzuki-Trotter decomposition [273,274]:

$$\hat{U}(z,0) = \lim_{\delta z\to 0}\prod_{j=0}^{z/\delta z}\exp\left(\mathrm{i}\,\hat{O}\,\delta z\right)\exp\left(\mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\hat{V}\delta z\right) = \lim_{\delta z\to 0}\prod_{j=0}^{z/\delta z}\exp\left(\mathrm{i}\,\hat{O}\,\delta z\right)$$
$$\times\left(\mathbb{1}+\mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\hat{V}\delta z\right) = \exp\left(\mathrm{i}\,\hat{O}z\right)+\mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\sum_{j=0}^{z/\delta z}\exp\left(\mathrm{i}\,\hat{O}\,(z-j\delta z)\right)\hat{V}\exp\left(\mathrm{i}\,\hat{O}\,j\delta z\right)$$
$$-\left(\frac{1}{L_{\mathrm{NL}}}\right)^2\sum_{j=0}^{z/\delta z}\sum_{j'=0}^{j}\exp\left(\mathrm{i}\,\hat{O}\,(z-j\delta z)\right)\hat{V}\exp\left(\mathrm{i}\,\hat{O}\,(j-j')\delta z\right)\hat{V}\exp\left(\mathrm{i}\,\hat{O}\,j'\delta z\right)+\ldots. \tag{4.30}$$

Replacing the sums by integrals, we obtain

$$\hat{U}(z,0) = \exp\Big(\mathrm{i}\,\hat{O}z\Big)+\mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int_0^z dz_1\exp\Big(\mathrm{i}\,\hat{O}\,(z-z_1)\Big)\hat{V}\exp\Big(\mathrm{i}\,\hat{O}\,z_1\Big)-\left(\frac{1}{L_{\mathrm{NL}}}\right)^2$$
$$\times\int_0^z dz_1\int_0^{z_1}dz_2\,\exp\Big(\mathrm{i}\,\hat{O}\,(z-z_1)\Big)\hat{V}\exp\Big(\mathrm{i}\,\hat{O}\,(z_1-z_2)\Big)\hat{V}\exp\Big(\mathrm{i}\,\hat{O}\,z_2\Big)+\ldots, \tag{4.31}$$

which is analogous to the Dyson series, a standard result of time-dependent perturbation theory.

Let us consider the terms in the expansion up to first order in $1/L_{\mathrm{NL}}$. Propagating the operators $\hat{a}_s(0,\omega)$ and $\hat{a}_i^\dagger(0,\omega)$, we obtain

$$\hat{a}_s(z,\omega) = e^{\mathrm{i}\,z\Delta k_s(\omega)}\hat{a}_s(0,\omega)+\int_0^\infty d\omega'\,U_{s,i}^{(1)}(z,\omega,\omega')\,\hat{a}_i^\dagger(0,\omega')+\mathcal{O}\left((1/L_{\mathrm{NL}})^2\right);$$
$$\hat{a}_i^\dagger(z,\omega) = e^{-\mathrm{i}\,z\Delta k_i(\omega)}\hat{a}_i^\dagger(0,\omega)+\int_0^\infty d\omega'\left[U_{i,s}^{(1)}(z,\omega,\omega')\right]^*\hat{a}_s(0,\omega')+\mathcal{O}\left((1/L_{\mathrm{NL}})^2\right), \tag{4.32}$$

where

$$U_{s,i}^{(1)}(z,\omega,\omega') = \mathrm{i}\,\frac{1}{L_{\mathrm{NL}}}\int_0^z dz_1\,\exp\big(\mathrm{i}\,\Delta k_s(\omega)[z-z_1]\big)\,f(\omega+\omega')\,\exp\big(-\mathrm{i}\,\Delta k_i(\omega')z_1\big)$$
$$=\mathrm{i}\,\frac{z}{L_{\mathrm{NL}}}\exp\Big(\mathrm{i}\,\frac{z}{2}\left[\Delta k_s(\omega)-\Delta k_i(\omega')\right]\Big)\,\mathrm{sinc}\Big(\frac{z}{2}\left[\Delta k_s(\omega)+\Delta k_i(\omega')\right]\Big)\,f(\omega,\omega'), \tag{4.33}$$

and $e^{-\mathrm{i}\,z\Delta k_s(\omega)}U_{s,i}^{(1)}(z,\omega,\omega') = e^{-\mathrm{i}\,z\Delta k_i(\omega')}U_{i,s}^{(1)}(z,\omega',\omega)$.

This first-order perturbation result is often employed in the context of photo-pair generation. The joint spectrum of SPDC photon-pairs can be computed by evaluating the expectation value $\langle 0|\hat{n}_s(z,\omega_s)\hat{n}_i(z,\omega_i)|0\rangle = \langle 0|\hat{a}_s^\dagger(z,\omega_s)\hat{a}_s(z,\omega_s)\hat{a}_i^\dagger(z,\omega_i)\hat{a}_i(z,\omega_i)|0\rangle$, where

$|0\rangle$ denotes the vacuum state at $z = 0$ (in a representation analogous to the Heisenberg picture of Hamiltonian evolution). To calculate this expression, we use the transformations given in Equation 4.32 and the equalities $\langle 0|\hat{a}_{s(i)}(0,\omega)\hat{a}^{\dagger}_{s(i)}(0,\omega')|0\rangle = \delta(\omega-\omega')$. Any other ordering of two mode-operators yields the expectation value 0. Expanding the expectation value, only one non-zero term survives up to second order in $1/L_{\mathrm{NL}}$:

$$\begin{aligned}
&\langle 0|\hat{a}^{\dagger}_s(z,\omega_s)\hat{a}_s(z,\omega_s)\hat{a}^{\dagger}_i(z,\omega_i)\hat{a}_i(z,\omega_i)|0\rangle \approx \left\langle 0 \left| \int d\omega' \left(U^{(1)}_{s,i}(z,\omega_s,\omega')\right)^* \hat{a}_i(0,\omega') \right.\right. \\
&\left.\left. \times e^{\mathrm{i}\, z\Delta k_s(\omega_s)}\hat{a}_s(0,\omega_s)\, e^{-\mathrm{i}\, z\Delta k_i(\omega_i)}\hat{a}^{\dagger}_i(0,\omega_i) \int d\omega''\, U^{(1)}_{i,s}(z,\omega_i,\omega'')\hat{a}^{\dagger}_s(0,\omega_i,\omega'') \right| 0 \right\rangle \\
&= \int d\omega' d\omega'' \left[e^{-\mathrm{i}\, z\Delta k_s(\omega_s)}U^{(1)}_{s,i}(z,\omega_s,\omega')\right]^* e^{-\mathrm{i}\, z\Delta k_i(\omega_i)}U^{(1)}_{i,s}(z,\omega_i,\omega'')\delta(\omega'-\omega_i)\delta(\omega_s-\omega'') \\
&= \left|U^{(1)}_{s,i}(z,\omega_s,\omega_i)\right|^2 = |z/L_{\mathrm{NL}}|^2\, |f(\omega_s+\omega_i)|^2\, \mathrm{sinc}^2\!\left(\frac{z}{2}\left[\Delta k_s(\omega_s)+\Delta k_i(\omega_i)\right]\right).
\end{aligned} \tag{4.34}$$

This expression is a well-known result that has proven very useful, due to the intuitive interpretation that it allows. It shows that, in the weak-gain limit, the joint spectral distribution of SPDC photon-pairs results from the product of two functions: a "pump function", which effectively enforces energy conservation, and a function that describes the effect of phase-matching. The pump factor leads to anti-correlation between the frequencies of photons generated as a pair, indicating spectro-temporal entanglement between the signal and idler. The phase-matching function can be engineered such as to minimise this entanglement, leading to the generation of photon-pairs that are almost separable in the spectral degree of freedom. For example, the KDP source used in the experiment reported in Chapter 2 satisfies $\Delta k_i(\omega_i) = 0$ (meaning that the idler propagates with the same group velocity as the pump). This leads to a nearly separable joint spectrum, as detailed in Section 2.3.2. Note that, due to the oscillation of the phase-matching function, we always have some remaining correlations.

Figure 4.5 depicts (in absolute value) the $U_{s,i}$ transfer function corresponding to a typical example of pulsed PDC. We use a Gaussian pump intensity spectrum of width $\sigma_p = 1$, the propagation distance $z = L_{\mathrm{NL}}$, and the phase-matching function defined by $\Delta k_s(\sigma_p)\, z = -2\Delta k_s(\sigma_p)\, z = 3$. We present both a numerical full-order result and the first-order approximation derived in Equation 4.32. For this relatively high interaction length (resulting in the generation of more than one photon-pair on average per pump pulse) the

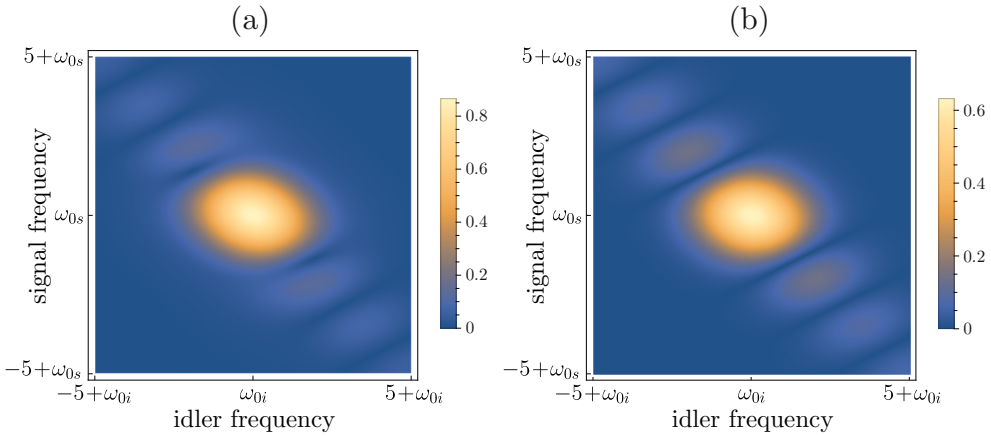


Figure 4.1: Transfer function describing the coupling between different frequencies in the signal and idler mode for a typical broadband PDC process. (a) Obtained by full-order numerical evaluation and (b) the first-order perturbation result (corresponding to the joint spectrum of SPDC photon-pairs). We depict the absolute values for the two functions. The frequency unit is given by the width of the Gaussian pump field intensity spectrum.

error of the first-order approximation is visible, in particular in the magnitude of the transfer function.

The first-order term in the series given by Equation 4.31 describes how a field in the idler (signal) mode stimulates the generation of a field in the signal (idler) mode. The second-order term corresponds to a process in which the generated field then stimulates the generation of a field in the initial mode. Thus, the second-order perturbation term contributes to the transfer functions $U_{s,s}$ and $U_{i,i}$. Applying the second order term, which we denote $\hat{U}^{(2)}(z, z_0)$, to $\hat{a}_s(0, \omega)$, we obtain

$$\begin{aligned}\hat{U}^{(2)}(z,0)\,\hat{a}_s(0,\omega) = &-\left(\frac{1}{L_{\mathrm{NL}}}\right)^2 \int_0^z dz_1 \int_0^{z_1} dz_2\, \exp\left(\mathrm{i}\,\Delta k_s(\omega)(z-z_1)\right) \int d\omega''\, f(\omega+\omega'') \\ &\times \exp\left(-\mathrm{i}\,\Delta k_i(\omega'')(z_1-z_2)\right) \int d\omega'\, f^*(\omega''+\omega') \exp\left(\mathrm{i}\,\Delta k_s(\omega')z_2\right) \hat{a}_s(0,\omega) \\ &\equiv \int d\omega'\, U^{(2)}_{s,s}(z,\omega,\omega')\hat{a}_s(0,\omega).\end{aligned} \tag{4.35}$$

Considering, for example, a Gaussian pump spectrum, the integration above can be performed analytically. Unfortunately, this approach yields an expression for $U^{(2)}_{s,s}(z, \omega, \omega')$ that does not invite simple interpretation. However, we find that, in absolute value, a good

approximation is provided by a Gaussian distribution:

$$\left|U_{s,s}^{(2,\mathrm{G})}(z,\omega,\omega')\right| = C_0(\xi_s,\xi_i)\left|\frac{z}{L_{\mathrm{NL}}}\right|^2 \exp\left(-\frac{1}{2\sigma_p^2}\left[\left(\frac{(\omega+\omega')/\sqrt{2}}{C_+(\xi_s,\xi_i)}\right)^2+\left(\frac{(\omega-\omega')/\sqrt{2}}{C_-(\xi_s,\xi_i)}\right)^2\right]\right) \tag{4.36}$$

where σ_p is the pump bandwidth and we introduced $\xi_s \equiv k_s(\sigma_p)z$ and $\xi_i \equiv k_i(\sigma_p)z$ (defining the bandwidth ratio between the pump and the phase-matching function). Expanding the analytical result for $U_{s,s}^{(2)}(z,\omega,\omega')$ around $\omega = 0$ and $\omega' = 0$, we find the following second-order expressions for the coefficients introduced in Equation 4.36:

$$\begin{aligned} C_0(\xi_s,\xi_i) &= \left(4+2\xi_i^2/3\right)^{-1}; \\ C_+(\xi_s,\xi_i) &= \left(6+\xi_i^2/4\right)\Big/\left[\sqrt{2}(\xi_i-\xi_s)\right]; \\ C_-(\xi_s,\xi_i) &= 240\Big/\left(240+(20+\xi_i^2)\xi_s^2\right). \end{aligned} \tag{4.37}$$

We find that $C_-(\xi_s,\xi_i)/C_+(\xi_s,\xi_i) \lesssim 0.62$, which means that $U_{s,s}^{(2,\mathrm{G})}(z,\omega,\omega')$ always shows positive correlations between ω and ω'.

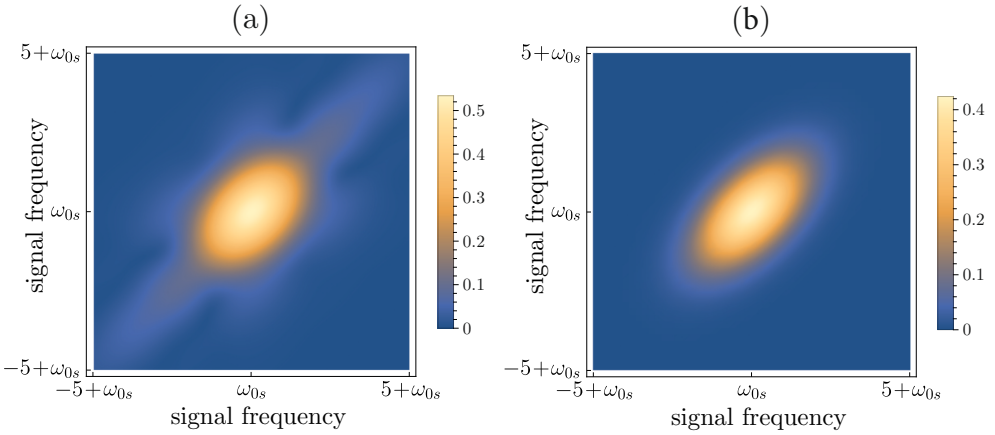


Figure 4.2: Nonlinear contribution to the transfer function describing the coupling between different frequencies in the signal mode, for a typical broadband PDC process. (a) Obtained by full-order numerical evaluation, and (b) Gaussian approximation to the second-order perturbation result. We depict the absolute values for the two functions. The frequency unit is given by the width of the Gaussian pump field intensity spectrum.

The $U_{s,s}$ transfer function takes the form

$$U_{s,s}(z,\omega,\omega') = e^{\mathrm{i}\,\Delta k_s(\omega)z}\delta(\omega-\omega') + U_{s,s}^{(2)}(z,\omega,\omega') + \ldots, \tag{4.38}$$

where the first term describes free propagation, while the second term, proportional to $(z/L_{\mathrm{NL}})^2$, and the corrections of higher order in z/L_{NL}, describe nonlinear frequency coupling. In Figure 4.2, we depict the nonlinear part of $U_{s,s}$, in absolute value, for the same parameters as those used to compute the functions displayed in Figure 4.5. We include both a numerical full-order evaluation and the Gaussian approximation to the second order perturbation term.

For the remainder of this chapter, we perform a basis transformation that corresponds to the free propagation of both input and output modes to the middle of the nonlinear interacting region, transforming the propagator to $e^{-\hat{O}z/2}\hat{U}e^{-\mathrm{i}\,\hat{O}z/2}$. This is just a passive transformation and it does not affect our results concerning nonlinear frequency generation. Then, the 0th order term of the signal-signal and idler-idler transfer functions becomes $U^{(0)}_{s,s}(\omega,\omega') = U^{(0)}_{i,i}(\omega,\omega') = \delta(\omega,\omega')$. We also drop the position z as input to the annihilation operators and transfer functions when considering only the input-output transformations for a nonlinear medium of a given length.

4.1.4 Schmidt mode decomposition

The Schmidt mode decomposition provides a natural way of interpreting the broadband structure of PDC[245–247,275,276]. This decomposition defines a set of broadband modes that are only coupled pair-wise by the PDC transformations described in Equation 4.20. The fact that it is always possible to find such modes shows that broadband PDC can always be modeled by a discrete set of independent squeezing interactions. The Schmidt modes are found using the singular value decomposition of the transfer functions. Using the conditions for preserving the commutation relations of mode operators, given by Equation 4.21, it can be shown that the singular value decompositions of the transfer functions are constrained

to take the form [247]

$$\begin{aligned} U_{s,s}(\omega,\omega') &= \sum_k \cosh(\zeta_k)\ \left[\phi_s^{(k)}(\omega)\right]^*\ \left[\psi_s^{(k)}(\omega')\right]^*; \\ U_{s,i}(\omega,\omega') &= \sum_k \sinh(\zeta_k)\ \left[\phi_s^{(k)}(\omega)\right]^*\ \psi_i^{(k)}(\omega'); \\ U_{i,i}(\omega,\omega') &= \sum_k \cosh(\zeta_k)\ \left[\phi_i^{(k)}(\omega)\right]^*\ \left[\psi_i^{(k)}(\omega')\right]^*; \\ U_{i,s}(\omega,\omega') &= \sum_k \sinh(\zeta_k)\ \left[\phi_i^{(k)}(\omega)\right]^*\ \psi_s^{(k)}(\omega'). \end{aligned} \tag{4.39}$$

The input and output Schmidt-mode annihilation operators are defined as $\hat{A}^{(k)}_{s(i)} \equiv \int d\omega\, \psi_{s(i)}(\omega)\, \hat{a}_{s(i)}(0,\omega)$ and $\hat{B}^{(k)}_{s(i)} \equiv \int d\omega\, \phi_{s(i)}(\omega)\, \hat{a}_{s(i)}(z,\omega)$, respectively. Using the orthogonality of eigenfunctions, we can write the PDC mode transformation in the simple form

$$\begin{aligned} \hat{B}_s^{(k)} &= \cosh(\zeta_k)\, \hat{A}_s^{(k)} + \sinh(\zeta_k)\, \hat{A}_i^{(k)\,\dagger}; \\ \hat{B}_i^{(k)\,\dagger} &= \cosh(\zeta_k)\, \hat{A}_i^{(k)\,\dagger} + \sinh(\zeta_k)\, \hat{A}_s^{(k)}, \end{aligned} \tag{4.40}$$

where the parameters ζ_k are referred to as squeezing parameters.

If we approximate the transfer functions by Gaussians, in the limit of weak squeezing, $\psi^{(k)}(\omega)$ and $\phi^{(k)}(\omega)$ become the Hermite-Gauss functions [275,276], and similar distributions are found for typical pulsed PDC processes.

4.2 Characterisation of a parametric down-conversion source using a seeding beam

The joint spectrum of SPDC photon pairs is an important characterisation tool for squeezed light sources. It was first measured using grating monochromators and APDs. The measurement required monitoring coincidence detections produced by photon pairs while scanning the detected wavelengths [40]. Later on, a more efficient measurement scheme was developed, using time-of-flight spectrometers based on highly dispersive optical fibers [276–278]. In Chapter 2, we introduced a similar method, based on chirped fiber Bragg gratings.

In 2013, Liscidini and Sipe proposed stimulated emission tomography [250], which uses a seeding beam to classically simulate SPDC, allowing the full reconstruction of the wave-

vector distribution of the down-conversion modes. For wave-guided sources, this method can be easily applied to the simulation of the joint spectrum of photon-pairs. The simulation has been implemented experimentally for an integrated PDC source [279], resulting in a characterisation with unprecedented spectral resolution.

In this section, we discuss seeded PDC, extending the measurement of joint spectra to all transfer functions defined in the previous section and proposing an application for this extension.

4.2.1 Seeded measurements

Seeding is modeled by a phase-space displacement applied to one of the two down-conversion modes prior to the squeezing transformation. The displacement is described by the unitary operator

$$\hat{D}^{\text{seed}}_{s(i)} = \exp\left(- \int d\omega \, \alpha(\omega) \, \hat{a}^{\dagger}_{s(i)}(\omega) - c.c. \right), \tag{4.41}$$

where $\alpha(\omega)$ is the spectral amplitude of the seeding light pulse. Introducing the squeezing operator $\hat{S}$, which produces the transformations given by Equation 4.20, we find the state generated by seeded PDC (seeded in the signal mode):

$$\hat{S} \, \hat{D}^{\text{seed}}_{s} \, |0\rangle = \hat{S} \, \hat{D}^{\text{seed}}_{s} \, \hat{S}^{\dagger} \hat{S} \, |0\rangle = \exp\left(- \int d\omega \, \alpha(\omega) \, \hat{S} \hat{a}^{\dagger}_{s}(\omega) \hat{S}^{\dagger} - c.c. \right) \hat{S} \, |0\rangle \, . \tag{4.42}$$

An intensity measurement on this state has a contribution from the field displacements, which is proportional to the seed amplitude, and a contribution from the spontaneous emission described by $\hat{S} \, |0\rangle$. The former can easily be made many orders of magnitude higher than the latter, which we can, therefore, neglect. The displacement given in Equation 4.42 is

$$\begin{aligned} \exp\left(- \int d\omega \, \alpha(\omega) \, \hat{S} \hat{a}^{\dagger}_{s}(\omega) \hat{S}^{\dagger} - c.c. \right) = \exp\left(- \int d\omega \, d\omega' \, \alpha(\omega) \, \left[U_{s,s}(\omega, \omega') \right]^{*} \hat{a}^{\dagger}_{s}(\omega') - c.c. \right) \\ \otimes \exp\left(\int d\omega \, d\omega' \, \alpha^{*}(\omega) \, U_{s,i}(\omega, \omega') \, \hat{a}^{\dagger}_{i}(\omega') - c.c. \right) . \end{aligned} \tag{4.43}$$

If the seed is a narrowband laser field, as modeled by $\alpha(\omega) \rightarrow \alpha_0 \, \delta(\omega - \omega_0)$, we find that the DFG signal at frequency ω' is $\alpha^{\text{DFG}}(\omega_0, \omega') = -\alpha_0^{*} \, U_{s,i}(\, \omega_0, \omega')$. Therefore, a spectrally

resolved intensity measurement of the DFG signal produced by narrowband seeding yields the distribution

$$I^{\mathrm{DFG}}(\omega_0,\omega') \sim |\alpha_0|^2 \left|U_{s,i}(\omega_0,\omega')\right|^2, \tag{4.44}$$

which, as discussed in Section 4.1.3, is proportional to the photon-pair joint spectrum.

In the signal mode, the displacement can be decomposed in two contributions, one from the seed and one from the field produced by nonlinear optical frequency generation (the factoring in two displacement operators introduces a global phase, which we ignore):

$$\begin{aligned}&\exp\left(-\int d\omega\, d\omega'\, \alpha(\omega)\ \left[U_{s,s}(\omega,\omega')\right]^*\, \hat{a}_s^\dagger(\omega') - c.c.\right) = \\ &= \exp\left(-\int d\omega\, \alpha(\omega)\, \hat{a}_s^\dagger(\omega) - c.c.\right) \exp\left(-\int d\omega\, d\omega'\, \alpha(\omega)\ \left[U_{s,s}^{NL}(\omega,\omega')\right]^*\, \hat{a}_s^\dagger(\omega') - c.c.\right),\end{aligned} \tag{4.45}$$

where $U_{s,s}^{(NL)}(\omega,\omega') \equiv U_{s,s}(\omega,\omega') - \delta(\omega-\omega')$ is the part of the transfer function corresponding to the nonlinear interaction. In order to distinguish these two contributions, we propose two methods.

The first method we propose is to access the nonlinear signal generated in the seed mode by performing a second displacement, which cancels the seed contribution. The final state would then be modeled by $\hat{D}_s^{\mathrm{seed}\dagger}\, \hat{S}\, \hat{D}_s^{\mathrm{seed}}\, |0\rangle$. In an experimental setup, this can be effectively implemented using the polarisation degree of freedom of the seeding field, assuming that the PDC process is non-degenerate (phase-matching occurs at different frequencies for the signal and idler). The method that we propose is to seed in diagonal polarisation, such that one of the polarisations of the seed mode takes part in the nonlinear interaction, while the other propagates freely, serving as a reference. The output of the source is then analysed in anti-diagonal polarisation, such that the seed field is blocked, and half of the intensity corresponding to the signal generated by PDC (with either horizontal or vertical polarisation) is transmitted. Note that birefringence compensation must be used in order to obtain high-visibility interference between the two polarisation modes of the seed, after propagation through the nonlinear medium.

A complementary method for accessing the nonlinear signal generated in the seed mode is to use a narrow-band seed. In this case, the nonlinear contribution can be separated from the seed using an optical spectrum analyser (OSA), for all wavelengths outside of a small

bandwidth occupied by the seed. In practice, this method requires both a CW laser source and an OSA with a high dynamic range. This requirement may be relaxed by implementing the polarisation method described above.

4.2.2 Application to self-referenced estimation of squeezing parameters

We now consider a single Schmidt mode, with the squeezing parameter ζ. The squezing parameter can be determined by measuring the number of photon-pairs generated per pulse by the spontaneous process. This can be calculated using Equation 4.40, and is equal to $\sinh^2(\zeta)$. The difficulty in measuring this precisely stems from the fact that the collection (coupling and detection) efficiency must be determined first. In the weak squeezing regime, this calibration can be performed using a scheme known as the Klyshko method [280–283], which relies on the fact that SPDC photons are always generated in pairs. If a photon is detected in one mode, this heralds the emission of a second photon in the conjugate mode, providing a signal that can be then used to calibrate the collection efficiency. Extensions to the Klyshko method that go beyond the single photon-pair regime have been proposed, using photon number resolving detectors, and including the possibility of statistics generated by multi-mode squeezed states [284–286].

Using classical simulation, by seeding the PDC process, we can detemine the squeezing parameter using the seeded joint spectrum given by Equation 4.44. This requires estimating the overlap between the seeding beam and the PDC modes. Our novel idea is to use the signal generated in the seeded mode (corresponding to the transfer function $U_{s,s}$ or $U_{i,i}$) as a precise reference, which allows for the estimation of squeezing parameters that is independent of any calibration for efficiencies or mode overlaps. We detail this proposal below.

Let us first note that the total intensity of the signal generated in the seeded mode does not contain any additional information compared to the intensity of the DFG. This is because the number of photons generated in the two downconversion modes must be equal. We may gain some intuition on how this occurs for an arbitrary seed spectrum, by considering a simple case with a single Schmidt mode. Let the corresponding spectral distributions be the real functions $\psi_s(\omega)$ and $\psi_i(\omega)$, and let the seeding field spectrum be $\alpha(\omega)$, normalised such that $\int d\omega \, |\alpha(\omega)|^2 = 1$. The DFG signal has the spectral amplitude $-\langle \alpha \, | \psi_s \rangle \sinh(\zeta) \psi_i(\omega)$, where $\langle \psi \, | \alpha \rangle \equiv \int d\omega \, \psi(\omega) \alpha^*(\omega)$. In the seeded

mode, the displacement generated by PDC (given by the transfer function $U_{s,s}^{(NL)}(\omega,\omega') = U_{s,s}(\omega,\omega') - \delta(\omega-\omega')$), is $\langle\psi|\alpha\rangle\,(\cosh(\zeta)-1)\,\psi_s(\omega)$. The total displacement (including seed) is $\langle\psi|\alpha\rangle\,(\cosh(\zeta)-1)\,\psi_s(\omega)+\alpha(\omega)$. Using $\int d\omega|\psi_s(\omega)|^2 = \int d\omega|\psi_i(\omega)|^2 = \int d\omega|\alpha(\omega)|^2 = 1$, we can check that the total intensities of the generated fields in the seeded and unseeded modes are equal:

$$\int d\omega\ \left|\langle\psi|\alpha\rangle\,[\cosh(\zeta)-1]\,\psi_s(\omega)+\alpha(\omega)\right|^2 - \int d\omega\ |\alpha(\omega)|^2 = \int d\omega\ |\langle\psi|\alpha\rangle|^2\,|\sinh(\zeta)\psi_i(\omega)|^2\,. \tag{4.46}$$

In the first term on the left-hand side of this equation, we see that the coherent interplay between the seed field and the signal generated by PDC plays an important role in defining the output of the seeded arm. In particular, if we take a narrowband seed, the intensity of the broadband signal generated by PDC is proportional to $(\cosh(r)-1)^2$, at all frequencies apart from the seed frequency. Furthermore, just enough light is generated at the seed frequency as to make the total intensity equal to that of the DFG signal. We can use the ratio between the intensity of the broadband PDC generated in the seeded arm by a narrowband seed and the DFG intensity, which is proportional to $\sinh^2(\zeta)$, to determine the squeezing parameter, ζ. Both the DFG signal and the frequency generation in the seeded mode are linearly dependent on the overlap between seed and down-conversion mode and on the losses that occur at different steps in the setup. Therefore, we can use their ratio with no prior efficiency calibration, to determine $(\cosh(\zeta)-1)^2/\sinh(\zeta)^2 = \tanh(\zeta/2)^2$. This is why we describe our proposed method as being self-referenced. The optimal sensitivity given by $\partial/\partial\zeta\,\tanh^2(\zeta/2)$ occurs for $\sinh^2(\zeta)=3$, and more than half of the optimal sensitivity is achieved for $0.3 \lesssim \sinh^2(\zeta) \lesssim 500$.

Of course, for the general case of a spectrally multi-mode PDC source, the contributions of different Schmidt modes overlap in the measured signals. We thus propose and present an implementation in the following section, which measures the full transfer functions by scanning the narrowband seed. We suggest that these functions can then be fitted to a model, using the theoretical tools developed in Section 4.1.

A practical issue is that the collection efficiencies for the two down-conversion modes, $\eta_{s(i)}$, can differ slightly. However, if we perform the seeded measurements for each of the two modes, the efficiencies contribute factors of η_s/η_i and η_i/η_s, respectively, to the relevant

ratios for our estimation. Therefore, taking the geometric average of the results obtained for the two cases, we can eliminate the effect of differential loss. A special role is played by loss in the nonlinear medium, which affects the seeded and spontaneous processes differently. However, the transmittance of nonlinear crystals is high enough for this effect to be negligible. In our implementation, presented in the following section, the transmittance of the nonlinear wave-guides is $> 98\%$ [275].

4.3 Preliminary experimental results

Here, we present a measurement of all four transfer functions defined and discussed in the previous sections, for a pulsed PDC process, in a wave-guided ppKTP type-II source. The interpretation of these experimental results is work in progress. We only discuss the results here, and present the future objectives of our work.

4.3.1 Experimental details

We use a source engineered to produce nearly-degenerate signal and idler at telecommunication wavelengths ($\sim 1550\,\text{nm}$). The nonlinear medium is an 8 mm ppKTP chip purchased from ADVR, made from a z-cut wafer (the z direction is depth, while x is the propagation direction), with waveguides written using a proton exchange process [276]. The waveguides are roughly $4\,\mu\text{m} \times 6\,\mu\text{m}$ in cross section. The periodic poling has a spatial period of 104 µm. The chip output is AR coated for 1550 nm.

We pump the source using a mode-locked Ti:Sapphire laser (Coherent Chameleon), with central wavelength 783 nm and 8 nm bandwidth, which we filter using two bandpass thin-film filters (Semrock LL01-785-12.5) to a bandwidth of 1.9 nm. We pump the ppKTP chip in the H polarisation, with pulse energy 100 pJ. We used a pulse characterisation device (APE LX Spider) to ensure that the pump pulses are Fourier-transform limited, minimising spectral chirp with a dispersion compensation setup. The signal is generated in the H polarisation, with the central-wavelength at 1548 nm, and the idler, in V polarisation, with the central wavelength at 1583 nm.

After many iterations, we converged on a simple setup for measuring the seeded PDC signal, scanning a CW seed and resolving the output using an OSA. We do not use both polarisations of the seed, as described in Section 4.2.1, because we found that a secondary

phase-matched process overlaps with the signal and idler at our chosen wavelengths. This additional process generates V polarised light at 1548 nm and H polarised light at 1583 nm. It is due to a higher-order spatial mode for the pump, with a different effective refractive index. We find that the specifications of our equipment are high enough to allow us to perform all measurements with a simple setup, which is depicted in Figure 4.3.

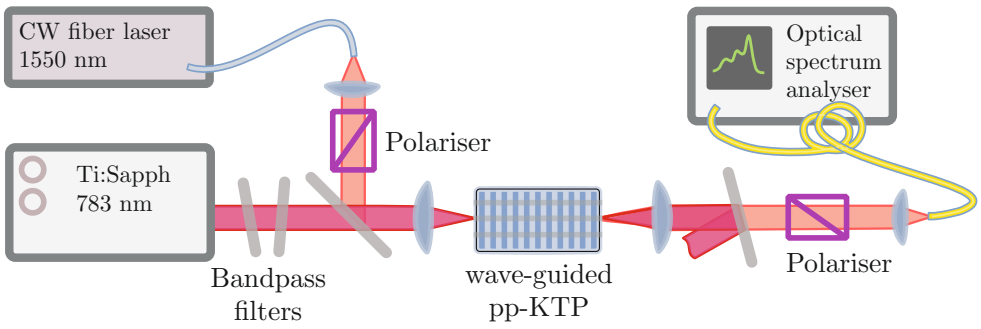


Figure 4.3: Experimental setup for seeded PDC measurements. We pump the ppKTP source with a mode-locked Ti:Sapphire laser and seed the process with a continuous wave fiber laser. We analyse the PDC signal using an OSA.

For seeding, we use a fiber laser (Yenista Tunics T100S-HP), with tunable wavelength, between 1500 nm and 1680 nm. The laser has a very high spontaneous noise suppression, of over 100 dB (measured over a bandwidth of 1 nm). The seed power was ≈ 2.5 mW, inside the waveguide. The OSA (Yokogawa AQ6370D) has an excellent dynamic range of 78 dB over 1 nm, ensuring that the weak PDC signal in the seed arm (< 1 nW) can be distinguished from the seed (2.5 mW). The OSA is blinded by the seed over a bandwidth of 1 nm. We simply interpolate the spectrum over this region, after performing background subtraction, as illustrated in Figure 4.4.

4.3.2 Seeded measurements

Figure 4.5 shows the main experimental results of this Chapter. It depicts a seeded measurement of all broadband PDC transfer functions defined in Section 4.1.2 (absolute-squared values). Panels (b) and (d) illustrate the transfer functions obtained by DFG, commonly associated to the joint spectra of photon pairs. These transfer functions resemble the example illustrated in Figure 4.5. As discussed in Section 4.1.3, two factors can be distinguished: the pump and the phase-matching functions, with the latter displaying characteristic oscillations. Panels (a) and (c) illustrate the PDC signals measured in the seeded mode,

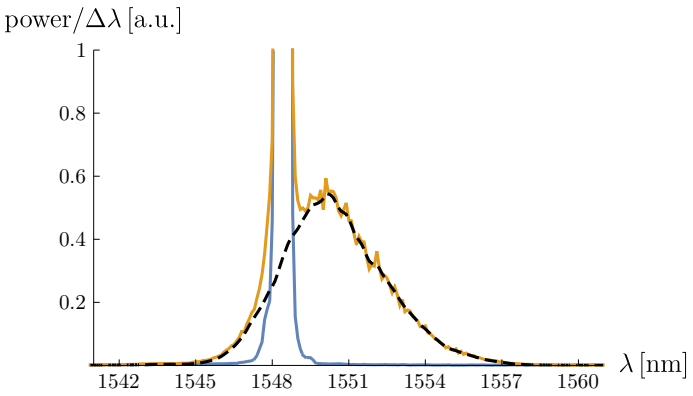


Figure 4.4: Example of a PDC spectrum measured in the seeded mode (orange), the background signal, taken with the pumped blocked (blue), and the spectrum after interpolation over the bandwidth blinded by the seed, and smoothing (black, dashed).

corresponding to the nonlinear terms in the signal-signal and idler-idler transfer functions. To our knowledge, these have not been measured before.

Our aim is to fit all measurements using the theoretical tools developed in Section 4.1, which will allow us to determine the absolute scale for the transfer functions. However, the measurements depicted in Figure 4.5 do not straight-away match our theoretical model, which suggests that an assumption made in our derivation does not correspond to the experimental setup. As an example, we depict, in Figure 4.6, the profile of the phase-matching function that corresponds to the joint spectrum depicted in panel (b) of Figure 4.5. For comparison, we display a sinc function, which should approximate well the profile of the phase-matching function according to our theoretical model. The data deviates significantly. This is due to a small, unwanted variation of the poling period and duty cycle along the wave-guide of our source. Our future efforts will concentrate on including this imperfection in the theoretical model and, at the same time searching for a different wave-guide, with a smaller deviation from the ideal poling.

The interpretation of the measurements presented in Figure 4.5 will be thus addressed in future research. As a rough preliminary analysis, we have performed singular value decompositions for all measured transfer functions (after taking the square root). We found that the principle eigenvectors are well overlapped, as illustrated in Figure 4.7, which suggests that the data agrees with the decomposition given by Equation 4.39. The ratios between the largest singular values corresponding to the transfer functions measured in the seeded arms and those corresponding to DFG suggest a squeezing parameter of ~ 0.8. This value

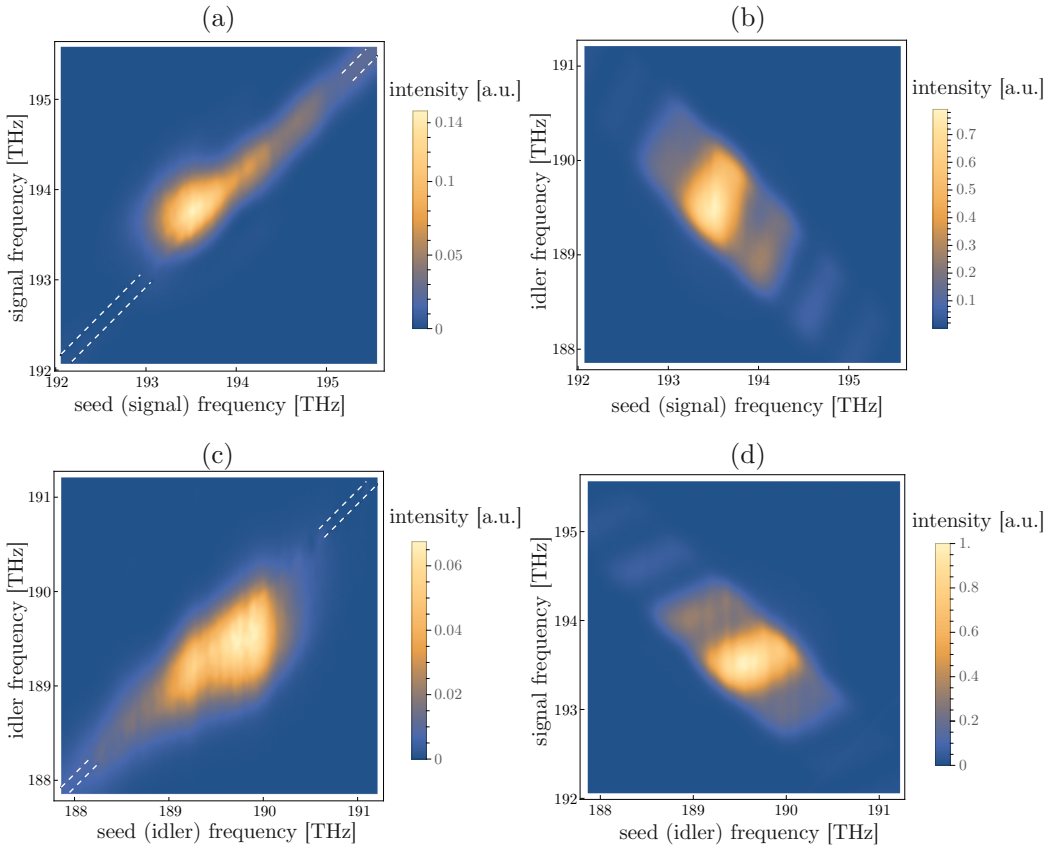


Figure 4.5: Measured transfer functions describing frequency generation in broadband PDC. (a) signal-signal, (b) signal-idler, (c) idler-idler and (d) idler-signal. The white dashed lines indicate the width of the diagonal region over which we interpolated the spectra, as the OSA was blinded by the seed.

falls in the expected range, which we can can check by estimating the squeezing parameter based on the DFG efficiency and APD count rates for the spontaneous process.

4.4 Summary and outlook

In this chapter, we presented a new experimental tool for the characterisation of broadband PDC sources which relies on a classical simulation method based on DFG. We described and analysed the broadband signal generated by PDC in the same mode as a seeding beam, linking it to the transfer functions describing the mode transformations of broadband squeezing. We determined that the intensity of this signal scales with the squeezing parameter in a different way than the DFG, and proposed an application in determining squeezing gain that does not require any calibration for efficiencies or mode overlaps. The method that we have

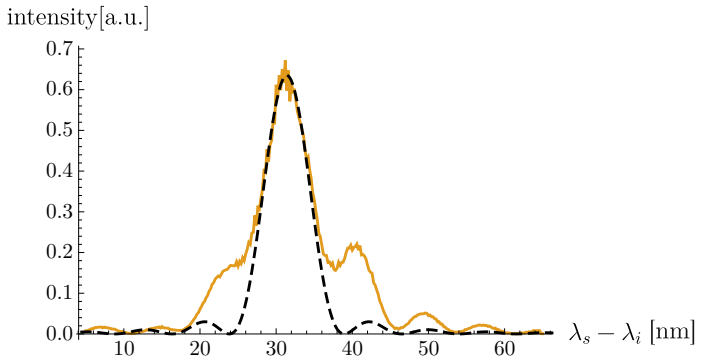


Figure 4.6: Measured phase-matching profile (orange) and, for comparison, a sinc function (dashed black line), as indicated by the theoretical model. We simply chose the sinc parameters such as to fit the main peak of the measured profile.

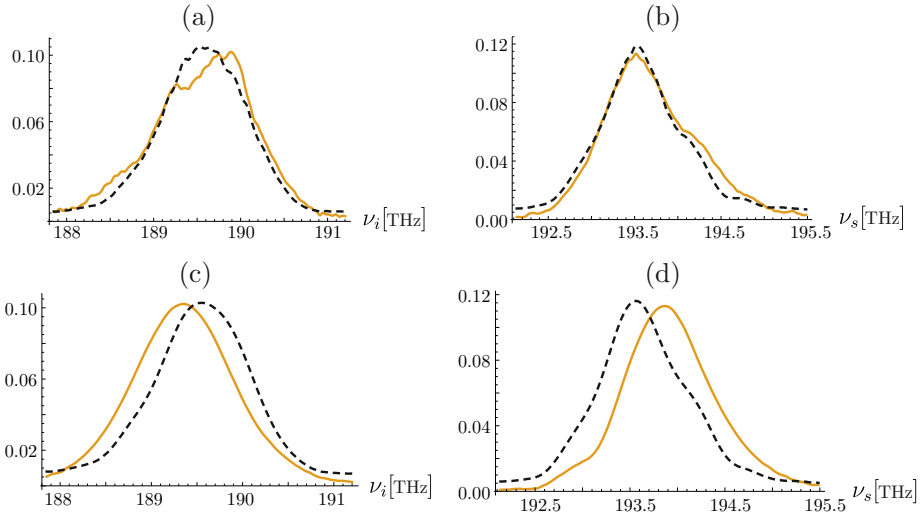


Figure 4.7: Principal eigenvectors of the measured transfer functions. These are obtained by performing a singular value decomposition of the square root of the measured transfer functions. The eigenvectors depicted in orange are determined from the PDC signal measured in the seeded mode, and those depicted with dashed black lines are determined from the DFG. (a) and (b) correspond to input modes, while (c) and (d) correspond to output modes. (a) and (c) correspond to idler modes, while (b) and (d) correspond to signal modes. The fact that these eigenvectors have similar shapes is a very rough indication that the experimental data corresponds to the expected decomposition of transfer functions given in Equation 4.39.

presented can be extended to the characterisation of bulk sources, where mode overlaps are more difficult to estimate.

Using a wave-guided PDC source, we measured all PDC signals generated by seeding. Our experimental results do not fit well with the derived theory. The disagreement is due to variable phase-matching conditions throughout the waveguide, but other effects could also play a role. By explaining the measured data, we will certainly gain new insights into the

workings of PDC sources. In this sense, the new type of measurement introduced here offers a useful experimental handle, making it a promising avenue of research.

Chapter 5

Conclusion

Through our research leading up to this thesis, we implemented three novel experimental approaches to optical quantum measurements.

In Chapter 2, we described the implementation of a phase estimation scheme that combines two-mode interferometry with double-homodyne measurement. We demonstrated its accuracy over the full 2π range of the estimated phase value, without implementing any adaptive phase technique. Our analysis showed that the estimation precision in the experimental realisation deviates by 5% from the fundamental limit (taking the finite efficiency into consideration). While, in principle, it is possible to use this scheme together with nonclassical interference in order to beat the shot noise limit, the required efficiencies are rather prohibitive. However, we demonstrated an interesting effect: that single-photon probe states can, for the same measurement, achieve better precision than weak coherent states. This is, in some ways, similar to the case of absorption estimation, where single-photon probes provide enhanced precision compared to classical light. The implication is that nonclassical signals can prove useful, even in the absence of nonclassical interference. Our findings may find applications in optical communications and imaging based on phase measurements.

Through our realisation of a photonic Maxwell's demon, presented in Chapter 3, we put forward experimental quantum optics as a testing platform for information thermodynamics. Future work is likely to concentrate on introducing nonclassical effects in setups that are similar to the one presented in this thesis. Our hope is that future research will also explore the implications of the information theoretic bound derived in this work. This result can be interpreted quite generally, as a relation between the information acquired about a system

and the extent by which that system can be perturbed (steered).

In Chapter 4, we presented an extension to the stimulated emission tomography method for the characterisation of parametric down-conversion sources. We reported experimental data corresponding to all four transfer functions that define the broadband mode transformations of a parametric down-conversion process. While the experimental results are somewhat difficult to interpret at the moment, we are confident that further analysis will shine some light on the subtleties of broadband parametric down-conversion. In this sense, the new type of measurement that we have described here will be a useful experimental tool.

Throughout the research presented in this thesis, we have highlighted the versatility of quantum optical measurements. We have shown that, by combining contrasting aspects of the electromagnetic field, we can obtain new experimental tools for the characterisation of optical processes, and advance towards practical and fundamental applications.

Bibliography

[1] A. Einstein. On a heuristic viewpoint concerning the production and transformation of light. *Annalen der Physik*, 17:132–148, 1905.

[2] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, 2014.

[4] J. F. Clauser. Experimental limitations to the validity of semiclassical radiation theories. *Phys. Rev. A*, 6:49–54, 1972.

[5] E. T. Jaynes. Survey of the present status of neoclassical radiation theory. In L. Mandel and E. Wolf, editors, *Coherence and Quantum Optics: Proceedings of the Third Rochester Conference on Coherence and Quantum Optics held at the University of Rochester, June 21–23, 1972*, pages 35–81. Springer US, Boston, MA, 1973.

[6] P. A. M. Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 114(767):243–265, 1927.

[7] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, Cambridge, 1995.

[8] C. Gerry and P. Knight. *Introductory Quantum Optics*. Cambridge University Press, Cambridge, 2005.

[9] K. Mølmer. Optical coherence: A convenient fiction. *Phys. Rev. A*, 55:3195–3203, 1997.

[10] S. Olivares. Quantum optics in the phase space. *The European Physical Journal Special Topics*, 203(1):3–24, 2012.

[11] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[12] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2010.

[13] E. A. Martinez and J. P. Paz. Dynamics and thermodynamics of linear quantum open systems. *Phys. Rev. Lett.*, 110:130406, 2013.

[14] R. Boyd. *Nonlinear Optics*, chapter 2. Nonlinear Optics Series. Elsevier Science, San Diego, CA, 2008.

[15] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich. Generation of optical harmonics. *Phys. Rev. Lett.*, 7:118–119, 1961.

[16] T. H. Maiman. Stimulated optical radiation in ruby. *Nature*, 187:493–494, 1960.

[17] M. Bass, P. A. Franken, J. F. Ward, and G. Weinreich. Optical rectification. *Phys. Rev. Lett.*, 9:446–448, 1962.

[18] C. E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55, 2001.

[19] E. Jaynes and G. Bretthorst. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.

[20] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics*, 2(1-4):157–168, 1968.

[21] J. V. Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in Physics. Princeton University Press, Princeton, NJ, 1955.

[22] M. Tribus and E. McIrvine. Energy and information. *Scientific American*, (224):178–184, 1971.

[23] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.

[24] W. Press. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2007.

[25] P. Hariharan. *Optical Interferometry*. Electronics & Electrical. Academic Press, San Diego, CA, 2003.

[26] M. Pitkin, S. Reid, S. Rowan, and J. Hough. Gravitational wave detection by interferometry (ground and space). *Living Reviews in Relativity*, 14(5), 2011.

[27] LIGO Scientific Collaboration and Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, 2016.

[28] C. M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23:1693–1708, 1981.

[29] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004.

[30] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46:R6797–R6800, 1992.

[31] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen. Squeezed atomic states and projection noise in spectroscopy. *Phys. Rev. A*, 50:67–88, 1994.

[32] J. Combes and H. M. Wiseman. States for phase estimation in quantum interferometry. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(1):14, 2005.

[33] M. Auzinsh, D. Budker, D. F. Kimball, S. M. Rochester, J. E. Stalnaker, A. O. Sushkov, and V. V. Yashchuk. Can a quantum nondemolition measurement improve the sensitivity of an atomic magnetometer? *Phys. Rev. Lett.*, 93:173002, 2004.

[34] A. Guillaume and J. P. Dowling. Heisenberg-limited measurements with superconducting circuits. *Phys. Rev. A*, 73:040304, 2006.

[35] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen. Biological measurement beyond the quantum limit. *Nature Photonics*, 7:229–233, 2013.

[36] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde. Entanglement-enhanced measurement of a completely unknown optical phase. *Nature Photonics*, 5:43–47, 2011.

[37] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington. Adaptive optical phase estimation using time-symmetric quantum smoothing. *Phys. Rev. Lett.*, 104:093601, 2010.

[38] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kolodynski. Quantum limits in optical interferometry. *ArXiv e-prints*, arxiv:1405.7703, 2014.

[39] M. Cooper, C. Söller, and B. J. Smith. High-stability time-domain balanced homodyne detector for ultrafast optical pulse applications. *Journal of Modern Optics*, 60:611–616, May 2013.

[40] P. J. Mosley, J. S. Lundeen, B. J. Smith, P. Wasylczyk, A. B. U'Ren, C. Silberhorn, and I. A. Walmsley. Heralded generation of ultrafast single photons in pure quantum states. *Phys. Rev. Lett.*, 100:133601, 2008.

[41] T. L. S. Collaboration. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, 2015.

[42] V. B. Braginski and Y. I. Vorontsov. Quantum-mechanical limitations in macroscopic experiments and modern experimental technique. *Soviet Physics Uspekhi*, 17(5):644, 1975.

[43] W. A. Edelstein, J. Hough, J. R. Pugh, and W. Martin. Limits to the measurement of displacement in an interferometric gravitational radiation detector. *Journal of Physics E: Scientific Instruments*, 11(7):710, 1978.

[44] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle. *Rev. Mod. Phys.*, 52:341–392, 1980.

[45] C. M. Caves. Quantum-mechanical radiation-pressure fluctuations in an interferometer. *Phys. Rev. Lett.*, 45:75–79, 1980.

[46] M. Ozawa. *Realization of Measurement and the Standard Quantum Limit*, pages 263–286. Springer US, Boston, MA, 1989.

[47] V. B. Braginsky, M. L. Gorodetsky, F. Y. Khalili, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin. Noise in gravitational-wave detectors and other classical-force measurements is not influenced by test-mass quantization. *Phys. Rev. D*, 67(8):082001, 2003.

[48] B. Yurke, S. L. McCall, and J. R. Klauder. SU(2) and SU(1,1) interferometers. *Phys. Rev. A*, 33:4033–4054, 1986.

[49] M. Jarzyna and R. Demkowicz-Dobrzański. Matrix product states for quantum metrology. *Phys. Rev. Lett.*, 110:240405.

[50] N. Killoran, M. Cramer, and M. B. Plenio. Extracting entanglement from identical particles. *Phys. Rev. Lett.*, 112:150501, 2014.

[51] P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta. Squeezed-light ˘ enhanced polarization interferometer. *Phys. Rev. Lett.*, 59:2153–2156, 1987.

[52] M. Xiao, L.-A. Wu, and H. J. Kimble. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.*, 59:278–281, Jul 1987.

[53] T. L. S. Collaboration. A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nature Physics*, 7(12):962–965, 2011.

[54] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel. Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector GEO 600. *Phys. Rev. A*, 88:041802, 2013.

[55] P. Meystre, M. Scully, and N. A. T. O. S. A. Division. *Quantum optics, experimental gravity, and measurement theory*. NATO advanced science institutes series: Physics. Plenum Press, 1983.

[56] H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin. Conversion of conventional gravitational-wave interferometers into quantum nondemolition interferometers by modifying their input and/or output optics. *Phys. Rev. D*, 65:022002, 2001.

[57] R. Schnabel, J. Harms, K. A. Strain, and K. Danzmann. Squeezed light for the interferometric detection of high-frequency gravitational waves. *Classical and Quantum Gravity*, 21(5):S1045, 2004.

[58] M. J. Holland and K. Burnett. Interferometric detection of optical phase shifts at the Heisenberg limit. *Phys. Rev. Lett.*, 71:1355–1358, 1993.

[59] S. Boixo, S. T. Flammia, C. M. Caves, and J. Geremia. Generalized limits for single-parameter quantum estimation. *Phys. Rev. Lett.*, 98:090401, 2007.

[60] R. Demkowicz-Dobrzański and L. Maccone. Using entanglement against noise in quantum metrology. *Phys. Rev. Lett.*, 113:250801, 2014.

[61] E. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Texts in Statistics. Springer New York, 2003.

[62] M. G. A. Paris. Quantum estimation for quantum technology. *ArXiv e-prints*, arxiv:0804.2981, 2008.

[63] H. Cramér. *Mathematical Methods of Statistics*. Princeton landmarks in mathematics and physics. Princeton University Press, 1999.

[64] D. Hinkley. *Theory of Statistical Estimation: The 1925 Paper*, pages 85–94. Springer New York, New York, NY, 1980.

[65] S. L. Braunstein. Quantum limits on precision measurements of phase. *Phys. Rev. Lett.*, 69:3598–3601, 1992.

[66] C. Helstrom. *Quantum Detection and Estimation Theory*. Mathematics in Science and Engineering : a series of monographs and textbooks. Academic Press, 1976.

[67] A. Holevo. *Probabilistic and statistical aspects of quantum theory*. North-Holland series in statistics and probability. North-Holland Pub. Co., 1982.

[68] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72:3439–3443, 1994.

[69] O. E. Barndorff-Nielsen and R. D. Gill. Fisher information in quantum statistics. *Journal of Physics A: Mathematical and General*, 33(24):4481, 2000.

[70] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman. How to perform the most accurate possible phase measurements. *Phys. Rev. A*, 80:052114, 2009.

[71] S. Boixo, C. M. Caves, A. Datta, and A. Shaji. On decoherence in quantum clock synchronization. *Laser Physics*, 16(11), 2006.

[72] R. Demkowicz-Dobrzański. Multi-pass classical vs. quantum strategies in lossy phase estimation. *Laser Physics*, 20(5):1197–1202, 2010.

[73] M. Jarzyna and R. Demkowicz-Dobrzański. Quantum interferometry with and without an external phase reference. *Phys. Rev. A*, 85:011801, 2012.

[74] L. Pezzé and A. Smerzi. Mach-zehnder interferometry at the heisenberg limit with coherent and squeezed-vacuum light. *Phys. Rev. Lett.*, 100:073601, 2008.

[75] G. Summy and D. Pegg. Phase optimized quantum states of light. *Optics Communications*, 77(1):75–79, 1990.

[76] Z. Hradil, R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch. Quantum phase in interferometry. *Phys. Rev. Lett.*, 76:4295–4298, 1996.

[77] D. W. Berry and H. M. Wiseman. Optimal states and almost optimal adaptive measurements for quantum interferometry. *Phys. Rev. Lett.*, 85:5098–5101, 2000.

[78] L. Pezzé, A. Smerzi, G. Khoury, J. F. Hodelin, and D. Bouwmeester. Phase detection at the quantum limit with multiphoton Mach-Zehnder interferometry. *Phys. Rev. Lett.*, 99:223602, 2007.

[79] S. M. Barnett and D. T. Pegg. Quantum theory of rotation angles. *Phys. Rev. A*, 41:3427–3435, 1990.

[80] U. Leonhardt, J. A. Vaccaro, B. Böhmer, and H. Paul. Canonical and measured phase distributions. *Phys. Rev. A*, 51:84–95, 1995.

[81] B. C. Sanders and G. J. Milburn. Optimal quantum measurements for phase estimation. *Phys. Rev. Lett.*, 75:2944–2947, 1995.

[82] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi. Beating the standard quantum limit with four-entangled photons. *Science*, 316(5825):726–729, 2007.

[83] M. Kacprowicz, R. Demkowicz-Dobrzanski, W. Wasilewski, K. Banaszek, and I. A. Walmsley. Experimental quantum-enhanced estimation of a lossy phase shift. *Nature Photonics*, 4(6):357–360, 2009.

[84] U. Dorner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley. Optimal quantum phase estimation. *Phys. Rev. Lett.*, 102:040403, 2009.

[85] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac. Improvement of frequency standards with quantum entanglement. *Phys. Rev. Lett.*, 79:3865–3868, 1997.

[86] M. G. Genoni, S. Olivares, and M. G. A. Paris. Optical phase estimation in the presence of phase diffusion. *Phys. Rev. Lett.*, 106:153603, 2011.

[87] B. M. Escher, R. L. de Matos Filho, and L. Davidovich. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. *Nature Physics*, 7(5):406–411, 2011.

[88] B. M. Escher, L. Davidovich, N. Zagury, and R. L. de Matos Filho. Quantum metrological limits via a variational approach. *Phys. Rev. Lett.*, 109:190404, 2012.

[89] R. Demkowicz-Dobrzanski, J. Kolodynski, and M. Guta. The elusive Heisenberg limit in quantum-enhanced metrology. *Nature Communications*, 3:1063+, 2012.

[90] M. Jarzyna and R. Demkowicz-Dobrzaski. True precision limits in quantum metrology. *New Journal of Physics*, 17(1):013010, 2015.

[91] S. I. Knysh, E. H. Chen, and G. A. Durkin. True limits to precision via unique quantum probe. *ArXiv e-prints*, ArXiv:1402.0495, 2014.

[92] K. Macieszczak. Upper bounds on the quantum Fisher information in the presence of general dephasing. *ArXiv e-prints*, ArXiv:1403.0955, 2014.

[93] J. Kołodyński and R. Demkowicz-Dobrzański. Phase estimation without a priori phase knowledge in the presence of loss. *Phys. Rev. A*, 82:053804, 2010.

[94] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín. Noisy metrology beyond the standard quantum limit. *Phys. Rev. Lett.*, 111:120401, 2013.

[95] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin. Quantum error correction for metrology. *Phys. Rev. Lett.*, 112:150802, Apr 2014.

[96] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus. Improved quantum metrology using quantum error correction. *Phys. Rev. Lett.*, 112:080801, 2014.

[97] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker. Increasing sensing resolution with error correction. *Phys. Rev. Lett.*, 112:150801, 2014.

[98] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg. Super-resolving phase measurements with a multiphoton entangled state. *Nature*, 429(6988):161–164, 2004.

[99] R. Okamoto, H. F. Hofmann, T. Nagata, J. L. O'Brien, K. Sasaki, and S. Takeuchi. Beating the standard quantum limit: phase super-sensitivity of n -photon interferometers. *New Journal of Physics*, 10(7):073033, 2008.

[100] D. Sych and G. Leuchs. Coherent state quantum key distribution with multi letter phase-shift keying. *New Journal of Physics*, 12(5):053019, 2010.

[101] M. Jachura, R. Chrapkiewicz, R. Demkowicz-Dobrzański, W. Wasilewski, and K. Banaszek. Mode engineering for realistic quantum-enhanced interferometry. *Nature Communications*, 7:11411+, 2016.

[102] J. W. Noh, A. Fougères, and L. Mandel. Measurement of the quantum phase by photon counting. *Phys. Rev. Lett.*, 67:1426–1429, 1991.

[103] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley. Joint estimation of phase and phase diffusion for quantum metrology. *Nature Communications*, 5:3532, April 2014.

[104] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde. Entanglement-free Heisenberg-limited phase estimation. *Nature*, 450:393–396, 2007.

[105] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde. Demonstrating heisenberg-limited unambiguous phase estimation without adaptive measurements. *New Journal of Physics*, 11(7):073023, 2009.

[106] H. M. Wiseman and R. B. Killip. Adaptive single-shot phase measurements: The full quantum theory. *Phys. Rev. A*, 57:2169–2185, 1998.

[107] D. W. Berry and H. M. Wiseman. Phase measurements at the theoretical limit. *Phys. Rev. A*, 63(1):013813, 2001.

[108] M. J. W. Hall and I. G. Fuss. Quantum phase detection and digital communication. *Quantum Optics: Journal of the European Optical Society Part B*, 3(3):147, 1991.

[109] H. M. Wiseman. Adaptive phase measurements of optical modes: Going beyond the marginal q distribution. *Phys. Rev. Lett.*, 75:4587–4590, 1995.

[110] H. M. Wiseman. Quantum trajectories and quantum measurement theory. *Quantum and Semiclassical Optics: Journal of the European Optical Society Part B*, 8(1):205, 1996.

[111] H. M. Wiseman and R. B. Killip. Adaptive single-shot phase measurements: A semiclassical approach. *Phys. Rev. A*, 56:944–957, 1997.

[112] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi. Adaptive homodyne measurement of optical phase. *Phys. Rev. Lett.*, 89:133602, 2002.

[113] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa. Quantum-enhanced optical-phase tracking. *Science*, 337(6101):1514–1517, 2012.

[114] K. Iwasawa, K. Makino, H. Yonezawa, M. Tsang, A. Davidovic, E. Huntington, and A. Furusawa. Quantum-limited mirror-motion estimation. *Phys. Rev. Lett.*, 111:163602, 2013.

[115] A. A. Berni, T. Gehring, B. M. Nielsen, V. Händchen, M. G. A. Paris, and U. L. Andersen. Ab initio quantum-enhanced optical phase estimation using real-time feedback control. *Nature Photonics*, 9:577–581, 2015.

[116] M. Szczykulska, T. Baumgratz, and A. Datta. Multi-parameter Quantum Metrology. *ArXiv e-prints*, arxiv:1604.02615, 2016.

[117] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley. Quantum enhanced multiple phase estimation. *Phys. Rev. Lett.*, 111:070403, 2013.

[118] P. J. D. Crowley, A. Datta, M. Barbieri, and I. A. Walmsley. Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry. *Phys. Rev. A*, 89:023845, 2014.

[119] C. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, 1967.

[120] R. D. Gill and S. Massar. State estimation for large ensembles. *Phys. Rev. A*, 61:042312, 2000.

[121] N. Li, C. Ferrie, J. A. Gross, A. Kalev, and C. M. Caves. Fisher-symmetric informationally complete measurements for pure states. *Phys. Rev. Lett.*, 116:180402, 2016.

[122] C. M. Caves and P. D. Drummond. Quantum limits on bosonic communication rates. *Rev. Mod. Phys.*, 66:481–537, 1994.

[123] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.*, 70:1244–1247, 1993.

[124] K. Vogel and H. Risken. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 40:2847–2849, 1989.

[125] A. I. Lvovsky and M. G. Raymer. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.*, 81:299–332, 2009.

[126] H. P. Yuen and V. W. S. Chan. Noise in homodyne and heterodyne detection. *Opt. Lett.*, 8(3):177–179, 1983.

[127] J. Graeme. *Photodiode Amplifiers: Op Amp Solutions*. Gain technology. McGraw Hill, New York, 1996.

[128] A. Ferraro, S. Olivares, and M. G. A. Paris. Gaussian states in continuous variable quantum information. *eprint arXiv:quant-ph/0503237*, 2005.

[129] H. Srivastava, H. Mavromatis, and R. Alassar. Remarks on some associated Laguerre integral results. *Applied Mathematics Letters*, 16(7):1131–1136, 2003.

[130] H. F. Hofmann. All path-symmetric pure states achieve their maximal phase sensitivity in conventional two-path interferometry. *Phys. Rev. A*, 79:033822, 2009.

[131] O. Morin, V. D'Auria, C. Fabre, and J. Laurat. High-fidelity single-photon source based on a type ii optical parametric oscillator. *Opt. Lett.*, 37(17):3738–3740, 2012.

[132] H. Hansen, T. Aichele, C. Hettich, P. Lodahl, A. I. Lvovsky, J. Mlynek, and S. Schiller. Ultrasensitive pulsed, balanced homodyne detector:application to time-domain quantum measurements. *Opt. Lett.*, 26(21):1714–1716, 2001.

[133] A. Zavatta, M. Bellini, P. L. Ramazza, F. Marin, and F. T. Arecchi. Time-domain analysis of quantum states of light: noise characterization and homodyne tomography. *J. Opt. Soc. Am. B*, 19(5):1189–1194, 2002.

[134] R. Okubo, M. Hirano, Y. Zhang, and T. Hirano. Pulse-resolved measurement of quadrature phase amplitudes of squeezed pulse trains at a repetition rate of 76 mhz. *Opt. Lett.*, 33(13):1458–1460, 2008.

[135] M. Cooper, L. J. Wright, C. Söller, and B. J. Smith. Experimental generation of multi-photon fock states. *Opt. Express*, 21(5):5309–5317, 2013.

[136] U. L. Andersen, J. S. Neergaard-Nielsen, P. van Loock, and A. Furusawa. Hybrid discrete- and continuous-variable quantum information. *Nature Physics*, 11:713–719, 2015.

[137] L. W. Nagel and D. Pederson. Spice (simulation program with integrated circuit emphasis). Technical Report UCB/ERL M382, EECS Department, University of California, Berkeley, 1973.

[138] M. Cooper, M. Karpinski, and B. Smith. Local mapping of detector response for reliable quantum state estimation. *Nature Communications*, 5:4332, 2014.

[139] M. Cooper, E. Slade, M. Karpiński, and B. J. Smith. Characterization of conditional state-engineering quantum processes by coherent state quantum process tomography. *New Journal of Physics*, 17(3):033041, 2015.

[140] M. F. Cooper. *Generation and manipulation of quantum states of travelling light fields*. PhD thesis, Oxford University, 2014.

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

[142] R. Kashyap, editor. *Fiber Bragg Gratings (Second Edition)*. Academic Press, Boston, 2010.

[143] H. Leff and A. Rex. *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing*. Taylor & Francis, London, 2002.

[144] K. Maruyama, F. Nori, and V. Vedral. *Colloquium* : The physics of Maxwell's demon and information. *Rev. Mod. Phys.*, 81:1–23, 2009.

[145] V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh. A molecular information ratchet. *Nature*, 445(7127):523–527, 2007.

[146] G. N. Price, S. T. Bannerman, K. Viering, E. Narevicius, and M. G. Raizen. Single-photon atomic cooling. *Phys. Rev. Lett.*, 100:093004, 2008.

[147] E. Narevicius, S. T. Bannerman, and M. G. Raizen. Single-photon molecular cooling. *New Journal of Physics*, 11(5):055046, 2009.

[148] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano. Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. *Nature Physics*, 6(12):988–992, 2010.

[149] A. Berut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, 2012.

[150] E. Roldan, I. A. Martinez, J. M. R. Parrondo, and D. Petrov. Universal features in the energetics of symmetry breaking. *Nat Phys*, 10(6):457–461, 2014.

[151] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin. Experimental realization of a Szilard engine with a single electron. *Proceedings of the National Academy of Sciences*, 111(38):13786–13789, 2014.

[152] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola. Experimental observation of the role of mutual information in the nonequilibrium dynamics of a Maxwell Demon. *Phys. Rev. Lett.*, 113:030601, 2014.

[153] J. V. Koski, A. Kutvonen, I. M. Khaymovich, T. Ala-Nissila, and J. P. Pekola. On-chip Maxwell's Demon as an information-powered refrigerator. *Phys. Rev. Lett.*, 115:260602, 2015.

[154] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley. Photonic Maxwell's Demon. *Phys. Rev. Lett.*, 116:050401, 2016.

[155] J. Maxwell. *Theory of Heat*. Longmans, London, 1872.

[156] L. Szilard. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, 9(4):301–310, 1964.

[157] L. Brillouin. Maxwell's Demon Cannot Operate: Information and Entropy. I. *Journal of Applied Physics*, 22(3):334–337, 1951.

[158] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961.

[159] J. A. Vaccaro and S. M. Barnett. Information erasure without an energy cost. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 467(2130):1770–1778, 2011.

[160] C. H. Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.

[161] C. H. Bennett. Notes on Landauer's principle, reversible computation, and Maxwell's Demon. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 34(3):501–510, 2003.

[162] W. H. Zurek. Algorithmic randomness and physical entropy. *Phys. Rev. A*, 40:4731–4751, 1989.

[163] C. M. Caves. Quantitative limits on the ability of a Maxwell demon to extract work from heat. *Phys. Rev. Lett.*, 64:2111–2114, 1990.

[164] C. M. Caves, W. G. Unruh, and W. H. Zurek. Comment on "Quantitative limits on the ability of a Maxwell demon to extract work from heat". *Phys. Rev. Lett.*, 65:1387–1387, 1990.

[165] K. Shizume. Heat generation required by information erasure. *Phys. Rev. E*, 52:3495–3499, 1995.

[166] B. Piechocinska. Information erasure. *Phys. Rev. A*, 61:062314, 2000.

[167] W. Zurek. *Complexity, Entropy, and the Physics of Information: The Proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information Held May-June, 1989, in Santa Fe, New Mexico.* Proceedings volume in the Santa Fe Institute studies in the sciences of complexity. Addison-Wesley, 1990.

[168] A. Peres. *Quantum Theory: Concepts and Methods.* Fundamental Theories of Physics. Springer, New York, 1995.

[169] A. S. Holevo. Some estimates for the amount of information transmittable by a quantum communications channel. *Problemy Peredači Informacii*, 9(3):3–11, 1973.

[170] M. B. Plenio and V. Vitelli. The physics of forgetting: Landauer's erasure principle and information theory. *Contemporary Physics*, 42:25–60, 2001.

[171] V. Vedral. Landauer's erasure, error correction and entanglement. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 456(1996):969–984, 2000.

[172] B. Groisman, S. Popescu, and A. Winter. Quantum, classical, and total amount of correlations in a quantum state. *Phys. Rev. A*, 72:032317, 2005.

[173] W. H. Zurek. Quantum discord and Maxwell's demons. *Phys. Rev. A*, 67:012320, 2003.

[174] K. Maruyama, F. Morikoshi, and V. Vedral. Thermodynamical detection of entanglement by Maxwell's demons. *Phys. Rev. A*, 71:012108, 2005.

[175] M. Smoluchowski. Experimentell nachweisbare, der blichen thermodynamik widersprechende molekularphnomene. *Pisma Mariana Smoluchowskiego*, 2(1):226–251, 1927.

[176] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics*, volume 1, chapter 46. Addison-Wesley, Boston, 1963.

[177] B. Huard et al. Experimental realisation of Maxwell's demon using a superconducting qubit. Conference presentation at the Statistical Mechanics of Quantum Dynamics workshop, Aalto University. 2016.

[178] F. Ritort. Single-molecule experiments in biological physics: methods and applications. *Journal of Physics: Condensed Matter*, 18(32):R531, 2006.

[179] Alemany, Anna and Ritort, Felix. Fluctuation theorems in small systems: extending thermodynamics to the nanoscale. *Europhysics News*, 41(2):27–30, 2010.

[180] F. Jülicher, A. Ajdari, and J. Prost. Modeling molecular motors. *Rev. Mod. Phys.*, 69:1269–1282, 1997.

[181] V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh. A molecular information ratchet. *Nature*, 445(7127):523–527, 2007.

[182] G. E. Crooks. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E*, 61:2361–2366, 2000.

[183] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012.

[184] R. Klages, W. Just, C. Jarzynski, and H. Schuster. *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*. Annual Reviews of Nonlinear Dynamics and Complexity (VCH). Wiley, Oxford, 2013.

[185] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.

[186] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, 296(5574):1832–1835, 2002.

[187] O. Karlström, H. Linke, G. Karlström, and A. Wacker. Increasing thermoelectric performance using coherent transport. *Phys. Rev. B*, 84:113415, 2011.

[188] H. Vazquez, R. Skouta, S. Schneebeli, M. Kamenetska, R. Breslow, L. Venkataraman, and M. Hybertsen. Probing the conductance superposition law in single-molecule circuits with parallel paths. *Nature Nanotechnology*, 7(10):663–667, 2012.

[189] S. V. Aradhya and L. Venkataraman. Single-molecule junctions beyond electronic transport. *Nature Nanotechnology*, 8(6):399–410, 2013.

[190] N. Lambert, Y.-N. Chen, Y.-C. Cheng, C.-M. Li, G.-Y. Chen, and F. Nori. Quantum biology. *Nature Physics*, 9:10–18, 2013.

[191] H. T. Quan, Y.-x. Liu, C. P. Sun, and F. Nori. Quantum thermodynamic cycles and quantum heat engines. *Phys. Rev. E*, 76:031105, 2007.

[192] M. O. Scully, M. S. Zubairy, G. S. Agarwal, and H. Walther. Extracting work from a single heat bath via vanishing quantum coherence. *Science*, 299(5608):862–864, 2003.

[193] T. D. Kieu. The second law, Maxwell's demon, and work derivable from quantum heat engines. *Phys. Rev. Lett.*, 93:140403, 2004.

[194] K. Brandner, M. Bauer, M. T. Schmid, and U. Seifert. Coherence-enhanced efficiency of feedback-driven quantum engines. *New Journal of Physics*, 17(6):065006, 2015.

[195] M. T. Mitchison, M. P. Woods, J. Prior, and M. Huber. Coherence-assisted single-shot cooling by quantum absorption refrigerators. *New Journal of Physics*, 17(11):115013, 2015.

[196] R. Uzdin, A. Levy, and R. Kosloff. Equivalence of quantum heat machines, and quantum-thermodynamic signatures. *Phys. Rev. X*, 5:031044, 2015.

[197] R. Kosloff and A. Levy. Quantum heat engines and refrigerators: Continuous devices. *Annual Review of Physical Chemistry*, 65(1):365–393, 2014.

[198] A. Levy and R. Kosloff. Quantum absorption refrigerator. *Phys. Rev. Lett.*, 108:070604, 2012.

[199] H. E. D. Scovil and E. O. Schulz-DuBois. Three-level masers as heat engines. *Phys. Rev. Lett.*, 2:262–263, 1959.

[200] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Phys.*, 48(2):119–130, 1976.

[201] A. Levy, R. Alicki, and R. Kosloff. Quantum refrigerators and the third law of thermodynamics. *Phys. Rev. E*, 85:061126, 2012.

[202] N. Brunner, N. Linden, S. Popescu, and P. Skrzypczyk. Virtual qubits, virtual temperatures, and the foundations of thermodynamics. *Phys. Rev. E*, 85:051117, 2012.

[203] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, 1996.

[204] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, 2009.

[205] M. Horodecki, P. Horodecki, and J. Oppenheim. Reversible transformations from pure to mixed states and the unique measure of information. *Phys. Rev. A*, 67:062104, 2003.

[206] G. Gour and R. W. Spekkens. The resource theory of quantum reference frames: manipulations and monotones. *New Journal of Physics*, 10(3):033023, 2008.

[207] M. Horodecki and J. Oppenheim. (Quantumness in the context of) resource theories. *International Journal of Modern Physics B*, 27(01n03):1345019, 2013.

[208] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens. Resource theory of quantum states out of thermal equilibrium. *Phys. Rev. Lett.*, 111:250404, 2013.

[209] P. Skrzypczyk, A. J. Short, and S. Popescu. Work extraction and thermodynamics for individual quantum systems. *Nature Communications*, 5:4185, 2014.

[210] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral. Inadequacy of von neumann entropy for characterizing extractable work. *New Journal of Physics*, 13(5):053015, 2011.

[211] M. Horodecki and J. Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4:2059, 2013.

[212] M. Lostaglio, D. Jennings, and T. Rudolph. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nature Communications*, 6, 2015.

[213] S. Deffner and C. Jarzynski. Information processing and the second law of thermodynamics: An inclusive, Hamiltonian approach. *Phys. Rev. X*, 3:041003, 2013.

[214] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings. The extraction of work from quantum coherence. *New Journal of Physics*, 18(2):023045, 2016.

[215] M. F. Frenzel, D. Jennings, and T. Rudolph. Quasi-autonomous quantum thermal machines and quantum to classical energy flow. *New Journal of Physics*, 18(2):023037, 2016.

[216] C. Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E*, 56:5018–5035, 1997.

[217] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2726, 1999.

[218] C. H. Bennett. Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics*, 22(2):245–268, 1976.

[219] T. Sagawa and M. Ueda. Generalized jarzynski equality under nonequilibrium feedback control. *Phys. Rev. Lett.*, 104:090602, 2010.

[220] Y. Morikuni and H. Tasaki. Quantum Jarzynski-Sagawa-Ueda relations. *Journal of Statistical Physics*, 143(1):1–10, 2011.

[221] P. Talkner, E. Lutz, and P. Hänggi. Fluctuation theorems: Work is not an observable. *Phys. Rev. E*, 75:050102, 2007.

[222] C. Jarzynski, H. T. Quan, and S. Rahav. Quantum-classical correspondence principle for work distributions. *Phys. Rev. X*, 5:031038, 2015.

[223] R. Chetrite and K. Mallick. Quantum fluctuation relations for the Lindblad master equation. *Journal of Statistical Physics*, 148(3):480–501, 2012.

[224] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.*, 89:050601, 2002.

[225] F. Douarche, S. Ciliberto, A. Petrosyan, and I. Rabbiosi. An experimental test of the Jarzynski equality in a mechanical experiment. *Europhysics Letters*, 70(5):593, 2005.

[226] S. An, J.-N. Zhang, M. Um, D. Lv, Y. Lu, J. Zhang, Z.-Q. Yin, H. Quan, and K. Kim. Experimental test of the quantum Jarzynski equality with a trapped-ion system. *Nature Physics*, 11(2):193–199, 2015.

[227] S. Bose. Planck's law and the light quantum hypothesis. *Journal of Astrophysics and Astronomy*, 15(1):3–7, 1994.

[228] R. H. Brown and R. Q. Twiss. Interferometry of the intensity fluctuations in light ii. an experimental test of the theory for partially coherent light. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 243(1234):291–319, 1958.

[229] R. Hanbury Brown. A test of a new type of stellar interferometer on Sirius. *Nature*, 178:1046–1048, 1956.

[230] F. T. Arecchi. Measurement of the statistical distribution of Gaussian and laser sources. *Phys. Rev. Lett.*, 15:912–916, Dec 1965.

[231] A. Valencia, G. Scarcelli, M. D'Angelo, and Y. Shih. Two-photon imaging with thermal light. *Phys. Rev. Lett.*, 94:063601, 2005.

[232] V. Parigi, A. Zavatta, and M. Bellini. Implementation of single-photon creation and annihilation operators: experimental issues in their application to thermal states of light. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 42(11):114005, 2009.

[233] K. O'Holleran, M. R. Dennis, F. Flossmann, and M. J. Padgett. Fractality of light's darkness. *Phys. Rev. Lett.*, 100:053902, 2008.

[234] J. Wenger, R. Tualle-Brouri, and P. Grangier. Non-Gaussian statistics from individual pulses of squeezed light. *Phys. Rev. Lett.*, 92:153601, 2004.

[235] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier. Generating optical Schrödinger kittens for quantum information processing. *Science*, 312(5770):83–86, 2006.

[236] V. Vedral. An informationtheoretic equality implying the jarzynski relation. *Journal of Physics A: Mathematical and Theoretical*, 45(27):272001, 2012.

[237] M. R. Vanner, M. Aspelmeyer, and M. S. Kim. Quantum state orthogonalization and a toolset for quantum optomechanical phonon control. *Phys. Rev. Lett.*, 110:010504, 2013.

[238] R. McConnell, H. Zhang, J. Hu, S. Ćuk, and V. Vuletić. Entanglement with negative Wigner function of almost 3,000 atoms heralded by one photon. *Nature*, 519(7544):439–442, 2015.

[239] M. Giustina et al. Significant-loophole-free test of bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, 2015.

[240] L. K. Shalm et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, 2015.

[241] A. Eckstein, A. Christ, P. J. Mosley, and C. Silberhorn. Highly efficient single-pass source of pulsed single-mode twin beams of light. *Phys. Rev. Lett.*, 106:013603, 2011.

[242] G. Harder, T. J. Bartley, A. E. Lita, S. W. Nam, T. Gerrits, and C. Silberhorn. Single-mode parametric-down-conversion states with 50 photons as a source for mesoscopic quantum optics. *Phys. Rev. Lett.*, 116:143601, 2016.

[243] Y. Eto, A. Koshio, A. Ohshiro, J. Sakurai, K. Horie, T. Hirano, and M. Sasaki. Efficient homodyne measurement of picosecond squeezed pulses with pulse shaping technique. *Opt. Lett.*, 36(23):4653–4655, 2011.

[244] J. B. Spring, P. S. Salter, B. J. Metcalf, P. C. Humphreys, M. Moore, N. Thomas-Peter, M. Barbieri, X.-M. Jin, N. K. Langford, W. S. Kolthammer, M. J. Booth, and I. A. Walmsley. On-chip low loss heralded source of pure single photons. *Opt. Express*, 21(11):13522–13532, 2013.

[245] W. Wasilewski, A. I. Lvovsky, K. Banaszek, and C. Radzewicz. Pulsed squeezed light: Simultaneous squeezing of multiple modes. *Phys. Rev. A*, 73:063819, 2006.

[246] A. I. Lvovsky, W. Wasilewski, and K. Banaszek. Decomposing a pulsed optical parametric amplifier into independent squeezers. *Journal of Modern Optics*, 54(5):721–733, 2007.

[247] A. Christ, B. Brecht, W. Mauerer, and C. Silberhorn. Theory of quantum frequency conversion and type-II parametric down-conversion in the high-gain regime. *New Journal of Physics*, 15(5):053038, 2013.

[248] T. Aichele, A. Lvovsky, and S. Schiller. Optical mode characterization of single photons prepared by means of conditional measurements on a biphoton state. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics*, 18(2):237–245, 2002.

[249] A. I. Lvovsky and T. Aichele. On the advanced wave model of parametric down-conversion. In V. Akulin, A. Sarfati, G. Kurizki, and S. Pellegrin, editors, *Decoherence, Entanglement and Information Protection in Complex Quantum Systems*, pages 41–48. Springer Netherlands, Dordrecht, 2005.

[250] M. Liscidini and J. E. Sipe. Stimulated emission tomography. *Phys. Rev. Lett.*, 111:193602, 2013.

[251] A. Eckstein, G. Boucher, A. Lemaitre, P. Filloux, G. Leo, J. E. Sipe, M. Liscidini, and S. Ducci. High-resolution measurement of the joint spectral density of quantum correlated photon pairs. In *CLEO: 2014*, page FTh4A.2. Optical Society of America, 2014.

[252] B. Y. Zel'Dovich and D. N. Klyshko. Field Statistics in Parametric Luminescence. *Soviet Journal of Experimental and Theoretical Physics Letters*, 9:40, 1969.

[253] D. N. Klyshko. Utilization of vacuum fluctuations as an optical brightness standard. *Soviet Journal of Quantum Electronics*, 7(5):591, 1977.

[254] D. C. Burnham and D. L. Weinberg. Observation of simultaneity in parametric production of optical photon pairs. *Phys. Rev. Lett.*, 25:84–87, 1970.

[255] R. E. Slusher, P. Grangier, A. LaPorta, B. Yurke, and M. J. Potasek. Pulsed squeezed light. *Phys. Rev. Lett.*, 59:2566–2569, 1987.

[256] N. Takei, H. Yonezawa, T. Aoki, and A. Furusawa. High-fidelity teleportation beyond the no-cloning limit and entanglement swapping for continuous variables. *Phys. Rev. Lett.*, 94:220502, 2005.

[257] C. K. Hong and L. Mandel. Experimental realization of a localized one-photon state. *Phys. Rev. Lett.*, 56:58–60, 1986.

[258] J. Rarity, P. Tapster, and E. Jakeman. Observation of sub-poissonian light in parametric downconversion. *Optics Communications*, 62(3):201–206, 1987.

[259] P. R. Tapster, J. G. Rarity, and J. S. Satchell. Use of parametric down-conversion to generate sub-poissonian light. *Phys. Rev. A*, 37:2963–2967, 1988.

[260] Y. H. Shih and C. O. Alley. New type of einstein-podolsky-rosen-bohm experiment using pairs of light quanta produced by optical parametric down conversion. *Phys. Rev. Lett.*, 61:2921–2924, 1988.

[261] T. E. Kiess, Y. H. Shih, A. V. Sergienko, and C. O. Alley. Einstein-podolsky-rosen-bohm experiment using pairs of light quanta produced by type-ii parametric down-conversion. *Phys. Rev. Lett.*, 71:3893–3897, 1993.

[262] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih. New high-intensity source of polarization-entangled photon pairs. *Phys. Rev. Lett.*, 75:4337–4341, 1995.

[263] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier. Quantum homodyne tomography of a two-photon fock state. *Phys. Rev. Lett.*, 96:213601, 2006.

[264] E. Bimbard, N. Jain, A. Macrae, and A. I. Lvovsky. Quantum-optical state engineering up to the two-photon level. *Nature Photonics*, 4:243–247, 2010.

[265] W. P. Grice and I. A. Walmsley. Spectral information and distinguishability in type-II down-conversion with a broadband pump. *Phys. Rev. A*, 56:1627–1634, 1997.

[266] A. B. U'Ren, C. Silberhorn, R. Erdmann, K. Banaszek, W. P. Grice, I. A. Walmsley, and M. G. Raymer. Generation of pure-state single-photon wavepackets by conditional preparation based on spontaneous parametric downconversion. *Laser Phys.*, 15:146, 2006.

[267] B. Brecht, A. Eckstein, A. Christ, H. Suche, and C. Silberhorn. From quantum pulse gate to quantum pulse shaperengineered frequency conversion in nonlinear optical waveguides. *New Journal of Physics*, 13(6):065029, 2011.

[268] R. Loudon. *The Quantum Theory of Light*, chapter 4. Oxford University Press, Oxford, 2000.

[269] M. I. Kolobov. The spatial behavior of nonclassical light. *Rev. Mod. Phys.*, 71:1539–1589, 1999.

[270] C. M. Caves and D. D. Crouch. Quantum wideband traveling-wave analysis of a degenerate parametric amplifier. *J. Opt. Soc. Am. B*, 4(10):1535–1545, 1987.

[271] M. M. Fejer, G. A. Magel, D. H. Jundt, and R. L. Byer. Quasi-phase-matched second harmonic generation: tuning and tolerances. *IEEE Journal of Quantum Electronics*, 28(11):2631–2654, 1992.

[272] L. E. Myers, R. C. Eckardt, M. M. Fejer, R. L. Byer, W. R. Bosenberg, and J. W. Pierce. Quasi-phase-matched optical parametric oscillators in bulk periodically poled LiNbO3. *J. Opt. Soc. Am. B*, 12(11):2102–2116, 1995.

[273] H. F. Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–551, 1959.

[274] K. L. Brown, W. J. Munro, and V. M. Kendon. Using quantum computers for quantum simulation. *Entropy*, 12:2268–2307, 2010.

[275] B. Brecht. *Engineering ultrafast quantum frequency conversion*. PhD thesis, Universität Paderborn, 2014.

[276] A. Eckstein. *Mastering quantum light pulses with nonlinear wave interactions*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2012.

[277] M. Avenhaus, A. Eckstein, P. J. Mosley, and C. Silberhorn. Fiber-assisted single-photon spectrograph. *Opt. Lett.*, 34(18):2873–2875, 2009.

[278] E. Y. Zhu, C. Corbari, P. G. Kazansky, and L. Qian. Self-calibrating fiber spectrometer for the measurement of broadband downconverted photon pairs. *ArXiv e-prints*, arxiv:1505.01226, 2015.

[279] A. Eckstein, G. Boucher, A. Lemaître, P. Filloux, I. Favero, G. Leo, J. E. Sipe, M. Liscidini, and S. Ducci. High-resolution spectral characterization of two photon states via classical measurements. *Laser Photonics Rev.*, 8(5):L76–L80, 2014.

[280] D. N. Klyshko. Use of two-photon light for absolute calibration of photoelectric detectors. *Sov J Quantum Electron*, 10(9):11121117, 1980.

[281] A. N. Penin and A. V. Sergienko. Absolute standardless calibration of photodetectors based on quantum two-photon fields. *Appl. Opt.*, (25):3582–3588.

[282] P. G. Kwiat, A. M. Steinberg, R. Y. Chiao, P. H. Eberhard, and M. D. Petroff. High-efficiency single-photon detectors. *Phys. Rev. A*, 48:R867–R870, 1993.

[283] G. Brida, S. Castelletto, C. Novero, and M. L. Rastello. Measurement of the quantum efficiency of photodetectors by parametric fluorescence. *Metrologia*, 35(4):397, 1998.

[284] M. Avenhaus, H. B. Coldenstrodt-Ronge, K. Laiho, W. Mauerer, I. A. Walmsley, and C. Silberhorn. Photon number statistics of multimode parametric down-conversion. *Phys. Rev. Lett.*, 101:053601, 2008.

[285] A. P. Worsley, H. B. Coldenstrodt-Ronge, J. S. Lundeen, P. J. Mosley, B. J. Smith, G. Puentes, N. Thomas-Peter, and I. A. Walmsley. Absolute efficiency estimation of photon-number-resolving detectors using twin beams. *Opt. Express*, 17(6):4397–4412, 2009.

[286] J. Lundeen, O. Cohen, P. Mahou, B. Smith, and I. A. Walmsley. Absolute calibration of optical detectors using two-mode squeezed light. In *Frontiers in Optics 2011/Laser Science XXVII*, page LTuH2. Optical Society of America, 2011.

Appendix A

Fisher information for double homodyne detection

In Section 2.2.2, we discuss phase estimation using a double homodyme detection scheme. In Equation 2.55, we find the following expression for the Fisher information:

$$F(p_\phi) = \int r\,dr\,d\varphi\ \left(\partial_\phi p_\phi(r,\varphi)\right)^2 \Big/ p_\phi(r,\varphi) = \int r\,dr\,d\varphi\ \left(2\,\partial_\phi\sqrt{p_\phi(r,\varphi)}\right)^2, \tag{A.1}$$

with

$$\begin{aligned} p_\phi(r,\varphi) &= \frac{1}{2\pi}\left|\sum_{i=0}^{n}\psi_i\,\xi_i^{(n)}(r)\,e^{\mathrm{i}\,(n/2-i)(\phi-2\varphi)}\right|^2; \\ \xi_i^{(n)}(r) &= \sqrt{2}\left\langle i\left|\hat{D}(r)\right|n-1\right\rangle \\ &= \sqrt{2}\left[\sqrt{i!}\Big/\sqrt{(n-i)!}\right]e^{-\frac{r^2}{2}}(-r)^{n-2i}L_i^{n-2i}(r^2). \end{aligned} \tag{A.2}$$

For path-symmetric states, we have $\psi_i = \psi_{n-i}^*\,e^{-2\mathrm{i}\,\chi_0}$. Let us write $\psi_i = |\psi_i|e^{\mathrm{i}\,\phi_i^{(0)}}$. Introducing this in Equation A.2, and using $\xi_i^{(n)}(r) = \xi_{n-i}^{(n)}(r)$, we obtain

$$\begin{aligned} p_\phi(r,\varphi) &= \frac{1}{2\pi}\left|e^{\mathrm{i}\,\chi_0}\frac{1}{2}\sum_{i=0}^{n}|\psi_i|\,\xi_i^{(n)}(r)\left[e^{\mathrm{i}\left((n/2-i)(\phi-2\varphi)-\chi_0+\phi_i^{(0)}\right)}+e^{-\mathrm{i}\left((n/2-i)(\phi-2\varphi)-\chi_0+\phi_i^{(0)}\right)}\right]\right|^2 \\ &= \frac{1}{2\pi}\left|\sum_{i=0}^{n}|\psi_i|\,\xi_i^{(n)}(r)\,\cos\left((n/2-i)(\phi-2\varphi)-\chi_0+\phi_i^{(0)}\right)\right|^2. \end{aligned} \tag{A.3}$$

The expression inside the modulus is real. For a real function $g(x)$, we can write $\left[\partial_x\sqrt{|g(x)^2|}\right]^2 = [\partial_x g(x)]^2$. We thus have

$$\left[2\,\partial_\phi\sqrt{p_\phi(r,\varphi)}\right]^2 = \frac{2}{\pi}\left[\sum_{i=0}^{n}|\psi_i|\,\xi_i^{(n)}(r)\,(n/2-i)\sin\left((n/2-i)(\phi-2\varphi)-\chi_0+\phi_i^{(0)}\right)\right]^2. \tag{A.4}$$

Let us define $\chi_i \equiv (n/2-i)(\phi-2\varphi)-\chi_0$. We now expand the sum in Equation A.4,

$$\left[2\,\partial_\phi\sqrt{p_\phi(r,\varphi)}\right]^2 = \frac{2}{\pi}\sum_{i,j=0}^{n}|\psi_i|\,|\psi_j|\,\xi_i^{(n)}(r)\,\xi_j^{(n)}(r)\,(n/2-i)(n/2-j)\sin(\chi_i)\sin(\chi_j). \quad \text{(A.5)}$$

Introducing this in Equation A.1, we can simplify the integration using

$$\begin{aligned}&\int_0^{2\pi} d\varphi\,\sin(\chi_i)\sin(\chi_j) = \pi\,\delta_{i,j} - \pi\,\delta_{i,n-j};\\ &\int_0^{\infty} r\,dr\,\left[\xi_i^{(n)}(r)\right]^2 = 1.\end{aligned} \quad \text{(A.6)}$$

This yields

$$\int r\,dr\,d\varphi\,\left(2\,\partial_\phi\sqrt{p_\phi(r,\varphi)}\right)^2 = 4\sum_{i=0}^{n}|\psi_i|^2(n/2-i)^2. \quad \text{(A.7)}$$

Appendix B

Predictions for the effect of the demon's measurements

Here, we calculate the effect that the measurement introduced in Chapter 3 has on a mode in a thermal state. These calculations are used for the theoretical predictions depicteded in Figures 3.6 and 3.8.

B.0.1 Effect of a lossy measurement

Let the beam splitter that implements this measurement have transmittance T and reflectance $R = 1 - T$ and let η be the efficiency of the avalanche click-detectors. A single mode thermal state can be written $(1-\lambda)\sum_i \lambda^i |i\rangle\langle i|$ in the photon number basis, where $\lambda = e^{-h\nu/(k_b T)}$. The second input to the beam splitter is the vacuum. The beam-splitter transforms the creation operators corresponding to the two input modes, $\hat{a}^\dagger$ and $\hat{b}^\dagger$, according to the linear transformation: $\hat{a}^\dagger \to \sqrt{T}\,\hat{a}^\dagger + \sqrt{R}\,\hat{b}^\dagger$; $\hat{b}^\dagger \to \sqrt{T}\,\hat{b}^\dagger - \sqrt{R}\,\hat{a}^\dagger$. Thus, the transformed input state is: $\rho = \hat{U}_{\rm bs}\left[(1-\lambda)\sum_i \lambda^i\ |i\rangle\langle i| \otimes |0\rangle\langle 0|\right]\hat{U}_{\rm bs}^\dagger = (1-\lambda)\sum_i \lambda^i\, \hat{U}_{\rm bs}\,|i\rangle\langle i| \otimes |0\rangle\langle 0|\ \hat{U}_{\rm bs}^\dagger$, where

$$\begin{aligned}\hat{U}_{\rm bs}\,|i\rangle\otimes|0\rangle &= \hat{U}_{\rm bs}\frac{\left(\hat{a}^\dagger\right)^i}{\sqrt{i!}}\,|00\rangle = \frac{1}{\sqrt{i!}}(\sqrt{T}\,\hat{a}^\dagger + \sqrt{R}\,\hat{b}^\dagger)^i\,|00\rangle = \\ \sum_k \binom{i}{k} \frac{1}{\sqrt{i!}} \left(\sqrt{T}\,a^\dagger\right)^{i-k} \left(\sqrt{R}\,b^\dagger\right)^k |00\rangle &= \sum_k \sqrt{\binom{i}{k} T^{i-k}R^k}\ |i-k\rangle\,|k\rangle\,.\end{aligned} \tag{B.1}$$

The state after the beam-splitter is

$$\hat{\rho} = (1-\lambda)\sum_{i,k,k'} \lambda^i \sqrt{\binom{i}{k}\binom{i}{k'} T^{2i-k-k'}R^{k+k'}}\ |i-k,k\rangle\left\langle i-k',k'\right|. \tag{B.2}$$

The effect of the measurement performed by an avalanche click-detector is described by two sets of measurement operators grouped according to the two possible outcomes. We denote the outcome with no photons detected $m = 0$. This outcome is registered if there are j photons in the incident mode and all of these are lost without producing an electrical signal. The measurement operators describing this possibility are $\left\{\hat{M}_j^{(0)} = (1-\eta)^j\,|j\rangle\langle j|\ \middle|\ j \geq 0\right\}$.

The measurement operators corresponding to the possibility that there are j photons that *do* produce an avalanche ($m = 1$) are $\left\{\hat{M}_j^{(1)} = (1-(1-\eta)^j)\,|j\rangle\,\langle j| \;\middle|\; j \geq 1\right\}$. The probability of the first outcome is $p^{(0)} = \mathrm{Tr}[\hat{\rho}\sum_j \hat{M}_j^{(0)}] = (1-\lambda)\sum_{i,k}\lambda^i \begin{pmatrix} i \\ k \end{pmatrix} T^{i-k}\,((1-\eta)R)^k = \frac{1-\lambda}{1-\lambda+R\eta\lambda}$. The average number of photons in the updated state corresponding to this measurement outcome is $n^{(0)} = \mathrm{Tr}\left[\hat{\rho}\left(\hat{n}\otimes\sum_k \hat{M}_j^{(0)}\right)\right]/p^{(0)}$ where $\hat{n}$ is the number operator, which yields $n^{(0)} = T\frac{1-\lambda}{1-\lambda+R\eta\lambda}\frac{\lambda}{1-\lambda}$. Noting that the expected number of photons in the initial state is $n_{\mathrm{th}} = \frac{\lambda}{1-\lambda}$ and approximating $T \to 1$, we get $n^{(0)} = p^{(0)}n_{\mathrm{th}}$. The probability of the second outcome is $p^{(1)} = 1-p^{(0)}$ and using the fact that the expected average number of photons after the measurement must be $p^{(0)}n^{(0)} + p^{(1)}n^{(1)} = Tn_{\mathrm{th}}$, we get (using $T \to 1$) $n^{(1)} = (1+p^{(0)})n_{\mathrm{th}} = (2-p^{(1)})n_{\mathrm{th}}$.

We have shown that, given high T, the expected number of photons in the states filtered according to the outcome of the demon's measurement is a simple function of the measurement's outcome probabilities, even when the implementation is based on lossy detectors. Moreover, these probabilities can be tuned by changing the efficiency η just as well as by changing the reflectivity R.

B.0.2 Energy imbalance between the two light modes produced by feed-forward

We can now calculate the expected photon number difference between the two thermal light modes when these are swapped according to outcomes of the measurement discussed above, performed on both modes. We denote $p_1^{(1)}$ and $p_2^{(1)}$ the probabilities for photon detection to occur in the two modes respectively. The average difference of photon number in the two arms has the following contributions: when neither detector fires, the average photon number difference is $(p_1^{(1)} - p_2^{(1)})n_{\mathrm{th}}$, with probability $(1-p_1^{(1)})(1-p_2^{(1)})$; when only the first detector fires, $(-1+p_1^{(1)} - p_2^{(1)})n_{\mathrm{th}}$ with probability $p_1^{(1)}(1-p_2^{(1)})$; when only the second detector fires, $(1+p_2^{(1)} - p_1^{(1)})n_{\mathrm{th}}$ with probability $(1-p_1^{(1)})p_2^{(1)}$; and when both detectors fire, $(p_1^{(1)} - p_2^{(1)})n_{\mathrm{th}}$ with probability $p_1^{(1)}p_2^{(1)}$. To obtain the average photon number difference taking into account the controlled swapping of the two modes, we add-up these contributions with either a plus of a minus sign. The resulting expression has a maximum value, $\langle n\rangle_2 - \langle n\rangle_1 = 16/27\langle n\rangle_0$ obtained when the modes are swapped in the second instance of the above enumeration and for photon detection probabilities $p_1^{(1)} = 2/3$ and $p_2^{(1)} = 1/3$.

B.0.3 Non-disturbance condition applied to the APD measurement

We now show that the non-disturbance condition introduced in Chapter 3,

$$\sum_i \frac{\sum_k \hat{M}_k^{(m)}|i\rangle\,\langle i|\,\hat{M}_k^{(m)\dagger}}{p(m|i)} = \hat{I} \tag{B.3}$$

holds for our measurement setup where $\left\{\hat{M}_k^{(m)}\right\}$ are the measurement operators corresponding to measurement outcome m. We refer to this as a non-disturbance condition due to the fact that $\frac{\sum_k \hat{M}_k^{(m)}|i\rangle\langle i|\hat{M}_k^{(m)\dagger}}{p(m|i)}$ is the normalized state after the measurement, given input

state $|i\rangle\langle i|$ and output m. The relation in Equation B.3 thus holds straight-forwardly for a measurement that does not disturb the initial states (a course energy measurement). Here, we consider this relation for an idealized, lossless measurement ($\eta \to 0$). Above we showed that the experimental implementation is equivalent to this model.
Let us start with the measurement outcome corresponding to no photons detected ($m = 0$). When this outcome is recorded given the input state $|i\rangle\langle i|$, the input is left unchanged: if no photons were detected in the reflected arm of the beam splitter, all photons must have been transmitted. Therefore, the condition stated above holds for $m = 0$. In the case of the outcome corresponding to a photon detection ($m = 1$), the situation is not as simple. Using the expressions above, we can write the measurement operators acting on the input thermal modes, describing the detection of k photons in the reflected port of the beam-splitter: $\hat{M}_k = \sum_i \sqrt{\binom{i}{k} T^{i-k}(1-T)^k}\,|i-k\rangle\langle i|$ with T the beam splitter transmittance. The probability of outcome $m = 1$ when the initial state is $|i\rangle\langle i|$ is $1 - T^i$ (with probability T^i all photons are transmitted and the measurement outcome is $m = 0$). We thus have

$$\begin{aligned}\sum_i \frac{\sum_k \hat{M}_k^{(1)}|i\rangle\langle i|\hat{M}_k^{(1)\dagger}}{p(1|i)} &= \sum_{i,k\geq 1} \hat{M}_k|i\rangle\langle i|\hat{M}_k^\dagger/(1-T^i) = \\ \sum_{i,k\geq 1}\binom{i}{k}\frac{T^{i-k}(1-T)^k}{1-T^i}|i-k\rangle\langle i-k| &= \sum_{i,k\geq 1}\binom{i+k}{k}\frac{T^i(1-T)^k}{1-T^{i+k}}|i\rangle\langle i|.\end{aligned} \tag{B.4}$$

We can bound the coefficients of $|i\rangle\langle i|$ from above using

$$\begin{aligned}\frac{1}{1-T^{i+k}} &\leq \frac{1}{1-T^{i+1}} \Rightarrow \\ \sum_{k\geq 1}\binom{i+k}{k}\frac{T^i(1-T)^k}{1-T^{i+k}} &\leq \sum_{k\geq 1}\binom{i+k}{k}\frac{T^i(1-T)^k}{1-T^{i+1}} = 1/T\end{aligned} \tag{B.5}$$

and from below, using the first term of the sum (all terms are positive):

$$\sum_{k\geq 1}\binom{i+k}{k}\frac{T^i(1-T)^k}{1-T^{i+k}} > \binom{i+1}{1}\frac{T^i(1-T)}{1-T^{i+1}}. \tag{B.6}$$

As $T \to 1$, the limit of both the upper and the lower bound is 1. This shows that

$$\lim_{T\to 1}\sum_i \frac{\sum_k \hat{M}_k^{(1)}|i\rangle\langle i|\hat{M}_k^{(1)\dagger}}{p(1|i)} = \hat{I}, \tag{B.7}$$

so for high transmittance beam splitters, the non-disturbance condition applies approximately, up to an error of the order $|1-T|$.

B.0.4 Computation of the expected mutual information

Here, we calculate the average mutual information characterizing the demon's measurement. We show that for high transmittance $T \approx 1$ of the measurement beam-splitter and high average photon population of the initial states, this quantity depends only on the photon detection probabilities.

The average mutual information is

$$I = \sum_{m,i} p(m,i)\left(\log(p(m|i)) - \log(p(m))\right) \tag{B.8}$$

where we use the same notation as above. The probability distribution of the initial state is $p(i) = (1-e^{-\beta\,\hbar\omega})e^{-\beta\,\hbar\omega\,i}$ where $\hbar\omega$ is the single photon energy. For high average occupation number, the distribution over i is smooth and we can convert the sum to an integral. We introduce the variable $x = \beta\,\hbar\omega\,i$.

$m = 0$ corresponds to the no-photons-detected outcome and $m = 1$ corresponds to the photons-detected outcome. Here, we denote the probability of the $m = 1$ outcome by ξ. We have for the joint probability of the initial state and measurement outcome $p(0,i) = (1 - e^{-\beta\,\hbar\omega})(1 - R\eta)^i e^{-\beta\,\hbar\omega\,i}$ and $p(1,i) = (1 - e^{-\beta\,\hbar\omega})(1 - (1 - R\eta)^i)e^{-\beta\,\hbar\omega\,i}$. Converting this to a probability density in terms of the new variable x, we get $\mathcal{P}(0,x) = \frac{1}{\beta\,\hbar\omega}(1 - e^{-\beta\,\hbar\omega})(1 - R\eta)^{\frac{x}{\beta\,\hbar\omega}} e^{-x}$ and $\mathcal{P}(1,x) = \frac{1}{\beta\,\hbar\omega}(1 - e^{-\beta\,\hbar\omega})(1 - (1 - R\eta)^{\frac{x}{\beta\,\hbar\omega}})e^{-x}$. For $\beta \ll 1$, $\frac{1}{\beta\,\hbar\omega}(1-e^{-\beta\,\hbar\omega}) \approx 1$, so we can write $\mathcal{P}(0,x) = e^{c\,x}$ and, since $1-\xi = \int_0^\infty \mathcal{P}(0,x)\,dx = \int_0^\infty e^{c\,x}\,dx = \frac{1}{c}$, we get $c = \frac{1}{1-\xi}$. Finally, we have

$$\begin{aligned} \mathcal{P}(0,x) &= e^{-\frac{1}{1-\xi}x} \\ \mathcal{P}(1,x) &= \left(1 - e^{-\frac{\xi}{1-\xi}x}\right) e^{-x} \end{aligned} \tag{B.9}$$

and we see that these probability densities only depend on ξ. We can now compute the mutual information in the limit of high occupation numbers and high beam-splitter transmittance, using $p(0|x) = e^{-\frac{\xi}{1-\xi}x}$ and $p(1|x) = 1 - e^{-\frac{\xi}{1-\xi}x}$:

$$\begin{aligned} I = \int & \mathcal{P}(0,x)\left[\log(p(0|x)) - \log(1-\xi)\right] + \mathcal{P}(1,x)\left[\log(p(1|x)) - \log(\xi)\right]\,dx = \\ & \int e^{-\frac{1}{1-\xi}x}\left(-\xi/(1-\xi)x - \log(1-\xi)\right)\,dx + \\ & \int \left(1 - e^{-\frac{\xi}{1-\xi}x}\right) e^{-x}\left(\log(1 - e^{-\frac{\xi}{1-\xi}x}) - \log(\xi)\right)\,dx. \end{aligned} \tag{B.10}$$

These integrals have an analytic solution, but this is a complicated expression, which we use only for plotting the bound in Figure 3.8.

B.0.5 Effect of multi-mode thermal statistics

The statistics of the light produced in our setup is not exactly that of a single-mode thermal state, which is indicated by the estimated second order autocorrelation function, $g^{(2)}(0) \approx 1.9$. When the thermal light has k independent modes, the expected second order autocorrelation is $g^{(2)}(0) = 1 + 1/k$. In the multi-mode case the average number of photons after the demon's measurement is different from that obtained above, for the single-mode case. Let the input state be a mixture of k thermal modes. We denote the probabilities of the detection outcomes when *only one* of the k thermal modes is probed $\left\{\xi^{(0)}, \xi^{(1)}\right\}$ and the probabilities of the detection outcomes when all k modes are probed $\left\{p^{(0)}(k), p^{(1)}(k)\right\}$. We denote the expected photon number in the initial state $n_{\mathrm{th}}(k) = kn_0$ where n_0 is the number of photons per mode. The probability for the detector not to fire when k modes are measured

is $p^{(0)}(k) = \left(\xi^{(0)}\right)^k$. Using the derivation above, we know the average number of photons created by this outcome: $n^{(0)}(k) = \xi^{(0)} k n_0 = \left(p^{(0)}(k)\right)^{\frac{1}{k}} n_{\text{th}}(k)$. The average energy is approximatively conserved by the measurement ($T \to 1$): $p^{(0)}(k)n^{(0)}(k) + p^{(1)}(k)n^{(1)}(k) = n_{\text{th}}$. Using this we can also find the average number of photons corresponding to a photon-detection outcome: $n^{(1)}(k) = \frac{1-\left(p^{(0)}\right)^{\frac{k+1}{k}}}{p^{(1)}} n_{\text{th}}(k)$.