Extracting Quantum Correlations from Gaussian States

Helen Mary Chrzanowski

A thesis submitted for the degree of Doctorate of Philosophy in Physics at The Australian National University

April 2014

Declaration

This thesis is an account of research undertaken between February 2009 and April 2014 at The Department of Physics, Faculty of Science, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Helen Chrzanowski
24th of April, 2014

Words are insufficient to describe my late mum's contribution to my life.
I dedicate this thesis to her.

Acknowledgements

First, I sincerely thank my two supervisors, Ping Koy Lam & Thomas Symul. Ping Koy's brilliance as a physicist and supervisor, combined with his relentless optimism ensured every problem was surmountable. Thomas has taught me so much of what I know, both in theory and in the lab, while introducing me to the world of quantum optics and quantum information.

I would like to thank Amanda White, for constantly fending off administrative foes whilst being a wonderful friend. I would like to thank Syed Assad for his inexhaustible patience, unabating enthusiasm, encyclopaedic knowledge of theoretical physics, and ability to code in *everything*. I will miss sharing a lab and office with you. I would like to thank Julien Bernu for sharing his expansive knowledge of experimental and theoretical physics. I am indebted to Ben Buchler, Boris Hage and Jiri Janousek. Their collective understanding of quantum optics demystified many an experimental mystery and solved my (numerous) problems.

I would like to thank the my colleagues, past and present in the quantum optics group: Michael Stefzsky, Geoff Campbell, Ben Sparkes, Seiji Armstrong, Mahdi Hosseini, Daniel Higginbottom, Jing Yan Haw, Jiao Geng, Giovanni Guiconne, Oliver Thearle, Jesse Everett, Alexandre Brieussel, Sheon Yong, Sarah Hosseini and others those I've forgotten. I will sincerely miss the beers, occasional rounds of tennis, and generally being surrounded by such a vibrant group of physicists. I would like to thank Andre for putting up with me: it was wonderful sharing an office with you and Assad. My friends in the department and research school, many of whom I have known since undergrad: Phil Threlfall, Richard Barry, Thanh Nyugen, Rose Ahlefeldt, and others I've surely forgotten. You are some of the fondest friends I made in Canberra. A special thanks to Nathan, Seiji, Geoff and Jing Yan who helped with proof reading.

I am very grateful for all the resources and support of the Department of Quantum Science, the greater Research School of Physical Sciences and Engineering, the Australian Research Council, and thus, the Australian tax payer (it could be worse).

I'd like to thank my wonderful collaborators Mile Gu, Nathan Walk, Austin Lund, and Tim Ralph. I have learnt so much from all of you. I would especially like to thank Tim, who welcomed me at the University of Queensland on numerous occasions.

I would like to thank my wonderful sister Tania for her constant love and support.

Finally to Nathan: whilst this doctorate has given me a great deal of trouble, it also gave me you. And you are without precedent. I cannot thank you sufficiently for all your unbounded love and kindness.

Abstract

The Gaussian toolbox of the continuous variables provides for deterministic, high-efficiency operations with non-classical states. Its very Gaussian nature, however, restricts its reach for quantum information and communication applications. This thesis comprises three experimental works, which seek to examine the strengths of this toolbox and address some of its weaknesses.

The measurement-based non-linearity of a conditional photon-counting measurement can be used to 'de-Gaussify a Gaussian state of light. Here, we propose a continuous variable analog of just such a 'heralding' measurement, replacing a non-deterministic photon-counting measurement with a deterministic measurement of the field quadratures. Such a technique cannot be used to prepare a non-Gaussian state, but it can, on average, yield the same non-Gaussian statistics. We demonstrate this technique by reconstructing the statistics of non-Gaussian photon-subtracted squeezed vacuum states.

We then consider the problem of noiseless linear amplification. We experimentally demonstrate that in certain scenarios, the requirement for a physical noiseless linear amplifier can be exchanged for a straightforward post-selection of the measurement record. We apply our 'virtual' noiseless amplifier to entanglement degraded by transmission loss of up to the equivalent of 100km of optical fibre. We extract an effective entangled resource stronger than even that achievable with a maximally entangled resource passively transmitted through the same channel. We also provide a proof-of-principle demonstration of the value of the measurement-based noiseless linear amplifier for quantum key distribution, extracting a secret key from an otherwise insecure regime.

Lastly, we turn to the recently popularised measure of all quantum correlations: quantum discord. Quantum discord has emerged as a measure of quantum correlations beyond entanglement, with significant ramifications for our understanding of Gaussian states. Here, we introduce a simple protocol that yields an operational interpretation of quantum discord: that discord describes information only accessible via coherent interactions. We first experimentally encode information within the discordant correlations of two separable Gaussian states. The amount of extra information recovered by coherent interactions is directly linked to the discord of the original state.

Contents

Chapter 1

Introduction

The notion of wave-particle duality to describe light emerged after many centuries of debate. Early in the 17th century, Declight popularised the notion that the behaviour of light could be accurately described as a wave travelling though a uniform medium he called the *plenum*. In the late 17th century, Isaac Newton aggressively championed his corpuscular hypothesis for light, which he documented in his book, *Optiks*. The work of Newton's contemporaries, Christian Huygens, Robert Hooke, and Augustin-Jean Fresnel presented a strong argument for the wave nature of light, accurately describing the refractive and diffractive properties of light, experimentally validated in the early 19th century by Thomas Young's double slit experiment. By the time Maxwell's equations emerged, the corpuscular hypothesis for light had been widely abandoned. James Maxwell's equations succinctly captured classical electromagnetism, predicting the nature of light to be the continuous propagation of energy in the electric and magnetic fields.

However, the turn of the 20th century gave rise to a paradigm shift. Planks' resolution of the ultraviolet catastrophe [1] and Einstein's description of the photoelectric effect [2] both relied on the quantisation of light. The rise of quantum mechanics saw the 'granular' nature of light widely promulgated, as it moved from a mere theoretical convenience in the mind of Planck, to a cornerstone of modern physics.

In many ways, however, modern quantum optics was actually inadvertently spawned by two radio astronomers. Hanbury Brown and Twiss set out to demonstrate a new technique to measure the angular size of stars: an intensity interferometer [3]. The pair had already demonstrated their intensity interferometer for radiowave sources, the successful extension to dim light sources was seen as controversial - even heretical - by many of the quantum mechanics establishment. The result, however, was well accommodated by classical electromagnetism, where, owing especially to the contributions of Wolf [4], a mature theory of coherence already existed. The Hanbury-Twiss result, and the ensuing controversy, was largely remedied by Roy Glauber in 1965 with his quantum theory of coherence. Glauber's seminal work was not precipitated by the Hanbury-Twiss result alone, but also the invention of the ruby laser by Ted Maiman in 1960.

Two distinct and concretely measurable observables arise from the quantised electromagnetic field: the energy and the electric field. These two observables have historically split quantum optics into two distinct camps. The first emphasises the particle-like *discrete variables* of light, constructed around measurements of the energy, or photon-number. The other probes the wave-like *continuous variables* of light, sampling the quadratures of the electric field. Though laser technologies improved throughout the sixties and seventies,

it was developments in non-linear optics that yielded a revised the toolbox for quantum optics. The availability of $\chi^{(2)}$ and $\chi^{(3)}$ non-linearities provided the discrete variables with entangled photon pairs, and the continuous variable quantum optics with squeezing.

Whilst the quantum mechanical revolution transpired, another emerged in the beginnings of the information revolution. For many, modern computer science began in 1936 when Alan Turing outlined his 'Turing Machine', the generalised computing primitive that formalised our modern concepts of 'algorithm' and 'computation' [5]. Only a decade later, Claude Shannon laid the foundations of classical information theory in his seminal paper "A Mathematical Theory of Communication", which provided a framework to understand and quantify the previously nebulous notions of information and communication [6]. With these two breakthroughs, the information revolution was well and truly underway. From the beginning, this information age would have remained a largely academic pursuit without the help of quantum mechanics. It is the uniquely quantum underpinnings of the fibre optic cable and the transistor that have propelled the ideas of Shannon and Turing into the modern world of internet connections and high speed computation.

However, perhaps due to the abstract splendour of classical information theory, it was many years before anyone made the realisation that forms the foundation the the field of quantum information: that information is physical. In essence, this idea is no more than the observation that the systems that we actually use to communicate and process information are ultimately governed by physics. The consequences of this become profound and, it turns out, useful when we consider that the relevant physics deviates from the classical. In 1982, Richard Feynman noted the real world - that described by quantum mechanics - cannot be simulated on a classical computer [7]. In 1985, David Deutsch built upon these ideas to outline his universal quantum computer - the quantum analog of the Turing Machine. The first quantum information protocol, however, was a quantum cryptography primitive outlined by Stephen Wiesner in 1970 (though unpublished until 1982) [8]. Though Weisner's uncounterfeitable *quantum money* was itself never implemented, it very much inspired the first quantum key distribution (QKD) protocol: Bennett and Brassard's seminal *BB84* [9].

Optics has consistently proved the favoured architecture for the first demonstrations of quantum information protocols: quantum teleportation in both its discrete and continuous formulations, super dense coding, and the numerous variations upon quantum cryptography. The seminal result of Knill, Laflamme, and Milburn [10], that linear-optics and single photons were sufficient for universal quantum computing, spawned the field of linear-optical quantum computing. Quantum optics has also provided a testbed for universal one-way quantum computation in both discrete [11, 12] and continuous variables [13]. Furthermore, optical quantum computing has produced numerous information processing demonstrations: from Shor's factoring algorithm [14], to Grover's search algorithm [15], and mixed-state quantum computing.

Although commonly the first successful architecture for quantum information processing demonstrations, it remains unlikely that optics alone will be extended to large-scale quantum computation. When it comes to quantum communication however, optics remains the medium of choice. Interestingly, Brassard has noted that he and Bennett spent many years mulling over quantum money, but their leap to BB84 came once they stopped trying to make photons "stay put" [16]. It is this feature of light: challenging to trap

and store, but able to retain its quantum coherence for long times at ambient temperatures, that make it the ideal candidate for the communication of quantum information over extremely large distances.

Continuous variable quantum optics possesses many highly desirable properties from a communications perspective. It is inherently broadband and compatible with standard telecom infrastructure. The workhorses of continuous variable quantum optics are the Gaussian states, so-called because they generally result in normally distributed measurement outcomes, the operations that preserve this property referred to as the set of Gaussian operations. These states operations enjoy the enviable position of being experimentally accessible in a completely deterministic manner, in stark contrast to much of discrete variable quantum optics. Furthermore, measurement of Gaussian states can also be undertaken with extraordinarily high efficiency and speed. Although these techniques have thus allowed the deterministic generation of entanglement [17], and quantum communication protocols including super-dense coding [18], quantum key distribution [19] and teleportation [20, 21], there is an important drawback. In recent years several "no-go" theorems have emerged showing that Gaussian states and operations alone preclude the possibility of entanglement distillation [22, 23, 24] and error correction [25]. Furthermore, although there exist proposals for universal quantum computation that largely utilise Gaussian tools, at least one non-Gaussian element remains indispensable [13].

All three areas of original research undertaken in this thesis are centred around the nature of quantum information and quantum correlations in Gaussian states, in which we seek to harness the strengths of this toolbox and address some of its weaknesses.

The first area considers an attempted to extract the corspuscular properties of light with measurements of the continuous variables. If an experimenter were to only have access to Gaussian measurements, one might naively think that they would be confined to probing the wave-like properties of light. In the first part of this thesis, we demonstrate homodyne and heterodyne measurements that allow us to mimic particle like measurements upon one mode of an entangled Gaussian state, and through appropriate conditional post-processing, extract strikingly non-Gaussian statistics.

The aforementioned no-go theorems for the distillation of Gaussian states with Gaussian operations require the inclusion of at least one non-Gaussian operation. In the absence of a large χ^3 nonlinearity, our most feasible option is to use the non-linearity of measurement, forgoing some of our cherished determinism in exchange for more exotic operations or stronger correlations. Ideally, this procedure would allow us to break out the Gaussian box in a controlled manner. Our second area of research is the experimental implementation of a proposal to perform just such a tradeoff. We show a carefully chosen post-selection upon the measurement record can provide access to a more strongly entangled Gaussian resource.

Our final area of research concerns the recently popularised measure of all quantum correlations, quantum discord. This information theoretic approach to further refining our understanding of where quantum and classical physics meet has a particular relevance for the Gaussian states. Consider the coherent state, the bread and butter of continuous variable quantum optics. Widely deemed the most classical of pure quantum states, for many single-mode tests of quantumness it *defines* the boundary of classicality. Many other Gaussian states are nothing more than statistical mixtures of coherent states, and

have obvious classical analogues. Nevertheless we are left with the fact that these states still exist as objects in a quantum theory and, perhaps more compellingly, have already been applied to distinctly non-classical communication protocols such as quantum key distribution. This motivates us to consider measures of correlations *between* Gaussian states as a path to uncovering the extent of their quantum nature. We demonstrate an operational relationship between the quantum discord and the extraction of information through the coherent interaction of correlated, but separable, Gaussian states.

Thesis Outline

Following this introduction, Chapter 2 provides the theoretical background and experimental techniques required for the rest of thesis. It comprises an small introduction to quantum optics, and introduces important results of quantum tomography, and quantum information.

Chapters 3 and 4 respectively address our theoretical and experimental results concerning our continuous variable analog of a photon counting measurement.

In Chapter 5 we present a measurement-based noiseless linear amplifier, a post-selective emulation of an noiseless linear amplifier operation. We examine its performance for entanglement distillation and present a proof-of-principle QKD demonstration.

Chapter 6 first provides a review of quantum discord. We then propose a new operational interpretation of quantum discord, providing a experimental demonstration with separable Gaussian states.

Finally, Chapter 7 summarises the main results of this thesis and present a brief outlook for future experimental work.

1.1 Publications

The majority of the research that appears in this thesis has been published in international peer-reviewed journals. The following is a list of academic publications:

1. B. Hage, J. Janousek, S. Armstrong, T. Symul, J. Bernu, H. M. Chrzanowski, P. K. Lam, and H. A. Bachor, "Demonstrating various quantum effects with two entangled laser beams," *The European Physical Journal D*, **63** 457461, (2011).

2. H. M. Chrzanowski, J. Bernu, B. M. Sparkes, B. Hage, A. P. Lund, T. C. Ralph, P. K. Lam, and T. Symul, "Photon-number discrimination without a photon counter and its application to reconstructing non-Gaussian states," *Physical Review A*, **84** 050302, (2011).

3. B. M. Sparkes, H. M. Chrzanowski, D. P. Parrain, B. C. Buchler, P. K. Lam, and T. Symul, "A scalable, self-analyzing digital locking system for use on quantum optics experiments," *Review of Scientific Instruments*, **82** 075113, (2011).

4. M. Gu, H. M. Chrzanowski, S. M. Assad, T. Symul, K. Modi, T.C. Ralph, V. Vedral, and P.K. Lam, "Observing the operational significance of discordconsumption," *Nature Physics* **8** 671675 (2012).

5. H. M. Chrzanowski, S.M. Assad, J. Bernu, B. Hage, A. P. Lund, T.C. Ralph, P.K. Lam, and T. Symul. "Reconstruction of photon number conditioned states using phase randomized homodyne measurements," *Journal of Physics B* (Special Issue on Quantum State Engineering) **46** 104009 (2013).

6. S Hosseini, S Rahimi-Keshari, J Y Haw, S. M. Assad, H. M. Chrzanowski, J Janousek, T Symul, T. C. Ralph and P. K. Lam, "Experimental verification of quantum discord in continuous-variable states,", *Journal of Physics B*, **47** 025503 (2014).

7. H. M. Chrzanowski, N. Walk, S. M. Assad, J. Janousek, S. Hosseini, T.C. Ralph, T. Symul, and P.K. Lam, "Measurement-based noiseless linear amplification for quantum communication," *Nature Photonics* **8**, 333 – 338 (2014).

Chapter 2

Theoretical Background

In this chapter I hope to provide a brief overview of the theoretical tools and experimental techniques required for the content of this thesis. There are numerous very comprehensive texts on this subject - my personal favourites being Leonhardt [26], Walls and Milburn [27], and Loudon [28].

2.1 The Quantum State

A pure quantum state is represented by its state vector $|\psi\rangle$, which is defined within a Hilbert space of a given dimension. As is the case with any vector space, the state vector can be decomposed into a linear combination of its basis vectors of the Hilbert space, such that

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle. \tag{2.1}$$

where the normalisation condition requires that $\sum_i |c_i|^2 = 1$. Here, $|c_i|^2$ is the probability that a measurement in the basis states will yield the state $|\phi_i\rangle$. The observable of any operator acting on the state will yield an expectation value

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \sum_i |\langle a_i | \psi \rangle|^2, \tag{2.2}$$

where $|a_i\rangle$ forms a basis for the operator $\hat{A}$ and $|\langle a_i|\psi\rangle|^2$ represents the probability of finding the state vector $|\psi\rangle$ in state $|a_i\rangle$. We can consider the extension of a single state vector to a system of many modes, $\{a, b, \ldots\}$, each described by a pure state vector $\{|\psi\rangle_a, |\psi\rangle_b, \ldots\}$. The state of the composite is given by the tensor product $|\psi_{ab\ldots}\rangle = |\psi\rangle_a \otimes |\psi\rangle_b \otimes \ldots$. Composite systems that can be represented in this form are *separable* and are known as *product* states. If a composite system cannot be factorised into its constitute states, it is known as *entangled* or *inseparable*. This is a uniquely quantum behaviour whereby it is impossible to independently attribute a pure state to each subsystem, even though the system as a whole can be described by one state vector.

2.1.1 Mixed States

Whilst a pure quantum state is described by a single state vector, the density operator describes a quantum system in a *mixed* state - that is, it is in a statistical ensemble of several pure quantum states, $|\psi_i\rangle$. The density operator is given by

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{2.3}$$

where the quantum system is described by the state vector $|\psi_i\rangle$ with a probability p_i. The fractional probabilities sum to 1. For a pure state $|\psi\rangle$ the density operator is simply $\hat{\rho} = |\psi\rangle\langle\psi|$, and a quantum state is pure if and only if $\hat{\rho} = \hat{\rho}^2$. The quantity $\mathrm{tr}(\hat{\rho}^2)$ is a scalar that defines the *purity* of a quantum state and can take values between 1 for a pure state and $\frac{1}{n}$ for a completely mixed state, where n is the dimension of the Hilbert space. By choosing an arbitrary basis ($\sum_j |b_j\rangle\langle b_j| = 1$) we define the density matrix with the elements

$$\rho_{mn} = \langle b_m|\hat{\rho}|b_n\rangle = \sum_i p_i \langle b_m|\psi_i\rangle\langle\psi_i|b_n\rangle. \tag{2.4}$$

The diagonal matrix elements of ρ correspond to the likelihood of finding the system in the state $|b_n\rangle$. The expectation value of any observable A of the system can be obtained from the density operator

$$\langle\hat{A}\rangle = \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle = \mathrm{tr}(\hat{\rho}\hat{A}). \tag{2.5}$$

We can also consider a composite system, where $\hat{\rho}_{ab}$ is the joint density operator of the multi-mode system comprised of systems a and b. The individual subsystems are now defined by their reduced density operator

$$\hat{\rho}_a = \mathrm{tr}_b(\hat{\rho}_{ab}), \tag{2.6}$$

where tr_b is the partial trace over system b. If subsystems $\hat{\rho}_a$ and $\hat{\rho}_b$ share no correlations, the composite system is the *product state* described by $\hat{\rho}_{ab} = \hat{\rho}_a \otimes \hat{\rho}_b$. The density operator is a powerful generalisation of the state vector that allows us to describe the role of operations such a measurement, in addition to an accurate description of physically realisable quantum states.

2.2 Quantum States of Light

2.2.1 Number or Fock states

A full quantisation of the electromagnetic field can be found in [28]. The Hamiltonian of a single mode, k, of the quantised electromagnetic field is given by

$$\hat{\mathbf{H}}_k = \hbar\omega(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}) \tag{2.7}$$

where $\hat{a}_k$, and its adjoint, $\hat{a}_k^\dagger$ are the quantised equivalents of the complex amplitudes describing the classical electromagnetic field. They are respectively dubbed the *annihilation* and *creation* operators, for reasons that will soon become apparent. They obey the usual bosonic commutation relations, identical to the commutation relations for the quantum harmonic oscillator,

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = (\hat{a}_k \hat{a}_{k'}^\dagger - \hat{a}_{k'}^\dagger \hat{a}_k) = \delta_{kk'}. \tag{2.8}$$

Photon number states, or Fock states, $|n_k\rangle$ are eigenstates of the Hamiltonian of the electromagnetic field with corresponding eigenvalues $\hbar\omega(n + \frac{1}{2})$, where n_k is a natural number. The Fock states are also eigenstates of the photon number operator $\hat{n}_k = \hat{a}_k^\dagger \hat{a}_k$,

$$\hat{a}_k^\dagger \hat{a}_k |n_k\rangle = n_k |n_k\rangle, \tag{2.9}$$

where the eigenvalue n_k corresponds to the number of quanta in the mode k. The creation and annihilation operators act on the corresponding mode k to give

$$\hat{a}_k^\dagger |n_k\rangle = \sqrt{n_k + 1}|n_k + 1\rangle \quad \text{and} \quad \hat{a}_k |n_k\rangle = \sqrt{n_k}|n_k - 1\rangle \tag{2.10}$$

respectively. This corresponds to the creation or destruction of one quanta of energy $\hbar\omega_k$, or a single photon in the mode of interest. The ground state, or vacuum state, $|0\rangle$ is defined as

$$\hat{a}_k |0_k\rangle = 0. \tag{2.11}$$

with a non-zero energy given by $\frac{\hbar\omega_k}{2}$. This small amount of energy is referred to as the zero-point energy. Whilst here we have only considered a single mode of the electromagnetic field, as there are an infinite number of frequency modes accessible, even in a finite volume, the energy associated with the vacuum state of the electromagnetic field is infinite. Any Fock state is accessible through successive applications of the creation operator on the vacuum state

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{2.12}$$

The number of photons in a Fock state $|n_k\rangle$, and thus its energy, is exactly defined, with the variance of the photon number of a Fock state equal to zero

$$\langle \Delta n^2 \rangle = \langle n|\hat{n}\hat{n}|n\rangle - \langle n|\hat{n}|n\rangle^2 = 0. \tag{2.13}$$

Fock states form a complete and orthonormal basis in which any arbitrary quantum state can be represented, and are typically favoured as the basis of choice for density matrix representations of quantum states. Fock states have no classical analogue, which is perhaps a little unsurprising as the notion of the photon itself is uniquely quantum. This inconsistency with notions of classical light is perhaps best illuminated by considering the so-called *quadrature operators*,

$$\hat{X} = \hat{a} + \hat{a}^\dagger \quad \text{and} \quad \hat{P} = i(\hat{a}^\dagger - \hat{a}). \tag{2.14}$$

This pair of operators represent the real and imaginary components of the complex amplitude, and are the quantised analogues of the phase and amplitude quadratures of the electromagnetic field. Unlike the annihilation and creation operators, they are Hermitian and thus correspond to observables. It follows from the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$ that $[\hat{X}, \hat{P}] = 2i$[1], and thus that the quadrature operators are conjugate observables, and cannot be perfectly determined simultaneously. The uncertainty principle constrains any attempt to simultaneously measure $\hat{X}$ and $\hat{P}$ of a state to a precision

$$\Delta\hat{X}\Delta\hat{P} \geq 1, \tag{2.15}$$

where $\Delta\hat{X}$ and $\Delta\hat{P}$ denotes the variance of $\hat{X}$ and $\hat{P}$. We can also consider a generalised quadrature operator, a linear combination of the two orthogonal quadratures

$$\hat{X}^\theta = \cos\theta\hat{X} + \sin\theta\hat{P} = \hat{a}^\dagger e^{i\theta} + \hat{a}e^{-i\theta}. \tag{2.16}$$

We can now ask what is the expectation value for the *amplitude* or *phase* of a Fock state? It is straightforward to see that $\langle n|\hat{X}^\theta|n\rangle = 0$, for all n. Regardless of the photon number of the Fock state, the average value of the quantised analogues to the amplitude and phase of light are always zero. It is straightforward to verify the uncertainty in the quadrature amplitudes for a Fock state is given by

$$\Delta\hat{X}_n^\theta = \langle n|(\hat{X}^\theta)^2|n\rangle - \langle n|\hat{X}^\theta|n\rangle^2 = (2n+1). \tag{2.17}$$

Using classically familiar quantities such as phase or amplitude to describe Fock states proves difficult; we can see that whilst the expectation value for the phase or amplitude of any given Fock state is zero, the uncertainty in this measurement scales as $2n$. With respect to the notions of classical electromagnetism, the states that form the most natural Eigenbasis for the quantised electromagnetic field seem wildly exotic. One notable result occurs for $n = 0$. Despite us having perfect knowledge that the vacuum state is 'empty' the field quadratures are still randomly fluctuating. Of course, Heisenberg's uncertainty principle requires they do fluctuate, if not, we could simultaneously obtain knowledge of both $\hat{X}$ and $\hat{P}$. The vacuum state $|0\rangle$ is a minimum uncertainty state, saturating the uncertainty principle for the quadrature observables (2.15) such that $\Delta\hat{X}\Delta\hat{P} = 1$. More precisely, this uncertainty is symmetrically distributed between the two quadratures and $\Delta\hat{X} = \Delta\hat{P} = 1$. The vacuum state in ubiquitous in our description of most every quantum optics experiment. And it is perhaps the only pure state that we regularly encounter in the lab.

2.2.2 Coherent States

The Fock states introduced in the previous section bear little resemblance to the classical light field, and prove impractical for a mathematical description of most of the light we encounter in the lab. A more appropriate basis for the description of many real electro-

[1]There are three widely used conventions for normalisation of the quadrature operators corresponding to the choice of $\hbar$: $\frac{1}{2}$, 1, and 2. Here, we choose $\hbar = 2$ which corresponds to a variance of the vacuum $\Delta\hat{X}_v^\theta = 1$.

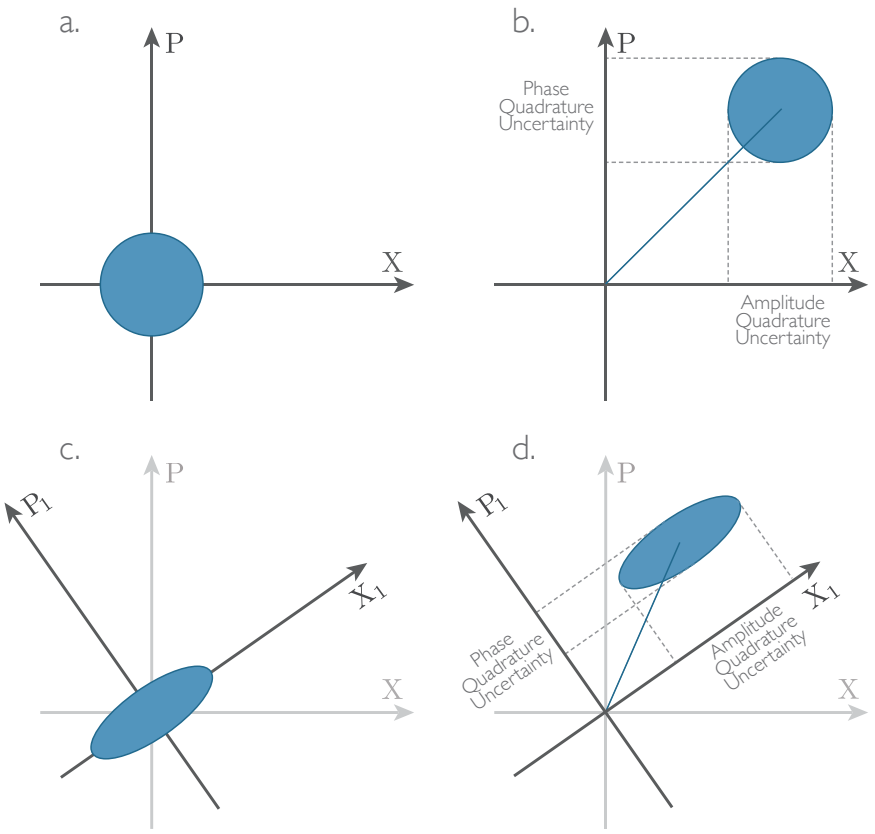

Figure 2.1: Ball-on-stick diagram for (a) vacuum state, (b) coherent state, (c) squeezed vacuum state, and (d) squeezed coherent state.

magnetic fields are the coherent states $|\alpha\rangle$. First described by Roy Glauber in 1967 - a result which in part constituted his Nobel prize - they provide the best quantum mechanical approximation to the light produced by a laser. The coherent states are generated by application of the unitary *displacement* operator to the vacuum state

$$|\alpha\rangle = \hat{D}(\alpha)|0\rangle \tag{2.18}$$

where

$$\hat{D}(\alpha) = \exp\left(\alpha\,\hat{a}^{\dagger} - \alpha^{z*}\,\hat{a}\right). \tag{2.19}$$

This corresponds to the displacement of the minimum uncertainty state, $|0\rangle$ by a distance α in phase space. In this sense, it is the best quantum mechanical approximation to a single point in phase space, and thus, the dynamics of a classical harmonic oscillator. Quantum optics textbooks often make reference to the *ball-on-stick* picture. The ball-on-stick picture or Caves diagram can be considered as a quantum mechanical generalisation of a classical phasor - where the stick is associated with the coherent amplitude and the ball with the associated quantum noise. This picture has no direct mathematical correspondence - unlike the Wigner function (which will be introduced later) or the P-function - and can be thought of as a slice through the Wigner function. The coherent states are eigenstates of the annihilation operator

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{2.20}$$

As $\hat{a}$ is a non-Hermitian operator, α is complex and often decomposed as $\alpha = |\alpha|e^{i\theta}$. The vacuum state is simultaneously a Fock state, and a coherent state with eigenvalue $\alpha = 0$. We can expand the coherent states in the Fock basis:

$$|\alpha\rangle = \sum_n |n\rangle\langle n|\alpha\rangle = e^{-|\alpha|^2/2}\sum_n \frac{\alpha^n}{(n!)^{1/2}}|n\rangle. \tag{2.21}$$

The coherent states are not pairwise orthogonal, with $|\langle\alpha|\beta\rangle|^2 = \exp\left(-|\alpha-\beta)\right|$, but they are complete and thus span the entire Hilbert space. They thus form an over-complete basis of the Hilbert space that can be used to decompose any state, $|\psi\rangle$:

$$|\psi\rangle = \frac{1}{\pi}\int |\alpha\rangle\langle\alpha|\psi\rangle d^2\alpha. \tag{2.22}$$

If we consider the expectation value of a general quadrature observable of (2.16) of a coherent state we have,

$$\langle\hat{X}^{\theta}\rangle = \langle\alpha|\hat{X}^{\theta}|\alpha\rangle = \alpha\,e^{i\theta} + \alpha^*\,e^{-i\theta} \tag{2.23}$$

If we consider the variances in the quadratures we find that the coherent states are not only minimum uncertainty states, but also occupy the unique position of symmetrically minimising all the quadrature variances simultaneously.

From Equation 2.20, the mean photon number for a coherent state is simply

$$\bar{n} = \langle \hat{a}^\dagger \hat{a} \rangle_\alpha = \alpha^* \, \alpha = |\alpha|^2. \tag{2.24}$$

Though this result coincides with that of a classical harmonic oscillator, in acquiring a more precise definition for the phase of our quantum light, we forgo our certainty regarding the energy of the state. As a result, a coherent state does not have a precisely defined photon number, rather its photon number distribution is Poissonian:

$$P(n) = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n} \, e^{-|\alpha|^2}}{n!}. \tag{2.25}$$

Provided n is very large, the uncertainty in n scales as $\sqrt{n}$, and we recover the behaviour of a classical laser.

2.2.3 Thermal States

In the previous section we briefly discussed coherent states as a useful basis to describe real light. One such state of light, the *thermal state* is a special mixed state, which exists on the boundary between quantum and classical light. Thermal states describe the light emitted from a black body of temperature, t. In the coherent state basis, the thermal state is a normally distributed statistical mixure,

$$\rho_{th} = \int d^2\alpha \, \frac{1}{\pi \bar{n}} e^{-|\alpha|^2/\bar{n}} \, |\alpha\rangle\langle, | \, \alpha \tag{2.26}$$

where $\bar{n}$ is the mean photon number [29]. In the Fock basis, the thermal states are described by a density matrix,

$$\rho_{th} = \frac{1}{1+\bar{n}} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{1+\bar{n}} \right)^n |n\rangle\langle n| \, . \tag{2.27}$$

The thermal states are diagonal in the Fock basis, and thus have no meaningful phase. The thermal states are symmetric in phase space, and the expectation value for the generalised quadrature operator $\langle \hat{X}^\theta \rangle = 0$, whilst the variance in $\hat{X}^\theta$ scales with $\bar{n}$ - reminiscent of the Fock states. Unlike a Fock state however, the photon number variance is not non-zero, but rather $\Delta n = \sqrt{\bar{n}^2 + \bar{n}}$ and thermal states exhibit *super-Poissonian* statistics.

2.2.4 Uncertainty and Squeezed states

So far, we have only been exposed to a class of minimum uncertainty states: those with their intrinsic uncertainty symmetrically distributed between the quadratures. However, we can consider a generalisation of the coherent states: where the uncertainty regarding one quadrature is smaller at the expense of complementary quadrature - whilst always the preserving the minimum uncertainty product $\Delta\hat{X}\Delta\hat{P} = 1$.[2] These are called the *squeezed*

[2] In reality, no experimentally prepared state can saturate the uncertainty principle. Experimental imperfections preclude us from ever preparing a pure squeezed state and thus the product $\Delta\hat{X}\Delta\hat{P}$ will always be greater than 1.

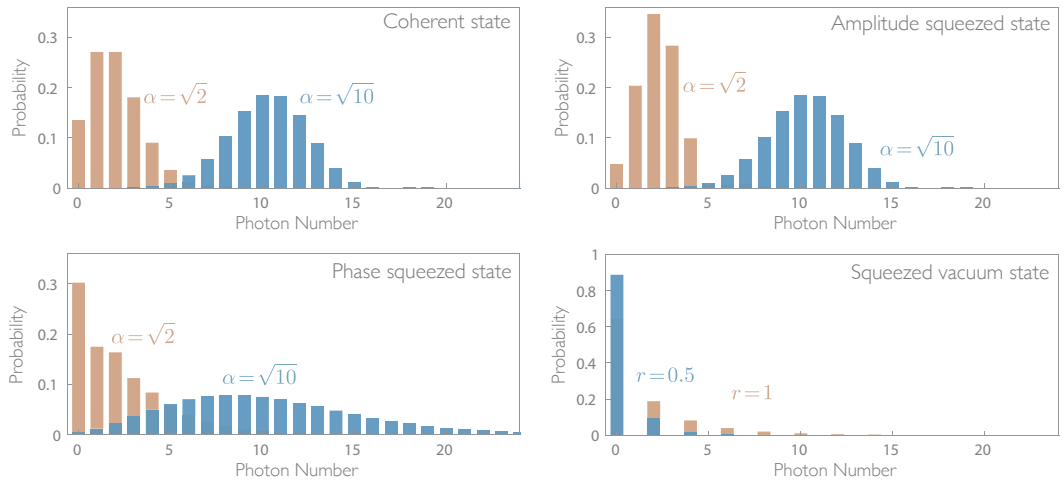

Figure 2.2: Photon number distributions for: (a) a coherent state, (b) a amplitude squeezed coherent state (with $r = 0.5$), (c) a phase squeezed coherent state (with $r = 0.5$), and (d) a squeezed vacuum state.

coherent states. For simplicity, we will first focus on the application of the *squeeze* operator on the vacuum state,

$$|\xi\rangle = \hat{S}(\xi)|0\rangle \quad \text{with} \quad \xi = r\exp(i2\phi), \tag{2.28}$$

where r represents the degree of squeezing and ϕ the orientation of the squeezing axis. The unitary squeezing operator is defined as

$$\hat{S}(\xi) = \exp\left(\tfrac{1}{2}\xi^*(\hat{a})^2 - \tfrac{1}{2}\xi(\hat{a}^\dagger)^2\right), \quad \hat{S}^\dagger(\xi) = \hat{S}^{-1}(\xi) = \hat{S}(-\xi) \tag{2.29}$$

The properties of the squeeze operator are best illuminated in the Heisenberg picture:

$$\hat{S}^\dagger(\xi)\hat{a}\hat{S}(\xi) = \hat{a}\cosh(r) - \hat{a}^\dagger\exp(-2i\phi)\sinh(r) \tag{2.30}$$

$$\hat{S}^\dagger(\xi)\hat{a}^\dagger\hat{S}(\xi) = \hat{a}^\dagger\cosh(r) - \hat{a}\,\exp(-2i\phi)\sinh(r). \tag{2.31}$$

Using the above results, the action of the squeezing operator on the generalised quadrature operator is given by [3]

$$\hat{S}^\dagger(\xi)\hat{X}^\theta\hat{S}(\xi) = \hat{X}^\theta\cosh(r) - \hat{X}^{\theta-\phi}\sinh(r). \tag{2.32}$$

To understand the noise properties of the squeezed vacuum state we introduce the quadratures $\hat{X}_1$ and $\hat{P}_1$, which are the standard quadrature operators $\hat{X}$ and $\hat{P}$ rotated by the orientation of the squeezing axis ϕ,

$$\hat{X}_1 + i\,\hat{P}_1 = (\hat{X} + i\,\hat{P})\,e^{-i\phi}. \tag{2.33}$$

[3] This can be shown by exploiting the operator identity $e^{\hat{A}}\hat{B}e^{\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2}[\hat{A},[\hat{A},\hat{B}]]\ldots$

The squeezing operation has transformed the symmetric variances of the quadratures to

$$\Delta\hat{X}_1 = \langle\xi|(\hat{X}^{\phi})^2|\xi\rangle - \langle\xi|\hat{X}^{\phi}|\xi\rangle^2 = e^{-2r} \tag{2.34}$$

$$\Delta\hat{P}_1 = \langle\xi|(\hat{X}^{\phi+\frac{\pi}{2}})^2|\xi\rangle - \langle\xi|\hat{X}^{\phi+\frac{\pi}{2}}|\xi\rangle^2 = e^{2r}. \tag{2.35}$$

We recover the uncertainty product $\Delta\hat{X}_1\Delta\hat{P}_1 = 1$. We have 'redistributed' the uncertainty of the original vacuum state; we have enhanced precision in the 'squeezed' quadrature, at the expense of the precision in the complementary, 'anti-squeezed' quadrature. A larger squeezing parameter r corresponds to a smaller variance in the squeezed quadrature. From an expansion of Equation 2.29 it is clear that the application of the squeezing operator to the vacuum state produces a superposition state of pairs of photons. The Fock basis decomposition of a squeezed state is given by

$$|\xi\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{2n!}}{n!} (\frac{1}{2} \sinh r e^{i\phi})|2n\rangle, \tag{2.36}$$

with the mean photon number

$$\bar{n}_{\xi} = \langle\xi|\hat{a}^{\dagger}\hat{a}|\xi\rangle = \sinh^2(r). \tag{2.37}$$

It is clear the squeezed vacuum is no longer a true vacuum, but rather a very dim state of light. The more squeezed a given vacuum state is, the higher the average photon number (Figure 2.2). Here we have discussed the squeezed vacuum, but one can straightforwardly generalise to squeezed coherent states. Recalling the coherent state as a class of minimum uncertainty states generated by displacing the vacuum state, we can also follow the same procedure to the generate squeezed coherent states,

$$|\alpha, \xi\rangle = \hat{D}(\alpha)\,\hat{\xi}(r)\,|0\rangle. \tag{2.38}$$

The mean photon number of the squeezed coherent state is simply

$$\langle\alpha, \xi|\hat{a}^{\dagger}\hat{a}|\alpha, \xi\rangle = |\alpha|^2 + \sinh^2(r). \tag{2.39}$$

While the modification of $\sinh^2(r)$ to the energy of the coherent state is typically small, the squeezed coherent state non-longer exibits Poisonnian statistics.

Squeezed states are a uniquely non-classical state of light. This can be simply argued from the fact they are a pure state, and thus cannot be written as a statistical mixture of coherent states. The first theoretical proposals for squeezed light emerged in the 1970's from Stoler [30] and later, Yuen and Shapiro [31, 32]. The first experimental demonstration of squeezed light followed in 1985, using four-wave mixing in a cavity with sodium vapour [33]. Shortly thereafter, Kimble's group achieved better results with an optical parametric amplifier (OPA) [34], which is now widely favoured as the tool for producing highly squeezed light. Whilst this will be discussed in greater detail in the next chapter, 'pair-production' processes in non-linear crystals form the workhorses of most quantum optics experiments. In continuous variables, squeezing provides access to a non-classical light resource, that, with linear optics, becomes an entanglement resource. In the discrete

variables domain, the same pair production processes realised as spontaneous parametric down conversion (SPDC) provide entangled photon pairs with myriad of demonstrated applications.

2.2.5 Two-Mode Squeezed Light

Another significant quantum state for this thesis is the two-mode squeezed state or EPR state [35]. The latter name emerges as two-mode squeezed light demonstrates the properties described in the seminal Einstein, Podolsky and Rosen paper [36]. The two-mode squeezed state is generated by combining two orthogonal squeezed vacuum states, or through a non-degenerate two-photon down conversion process. Mathematically, this corresponds to the application of the unitary *two-mode squeeze operator* on the two mode vacuum,

$$|\xi_1, \xi_2\rangle = \hat{S}_{12}|0,0\rangle = \exp\left(\xi \hat{a}_1^\dagger \hat{a}_2^\dagger - \xi^* \hat{a}_1 \hat{a}_2\right)|0,0\rangle. \tag{2.40}$$

The resulting state (2.40) can be decomposed in the Fock basis of the two modes

$$|\xi_1, \xi_2\rangle = \frac{1}{\cosh r} \sum_{n=0}^{\infty} e^{in\phi} (\tanh r)^n |n\rangle_1 |n\rangle_2. \tag{2.41}$$

From (2.41) we can see the two-mode squeezed vacuum is a highly correlated superposition of Fock states, where the two modes contain the same photon number. Each mode however, when considered locally, is thermal. This can be seen by tracing over mode 2, Equation (2.41) becomes

$$\rho_1 = \mathrm{Tr}_2(|\xi_1, \xi_2\rangle\langle\xi_1, \xi_2|) = \frac{1}{\cosh^2 r} \sum_{n=0}^{\infty} (\tanh r)^{2n} \, |n\rangle_1\langle n|_1 \tag{2.42}$$

$$= \frac{1}{1+\bar{n}} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{1+\bar{n}}\right)^n |n\rangle_1\langle n|_1, \tag{2.43}$$

where we have used $\bar{n} = \sinh r$. The individual states are simply thermal states and the resulting reduced density matrix is diagonal in the Fock basis. Accordingly, an independent measurement of the variance of any quadrature of one subsystem $\Delta X^\theta = \cosh 2r = 2\bar{n} + 1$, the result for a thermal state. Whilst the individual quadratures appear noisy, the highly correlated nature of the joint system becomes apparent when considering two new joint operators of subsystems 1 and 2:

$$\Delta(\hat{X}_1 - \hat{X}_2) = \Delta(\hat{P}_1 + \hat{P}_2) = e^{-2r} \tag{2.44}$$

$$\Delta(\hat{X}_1 + \hat{X}_2) = \Delta(\hat{P}_1 - \hat{P}_2) = e^{-2r}. \tag{2.45}$$

Whilst the individual quadratures are not squeezed, the above linear combinations of the two are. The two mode squeezed state is fully characterised by its covariance matrix,

$$C(\xi) = \begin{pmatrix} \cosh 2r & 0 & -\sinh 2r & 0 \\ 0 & \cosh 2r & 0 & \sinh 2r \\ -\sinh 2r & 0 & \cosh 2r & 0 \\ 0 & \sinh 2r & 0 & \cosh 2r \end{pmatrix}. \tag{2.46}$$

The highly correlated nature of the state manifests as an example of the Einstien, Podolsky and Rosen paradox [36]. The EPR entanglement of the two mode squeezed state was first experimentally demonstrated in 1992[17], and now forms an indispensable resource for the fundamental CV quantum information protocols, including teleportation, quantum dense coding and certain variations of quantum cryptography.

2.3 Phase-space representations

Quantum mechanics typically describes a microscopic system in terms of a state vector, $|\psi\rangle$ or a density operator, $\hat{\rho}$. While well suited for a great many tasks in quantum optics, these descriptions are often not particularly well adapted to provide an intuitive picture of states occupying an infinite dimensional Hilbert spaces. The phase space formulation of quantum mechanics provides an important tool for continuous variable quantum optics and quantum information. It provides an alternative framework for quantum mechanics without reference to wave functions, density matrices or a Hilbert space. Instead, the quantum state is described by a *quasi-probability distribution* in phase space. These phase-space distributions are not unique, and for every quantum system there exist infinitely many formulations of the quasi-probability distribution - the *Wigner representation* (or Wigner function) is perhaps the most identifiable. The *Husmi Q representation* and the *Glauber-Sudarshan P representation* also make regular appearances in quantum optics and information.

2.3.1 The Wigner representation

The power of the phase-space formulation arises in its correspondence with classical mechanics. Instead of confining ones description of their quantum system in either the *position* or *momenta* space, the phase space representation considers both equally and symmetrically. The Wigner quasi-probability distribution behaves like a precisely defined function in x and p without reference to any simultaneous measurement of x and p. This appears at odds with the Heisenberg uncertainty principle, which restricts our simultaneous knowledge of these conjugate observables. Harmony is restored however, once we realise that the Wigner representation does not in general permit a standard classical joint-probability distribution for the incompatible observables. Instead we obtain a 'quasi-probability' distribution, where the marginals describe real probability distributions for x and p, but negative "probabilities" may arise. The Wigner distribution for a general

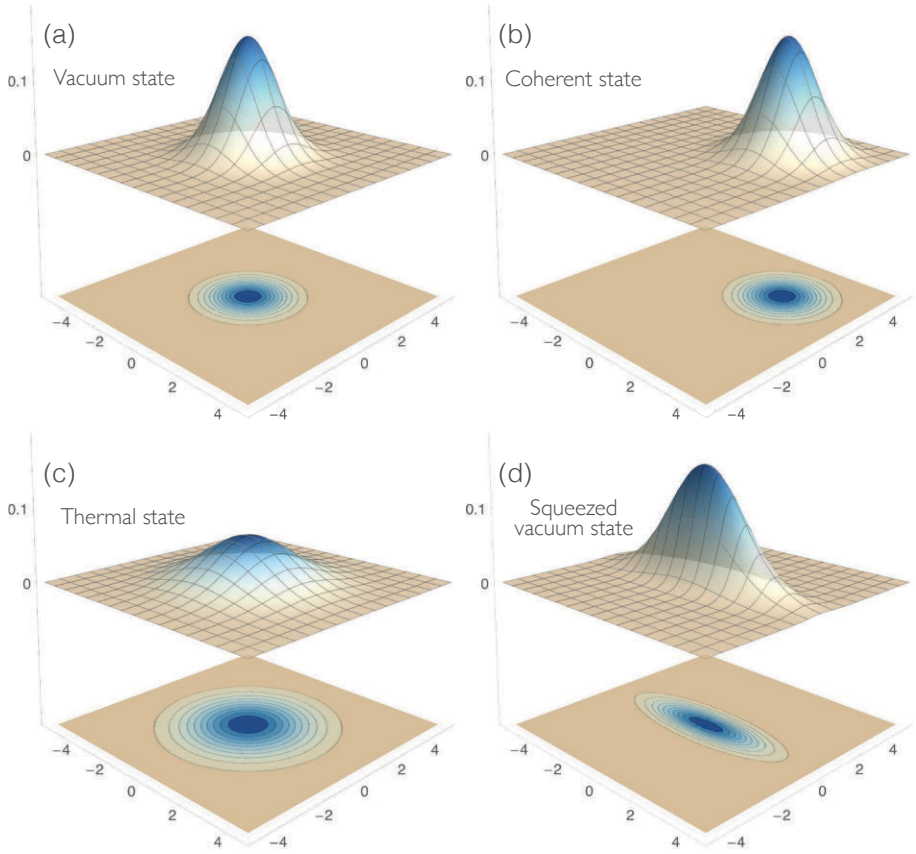

Figure 2.3: The Wigner functions for (a) the vacuum state, (b) a coherent state with $\alpha = 2$, (c) a thermal state with $\bar{n} = 2$, and (d) a squeezed vacuum state with squeezing parameter $\xi = 3$.

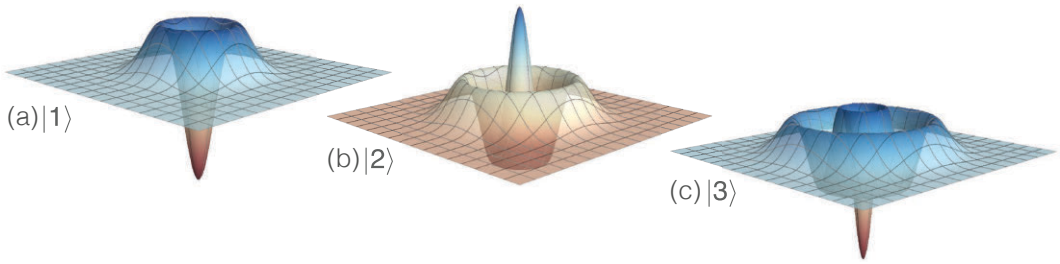

Figure 2.4: The Wigner functions for the (a) $|1\rangle$, (b) $|2\rangle$ and (c) $|3\rangle$ photon Fock states.

density operator, ρ

$$W(x,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} dq\, \langle x-q|\hat{\rho}|x+q\rangle\, e^{(\frac{2ipq}{\hbar})}, \tag{2.47}$$

where x and p are the position and momenta - but could be any pair of conjugate variables, for instance the phase and amplitude quadratures, X and P, of a light field. The Wigner distribution is a one-to-one mapping of the density matrix to a real function in phase space. The classical probability distributions describing a measurement of position or momentum are accessible via the marginals,

$$\int_{-\infty}^{\infty} dp\, W(x,p) = \langle x|\hat{\rho}|x\rangle \;\text{ and }\; \int_{-\infty}^{\infty} dx\, W(x,p) = \langle p|\hat{\rho}|p\rangle. \tag{2.48}$$

The trace of the density matrix is given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\, dp\, W(x,p) = \mathrm{Tr}(\rho) = 1, \tag{2.49}$$

and perhaps of more interest, the purity of the state concerned is

$$2\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\, dp\, W(x,p)^2 = \mathrm{Tr}(\rho^2) = p, \tag{2.50}$$

where $p = 1$ for a pure state, and $p < 1$ for a mixed state. Operator expectation values are given by

$$\langle\psi|\hat{\mathbf{O}}|\psi\rangle = \mathrm{Tr}(\hat{\mathbf{O}}\rho) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\, dp\, W(x,p)O(x,p)\ . \tag{2.51}$$

The resemblance between this quasi-probability distribution and its classical counterpart deteriorates for quantum states that have no classical analogy. All pure states that are neither coherent states or squeezed states will have a Wigner function that is negative somewhere - Fock states for instance (Figure 2.4). Of course, negative probabilities have no classical interpretation, and "negativity" of Wigner function is a widely used metric of non-classicality.

2.3.2 The Glauber-Sudarshan P Representation

Though the Wigner Representation preceded it, the P Representation introduced independently by Glauber [37] and Sudarshan [38] is perhaps the true phase space primitive. The coherent states form an over-complete basis in which any density matrix $\hat{\rho}$ can be diagonalised in the form

$$\hat{\rho} = \int d^2\alpha \; P(\alpha)|\alpha\rangle\langle\alpha|, \tag{2.52}$$

where $P(\alpha)$ is the Glauber-Sudarshan P distribution. Whilst at first glance equation 2.52 looks to be relatively innocuous decomposition of $\hat{\rho}$ into a statistical mixture of coherent states. However, owing to the non-orthogonality of the coherent state basis, the resulting $P(\alpha)$ cannot be interpreted as a genuine probability distribution. The non-orthogonality of the basis states also manifests in $P(\alpha)$ being more often than not either ill behaved or ill defined. Consider a pure coherent state $\hat{\rho} = |\beta\rangle\langle\beta|$, the corresponding P-function of a pure coherent state $|\psi\rangle = |\alpha\rangle$ is singular, a Dirac delta function at the complex amplitude, α. And the only pure states for which it is positive are the coherent states.

Though often mathematically pathological, the P-function provides a straightforward litmus test of non-classicality unique amongst all other phase-space distributions. For quantised fields with a classical 'analog', that is, can expressed as a statistical mixture of coherent states, it will be non-negative everywhere. For systems without classical 'analog' $P(\alpha)$ will be negative somewhere, or more singular than a delta function. In this sense, the singularity of the P-function is proves to be a feature. In smoothing the P-function to obtain a something more well behaved we forgo this classical-quantum dichotomy. For example, the Wigner function is accessible via a Gaussian convolution of the P-function,

$$W(\alpha) = \frac{2}{\pi} \int d^2\beta \, P(\beta) \exp\left(-2|\beta - \alpha|^2\right) \tag{2.53}$$

whilst the negativity of the Wigner function is successfully used as a metric for the 'non-classicality' of states, it fails to capture important quantum states without classical analog, such as squeezed states.

2.3.3 The Husimi Q Representation

In the same manner that the Wigner function is related to the P function via a Gaussian convolution, the *Husimi Q Representation* can be derived from the Wigner function via the same Gaussian convolution,

$$Q(\alpha) = \frac{2}{\pi} \int d^2\beta W(\beta) \exp\left(-2|\beta - \alpha|^2\right). \tag{2.54}$$

In the same way the Gaussian convolution tempered the P-function to yield typically well-behaved state representation – albeit with negative 'probabilities' – the Gaussian filtering of the Wigner function produces a *real* probability distribution that is positive everywhere. The Q-Function describes the projection of the quantum state onto the coherent state

basis,

$$Q(\alpha) = \langle\alpha|\hat{\rho}|\alpha\rangle. \tag{2.55}$$

Physically, the Q-function describes the probability distribution obtain from heterodyne detection of a quantum state.

2.4 Correlations, Quantum Correlations and Entanglement

In this chapter we have met already number of single mode states that have statistical or physical properties without analog in classical described light. If there is one subject I hope the results in this thesis address a little, however, it is the non-classical nature of correlations. Whilst there is little that is too surprising about correlations in the classical world, the implications for correlated quantum systems are very different.

Imagine a *bipartite* system which describes two spatially separated components, a and b. The two subsystems interacted in the past, but are now well separated, such that one can locally measure one subsystems without physically perturbing the other. Consider a measurement of the amplitude quadrature operators of both modes a and b. The dependence of the two quadrature operators $\hat{X}_a$ and $\hat{X}_b$ is described by the correlation coefficient

$$c_{ab}^{xx} = \frac{C_{ab}^{xx}}{\sqrt{\Delta\hat{X}_a\,\Delta\hat{X}_b}}, \tag{2.56}$$

where C_{ab}^{xx} denotes the covariance, and $\Delta\hat{X}_a$ and $\Delta\hat{X}_b$ the variance. This quantity varies between -1 (perfectly anti-correlated) and 1 (perfectly correlated) through 0 (a and b are independent). Imagine subsystems a and b are not perfectly correlated. A measurement of one subsystem does not allow us to perfectly infer the state of system b - but we do gain some information. This reduced uncertainty on b following a measurement of a is characterised by the conditional variance, defined as

$$V_{a|b}^{xx} = \Delta\hat{X}_a(1 - (c_{ab}^{xx})^2) = \Delta\hat{X}_a - \frac{(C_{ab}^{xx})^2}{\Delta\hat{X}_b}. \tag{2.57}$$

Provided the systems a and b are perfectly correlated, $V_{a|b}^{xx} = 0$ and there is no uncertainty regarding the inference of $\hat{X}_a$ from a measurement of $\hat{X}_b$. If the two bi-partitions share no correlations, the uncertainty regarding each remains unchanged after a measurement of the other.

2.4.1 Correlations and Gaussian states

Any bipartite Gaussian state is fully characterised by its coherent amplitudes $\langle \hat{X}_a \rangle$, $\langle \hat{P}_a \rangle$ and $\langle \hat{X}_b \rangle$, $\langle \hat{P}_b \rangle$, and its covariance matrix

$$C(\hat{X}_a, \hat{X}_b, \hat{P}_a, \hat{P}_b) = \begin{pmatrix} C_{aa}^{xx} & C_{aa}^{xp} & C_{ab}^{xx} & C_{ab}^{xp} \\ C_{aa}^{px} & C_{aa}^{pp} & C_{ab}^{px} & C_{ab}^{xx} \\ C_{ba}^{xp} & C_{ba}^{xp} & C_{bb}^{xx} & C_{bb}^{xp} \\ C_{ba}^{px} & C_{ba}^{pp} & C_{bb}^{px} & C_{bb}^{pp} \end{pmatrix}, \tag{2.58}$$

where the matrix coefficients are given by

$$C_{ab}^{mn} = \tfrac{1}{2}\langle \hat{M}_a \hat{N}_b + \hat{N}_b \hat{M}_a \rangle - \langle \hat{M}_a \rangle \langle \hat{N}_b \rangle. \tag{2.59}$$

If we only consider bipartite systems with zero coherent amplitude - and we generally will - the above expression further simplifies to

$$C_{ab}^{mn} = \tfrac{1}{2}\langle \hat{M}_a \hat{N}_b + \hat{N}_b \hat{M}_a \rangle. \tag{2.60}$$

The symmetry of the form of C_{ab}^{mn} requires that in general, $C_{ab}^{mn} = C_{ba}^{nm}$. The entire covariance matrix is thus described by 10 matrix coefficients [39].

2.4.2 Quantum Correlations, Inseparability and Entanglement

In the previous sections I have briefly introduced correlations in bipartite systems. I have made no reference to the quantum nature of these correlations, and have addressed them classically. In reality, even the presence of strong correlations in quantum systems does not ensure their 'quantumness'. Classical correlations can be embedded in quantum states and recent results argue most every correlated bi-partite system will possess both quantum and classical correlations. It is now widely accepted that states need not be entangled to demonstrate uniquely quantum correlations, and that nearly any quantum state selected at random will satisfy some metric for quantum correlations. Historically however, especially in the framework of pure states, the notion of quantum correlations is synonymous with the notion of inseparability. If we consider a general bi-partite pure state, described by a state vector

$$|\psi\rangle_{AB} = \sum_{i,j} c_{i,j} |i\rangle_A \otimes |j\rangle_B, \tag{2.61}$$

where $|i\rangle_A$ and $|j\rangle_B$ are defined in their respective Hilbert spaces H_A and H_B. The Schmidt decomposition states that for every $|\psi\rangle$ there exists bases $|u\rangle_A$ and $|v\rangle_B$ defined in H_A and H_B respectively such that

$$|\psi\rangle_{AB} = \sum_{i=1}^{n} \lambda_i |u\rangle_A \otimes |v\rangle_B, \tag{2.62}$$

where λ_i are non-negative real numbers satisfying $\sum_i^n |\lambda_i|^2 = 1$. The number n of non-zero values λ_i is referred to as the Schmidt number for the state $|\psi\rangle_{AB}$. The Schmidt

number allows us to characterise the separability of a given composite state: if $n = 1$ the state is separable, and if $n > 1$ the state is entangled. If all the coefficients λ_i are non-zero and equal, the state $|\psi\rangle_{AB}$ is maximally entangled. When restricted to pure states, any correlations apparent between local measurements of subsystems A and B guarantee the state is inseparable. Any correlations shared between the components of a pure bipartite state are quantum in character, and for pure states, entanglement and correlations are synonymous. Complications arise when we move away from pure states. A general density matrix ρ_{AB} is considered separable if it can be expressed as a mixture of product states

$$\rho_{AB} = \sum_{i=1}^{\lambda} p_i \, \rho_i^A \otimes \rho_i^B, \tag{2.63}$$

where $p_i \geq 0$ and $\sum_i p_i = 1$. If a given ρ does not satisfy (2.63) it is entangled, but subtleties arise in the nature of the correlations involved. In general, determining the separability of a given bipartite state ρ is difficult, as there are infinitely many ways express a given ρ as a mixture of pure states.

2.4.3 The Inseparability Criterion for Gaussian States

The problem of developing a criterion for inseparability is dramatically simplified when restricted to only a class of quantum states, and can become a tractable problem. The *inseparability criterion* introduced by Duan *et al.* [40] provides a necessary and sufficient condition for the inseparability of two-mode Gaussian states. This is the first of two entanglement witnesses for Gaussian two-mode states introduced in this thesis - the second being the EPR-paradox criterion introduced by Reid [41].

Consider the form of the general covariance matrix of (2.58) describing any two-mode Gaussian state. Duan *et al.* [40] have shown that (2.58) can be transformed through a series of local linear unitary Bogoliubov operators (LLUBOs) into *Standard Form I*, satisfying

$$C(\hat{X}_a, \hat{X}_b, \hat{P}_a, \hat{P}_b) = \begin{pmatrix} C_{aa}^{xx} & 0 & C_{ab}^{xx} & 0 \\ 0 & C_{aa}^{pp} & 0 & C_{ab}^{pp} \\ C_{ab}^{xx} & 0 & C_{bb}^{pp} & 0 \\ 0 & C_{ab}^{pp} & 0 & C_{bb}^{pp} \end{pmatrix}, \tag{2.64}$$

where the matrix components satisfy the following expressions

$$\frac{C_{aa}^{xx} - 1}{C_{bb}^{xx} - 1} = \frac{C_{aa}^{pp} - 1}{C_{bb}^{pp} - 1} \tag{2.65}$$

and

$$\sqrt{(C_{aa}^{xx} - 1)(C_{bb}^{xx} - 1)} - |C_{ab}^{xx}| = \sqrt{(C_{aa}^{pp} - 1)(C_{bb}^{pp} - 1)} - |C_{ab}^{pp}|. \tag{2.66}$$

The transformation consists of the application of a squeezing operation to ensure modes a and b have identical variances, and a quadrature rotation to eliminate any cross-quadrature

correlations. As the unitary operations applied to transform the state to the form of (2.64) are local, they have no effect on the separability or inseparability of the state. Duan *et al.* [40] showed that any general state with Glauber-Sudarshan P-representation that is positive everywhere (and thus is a quantum state with a classical analog) will satisfy the will satisfy the criterion

$$\langle(\hat{X}_I)^2\rangle + \langle(\hat{P}_I)^2\rangle \geq 2\left(k^2 + \frac{1}{k^2}\right), \tag{2.67}$$

where

$$\langle\hat{X}_a^2\rangle = \left\langle \left(k\hat{X}_a - \frac{1}{k}\frac{C_{ab}^{xx}}{|C_{ab}^{xx}|}\hat{X}_b\right)^2 \right\rangle, \tag{2.68}$$

and accordingly for $\langle\hat{P}_a^2\rangle$. The parameter k compensates for any asymmetry arising between the two subsystems a and b and is given by

$$k = \left(\frac{C_{bb}^{xx} - 1}{C_{aa}^{xx} - 1}\right)^{\frac{1}{4}} = \left(\frac{C_{bb}^{pp} - 1}{C_{aa}^{pp} - 1}\right)^{\frac{1}{4}}. \tag{2.69}$$

Violation of the *inseparability criterion* equality of (2.67) is a necessary and sufficient condition for the entanglement of a bipartite Gaussian state.

2.4.4 The EPR Paradox

In their seminal paper [36] of 1935, Einstein, Podolsky and Rosen introduced a *gedenken-experiment* that proved to be an illuminating critique of the formalism of quantum mechanics. They presented an apparent violation of the Heisenberg uncertainty principle could be achieved, with a pair of non-commuting observables simultaneously known to perfect precision.

The original thought experiment considers a system of two entangled particles, denoted a and b, with position and momenta q_a, q_b, p_a and p_b. The pair of particles interacted at some point in the past, but are now well separated. The authors argue that the formalism of quantum mechanics permits a wave-function describing the bi-partite system that is simultaneously an eigenstate of the linear operators $\hat{q}_a - \hat{q}_b$ and $\hat{p}_a + \hat{p}_b$. In this state, a measurement of position on particle a - which cannot physically affect particle b - allows one obtain perfect knowledge of the position of particle b. With a direct measurement of momentum on particle b, and the position of particle a known, particle b can seemingly have simultaneously well defined position and momentum. Outwardly, this appeared to be in direct conflict with the formalism of quantum mechanics, which demands the two conjugate observables cannot be known both precisely and simultaneously. The authors arguing that the wavefunction alone "does not provide a complete description of the physical reality" [36].

2.4.5 EPR Paradox Criterion for Continuous Variables

Whilst the original paper of Einstein, Podolsky and Rosen discussed the position and momentum of a pair of entangled particles. In 1988, Reid [41, 42] introduced a measure for the continuous variables analog of the paradox, where an argument can be constructed for the same apparent violation. The *EPR paradox criterion* is defined as the product of the two conditional variances for the amplitude and phase quadratures. Unlike the Inseparability criterion introduced in §2.4.3, the EPR paradox criterion is an inherently directional quantity. As a result, there are two criteria, depending on the direction of inference

$$\epsilon_{ab} = \Delta\hat{X}_{a|b}\Delta\hat{P}_{a|b} < 1 \tag{2.70}$$

$$\epsilon_{ba} = \Delta\hat{X}_{b|a}\Delta\hat{P}_{b|a} < 1 \tag{2.71}$$

where the quadrature conditional variances are defined as follows

$$\Delta\hat{X}_{a|b} = \Delta\hat{X}_a - \frac{|C^{xx}_{ab}|^2}{\Delta\hat{X}_b}, \tag{2.72}$$

and accordingly for $\Delta\hat{P}_{a|b}$. Provided the product of the conditional variances of the two orthogonal quadratures is below one, the state is said to demonstrate the EPR paradox. Of course, this is only an apparent violation, as the inequality is concerned with conditional variances, not the variance itself. An EPR criterion ϵ less than one is a *sufficient* but not *necessary* condition for entanglement.

2.5 Quantum State Tomography

The uncertainty principle of quantum mechanics, or equivalently, the *no-cloning theorem* require you cannot infer the quantum state or density matrix of a single quantum system without some prior knowledge of that system. A single measurement of some observable yields a single outcome corresponding to the projection of the quantum state into an eigenstate with some non-zero probability. The resulting measurement back-action precludes subsequent measurements. However, given several identical preparations of the same unknown quantum system, one can choose an set of measurements that allow characterisation of the density matrix. A recipe to do so was first provided by Fano, who defined the *quorum* - the minimum set of operators sufficient for determination of the density matrix.

For this problem quantum optics is itself uniquely placed; a balanced homodyne detector can measure all linear combinations of position and momentum, specifying a single mode of the electromagnetic field. The field of optical homodyne tomography began with the observation of Vogel and Risken that, as probability distributions describing the homodyne observables are given by the marginals of the Wigner function [43], homodyne observables are related to the Wigner function via the Radon transform. Therefore, akin to established classical imaging techniques, one can reconstruct the Wigner function via the measured distributions of the homodyne observables by inverting the Radon transform.

Here, we will briefly discuss three different approaches to quantum tomography: the

inverse-Radon transform method, the pattern function sampling, and the Maximum Entropy principle. I would also like to direct the reader to the excellent review of Lvovsky and Raymer[44], and also the text *Quantum State Estimation*[45].

2.5.1 The Inverse-Radon transform

The very first experimental demonstration of optical homodyne tomography was provided by Smithey *et al.*[46], reconstructing the Wigner function (and density matrix) via the inverse Radon transformation.

Recall the general homodyne observable,

$$\hat{x}_\theta = \hat{x}\cos\theta + \hat{p}\sin\theta, \tag{2.73}$$

described by a marginal distribution $w(x_\theta, \theta) = \langle x_\theta|\hat{\rho}|x_\theta\rangle$. The Wigner function $W(q,p)$ of the quantum state, $\hat{\rho}$ and the marginals $w(x_\theta,\theta)$ are related via the Radon transform,

$$w(x_\theta,\theta) = \int W(q,p)\,\delta(x_\theta - q\cos\theta - p\sin\theta)\,dq\,dp \tag{2.74}$$

$$= \int W(x_\theta\cos\theta - x_{\theta+\frac{\pi}{2}}\sin\theta, x_\theta\sin\theta - x_{\theta+\frac{\pi}{2}}\cos\theta)\,dx_{\theta+\frac{\pi}{2}}. \tag{2.75}$$

Or more simply, the marginal, $w(x_\theta,\theta)$, is simply the projection of the Wigner function, $W(q,p)$, onto the vertical plane oriented at an angle, θ. Vogel and Risken [43] showed that knowledge of $w(x_\theta,\theta)$ for all values of θ was equivalent to knowledge of the Wigner function itself.

Given our set of marginal distributions, $w(x_\theta,\theta)$, the Wigner function can be obtained by 'inverting' the Radon transform,

$$W(q,p) = \frac{1}{\pi^2}\int_0^\pi\int_{-\infty}^\infty w(x_\theta,\theta)\times \mathrm{K}(q\cos\theta + p\sin\theta - x_\theta)\,dx_\theta\,d\theta, \tag{2.76}$$

with the integration kernel,

$$\mathrm{K}(x) = \int_{-\infty}^\infty |\xi|\exp i\xi x\,d\xi. \tag{2.77}$$

The kernel is infinite at $x=0$ and needs to be regularised. The usual approach is to restrict the integration limits to some frequency, $\pm k_c$, corresponding to a low-pass filtering of the Wigner function. The choice of k_c is an open problem. For a choice of k_c too high, unphysical high frequency components associated with statistical noise can dominate the reconstruction. Set k_c too low and finer details of Wigner function will be smoothed away. Problematically, quantum mechanics provides no guidance regarding the choice of k_c; it is a somewhat arbitrary choice left to the judgement of the experimentalist. The tell-tale ripples of the inverse Radon transform are evident in our previous work[47].

2.5.2 Pattern Functions

First introduced by dAriano *et al.* [48] and Kühn [49], and further developed by Leonhardt *et al.* [50, 51] and Richter [52, 53], the *pattern functions* specify a set of functions that allow direct sampling of the density matrix of a single mode of the electromagnetic field.

The density matrix, $\hat{\rho}$ in the Fock basis can be reconstructed directly from the measured quadrature distributions, $w(x_\theta, \theta)$, by sampling the individual density matrix elements, ρ_{mn} via

$$\rho_{mn} = \int_0^\pi \int_{-\infty}^{\infty} F_{mn} w(x_\theta, \theta) dx_\theta d\theta. \tag{2.78}$$

where F_{mn} are a set of *sampling* functions. This corresponds to simply averaging the sampling function over the measured quadrature distribution. The Wigner function is then accessible via the relation,

$$W(x_\theta, \theta) = \sum_{m,n} \rho_{mn} W_{|m\rangle|n\rangle}(x_\theta, \theta). \tag{2.79}$$

The set of sampling functions required for (2.78) are specified by [50, 54]

$$F_{mn} = f_{mn}(x_\theta) e^{i(m-n)\theta}, \tag{2.80}$$

where f_{mn} are the so-called *pattern functions*. Note that there is only a dependence on θ for the off-diagonal elements ($m \neq n$) corresponding to the coherences. The method for computing these pattern functions for the Fock basis (and also the coherent state basis) is provided in [50]. The pattern function approach is more efficient than the customary inverse-Radon transformation[51] and also bypasses the somewhat arbitrary filtering the inverse-Radon transform requires. Instead, the dimension of Hilbert space for the reconstruction must be truncated, though this a constraint with a clear physical interpretation, and is also a condition of the Maximum Likelihood and Maximum Entropy principles. Excepting the truncation of the Hilbert space, the pattern function method involves no additional restrictions or assumptions regarding the nature of the unknown state. As a result, the statistical noise associated finite measurement ensemble can manifest in unphysical contributions (negative elements), and as each element of the density matrix is sampled independently, nothing constrains the reconstructed state to be normalised.

2.5.3 Maximum Entropy Principle

The *maximum entropy principle* was first applied to the problem of quantum tomography by Bužek and Drobný [55]. Like its more popular cousin, *maximum likelihood estimation* [56], it provides a robust approach to tomography by using variational principles within a framework defined by quantum mechanics. Unlike maximum likelihood estimation, which seeks to find the most-probable state described by the measurement outcomes, the principle of maximum entropy minimises the knowledge obtained from the measurement record. As such, it reconstructs the least-biased state consistent with the measurement record, and is well suited to applications where the measurement record is tomographically incomplete.

Consider our usual problem of optical homodyne tomography, where we sample the quadrature distributions, $w(x_\theta, \theta)$. Our final measurement record is binned into a rectangular array of dimension $M_x \times N_\theta$. The population of the (m, n) bin, p_{mn}, associated with the quadrature value, x_θ^m ,and angle, θ_n, is proportional to the expectation value of the observable, $\hat{\Pi}_{mn} = \left|x_{\theta_n}^m\right\rangle\!\left\langle x_{\theta_n}^m\right|$. As the name suggests, the maximum entropy principle seeks to find an estimate of $\hat{\rho}$ that maximises the von Neumann entropy (§2.7.1)

$$S(\hat{\rho}) = -\mathrm{tr}(\hat{\rho}\,\log_2\hat{\rho}). \tag{2.81}$$

whilst fulfilling the conditions,

$$\mathrm{Tr}\hat{\rho} = 1 \quad \text{and} \quad \mathrm{Tr}\hat{\rho}\,\Pi_{mn} = p_{mn}. \tag{2.82}$$

The state satisfying these conditions is given by

$$\hat{\rho}_{\mathrm{Max}} = \mathcal{N}\left[\exp\left(-\lambda_0\hat{n} - \sum_{m=1}^{M_\theta}\sum_{n=1}^{N_\theta}\lambda_{mn}\hat{\Pi}_{mn}\right)\right], \tag{2.83}$$

where $\hat{n}$ is the usual photon number operator, and λ_{mn} are the Lagrange multipliers which allow us to constrain the state to the conditions of (2.83). Once can solve for the Lagrange multipliers by minimising the deviation function,

$$\Delta x_\theta = (\bar{n} - \mathrm{Tr}(\hat{\rho}_{\mathrm{Max}}\hat{n}))^2 - \sum_{m=1}^{M_\theta}\sum_{n=1}^{N_\theta}\left[p_{mn} - \mathrm{Tr}\left(\hat{\rho}_{\mathrm{Max}}\hat{\Pi}_{mn}\right)\right]^2, \tag{2.84}$$

where $\bar{n}$ is the mean photon number.

2.6 Classical Information Theory

2.6.1 Shannon Entropy

> *"You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage."*
> – John von Neumann to Claude Shannon regarding what to name the attenuation experienced by phone line signals.

With his landmark paper of 1948 [6], Claude Shannon provided the founding text for the field of information theory, the mathematical backbone of which is its variant of entropy. The Shannon entropy, $H(X)$, is a measure of the *unpredictability* associated with the outcome of a random variable, X before the outcome is known. Alternatively, and perhaps more intuitively, it can thought of as the average amount of information gained about a random variable, X after the outcome is broadcast. The Shannon entropy is a

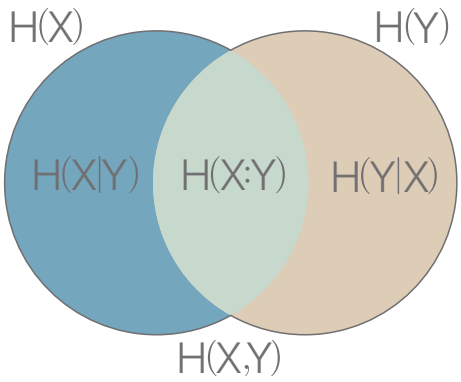

Figure 2.5: A Venn diagram capturing the relationships between the different entropies: the Shannon entropies $H(X)$ and $H(Y)$; the joint entropy $H(X,Y)$; the conditional entropies, $H(X|Y)$ and $H(Y|X)$; and the mutual information $H(X:Y)$.

familiar recasting of the Boltzmann entropy of statistical mechanics,

$$H(X) = -\sum_i P(x_i) \log_b P(x_i) \tag{2.85}$$

where $P(x_i)$ is the probability of obtaining an outcome x_i for variable X, and b is the base of the logarithm (usually 2). Equation 2.85 quantifies the *unevenness* of the probability distribution P describing your random variable X. This concept of *unevenness* often best elucidated with the example of a coin toss. Consider the coin toss of a fair coin: there are two possible outcomes, each occurring with equally probability. The best we can do is make a guess as to the outcome of the coin toss, and 50% of the time we will be correct. As the uncertainly is symmetrically distributed between the two outcomes, the entropy of the system is maximal. The entropy associated with a fair coin toss is one bit, and we obtain one bit of information when we learn the outcome of the toss. But if we now consider tossing a biased coin - perhaps biased to an outcome of heads - because we possess some information regarding the likeliness of each outcome the entropy of the coin toss is reduced. Consider the extreme of a coin of two heads: as the outcome of each coin toss can be predicted perfectly the entropy of the system is zero.

The english text is a (popular) example of a low entropy system. If we consider we have 26 possible characters in the english alphabet, if all letters are equally probably, their entropy would be $\log_2 26$, or approximately 8 bits. Probabilistically however, all letters of the english alphabet are not created equal. For a randomly chosen word, we intuitively know we are more likely to find an 'e' than we are to encounter a 'z'. If we are given an unknown word of an unknown 5 letters, uncovering a 'z' provides us more information than uncovering an 'e'. Or if we encounter a 'q', we know it is extremely probable the next letter will be a'u'. The entropy of the alphabet is estimated at between $1 - 1.5$ bits per letter.

2.6.2 Joint Entropy

In the last section we considered the entropy of a single random variable, but consider two random variables, X and Y, which may or may not be independent. The pair $\{X, Y\}$ is described by a set of possible outcomes $\{x, y\}$ in two-dimensional space occurring with a probability $P(x, y)$. The *joint Shannon entropy* of X and Y is given by

$$H(X, Y) = -\sum_x \sum_y P(x, y) \log_2 P(x, y), \tag{2.86}$$

$$H(X_1, \ldots, X_2) = -\sum_{x_1} \ldots \sum_{x_n} P(x_1, \ldots, x_n) \log_2 P(x_1, \ldots, x_n). \tag{2.87}$$

The joint entropy is bounded above by the sum of all the individual entropies, $H(X_1, \ldots, X_n) \leq H(X_1) + \ldots + H(X_n)$ and from below by the maximum individual entropy of the set of individual entropies, $H(X_1, \ldots, X_n) \geq \max(H(X_1), \ldots, H(X_n))$.

2.6.3 Conditional Entropy

Suppose the outcome of Y is known, and thus we have acquired $H(Y)$ bits of information about the pair $\{X, Y\}$. Any remaining uncertainty regarding the pair $\{X, Y\}$ is associated with our residual ignorance of X given our knowledge of Y. The entropy of X *conditional* on our knowledge of Y is given by

$$H(X|Y) \equiv H(X, Y) - H(Y). \tag{2.88}$$

The conditional entropy captures our average uncertainty regarding X given knowledge of Y. Consider the scenario where X and Y are completely independent random variables: knowledge of Y does not surrender any information regarding X, and thus the conditional entropy of X given Y is simply the original entropy of $H(X)$. If instead X and Y were perfectly correlated, knowledge of Y perfectly specifies X and the conditional entropy $H(X|Y)$ is zero.

2.6.4 Mutual Information

The entropy of the joint distribution comprising the pair X and P is simply the joint entropy defined §2.6.2. But suppose instead we crudely add the individual entropies of X and Y; any information common to both X and Y would be counted twice, whilst any independent information is counted once. The subtracting the joint entropy from this quantity we obtain the common or *mutual information* between X and Y,

$$H(X:Y) = H(X) + H(Y) - H(X, Y). \tag{2.89}$$

The mutual information is a measure of the inherent dependence of two random variables, X and Y. Using (2.88) the mutual information can also be expressed in terms of the

conditional entropy

$$H(X:Y) = H(X) - H(X|Y), \tag{2.90}$$

or equivalently

$$H(X:Y) = H(Y) - H(Y|X). \tag{2.91}$$

Classically, all three expressions for the mutual information are equivalent. Whilst this may seem obvious as we a describing quite non-mysterious correlations of classical probability distributions, subtleties arise when we will later consider the generalisation of the mutual information to quantum mechanical systems.

2.7 Quantum Information Theory

2.7.1 von Neumann Entropy

Until now we have considered the Shannon entropy and its variants within the framework of classical information theory. Though quantum mechanics was a largely matured and fully formed theory by the time Shannon's 'A Mathematical Theory of Communication' appeared in 1948, the field of quantum information - where information protocols are formed around the idea of encoding in the state of a quantum mechanical system - emerged in the 1980's. The theoretical backbone of quantum information theory is the von Neumann entropy. Named after John von Neumann - one of the great polymaths of the 20th century - the von Neumann entropy is the generalisation of classical entropy to quantum mechanical systems. The von Neumann entropy for a system described by a density operator ρ is given by,

$$S(\rho) = -\mathrm{tr}(\rho \log_2 \rho) = -\sum_i \lambda_i \log_2 \lambda_i, \tag{2.92}$$

where $\{\lambda_i\}_i$ are the eigenvalues of ρ.[4] As $0 \leq \lambda_i \leq 1$, the von Neumann entropy is strictly non-negative, and is only zero if the quantum system is described by a pure state ($\lambda_i = 1$). For a state, ρ, described in a Hilbert space of dimension d, the entropy is at most $\log d$, corresponding to a maximally mixed state. If one can find an orthonormal basis, $|x\rangle$, that diagonalises ρ in the form,

$$\rho = \sum_x \lambda_x |x\rangle\langle x| \tag{2.93}$$

then the von Neumann entropy reduces to the Shannon entropy, $S(\rho) = H(X)$, where $H(X)$ is the Shannon entropy of the classical ensemble X with a set of outcomes $\{x\}$ occurring with a probabilities $\{\lambda_x\}$. For the state decomposed in an orthogonal basis, all outcomes can be distinguished perfectly and the problem reduces to a classical one.

[4]The logarithm is near-universally taken to the base 2, owing to the majority of quantum information applications considering binary encoding of *qubits*. Occasionally the natural log appears in information theory as it sometimes proves more amenable for calculations. The corresponding information unit is *nats*.

Though for a single quantum state solving the eigenvalue problem. In general, however - especially for composite quantum systems - this is often not the case, and the implications of this will be discussed in the context of 'quantum discord'.

Before we consider other quantum generalisations of entropic quantities, we need to establish notation for the von Neumann entropy of composite systems and their parts. Given a bipartite system ρ_{AB}, we generalise the joint entropy in the obvious way,

$$S(\rho_{AB}) = -\mathrm{tr}(\rho_{AB} \log_2 \rho_{AB}). \tag{2.94}$$

The entropy of the reduced subsystems is given by $S(\rho_A) = S(\mathrm{tr}_B(\rho_{AB})$, and equivalently for $S(\rho_B)$. Analogous to the classical scenario, the von Neumann entropy is additive for independent systems (product state) $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$, and this provides with an upper bound on $S(\rho_{AB})$ for any bipartite quantum state.

2.7.2 Quantum Conditional Entropy

With the von Neumann entropy providing an generalisation of the Shannon entropy to quantum systems, we can also consider quantum formulations of the conditional entropy (§2.6.3) and mutual information (§2.6.4). By analogy with the classical equivalent of (2.88) the quantum conditional entropy is defined as

$$S(\rho_A|\rho_B) \equiv S(\rho_{AB}) - S(\rho_B). \tag{2.95}$$

First defined in [57] it was immediately noticed that quite unlike its classical counterpart, the quantum conditional entropy could be negative. Even though the Von Neumann entropy of any individual quantum variable is strictly non-negative, counter-intuitively, the entropy of the entire quantum system can be, and often is, smaller than the individual entropies of its reduced subsystems. That is to say: our uncertainty regarding the entire quantum system is sometimes smaller than our uncertainty regarding its individual subsystems.

Consider, for instance, a pure two-mode squeezed state (§2.2.5). Whilst the entropy of the entire system is zero, the individual subsystems are locally thermal states with a positive entropy, and accordingly, the quantum conditional entropy would be negative. A operational meaning for this 'negative' information was provided in [58] by introducing a new quantum information primitive called "quantum state merging".

2.7.3 Quantum Mutual Information

In §2.6.4 we introduced three equivalent expressions for the classical mutual information. Using (2.95) we can write down the quantum generalisations of the mutual information,

$$\mathcal{I}(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \tag{2.96}$$

$$= S(\rho_A) - S(\rho_A|\rho_B) \tag{2.97}$$

$$= S(\rho_B) - S(\rho_B|\rho_A). \tag{2.98}$$

The quantum mutual information captures all the information shared between the two partitions of a bipartite state. Where the Shannon mutual information describes the correlations shared between the bi-partitions if the system was described by a classical probability distribution, the quantum mutual information encapsulates *all* correlations within the system, whether classical or quantum in origin. Though all three expressions are equivalent, earlier I hinted at subtleties that arise when one considers the quantum generalisation of the mutual distinction. This distinction arises in the notion of accessible information and will be discussed in the context of its relevance to quantum discord in Chapter 6.

2.7.4 Holevo's Bound

Holevo's bound [59] is one of the earliest and most significant results in quantum information. It provides an immensely useful upper bound on the amount of information that can be known about a quantum state. Holevo's bound is usually best elucidated by considering two-party communication. Consider two parties, Alice and Bob.[5] Alice possesses a classical random variable X from which she draws values x with a probability p_x. Based on her outcome x, Alice prepares a mixed state ρ_x which she transmits to Bob. The total density matrix of the system is described by the mixture of mixed states,

$$\rho = \sum_x p_x \rho_x. \tag{2.99}$$

Bob is tasked with discovering X. Upon receiving ρ_x Bob preforms a measurement, obtaining a classical value Y. We then pose the question: is there any fundamental limit to the amount of information Bob can obtain about Alice's random variable X? Or more formally, is there an upper bound on the mutual information $S(X:Y)$ between Alice's variable X and Bob's measurement record Y? This upper bound on the accessible information is given by:

$$S(X:Y) \geq S(\rho) - \sum_i p_i S(\rho_i) \tag{2.100}$$

where $\rho = \sum_i p_i \rho_i$. This result summarises perhaps a surprising difference between quantum systems and classical systems. In classical information theory the notion of accessible information is not particularly interesting - one should, in theory, always be able to differentiate between two classical information states. However in quantum mechanics this is not always the case. Given two arbitrary quantum states, it is not always possible to find a measurement that will permit you to distinguish between the two. This scenario is not a consequence of lacking the optimal measurement, rather, it is a fundamental property of the quantum system.

[5] This also marks the debut of the nomenclature of information theory: where Alice and Bob are the quintessential placeholder names for parties A and B for communication and information protocols.

2.8 From discrete to continuous modes

So far we have only considered idealised single modes, we now briefly transition to a continuous mode description for our annihilation and creation operators. A formal transition from the discrete mode formalism to the continuous mode formalism can be found in Loudon [60]. A mathematically rigorous extended formalism to describe continuum multimode quantum states was developed by Caves and Schumaker in two successive publications [61, 62]. From these, the continuous mode creation and annihilation operators emerge, which are related to their discrete-mode counterparts (introduced in §2.2.1) by

$$\hat{a}_k \to \sqrt{\Delta\omega}\,\hat{a}(\omega) \;\text{ and }\; \hat{a}_k^\dagger \to \sqrt{\Delta\omega}\,\hat{a}^\dagger(\omega). \tag{2.101}$$

The discrete Kronecker delta and the continuous Dirac delta-function are related by $\delta_{k,k'} \to \Delta\omega\,\delta(\omega-\omega')$, giving the new, but hopefully reminiscent, continuous-mode commutation relation,

$$\left[\hat{a}(\omega), \hat{a}^\dagger(\omega')\right] = \delta(\omega-\omega'). \tag{2.102}$$

Namely, they commute unless they describe the same mode.

2.8.1 Fourier domain operators

In most practical applications, the frequency bandwidth is much smaller than the central frequency. This narrow-band assumption allows us to extend the lower limit of frequency integration to $-\infty$ without significant error. This motivates a time-domain definition of the creation and annihilation operators

$$\hat{a}(t) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\hat{a}(\omega)\exp\left(-i\omega t\right) \tag{2.103}$$

$$\hat{a}^\dagger(t) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,\hat{a}^\dagger(\omega)\exp\left(i\omega t\right), \tag{2.104}$$

where the above has been chosen to be consistent with $\hat{a}^\dagger(t) = [\hat{a}(t)]^\dagger$. Through an inverse Fourier transform we obtain our frequency-domain creation and annihilation operators,

$$\hat{a}(\omega) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{a}(t)e^{i\omega t}dt \tag{2.105}$$

$$\hat{a}^\dagger(\omega) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{a}^\dagger(t)e^{i\omega t}dt, \tag{2.106}$$

and also the relevant Hermitian conjugates,

$$[\hat{a}(\omega)]^\dagger \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{a}(t)e^{-i\omega t}dt \tag{2.107}$$

$$[\hat{a}(-\omega)]^\dagger \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{a}^\dagger(t)e^{i\omega t}dt \tag{2.108}$$

$$= \hat{a}^\dagger(\omega). \tag{2.109}$$

Denoting $\mathcal{F}[\cdot]$ to be a Fourier transform, we can now consider the frequency domain generalisation of the familiar generalised quadrature operator,

$$\hat{X}^{\theta}(\omega) = \mathcal{F}[e^{i\theta}a^{\dagger}(t) + e^{-i\theta}a(t)] \tag{2.110}$$

$$= (\cos\theta + i\sin\theta)\hat{a}(\omega) + (\cos\theta - i\sin\theta)\hat{a}^{\dagger}(\omega) \tag{2.111}$$

$$= \cos\theta(\hat{a}(\omega) + \hat{a}^{\dagger}(\omega)) + i\sin\theta(\hat{a}^{\dagger}(\omega) - \hat{a}(\omega)) \tag{2.112}$$

$$= \cos\theta\hat{X}(\omega) + \sin\theta\hat{P}(\omega). \tag{2.113}$$

where, using (2.109) we have defined the frequency domain amplitude and phase quadrature operators,

$$\hat{X}(\omega) = \hat{a}(\omega) + [\hat{a}(\omega)]^{\dagger} \quad \text{and} \quad \hat{P}(\omega) = i\left([\hat{a}(\omega)]^{\dagger} - \hat{a}(\omega)\right). \tag{2.114}$$

2.8.2 Linearised decomposition of the operators

A more intuitive formalism for continuous variable quantum optics comes from a linearised decomposition of the operators. The first use of this method in the context of quantum optics dates back to Yurke [63]. For the purposes of most experimental situations, the fluctuations of the field are negligible when compared the average intensity of the field, and higher order fluctuation terms are negligible for most scenarios. The linearisation procedure decomposes the annihilation operator into two contributions: a steady-state term associated with the expectation value, or amplitude α, of the single frequency, and a fluctuating term $\delta\hat{a}(t)$, describing the fluctuations in the continuum of modes surrounding the carrier. We thus write,

$$\hat{a}(t) = \alpha + \delta\hat{a}(t) \text{ and } \hat{a}^{\dagger}(t) = \alpha^{*} + \delta\hat{a}(t). \tag{2.115}$$

We also require two assumptions,

$$\langle\delta\hat{a}(t)\rangle = 0 \text{ and } |\delta\hat{a}(t)| \ll |a|. \tag{2.116}$$

The first, that the fluctuation term has no net contribution to the field amplitude and is perfectly centred around zero, and second, that any fluctuations are much smaller than the steady-state amplitude, α. These two assumptions allow us to neglect any contribution from higher-order quantum fluctuation terms (e.g. $\delta\hat{a}^{\dagger}(t)\delta\hat{a}(t)$). In the same manner, we can also simplify our description of the quadrature operators via quadrature fluctuation operators,

$$\delta\hat{X}(t) = \delta\hat{a}^{\dagger}(t) + \delta\hat{a}(t) \tag{2.117}$$

$$\delta\hat{P}(t) = i(\delta\hat{a}^{\dagger}(t) - \delta\hat{a}(t)). \tag{2.118}$$

2.8.3 Phase and amplitude modulation

Amplitude modulation is a direct modulation of the intensity of a light field at some modulation frequency, ω_m. The amplitude modulation of an initial field $\hat{a}_0(t)$ with carrier

frequency, Ω is given by

$$\hat{a}(t) = \hat{a}_0(t)(1 - \xi + \xi \cos \omega_m t), \tag{2.119}$$

where ω_m is the modulation frequency, and ξ represents the modulation depth. Equation (2.119) can be rewritten in the perhaps more illuminating form of,

$$\hat{a}(t) = \hat{a}_0(t)(1 + \tfrac{\xi}{2}(e^{i\omega_m t} + e^{-i\omega_m t})), \tag{2.120}$$

with the additional assumption of a small modulation depth, $\xi \ll 1$. Equation (2.120) decomposes the modulated field $\hat{a}(t)$ into the original carrier $\hat{a}_0(t)$ at frequency Ω, and two sidebands of equal amplitude $\frac{\hat{a}_0(t)\xi}{2}$ at optical frequencies $\Omega + \omega_m$ and $\Omega - \omega_m$. The modulation depth, ξ, relates the proportion of energy transferred from the carrier mode to the two generated sideband modes. Taking the Fourier transform we obtain

$$\hat{a}(\Omega) = \hat{a}_0(\omega) + \frac{\xi}{2} \int_{-\infty}^{\infty} \left[\hat{a}_0(t) e^{i(\Omega+\omega_m)t} + \hat{a}_0(t) e^{i(\Omega-\omega_m)t} \right] dt \tag{2.121}$$

$$= \hat{a}(\Omega) + \frac{\xi}{2}\hat{a}(\Omega + \omega_m) + \frac{\xi}{2}\hat{a}(\Omega - \omega_m), \tag{2.122}$$

where the amplitude modulation is comprised of an upper $(\Omega + \omega_m)$ and lower $(\Omega + \omega_m)$ sideband.

As the name suggests, *phase modulation* is a modulation of the phase of a light field. The phase modulation of some initial field, $\hat{a}_0(t)$ is described by

$$\hat{a}(t) = \hat{a}_0(t)\, e^{i\xi \cos \omega_m t}. \tag{2.123}$$

For $\xi \ll 1$, we consider only the first order expansion,

$$\hat{a}(t) = \hat{a}_0(t)(1 + i\xi \cos \omega_m t) \tag{2.124}$$

$$= \hat{a}_0(t)(1 + \frac{i\xi}{2}(e^{i\omega_m t} + e^{-i\omega_m t})). \tag{2.125}$$

Taking the Fourier transform we obtain,

$$\hat{a}(\Omega) = \hat{a}_0(\Omega) + \frac{i\xi}{2} \int_{-\infty}^{\infty} \left[\hat{a}_0(t) e^{i(\Omega+\omega_m)t} + \hat{a}_0(t) e^{i(\Omega-\omega_m)t} \right] dt \tag{2.126}$$

$$= \hat{a}(\Omega) + \frac{i\xi}{2}\hat{a}(\Omega + \omega_m) + \frac{i\xi}{2}\hat{a}(\Omega - \omega_m). \tag{2.127}$$

Phase modulation produces an upper $(\Omega+\omega_m)$ and lower $(\Omega+\omega_m)$ sideband with imaginary amplitudes.

Phase modulation is straightforward to realise in the laboratory. Simply varying the path length by dithering the position of a mirror is sufficient for low frequencies. Higher frequency phase modulation usually employs the electro-optic effect, with a sinusoidally varying electric field applied across a crystal to modulate the refractive index, and thus the optical path length. Amplitude modulation is usually tricker, here we again use the electro-optic effect. Applying an electric field to a suitable electro-optic crystal can produce

a field dependent birefringence, and thus a field dependent polarisation. Selecting one of the components of the output field with a polariser results in an amplitude modulation of the field.

2.9 Linear Optics, Losses and Detection

2.9.1 The Beam-Splitter

The humble beam splitter is ubiquitous in quantum optics. Both theoretically and experimentally, it is one of the most simple and powerful tools we have available. The beam-splitter allows us to interfere two input modes $\hat{a}_1$ and $\hat{a}_2$ which share a frequency, polarisation and transverse spatial profile. Figure 2.6 shows a schematic diagram of the beamsplitter with transmissivity η with inputs $\hat{a}_1$ and $\hat{a}_2$. On transmission the phase of

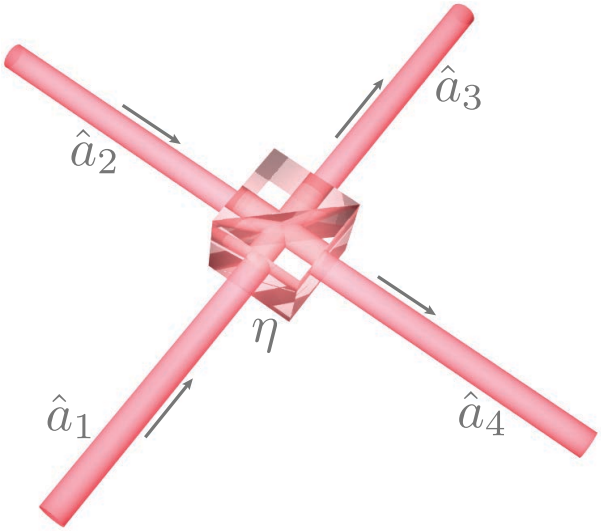

Figure 2.6: The schematic diagram of a beamsplitter. $\hat{a}_1$ and $\hat{a}_2$ are the fields incident, $\hat{a}_3$ and $\hat{a}_4$ are the resulting outputs, and η is the transmissivity.

each field remains unchanged, however, on reflection one field obtains a π phase shift. The following output fields are therefore given by

$$\begin{aligned}\hat{a}_3 &= \sqrt{\eta}\hat{a}_1 + \sqrt{1-\eta}\hat{a}_2 \\ \hat{a}_4 &= \sqrt{1-\eta}\hat{a}_1 - \sqrt{\eta}\hat{a}_2.\end{aligned} \tag{2.128}$$

The two incident fields are coupled together with a strength dependent on the transmissivity, η. There are two approaches to our understanding of the beamsplitter transformation in quantum optics. The first considers the beamsplitter as a randomisation device, whereby an incident photon is probabilistically directed to one of the two ports. The continuous field perspective considers that even in the absence of light at one of the inputs, vacuum fluctuations are coupled into the system through the empty port. One must consider the role of the vacuum mode in both scenarios to accurately model the physics of the system. It is this tidy concept that allows us to utilise the simple beam splitter to model all

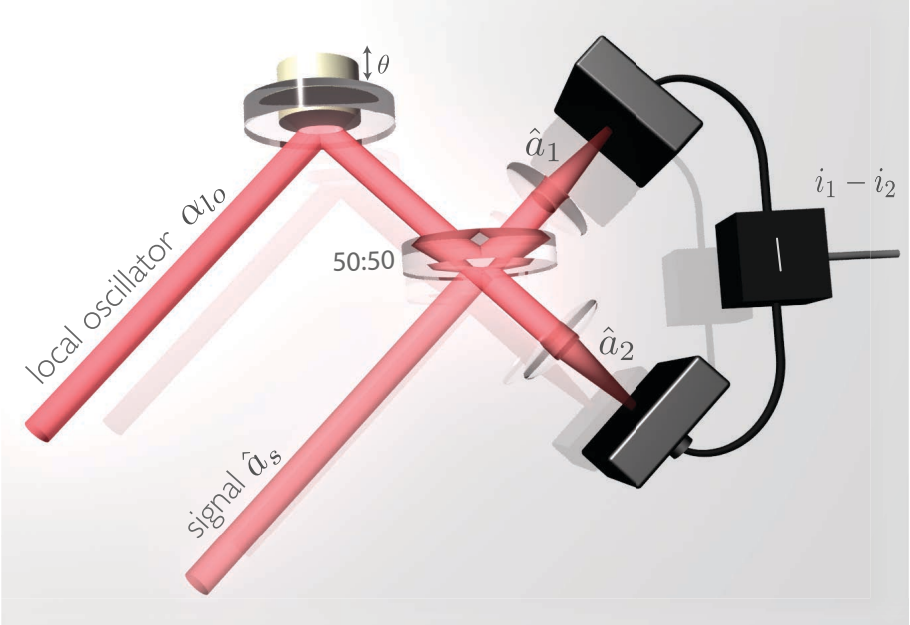

Figure 2.7: Schematic of balanced homodyne detection. The difference current is proportional to the quadrature observable of the input signal.

manner of losses in optical systems. Factors such as inefficient photo-detection, attenuation through optical elements, and imperfect spatial mode matching all attenuate the field concerned and introduce undesirable vacuum fluctuations. These effects can be modelled by simple beam splitter transformation of the field, with the vacuum mode entering the empty port.

2.9.2 Direct detection

Perhaps the simplest measurement of a quantum state of light is a direct measurement of intensity, that is, a measurement of $\hat{a}^\dagger\hat{a}$. In actuality, this measurement is essentially all an individual photodetector can ever do, producing a photocurrent, $i(t)$, proportional to the number of photons in the optical field,

$$i(t) \propto \hat{a}^\dagger(t)\,\hat{a}(t) \tag{2.129}$$

$$\propto (\alpha^* + \delta\hat{a}^\dagger(t))(\alpha + \delta\hat{a}(t)) \tag{2.130}$$

$$\propto |\alpha|^2 + \alpha\,\delta\hat{a}^\dagger(t) + \alpha^*\,\delta\hat{a}(t) + \delta\hat{a}^\dagger(t)\,\delta\hat{a}(t). \tag{2.131}$$

Considering only the first order terms (as $|\delta\hat{a}(t)| \ll |a|$), and assuming α is real we obtain

$$i(t) \propto \hat{a}^\dagger(t)\,\hat{a}(t) \approx\ |\alpha|^2 + \alpha\,\delta\hat{X}^+(t). \tag{2.132}$$

For our linearised decomposition of the field, the photocurrent is comprised of two terms: a DC term associated with the average optical intensity $|\alpha|^2$, and time dependent term, $\delta\hat{X}^+(t)$, describing the fluctuations in the amplitude quadrature.

2.9.3 Homodyne detection

As a solitary photodetector is insensitive to the quadrature amplitudes of an optical field, probing the phase information requires that we introduce a phase reference. Homodyne detection utilises interference of two phase-coherent fields to give a measurement of an arbitrary quadrature amplitude, $\hat{X}^\theta$.

The principle for optical balanced homodyne detection was developed in Yuen and Shapiro [64] in 1980. A basic schematic for balanced homodyne detection is given in Figure 2.7. The signal field, $\hat{a}_s$ of interest is interfered on an (ideally) $50:50$ beam splitter with comparatively very intense coherent beam, α_{lo}. This *local oscillator* provides a phase reference for our quadrature measurement, and it is sufficient for us to assume it is bright enough to be treated classically. Consider the measured photocurrents i_1 and i_2 are proportional to the photon numbers, $\hat{n}_1$ and $\hat{n_2}$. Using the relation of (2.128), the modes emerging from the beamsplitter are

$$\begin{aligned}\hat{a}_1 &= \frac{1}{\sqrt{2}}(\hat{a}_s - \alpha_{lo}) \\ \hat{a}_2 &= \frac{1}{\sqrt{2}}(\hat{a}_s + \alpha_{lo}),\end{aligned} \tag{2.133}$$

where $\hat{a}_s$ denotes the annihilation operator of the signal field, and α_{lo} the complex amplitude of the local oscillator. The difference current, i_- is proportional to the difference photon number,

$$\hat{n}_- = \hat{n}_1 - \hat{n}_2 \tag{2.134}$$

$$= \hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2 \tag{2.135}$$

$$= |\alpha_{lo}|(e^{i\phi}\hat{a}_s + e^{-i\phi}\hat{a}_s^\dagger), \tag{2.136}$$

where we have used $\alpha_{lo} = |\alpha_{lo}|e^{i\phi}$, where ϕ is the phase of the local oscillator relative to the signal field. From (2.16) we can see that the difference photocurrent is directly proportional to the generalised quadrature operator,

$$i_- \propto |\alpha_{lo}|\hat{X}_s^\phi. \tag{2.137}$$

Homodyne detection allows us to sample any arbitrary quadrature amplitude by varying the relative phase between the two fields. The local oscillator itself acts as an amplifier, 'boosting' the measured quadrature by a scaling factor proportional to the local oscillator amplitude, $|\alpha_{lo}|$. This enables us to sample the quantum attributes of states that may, on average, contain less than a photon without being defeated by the inherent electronic noise of our measuring device. Nevertheless, (2.136) is only true for a 'balanced' homodyne detection, as we require cancellation of the large $|\alpha|^2$ terms. Experimentally, imperfect balance can be remedied though electronic attenuation of one photocurrent relative to the other, allowing close to ideal cancellation of the classical noise associated with the local oscillator. While the analysis of (2.136) is rather crude, it recovers the same result as the more sophisticated analyses of [64, 65, 66, 67].

The inclusion of a fictitious beam-splitter is all that is usually required in quantum

optics to model passive loss. It is no different for the usual inefficiencies that arise in homodyne detection. Thus we can model a lossy detector as a perfect device preceded by a beamsplitter with a transmission equal to the detection efficiency.

2.10 Summary

In this chapter we have provided the theoretical background and experimental techniques required for this thesis. We provided a brief overview of quantum optics and its experimental techniques, classical and quantum information theory, and quantum tomography.

Chapter 3

A Continuous Variable Analog of a Photon Counting Measurement: Part I

3.1 Introduction

Central to the weirdness of quantum mechanics is the notion of wave-particle duality, where classical concepts of particle or wave behaviour alone cannot provide a complete description of quantum objects. When investigating quantum systems, information concerning one description is typically sacrificed in favour of the other, depending on which description suits your endeavour. Probing the continuous variables of an infinite Hilbert space, such as the amplitude and phase of a light field, is often viewed as less interesting than probing the quantised variables of a quantum system. This is largely due to the fact that, given current technology, when probing the continuous variables (CV) of a quantum system alone, one is restricted to transformations that map Gaussian states onto Gaussian states. This restriction ensures that computing protocols involving only Gaussian states and Gaussian operations can always be efficiently simulated on a classical computer.

Nevertheless, the idea of measuring the quantised nature of light with only CV techniques has been theoretically [50, 68, 69, 70] and experimentally [71, 72, 73, 74] investigated. The usual CV toolbox of Gaussian transformations, comprising beam splitters, displacements, rotations, squeezing, homodyne and heterodyne detection allows for deterministic manipulation of quantum optical states that can be experimentally realised with typically very high efficiency. However, the absence of a strong non-linearity within this toolbox severely handicaps the reach of CV techniques for quantum information processing applications [22, 24]. Conversely, discrete variables (DV) is implicitly non-linear — forgoing determinism to harness the measurement-induced non-linearity of a photon-counting measurement. Recently, there has been a move to hybridise both CV and DV techniques for quantum information purposes, as one non-Gaussian operation, when combined with Gaussian resources and operations, is sufficient to realise universal quantum computing [75].

Here we present the CV analog of the photon counting measurement, whereby we replace a non-deterministic photon counting measurement with a deterministic phase randomised measurement of the field quadratures. This extends the ideas reported in [70] to

show how the requirement of a photon counting measurement can be replaced by CV measurements for the reconstruction of the statistics of non-Gaussian states. This approach forgoes the shot by shot nature of DV photon counting in favour of ensemble measurements, and consequently cannot be appropriated for state preparation. It does, however, still permit access to the same non-Gaussian statistics that previously only accessible with the requirement of a projective photon counting measurement. Using this method, we have successfully reconstructed the non-Gaussian 1, 2 and 3 photon subtracted squeezed vacuum (PSSV) states.

3.1.1 Schrödinger Kitten States

The work of this Chapter and the next focus on the continuous variable analog of a photon number discriminating measurement. We then consider the use of this technique for extracting the statistics of a non-Gaussian state from a system. The non-Gaussian states of our choosing are the photon subtracted squeezed vacuum states, which bear close relation to the Schrödinger cat states.

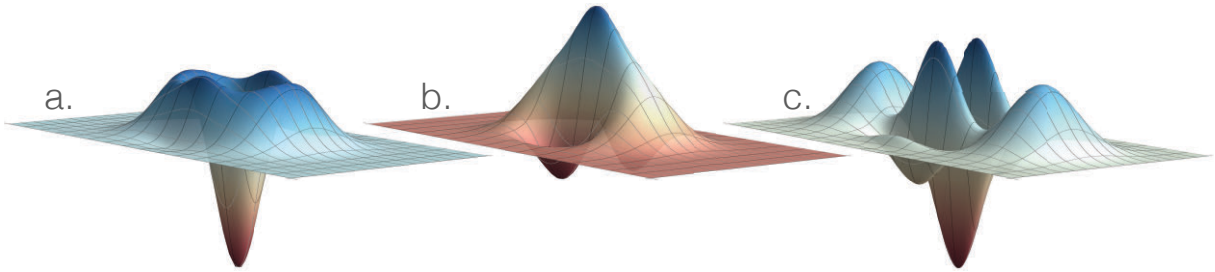

Figure 3.1: The Wigner functions of the ideal Schrödinger cat state of Equation (3.1) for: **a.** $\alpha = 1$ and $\phi = \pi$, **b.** $\alpha = 1$ and $\phi = 0$, and **c.** $\alpha = 2$ and $\phi = \pi$.

In itself, "Schödinger cat state" is not a precisely defined term and is broadly used within many quantum architectures to describe a quantum superposition of macroscopic states, alluding to the original paradox. Within quantum optics its usual incarnation is a coherent superposition of two coherent states with wholly opposite phase defined by,

$$|\psi_{cat}\rangle \equiv \frac{|\alpha\rangle + e^{i\phi}|-\alpha\rangle}{\sqrt{2(1+\cos\phi\, e^{-|\sqrt{2}\alpha|^2})}}. \tag{3.1}$$

Two special cases of (3.1) occur for $\phi = 0$ and $\phi = \pi$, which we label the *even* and *odd* cat states respectively. This choice of terminology becomes apparent when one considers the expansion of (3.1) in the Fock basis

$$|\psi_{cat}\rangle = \frac{1}{N} e^{-|\alpha|^2} \sum_n (1 + e^{i\phi + in\pi}) \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{3.2}$$

The term $(1 + e^{i\phi+in\pi})$ ensures the even cat state ($\phi = 0$) only occupies the even Fock states, and the odd cat state only occupies the odd Fock states. Within the literature 'cat state' is somewhat of a 'pet' name, often used interchangeably with the more precise

'coherent state superposition'. The premise for the 'Schrödinger cat' label emerges from the view of coherent states as macroscopic objects. Coherent states define a boundary of 'classicality' in quantum optics; they recover classical behaviour for large number of photons, and two coherent states well separated in phase space are distinguishable by a macroscopic measurement, *i.e.* they can be efficiently discriminated via homodyne detection without requiring photon number resolution. As such, a large optical Schrödinger cat would provide the archetypal system for investigating phenomena at the boundary of quantum and classical regimes. However, providing the ideal testbed for studying the decoherence necessitates such states be vulnerable to the environment. As such, all experimental investigations have either considered systems exceptionally isolated from the environment[76], or small states that are often referred to as "kitten states"[77, 78, 79, 80].[1]

Yurke and Stoler [81] showed that a coherent state interacting with a Kerr-like Hamiltonian could evolve to a superposition state, but this requires Kerr non-linearities orders of magnitudes larger than what is currently accessible. Macroscopic Schrödinger cats have been realised in cavity quantum electrodynamics (QED) systems. These experiments interacted a coherent state with Rydberg atom in a remarkably high Q-factor cavity, creating macroscopic superpositions strongly isolated from the environment [76]. However, the majority of the optical implementations of the Schrödinger Cat states have focused on the a protocol introduce by Dakna *et al.* [82].

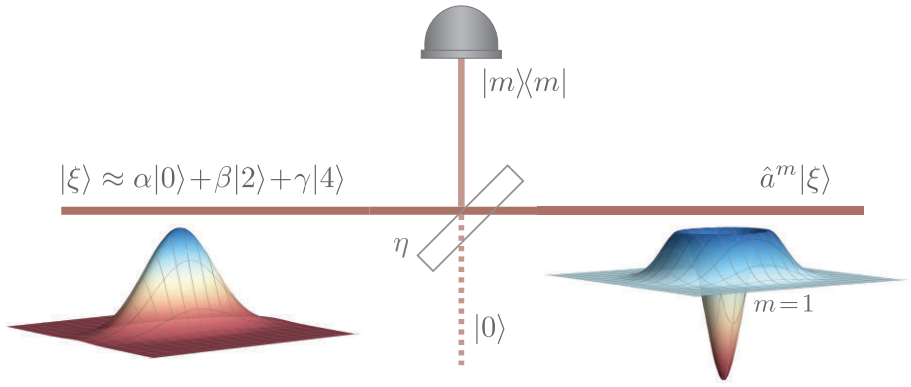

Figure 3.2: A squeezed vacuum state is incident on a weakly reflective beamsplitter. A measurement of $|m\rangle\langle m|$ heralds the preparation of an m-photon subtracted squeezed vacuum state.

3.1.2 Photon-subtracted squeezed vacuum states

In 1997, Dakna *et al.* published a rather straightforward theoretical proposal to create small amplitude Schrödinger cat states [82]. The annihilation of a single photon from a squeezed vacuum state with appropriate variance produces a quantum state with potentially very high fidelity to a small Schrödinger cat state.[2] Larger Schrödinger cat states are accessible via successive applications of the annihilation operation.

[1] Convention within the literature would suggest 'kitten' denotes a cat state with an amplitude $\alpha \approx 1$ or smaller.

[2] Equivalently, one can also squeeze a Fock state, though this is experimentally more ambitious.

The photon-subtraction protocol of Danka *et al.* is outlined in Figure 3.2. Consider a squeezed vacuum state split on a beamsplitter of low reflectivity, $1-\eta \ll 1$. A vacuum state occupies the other input port. Measurement of m photons in the reflected mode projects the transmitted state into a squeezed vacuum state with m-photon subtracted. There exist numerous theoretical works considering single mode [82, 83] and multi-mode [84, 85] descriptions of the original proposal of Danka *et al.*. Here, I shall consider a simple theoretical model for the photon-subtracted squeezed vacuum state.

Consider our initial squeezed vacuum state incident on a beamsplitter

$$|\psi_{\text{in}}\rangle = \hat{B}(\eta)\hat{\xi}(r)|0,0\rangle \tag{3.3}$$

where,

$$\hat{\xi}(r)|0\rangle = |\xi\rangle = \frac{1}{\sqrt{\cosh s}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n!)}}{n!} \left(-\tfrac{1}{2}\tanh s\right)^n |2n\rangle \tag{3.4}$$

The action of the beamsplitter operator is given by [28]

$$\hat{B}(\eta)|n,0\rangle = \sum_{m=0}^{n} \sqrt{B_m^n(\eta)}|m,n-m\rangle, \tag{3.5}$$

where

$$B_m^n(\eta) = \sqrt{\frac{n!}{k!(n-k)!}}\,\eta^{n-m}(1-\eta)^m, \tag{3.6}$$

and η is the transmissivity of the beamsplitter. The m-photon subtracted squeezed vacuum state is given by

$$|\psi_{\text{out}}\rangle = \langle m|\psi_{\text{in}}\rangle = \langle m|\hat{B}(\eta)\hat{\xi}(s)|0\rangle|0\rangle. \tag{3.7}$$

Introducing a closure relation on the squeezed vacuum mode and using (3.5)

$$\begin{aligned} |\psi_{\text{out}}\rangle &= \sum_{n=0}^{\infty} \langle n|\hat{\xi}(s)|0\rangle\langle m|\hat{B}(\eta)|n,0\rangle \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \langle n|\hat{\xi}(s)|0\rangle B_k^n(\eta)\delta_{n-k}^m|k\rangle. \end{aligned} \tag{3.8}$$

After making the change of indices $n \to m+k$, (3.8) simplifies to

$$|\psi_{\text{out}}\rangle = \sum_{k=0}^{\infty} \langle m+k|\hat{\xi}(s)|0\rangle B_k^{m+k}(\eta)|k\rangle. \tag{3.9}$$

Using (3.4) and (3.6) we obtain a final expression,

$$|\psi_{\text{out}}\rangle = \frac{1}{\cosh s}\sum_{k=0}^{\infty}\Phi(m+k)\frac{(k+m)!}{(\frac{1}{2}(k+m))!}\sqrt{\frac{(\tanh s)^{k+m}}{2^{k+m}m!k!}}\eta^k(1-\eta)^m|k\rangle, \tag{3.10}$$

where $\Phi(l) = (1+e^{il\pi})/2$ enforces the condition imposed by (3.4) that $(m+n)$ be an even integer. As the (ideal) squeezed vacuum state is a superposition state of the even Fock

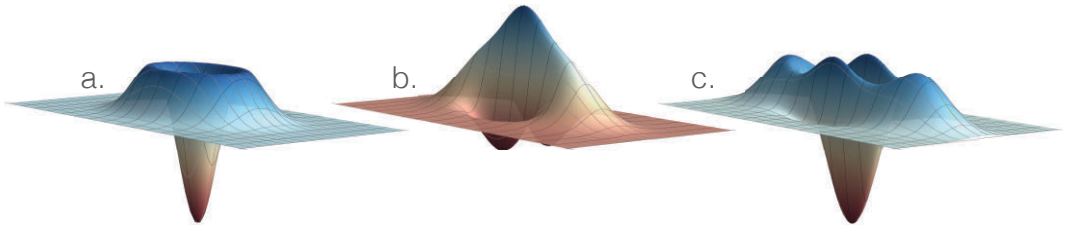

Figure 3.3: The Wigner functions of the: (a) one photon subtracted squeezed vacuum states ($r = 0.3$), (b) two photon subtracted squeezed vacuum state ($r = 0.5$) , and (c) three photon subtracted squeezed vacuum state ($r = 0.7$).

states, the number of photons subtracted determines whether the reconstructed state resembles the *odd* or *even* Schrödinger kitten state. Experimentally, the subtraction is usually implemented by tapping off 5-10% of a squeezed vacuum state and measuring via a regular (not photon number resolving) avalanche photo-diode[77, 78, 79]. Photon-number-resolving capabilities of transition edge sensor photo-detectors have also been used to reconstruct up the the 3 photon-subtracted squeezed vacuum state. [80].

3.1.3 Hybrid experiments

Our discussion of photon subtracted squeezed vacuum states and Schrödinger kitten states, whilst important for understanding the results of this and the following Chapter, is somewhat of an aside from the theory presented in this Chapter. The 'photon-subtracted squeezed vacuum state', however, is the prototypical example of a hybrid experiment: where both the discrete and continuous degrees of freedom are simultaneously exploited. These experiments combine two historically distinct camps of quantum optics, requiring the marriage of experimental techniques that address different regimes: with discrete-variables implementations broadly focusing on the time domain, while continuous-variables implementations are typically framed within the frequency domain. While combining these two domains is experimentally very challenging, it allows one to combine the weak, deterministic non-linearities readily accessible in continuous variable architectures, with the non-deterministic, strong measurement-based non-linearities accessible in the discrete variables. This has provided new research direction in quantum optics, with excellent prospects for quantum information and communication[86, 87, 88, 89, 90, 91, 92].

3.2 Theory

In the previous section I introduced the photon-subtraction protocol of Danka *et al.* [82] and outlined how a photon counting measurement is used to herald the preparation of a non-Gaussian state of light. Instead of heralding the preparation of a photon subtracted state with a 'photon counting' measurement, we construct a continuous variable analog of such a heralding measurement.

If ρ_{ab} originates from a squeezed vacuum mode passing through a weakly reflecting beam splitter, the resulting mode at b conditioned on finding n photons at a will be an n-PSSV state (see Figure 3.2). Of course, to prepare such a state, a projective measurement on the Fock basis is required. Instead, we demonstrate how we can imitate this conditioning with quadrature measurements to reconstruct the statistics of the n-PSSV states. In doing so we forgo any designs we have on preparing states, and instead take an ensemble approach, hoping to recover the same statistics from the final measurement ensemble. We will consider two scenarios: a phase randomised homodyne detection on mode a, and a *heterodyne* detection on mode a.

For the first scenario the conditioning measurement consists of sampling the homodyne observable $\hat{X}_a^\phi$ in a phase randomised manner such that each quadrature angle, ϕ contributes equally (Figure 3.4.b.). Here, $\hat{X}_a^\phi = e^{-i\phi}\hat{a}_a + e^{i\phi}\hat{a}_a^\dagger$, where $\hat{a}_a$ and $\hat{a}_a^\dagger$ are the annihilation and creation operators in mode a, respectively, and ϕ is the field quadrature angle. The mode b is characterised via homodyne tomography, with a measurement of the field quadrature $\hat{X}_b^\theta$. We will later consider a second scenario, where we exchange our phase-randomised homodyne detection of mode a for a heterodyne detection of mode a (Figure 3.4.c.).

3.2.1 Transformation Polynomials

In the first example, we will attempt to condition the output state b on the measurement outcome of operator $\hat{n}_a$. To do so, we want to estimate the expectation value

$$\langle X_b^\theta \rangle_n = \mathrm{Tr}\left[\rho_{ab}\, \hat{n}_a \otimes \left|X_b^\theta\right\rangle\!\left\langle X_b^\theta\right|\right] \ , \tag{3.11}$$

where ρ_{ab} is the joint state at modes a and b. Expanding the operator $\hat{n}_a$, (3.11) can be rewritten as,

$$\langle X_b^\theta \rangle_n = \sum_n n\, \mathrm{pr}(n_a) \mathrm{Tr}\left[\rho_{b|n_a} \left|X_b^\theta\right\rangle\!\left\langle X_b^\theta\right|\right] \tag{3.12}$$

$$= \sum_n n_a\, \mathrm{pr}(n_a)\, \mathrm{pr}(X_{b|n_a}^\theta) \tag{3.13}$$

$$= \sum_n n_a\, \mathrm{pr}(X_{b|n_a}^\theta) \ . \tag{3.14}$$

where $\mathrm{pr}(n)$ denotes the probability of measuring the eigenvalue n at a, and $\rho_{b|n}$ is the state at b conditional on measuring an outcome n at a.

To illuminate this problem consider ρ_{ab} is a weakly squeezed vacuum state passing through a low reflectivity beam-splitter (with vacuum occupying the other input port).

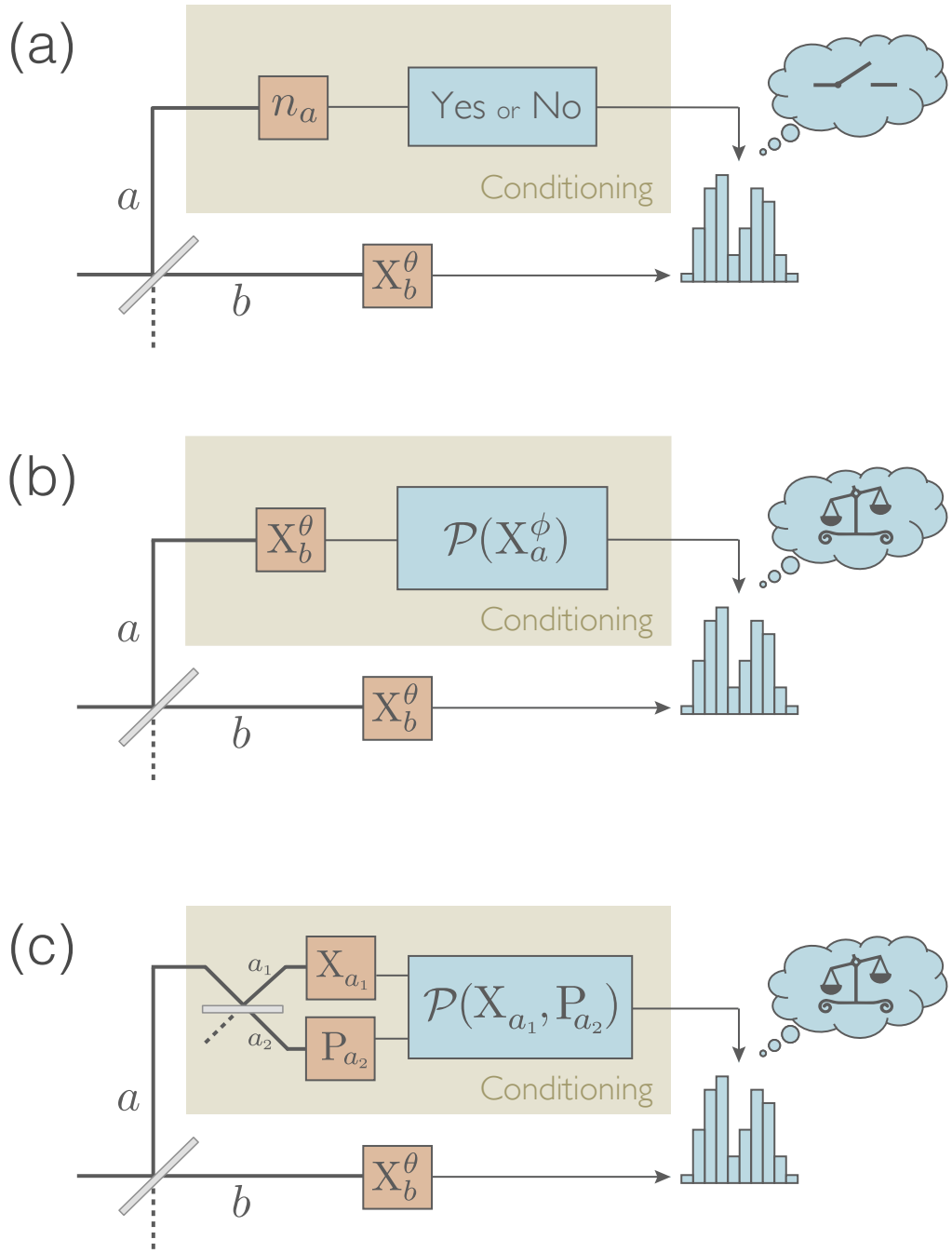

Figure 3.4: Reconstruction of the output state at b on $\hat{n}_a$. (a) A photon number discriminating detector is used to achieve the conditioning. The outcome of the photon counting measurement on a heralds the correct preparation of the state. The statistics of the non-Gaussian state are then reconstructed by keeping only samples X_b^θ that correspond to successful preparations. (b) The same statistics can be obtained by replacing the photon number discriminating detector with a phase-randomised homodyne detection, or equivalently, (c) heterodyne detection at mode a. Each sample at X_b^θ is then weighted by a continuous function of the outcome X_a^ϕ.

Ignoring higher-order terms, a weakly squeezed vacuum state can be approximated by $|\psi\rangle = |0\rangle - \gamma|2\rangle$ where $\gamma \ll 1$. The beam-splitter transforms this state to

$$|0,0\rangle + \gamma\left(|1,1\rangle\sqrt{2\eta^2(1-\eta^2)} + |2,0\rangle(1-\eta^2) + |0,2\rangle\eta^2\right), \tag{3.15}$$

where the beamsplitter transmissivity is $\eta \sim 1$, and $|n,m\rangle$ describes the state with n photons in mode a and m photons in mode b.

For this state, the expectation value of (3.11) becomes

$$\mathrm{Tr}\left[\left(|1\rangle\langle 1|\, 2\eta^2(1-\eta^2) + |0\rangle\langle 0|\, 2(1-\eta^2)^2\right)\left|X_b^\theta\right\rangle\left\langle X_b^\theta\right|\right]\gamma^2 \,. \tag{3.16}$$

where the second term arises from the probability of reflecting two photons. For $\eta \sim 1$ we can assume this probability to be very small and the output expectation value gives the statistics corresponding to a single photon Fock state.

To realise this conditioning, we would typically measure mode a in the Fock basis, $|n\rangle\langle n|$, with the measurement outcomes n_a informing our decision of whether to keep or reject the corresponding outcomes of b. But consider instead, a futile attempt to measure $\hat{n}_a$ by probing the quadratures. Expressing $\hat{n}$ in terms of the quadrature operators $\hat{X}$ and $\hat{P}$ we obtain

$$\hat{n} = \hat{a}^\dagger\hat{a} = \frac{1}{4}\left(\hat{X}^2 + \hat{P}^2 - 2\right) \tag{3.17}$$

where $\hat{X}$ and $\hat{P}$ share the commutation relation $[\hat{X},\hat{P}] = 2i$. Although this ensures the perfect simultaneous measurement of conjugate observables $\hat{X}$ and $\hat{P}$ is forbidden, (3.11) can nevertheless can be written as the sum

$$\langle X_b^\theta\rangle_n = \mathrm{Tr}\left[\rho_{ab}\,(\frac{1}{4}\hat{X}^2 - 1)\otimes\left|X_b^\theta\right\rangle\left\langle X_b^\theta\right|\right] \tag{3.18}$$

$$+\,\mathrm{Tr}\left[\rho_{ab}\,(\frac{1}{4}\hat{P}^2 - 1)\otimes\left|X_b^\theta\right\rangle\left\langle X_b^\theta\right|\right] \,. \tag{3.19}$$

The expectation value $\langle X_b^\theta\rangle_n$ is accessible by combining the outcomes of the two independent measurements.

3.2.2 Phase-randomised homodyne detection

The central relationship (3.20) between the familiar field quadrature operators holds for any pair of orthogonal quadrature operators,

$$\hat{n} = \hat{a}^\dagger\hat{a} = \frac{1}{4}\left((\hat{X}^\phi)^2 + (\hat{X}^{\phi+\pi/2})^2 - 2\right), \tag{3.20}$$

where $\hat{X}^\phi = e^{-i\phi}\hat{a} + e^{i\phi}\hat{a}^\dagger$. As a result, the global phase θ is immaterial and does not need to be controlled. Noting this, (3.20) can be written as an integration over the global

phase

$$\hat{n} = \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{4}\left[(\hat{X}^{\phi})^2 + (\hat{X}^{\phi+\frac{\pi}{2}})^2 - 2\right] d\phi \tag{3.21}$$

$$= \frac{1}{2}\left(\bar{X}^2 - 1\right), \tag{3.22}$$

where,

$$\bar{X}^n = \frac{1}{2\pi}\int_0^{2\pi} (\hat{X}^{\phi})^n d\phi \,. \tag{3.23}$$

is the phase-averaged quadrature moment operator.

Substituting this into (3.11) we obtain

$$\langle X_b^{\theta}\rangle_n = \mathrm{Tr}\left[\hat{\rho}_{ab}\,\frac{1}{2}\left(\bar{X}_a^2 - 1\right) \otimes \left|X_b^{\theta}\right\rangle\!\left\langle X_b^{\theta}\right|\right]. \tag{3.24}$$

The expectation value $\langle X_b^{\theta}\rangle_n$ can be obtained by a phase randomised sampling of the quadratures and weighting the outcomes at b by the result $\frac{1}{2}(\bar{X}_a^2 - 1)$ at a given an outcome X_a^{θ}.

Example: Conditioning on $\hat{n}_a(\hat{n}_a - 2)$

To obtain a more faithful reproduction of the single photon Fock state from the weakly squeezed state described above, consider that we instead condition on a different polynomial, $\hat{n}_a(\hat{n}_a - 2)$. This further isolates the $n_a = 1$ eigenvalues removing the contribution of two photon outcomes at mode a, in addition to the $n_a = 0$ vacuum contribution. For a weakly squeezed vacuum input state (neglecting four photon terms), the analogue of (3.16) for this new conditioning is

$$\langle X_b^{\theta}\rangle_n = \mathrm{Tr}\left[|1\rangle\langle 1|\, 2\eta^2(1-\eta^2)\left|X_b^{\theta}\right\rangle\!\left\langle X_b^{\theta}\right|\right]\gamma^2 \,. \tag{3.25}$$

To realise this conditioning via our homodyne measurements, we repeat the recipe as before and cast $\hat{n}$ in terms quadrature operators $\hat{X}$ and $\hat{P}$.

$$\begin{aligned}\hat{n}(\hat{n}-2) &= \frac{1}{16}\left(\hat{X}^2 + \hat{P}^2 - 2\right)\left(\hat{X}^2 + \hat{P}^2 - 4\right) \\ &= \frac{1}{16}\left(2\bar{X}^4 - 12\bar{X}^2 + 8 + \hat{X}^2\hat{P}^2 + \hat{P}^2\hat{X}^2\right) \,.\end{aligned} \tag{3.26}$$

The terms involving products of $\hat{X}$ and $\hat{P}$ cannot be evaluated directly through a phase randomised homodyne measurement. In order to make them accessible, we reexpress

$\hat{X}^2\hat{P}^2 + \hat{P}^2\hat{X}^2$ as a function of $\bar{X}$, which can be done as follows:

$$\hat{X}^2\hat{P}^2 + \hat{P}^2\hat{X}^2 = \frac{1}{2\pi}\int_0^{2\pi} 2(\hat{X}^{\phi})^2(\hat{X}^{\phi+\frac{\pi}{2}})^2 d\phi \tag{3.27}$$

$$= \frac{1}{\pi}\int_0^{2\pi} (2\hat{a}_\phi^\dagger\hat{a}_\phi\hat{a}_\phi^\dagger\hat{a}_\phi + 2\hat{a}_\phi^\dagger\hat{a}_\phi - 1 - \hat{a}_\phi^4 - (\hat{a}_\phi^\dagger)^4) d\phi \tag{3.28}$$

$$= \frac{1}{\pi}\int_0^{2\pi} (2\hat{a}_\phi^\dagger\hat{a}_\phi\hat{a}_\phi^\dagger\hat{a}_\phi + 2\hat{a}_\phi^\dagger\hat{a}_\phi - 1) d\phi \tag{3.29}$$

$$= \frac{1}{\pi}\int_0^{2\pi} \left(\frac{(\hat{a}_\phi^\dagger + \hat{a}_\phi)^4}{3} - 2\right) d\phi \tag{3.30}$$

$$= \frac{2\bar{X}^4}{3} - 4\,, \tag{3.31}$$

where we define $\hat{a}_\phi = \hat{a}\exp(-i\phi)$. Substituting this into Equation (3.26), we obtain the sampling polynomial,

$$\hat{n}(\hat{n}-2) = \frac{\bar{X}^4}{6} - \frac{3\bar{X}^2}{2} + 1\,. \tag{3.32}$$

With this, the expectation value becomes

$$\langle X_b^\theta\rangle_{n(n-2)} = \mathrm{Tr}\left[\rho_{ab}\left(\frac{\bar{X}_a^4}{6} - \frac{3\bar{X}_a^2}{2} + 1\right) \otimes \left|X_b^\theta\right\rangle\!\left\langle X_b^\theta\right|\right] \tag{3.33}$$

which can be sampled via a randomised phase quadrature measurement.

3.2.3 Arbitrary conditioning in $\hat{n}_a$

The first sampling polynomial we considered, $\mathcal{Q} = \hat{n}_a$, conditioned the reconstructed state at b on the statistical mixture

$$\hat{\rho}_a = |1\rangle\langle 1| + 2\,|2\rangle\langle 2| + 3\,|3\rangle\langle 3| + \ldots\,, \tag{3.34}$$

a state diagonal in the Fock basis, where the relevance of each term is accordingly weighted by their individual eigenvalues. The example of (3.26) then considers a correction to erroneous 2 photon 'events', removing the $|2\rangle\langle 2|$ term and essentially 'purifying' the conditioning. To affect a perfect conditioning at a we would ideally like to isolate the $|1\rangle\langle 1|$ term, which would require an infinite order polynomial. However, we can construct a higher-order polynomial of $\hat{n}$ to finite degree in a similar way to that presented in §3.2.2. We provide two algorithms in Appendix A. These polynomials provide a simple construction for a k photon subtracted state by conditioning on

$$\mathcal{P}(\hat{n}) = \frac{1}{\hat{n}-k}\prod_{j=0}^{j_{max}} \hat{n} - j\,, \tag{3.35}$$

with $j_{max} > k$. The choice of j_{max} allows you to specify the maximum photon correction for higher photon number contributions up to j_{max}. However, this will be at the expense of

introducing larger weighting on the residual components with photon numbers greater than j_{max}. If the probabilities associated with these residual photon numbers are not vanishing sufficiently rapidly to negate any growing weighting, one cannot find a polynomial to provide an arbitrary purity.

As an example, to obtain a 2-PSSV state, we could implement the conditioning polynomial with $k = 2$ and $j_{max} = 6$:

$$\mathcal{Q}(\hat{n}) = \hat{n}(\hat{n}-1)(\hat{n}-3)(\hat{n}-4)(\hat{n}-5)\,. \tag{3.36}$$

Expanding in the Fock basis,

$$\mathcal{Q}(\hat{n}) = -12\,|2\rangle\langle 2| + 180\,|7\rangle\langle 7| + 1008\,|8\rangle\langle 8| + \ldots \tag{3.37}$$

In this example, we see that the seven photon and eight photons events are weighted by a factor of 15 and 84 respectively compared to the two photon events. For squeezed states however, these high photon number states would have exponentially vanishing probabilities in almost all applications.

3.2.4 Pattern functions

The resolution to the purification issue described above emerges from the pattern functions (described in detail in §2.5.2). The pattern functions [48, 49, 50] specify the link between the homodyne observables and the density matrix. In optical homodyne tomography, they enable the direct sampling of the density matrix, bypassing the need to reconstruct the Wigner function.

We want to characterise the state at a conditioned on an n photon event at b. Ideally, we would choose an appropriate polynomial in X_a^{ϕ} that corresponds to $|n\rangle\langle n|$. Practically, however, we can only realise a polynomial of a limited order—correcting for the finite undesired photon number events that may prove statistically significant. The pattern functions however permit a perfectly isolating characterisation that removes all unwanted photon number events.

We start with the general problem of reconstructing the statistics of the post-selected state at b, $\hat{\rho}_b$, conditioned on the event of having a state $\tilde{\rho}_a$ at a. This conditioning can be achieved by means of a measurement apparatus at a having two outcomes:

$$\pi_1 = \tilde{\rho}_a \tag{3.38}$$

$$\pi_2 = 1 - \tilde{\rho}_a. \tag{3.39}$$

The output at b conditioned on the outcome π_1 would be

$$\hat{\rho}_b = \frac{1}{\mathrm{pr}_1}\mathrm{Tr}_a[\rho_{ab}\,\pi_1]\,. \tag{3.40}$$

where

$$\mathrm{pr}_1 = \mathrm{Tr}\,[\rho_{ab}\,\pi_1] \tag{3.41}$$

is the probability of obtaining outcome π_1. We decompose the conditioned state $\tilde{\rho}_a$ in the

Fock basis with coefficients c_{mn}

$$\rho_a^{cond} = \sum_{mn} c_{mn} |n_a\rangle\langle m_a| \tag{3.42}$$

so that the post-selected state at b can be written as the sum

$$\tilde{\rho}_b = \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \mathrm{Tr}_a[\rho_{ab} \ |n_a\rangle\langle m_a|] \ . \tag{3.43}$$

To be able to reconstruct the post-selected state, we perform a quadrature tomography by measuring X_b^θ at b. The probability of getting an outcome X_b^θ on the post-selected state is

$$\tilde{\mathrm{pr}}\left(X_b^\theta\right) = \langle X_b^\theta|\tilde{\rho}_b|X_b^\theta\rangle \tag{3.44}$$

$$= \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \langle m_a, X_b^\theta|\rho_{ab}|n_a, X_b^\theta\rangle \tag{3.45}$$

$$= \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \langle m_a|\mathrm{Tr}_b\left[\rho_{ab} \ \left|X_b^\theta\right\rangle\left\langle X_b^\theta\right|\right] |n_a\rangle \tag{3.46}$$

$$= \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \langle m_a|\rho_a\left(X_b^\theta\right) |n_a\rangle \mathrm{pr}\left(X_b^\theta\right) \tag{3.47}$$

where $\rho_a\left(X_b^\theta\right)$ is the state at a when we obtain outcome X_b^θ at b. The probability of getting this outcome is denoted as $\mathrm{pr}\left(X_b^\theta\right)$. We want to write the matrix elements $\langle m_a|\rho_a\left(X_b^\theta\right)|n_a\rangle$ in terms of quadrature value measurements. For this we utilise the Fock basis pattern function [50] to write

$$\langle m_a|\rho_a\left(X_b^\theta\right)|n_a\rangle = \int_0^\pi \int_{-\infty}^{+\infty} \mathrm{pr}\left(X_a^\phi|X_b^\theta\right) F_{mn}(X_a^\phi) dX_a d\phi \,, \tag{3.48}$$

where the F_{mn} are the pattern functions of the Fock basis. They are given by

$$F_{mn}(X_a^\phi) = \frac{1}{\pi} \exp(i(m-n)\phi) \frac{\partial}{\partial x} \left[\psi_m(X_a)\varphi_n(X_a)\right] \tag{3.49}$$

where $\psi_m(X_a)$ and $\varphi_m(X_a)$ are the m-th regular and irregular eigenfunctions, respectively, of the Schrödinger equation in a harmonic potential [3]. Substituting this into (3.47), we find

$$\tilde{\mathrm{pr}}\left(X_b^\theta\right) = \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \int_0^\pi \int_{-\infty}^{+\infty} \mathrm{pr}\left(X_b^\theta\right) \mathrm{pr}\left(X_a^\phi|X_b^\theta\right) F_{mn}(X_a^\phi) dX_a d\phi \tag{3.50}$$

$$= \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} \int_0^\pi \int_{-\infty}^{+\infty} \mathrm{pr}\left(X_a^\phi, X_b^\theta\right) F_{mn}(X_a^\phi) dX_a d\phi \tag{3.51}$$

[3] Any second order differential equation such as the Schrödinger equation must have linearly independent eigenfunctions. The first, normalisable, solution is the standard wave function whilst the second is called the irregular eigenfunction. It is not normalisable and must be discarded as a physical state, but nonetheless can be a valuable tool in calculations[93].

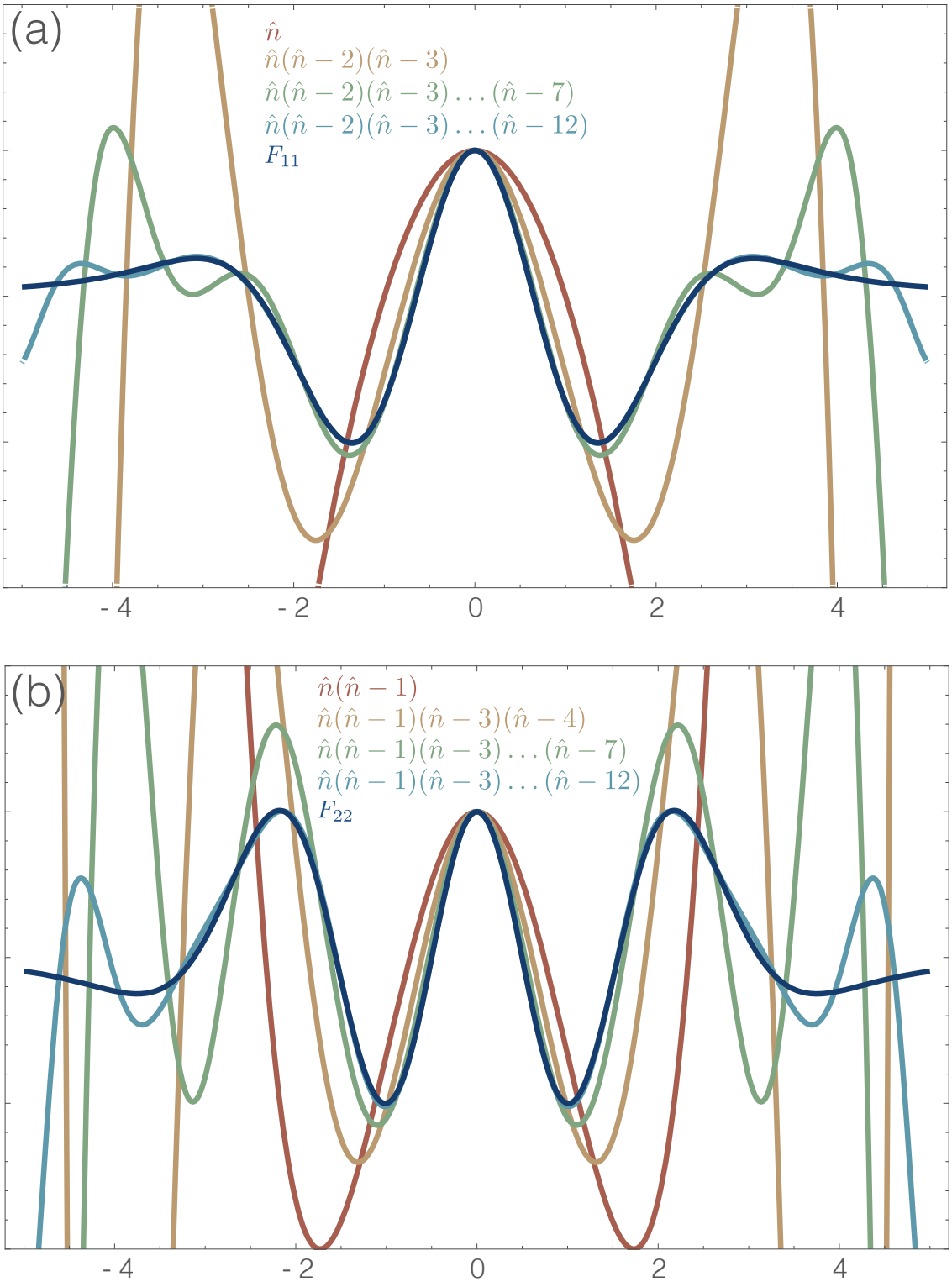

Figure 3.5: The convergence of the $\hat{n}$ polynomials to their corresponding pattern functions for a photon number measurement of (a) $n = 1$ and (b) $n = 2$.

where $\mathrm{pr}(X_a^\phi, X_b^\theta)$ is the unconditioned probability of obtaining outcomes X_a^ϕ and X_b^θ when we measure a and b in quadrature at angles ϕ and θ. Introducing the weighting function

$$w\left(X_a^\phi\right) = \frac{1}{\mathrm{pr}_1} \sum_{mn} c_{mn} F_{mn}\left(X_a^\phi\right) \ , \tag{3.52}$$

we can write

$$\tilde{\mathrm{pr}}\left(X_b^\theta\right) = \int_0^\pi \int_{-\infty}^{+\infty} \mathrm{pr}\left(X_a^\phi, X_b^\theta\right) w\left(X_a^\phi\right) dX_a d\phi \ . \tag{3.53}$$

From this expression, we see that the conditioned distribution $\tilde{\mathrm{pr}}\left(X_b^\theta\right)$ can be obtained by sampling the distribution $\mathrm{pr}\left(X_a^\phi, X_b^\theta\right)$ and weighting the outcomes by $w\left(X_a^\phi\right)$.

As an example, to obtain $\tilde{\rho}_b$ conditioned on a one photon event at a, we condition on $\tilde{\rho}_a = |1\rangle\langle 1|$. This sets $c_{11} = 1$ and all other $c_{mn} = 0$. To condition on the superposition state $\tilde{\rho}_a = \frac{1}{2}(|1\rangle + |2\rangle)(\langle 1| + \langle 2|)$, we require $c_{00} = c_{01} = c_{10} = c_{11} = \frac{1}{2}$ and all other $c_{nm} = 0$. For a conditioned state $\tilde{\rho}_a$ that is diagonal in the Fock basis, the weighting function $w(X_a^\phi)$ is a sum of $F_{mn}(X_a^\phi)$ with $m = n$ which does not depend on the angle ϕ. Hence the probability

$$\tilde{\mathrm{pr}}\left(X_b^\theta\right) = \int_0^\pi \int_{-\infty}^{+\infty} \mathrm{pr}\left(X_a^\phi, X_b^\theta\right) w\left(X_a\right) dX_a d\phi \tag{3.54}$$

can be obtained by doing a phase randomised sample of the quadratures of a. To extract the statistics of an off-diagonal element, one needs to keep track of the phase, ϕ. As one would anticipate, the polynomials we set about constructing earlier appear to converge to the pattern functions the order grows (Figure 3.5).

3.2.5 Heterodyne Detection

Exactly the same conditioning that we have demonstrated via a scanned homodyne detection is accessible via a heterodyne, or dual-homodyne detection. From an information perspective, this is somewhat unsurprising, as both are *informationally complete* positive-operator valued measures (POVMs) [94] [4]. The heterodyne approach, however, is not nearly as mathematically elegant. The crucial reason for this being the vacuum penalty paid for the attempted simultaneous measurement of $\hat{X}$ and $\hat{P}$. Consider again, our expression for the photon number operator in terms of the quadratures,

$$\hat{n}_a = \frac{1}{4}(\hat{X}_a^2 + \hat{P}_a^2 - 2). \tag{3.55}$$

Our attempt to measure $\hat{X}$ and $\hat{P}$ results in the actual observables,

$$\hat{X}_{a_1} = \frac{1}{\sqrt{2}}(\hat{X}_a + \hat{X}_v) \tag{3.56}$$

$$\hat{P}_{a_2} = \frac{1}{\sqrt{2}}(\hat{P}_a - \hat{P}_v), \tag{3.57}$$

[4] Provided one keeps a record of the homodyne phase for the scanned homodyne detection. This is not a problem here because our 'projection' on mode a is only concerned with mean amplitudes.

where $\hat{X}_v$ and $\hat{P}_v$ are the vacuum fluctuations. Whilst this noise penalty precludes any shot-by-shot estimation of n_a, because we are focused on average outcomes any contamination from the vacuum can be perfectly accounted for. As the vacuum is necessarily uncorrelated to our measurements and its variance is perfectly specified by quantum mechanics, our simplest conditioning transforms to

$$\hat{n}_a = \frac{1}{2}(\hat{X}_{a_1}^2 + \hat{P}_{a_2}^2 - 2). \tag{3.58}$$

To obtain more sophisticate polynomials in $\hat{n}_a$, we essentially follow the same procedure as in §3.2.3 and Appendix A. We first note that the polynomials of degree d in $\hat{n}_a$ are of the form,

$$\mathcal{Q}(\hat{n}_a) = R(\hat{X}_a^2 + \hat{P}_a^2). \tag{3.59}$$

We thus anticipate the polynomials of our measured heterodyne quadratures will take the general form,

$$\mathcal{Q}(\hat{n}_a) = P(\hat{X}_{a_1}^2 + \hat{P}_{a_2}^2). \tag{3.60}$$

For a polynomial $\mathcal{Q}(\hat{n}_a) = \mathcal{R}(\hat{X}_a, \hat{P}_a)$ of dimension d we 'guess' the general form,

$$\mathcal{P}(\hat{X}_{a_1}, \hat{P}_{a_2}) = \sum_{k=0}^{d} \alpha_k (\hat{X}_{a_1}^2 + \hat{P}_{a_2}^2)^k \tag{3.61}$$

$$= \sum_{k=0}^{d} \alpha_k \left(\frac{(\hat{X}_a + \hat{X}_v)^2}{2} + \frac{(\hat{P}_a + \hat{P}_v)^2}{2} \right)^k . \tag{3.62}$$

$$\tag{3.63}$$

Requiring that $\langle \mathcal{P}(\hat{X}_{a_1}, \hat{P}_{a_2}) = \langle \mathcal{R}(\hat{X}_a, \hat{P}_a) \rangle$ allows one to solve for the coefficients α_k.

Historically, the *heterodyne* experiment was the first version we experimentally implemented, emerging from the ideas of Ralph *et al.* [70]. That first theoretical work demonstrated a method to reconstruct the Wigner function of single photon Fock state by first constructing the conditioned moments of the homodyne observable $\langle n_a (X_b)^k \rangle$. While that result evolved into the work presented here, it also recalls some results from §3.2.1. For the estimation of the simplest polynomial, $\mathcal{P} = \hat{n}_a$, one only requires two independent measurements of $\hat{X}$ and $\hat{P}$, *i.e.* one could set their homodyne detector at a to first measure $\hat{X}$, and then $\hat{P}$. For the monomial of order 2 we encounter cross terms in $\langle \hat{X}\hat{P} \rangle$. Estimation of the cross term $\langle \hat{X}\hat{P} \rangle$ only requires two additional homodyne angles, $\hat{X}^{\pi/4}$ and $\hat{X}^{3\pi/4}$. As the order of the polynomial in $\hat{n}_a$ increases, the size of the cross terms in $\hat{X}$ and $\hat{P}$ increases and accordingly the reconstruction requires an increasing number of independent homodyne slices. This is precisely the limit in which in the series of independent homodyne measurements becomes informationally complete [94].

As another aside, one might anticipate because of the informational completeness of heterodyne detection, an analog of the sampling functions should exist for heterodyne detection. They were provided by Paris in 1996 [95], and allow for the same perfect

discrimination demonstrated in §3.2.4.

3.3 Discussion & Summary

We must emphasise this technique is not shot-by-shot and cannot be used in any meaningful way to prepare states. There is no way to talk sensibly about 'events', and as such, it is not post-selective. For it to work you need the *entire* measurement record at your disposal, with all measurements required to perfectly 'cancel' the Gaussianity and yield the correct statistics. This is quite different to the post-selective technique that will be discussed in Chapter 5, where a post-selection informed by a heterodyne measurement allows one to actually herald a more-entangled system.

This technique requires that the statistical relevance for *each* measurement you obtain at mode b to the final ensemble is decided by the outcome of some polynomial $\mathcal{P}(\hat{X}_a^{\phi})$. Analogies can be drawn with the aforementioned hybrid experiments, where the measurement-based non-linearity of a *real* photon-counting measurement allows one to de-Gaussify a state. In the analogous hybrid system, the outcome of the conditioning measurement at b heralds the correct preparation of the state - informing the experimenter whether the state at a should be kept or rejected. For the technique presented here, the experimenter has no such information, just a seemingly random string of continuous numbers that he uses to weight his measurement result at b. The non-Gaussian statistics emerge from this post-processing of the data. It could be argued it arises as the polynomials themselves are usually (but not always) greater than quadratic in the annihilation and creation operators - but this is not always the case. A more satisfying resolution is that the non-linearity emerges from the weighting procedure itself.

Chapter 4

A Continuous Variable Analog of a Photon Counting Measurement: Part II

This chapter will address the experimental implementation and results of the theory outlined in the previous chapter. It concerns the experimental details of two very similar experiments that were done approximately one year apart using (mostly) the same infrastructure. As such, the experimental details and results have been condensed into one chapter, noting the distinctions between both implementations as they arise.

The historically first experiment was the implementation of the 'heterodyne' conditioning described in §3.2.5. Here, the term 'heterodyne' is used interchangeably with 'dual-homodyne'. Though equivalent, heterodyne detection is the standard theoretical term for the simultaneous measurement of $\hat{X}$ and $\hat{P}$, while dual-homodyne detection best describes what we actually did.

The second experiment replaced the dual-homodyne conditioning measurement with a single, phase randomised homodyne detection. The experimental schematic for the single-homodyne implementation is provided in Figure 4.1. The distinction between the two conditioning measurements is provided in Figure 4.3.

4.1 Experimental Generation of Squeezed Light

4.1.1 Preparation of Seed and Pump Light

The light resource for this experiment was an *Innolight Diablo* ND:Yag laser producing continuous wave single mode light at 1064 nm. The laser also featured an internal frequency doubler that frequency doubled a portion of the available 1064 nm light to its second harmonic at 532 nm. The maximum available power was 400 mW at 1064 nm, and 800 mW at 532 nm. The laser FWHM linewidth was specified by the manufacturer to be ≈ 1 KHz. The laser also featured an internal 'noise eater' option which provided 30 dB of suppression of the 900 kHz natural relaxation oscillation of the laser. Even with the noise eater on, the roll off of the relaxation oscillation is still visible up to approximately 4 MHz.

The 1064 nm light were then passed through an Faraday isolator as a precaution against optical feedback via unintended backscatter. The 1064 nm and 532 nm light was

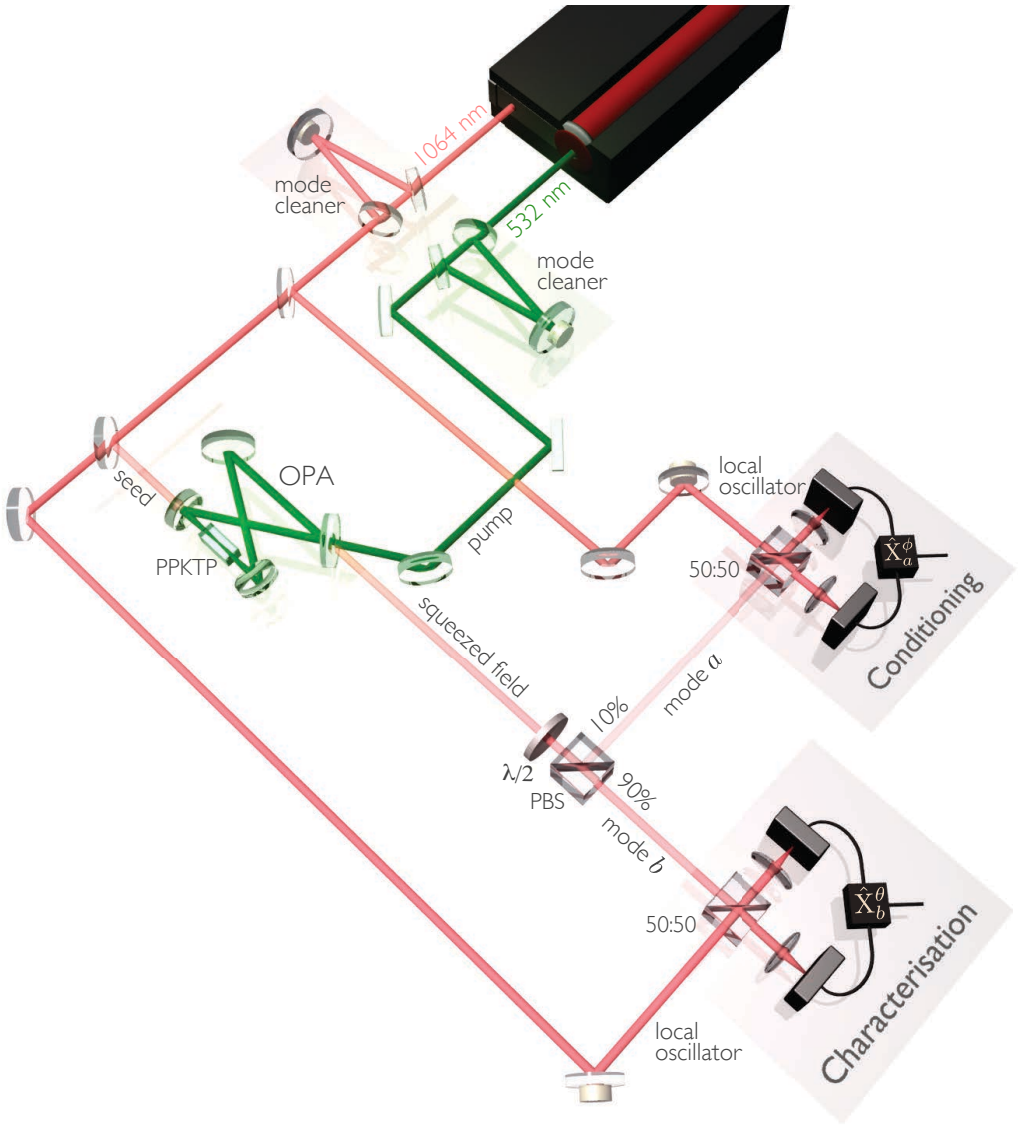

Figure 4.1: Experimental Setup A CW Nd:YAG laser at 1064 nm provides the light resource for this experiment. An internal second harmonic generation (SHG) cavity frequency doubles a portion of the 1064nm light. Both the 1064 nm and 532 nm fields undergo spatial and frequency filtering before providing seed and pump resources respectively for a doubly-resonant optical parametric amplifier (OPA). A small portion of the resulting squeezed coherent state is then reflected for 'conditioning' by a variable beam-splitter - implement with a $\lambda/2$ wave-plate and a polarising beam splitter (PBS). The reflected light (mode a) is subsequently sampled via a phase randomised homodyne detection. The remaining transmitted light (mode b) is characterised by a tomographic homodyne detection, sampling X_b^θ for $\theta = 0 \ldots 165°$ in intervals of 15°.

then passed through their respective high finesse optical resonators, or *mode cleaners*. The value of this step was twofold: first, it provided a well-defined single TEM-00 spatial mode for the entire experiment, and second, it provided additional filtering of the intensity and frequency noise of the laser above the cavity bandwidth. The two mode cleaners were both of the same design, consisting of a 3 mirror triangular ring resonator, with an optical path length of 800 mm. The 1064 and 532 nm mode-cleaners had respective cavity linewidths of 0.4MHz and 1.0MHz. This additional suppression of the remnant relaxation oscillation provides a shot noise limited laser field at frequencies above 4 MHz.

Both of the mode cleaners were controlled using the Pound-Drever-Hall (PDH)[96, 97] with a sideband frequency well outside the resonator linewidth. For this purpose we recycle the 40 MHz phase modulation sidebands used for control of the internal frequency doubling cavity within the laser unit. An error signal is obtained via analog demodulation before PID control is implemented using software written in *National Instruments LabView*. More details on this control are given in §4.2.3.

After filtering, the 1064 nm light was divided for three main purposes. A small portion of the 1064 nm light provided a seed field for the optical parametric amplifier (OPA) cavity. The rest was provided bright phase reference for measurement, and a field used for displacement of the prepared squeezed light. For the purposes of the control of both the OPA cavity length and relative phase between the seed and pump fields, a phase sideband was encoded on the seed field at 11.25 MHz.

4.1.2 Optical Parametric Amplifier

The optical parametric amplifier is the most complicated aspect of these two experiments. The same OPA cavity was the source of squeezing for both experiments, however a superior quality crystal was used in the single homodyne experiment. This was one of two major improvements made - the other being a dramatic improvement to photodiode efficiency.

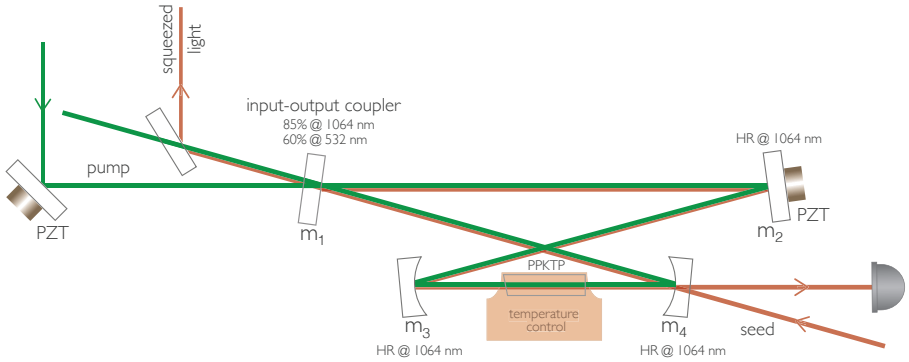

Figure 4.2: Detailed schematic of the optical parametric amplifier cavity.

Cavity specifications

The OPA cavity was originally designed by Nicolai Grosse and the full details of its original iteration are available in his thesis [98]. The cavity is a travelling-wave cavity comprised of four mirrors in a bow tie geometry. A detailed schematic of the OPA cavity is provided

in Figure 5.3. The cavity consisted of two inner concave mirrors (m_3 and m_4) with radii of curvature of 38 mm spaced 44 mm apart, and two flat mirrors spaced 90 mm apart. The total optical path length was 285 mm with an angle of incidence of 6°. The angle of incidence was made as small as possible to reduce astigmatism in the cavity mode. As the cavity was designed to be simultaneously resonant at both the fundamental and second harmonic, mirrors m_2, m_3 and m_4 were HR for both 1064 nm and 532 nm. The 10 mm long crystal was centred between the two convex mirrors, focusing the light to a waist of approximately 40 μm inside the crystal.

When optimising the squeezing performance of an OPA cavity a parameter of importance is the 'escape efficiency', η_{esc}. As the name suggests, the escape efficiency is a measure of the efficiency with which the squeezed light can exit the cavity mode, and is defined as $\eta_{\text{esc}} = T/(T + A)$, where T is the transmission through the input/output coupler, and A the total intra-cavity losses. There are two ways to improve the escape efficiency. The most desirable is to maximise η_{esc} by minimising any intra-cavity losses, A. This requires that we both reduce the number of surfaces the intra-cavity field interacts with, and also improve the quality of those surfaces. Alternatively, one can also reduce the reflectivity of the input-output coupler. For bow-tie geometries where the number of intra-cavity surfaces is inherently higher, the reflectivity of the input-output coupler provides flexibility. Increasing the transmissivity of the input-output coupler can have undesirable consequences as reducing the reflectivity will lower the the finesse of the cavity, and thus the threshold of the system. This was not a concern for our implementation however, as we had an abundance (700 mW) of available pump light (even less of a concern for a doubly-resonant system). The final choice of input-output coupler was 85% for 1064 and 70% for 532. For the 1064 nm cavity, the choice of output coupler gave a linewidth and finesse of 32 MHz and 35 respectively. The 532 nm cavity had a linewidth of 65 MHz and a finesse of 16.

The non-linear crystal

Both experiments used a periodically-poled Potassium Titanyl Phosphate (PPKTP) crystal with dimensions of 10×1 mm^3. Potassium Titanyl Phosphate is immensely popular in non-linear optics, owing to its excellent non-linearity and relatively high damage threshold (when compared to Lithium Niobate).

The KTP crystal is periodically poled to provide quasi-phase matching. Traditional phase matching ensures the phase relationship between two interacting fields is maintained across the non-linear medium. This technique usually exploits the birefringence of the non-linear medium to find a critical phase matching temperature where the refractive indices for both fields coincide. Quasi-phase matching, however, relaxes the requirement for a constant phase relationship across the interaction length, permitting phase mismatch over some propagation distance. The material is 'periodically poled', inverting the sign of the non-linearity where - due to acquired phase mismatch - the non-linear conversion would take place in the wrong direction. This provides a net accumulation of the non-linear interaction over the length of the crystal. This technique also provides some flexibility in choosing the temperature at which quasi-phase matching occurs, limited by the minimum domain size for poling. The quasi-phase matched temperature of the PPKTP crystal was

specified to be around 35°c.

Our choice of a doubly-resonant system also requires that we need to compensate for the effects intra-cavity dispersion. As all the intra-cavity components are dispersive, the fundamental and second harmonic fields will accumulate a different relative phase shift through a round trip. Consequently, controlling the cavity path length to be resonant at the seed frequency will not guarantee co-resonance at the pump frequency. To compensate for intra-cavity dispersion a 'wedged' crystal was used, with one surface cut at an angle of approximately 1°. As the crystal is quasi-phase matched, the fundamental and second harmonic fields experience different refractive indices. With the addition of a small wedged section at the end of the crystal, one can vary the intra-cavity path lengths of the fundamental and second harmonic fields by translating the crystal laterally to the beam propagation direction until their individual resonance conditions coincide. This a particularly elegant solution to the dispersion problem as it does not introduce additional intra-cavity surfaces.

The two experiments used two different crystals. Both crystals were dual band anti-reflection AR coated at 532 nm and 1064 nm, however the AR coatings were of different qualities. For the first (dual-homodyne) experiment the AR coatings were provided by the crystal manufacturer *Raicol* and specified to be AR coated to $< 0.1\%$. The AR coatings were seen as a potential bottleneck for the quality of the squeezing, given the observably high non-linearity of KTP. For the second experiment we threw money at the problem. The non-linear crystals were manufactured by *Raicol*, the incident surfaces were polished by *LaserOptik* and the dual-band AR coating was provided by *Advanced Thin Films*. The improved AR coating quality was specified to be 0.01% reflective.

OPA operation and control

Our doubly resonant system can be decomposed into its 1064 nm (seed) cavity, and a 532 nm (pump) cavity. The physical cavity length was defined by the resonance condition of the fundamental seed field, and then co-resonance with the pump field was established through a combination of adjustments of the crystal temperature and translations of the crystal. The relative phase between the seed field and the pump field defines the angle of the produced squeezed light. We work in the de-amplification regime, where the relative phase between the pump and seed is maintained at π, such that the seed field is 'de-amplified' by the pump field. This results in a state squeezed in the amplitude quadrature and anti-squeezed in the phase quadrature.

There are three active control loops required for the operation of the OPA. The first controls the OPA cavity length and is defined by the resonance condition of the 1064nm seed light. The second controls the phase relationship between the seed and pump fields to define the angle of the squeezing. The third controls the non-linear crystal temperature to ensure quasi-phase matching.

The seed field was injected into the cavity through a HR back mirror (Figure 5.3) with the reflected/transmitted light detected for the purposes of extracting an error signal for both the OPA cavity length and the relative phase between the seed and pump fields. The cavity length was controlled by actuating the position of mirror (m_2) using a piezo-electric transducer (PZT). Our choice of sideband modulation for locking the OPA was

complicated by a few considerations. To ensure a reasonably wide detection band for acquisition free of classical noise, we needed to take care with our choice of sideband modulations. This meant introducing as few classical sideband modulations as possible. In addition to providing an error signal for the OPA cavity length and relative phase, the phase modulation at 11.25 MHz was also recycled to control the displacement of the squeezed state.

The technique for extracting an error signal for both the OPA cavity length, and the phase relationship between the seed and pump fields, from the same photocurrent was introduced to me by Boris Hage - who briefly discusses it in his thesis[99]. The technique requires the sideband modulation to be (preferably well) within the cavity linewidth. An error signal obtained via the PDH locking technique [96, 97] is formed of two components: a real term and an imaginary term. These two terms are individually accessible via demodulation of the photocurrent at $\phi = 0$ and $\phi = \pi/2$ respectively, but in almost all PDH locking scenarios the choice of the modulation frequency (either much larger or much smaller than the resonator linewidth) ensures that one term is dominant (and thus only one choice of ϕ is required). It is perhaps more commonplace to choose a modulation frequency well outside the resonator linewidth, but here we consider a modulation within the resonator linewidth. Demodulating of the aforementioned photocurrent at $\phi = \pi/2$ produces the desired error signal for the cavity length. Demodulation of the same photocurrent at $\phi = 0$ recovers an error signal for the relative phase between the seed and pump fields, with zero crossings at $\theta = 0$ (seed amplification) and $\theta = \pi$ (seed de-amplification). The $\pi/2$ difference in the demodulation phase also ensures that the two error signals should be uncoupled. As the seed field was injected through a HR mirror, it experienced a very under-coupled cavity, and as such, the error signal was especially sensitive to any parasitic amplitude modulations on the seed field. A resonant phase modulator circuit was used to enhance the modulation depth.

The crystal temperature was stabilised in a conventional manner. The crystal was seated within a small copper coffin mounted on a Peltier. The Peltier itself was mounted on a comparatively large copper block, which provided a large thermal mass for additional temperature stability. The temperature was controlled using a high resistivity thermistor mounted within the copper coffin, with the feedback implemented by a commercial temperature controller.

Doubly-resonant system

One of the primary requirements of the experiment was the long-term stability of the squeezing source. This required the long-term stability of the co-resonance condition. As the quasi-phase matching temperature is defined by the periodic poling domain, the manufacturer specifies an approximate quasi-phase matching temperature. An estimate of the temperature can be further improved by coupling a bright seed field into the cavity and measuring the second harmonic generation efficiency as a function of temperature. In a doubly resonant design, however, achieving sensitive measurements of the SHG efficiency is complicated by the requirement of co-resonance. For periodically poled crystals however, the sinc-squared dependence of the conversion efficiency is substantially larger (5°C) than the co-resonance condition (0.1°C). As such, it was more critical to meet the co-resonance

condition than satisfy the exact quasi-phase matching temperature.

The wedged crystal angle of 1° gave approximately five points of co-resonance (where the effects of intra-cavity dispersion were compensated) across the diameter of the crystal. Given an initial crystal temperature close to the ideal phase matching temperature, the crystal was then translated laterally until the two fields were approximately co-resonant. Stabilising the cavity on resonance, the crystal temperature was further tuned to maximise the observed parametric gain, corresponding to achieving the co-resonance condition. As this co-resonance condition was dependent on the localised heating due to green absorption, the temperature was again optimised once the relative phase between the seed and pump was stabilised. Provided this temperature was set correctly, theoretically the system should 'self-stabilise'. If the system is perturbed away from the co-resonance condition - say, the cavity loses lock - the loss of localised heating required that upon reacquiring the cavity lock, the condition for co-resonance will no longer be met. However, the effects of absorption from the slightly off-resonant green light should push the system closer to co-resonance, producing more green absorption and iteratively driving the system towards co-resonance.

As is usually the case, however, things are experimentally more subtle. Small perturbations around the co-resonance condition - usually driven by fluctuations in the pump power or temperature controller stability - have an observable effect on the error signal controlling the relative phase between the seed and pump fields. Jitter in the phase angle between the seed and pump field essentially shakes the squeezing ellipse, manifesting as non-Gaussian noise that is proportional not only to the size of the rotation, but also the size of the initial squeezed state [100]. This effect can be largely overcome with optimisation of the control systems, but over the timescales required for this experiment it proved problematic. Over timescales of 10 to 20 minutes fluctuations in the pump power and the temperature control lead to drifts in the squeezing and the squeezing angle. This is most problematic for tomographic aspect of the experiment, where a fundamental assumption is the *identical preparation* of the unknown state over the measurement time. These effects were partially mitigated by choosing a input coupler with reflectivities 70% for 532 nm and 85% for 1064 nm, that, with the additional absorption for visible light in the bulk material, produced a 532 nm cavity linewidth over twice that of the 1064 nm cavity. This ensured that small changes in the co-resonance condition would not have such pronounced consequences for the circulating green power, thus improving the overall stability. In hindsight however, moving to a singly resonant system would have been a smarter choice.

4.2 State Reconstruction

After preparation of the squeezed resource, our state is divided between two measurement stages: a conditioning measurement and a reconstruction measurement.

Displacement stage

Homodyne tomography requires we reconstruct the marginals $\mathrm{pr}(\hat{X}^\theta)$ for several quadrature angles θ. Alternatively, one can continuously sample $\hat{X}^\theta$ whilst scanning the measurement angle, θ, reconstructing the phas. Our technique to define and control the homodyne

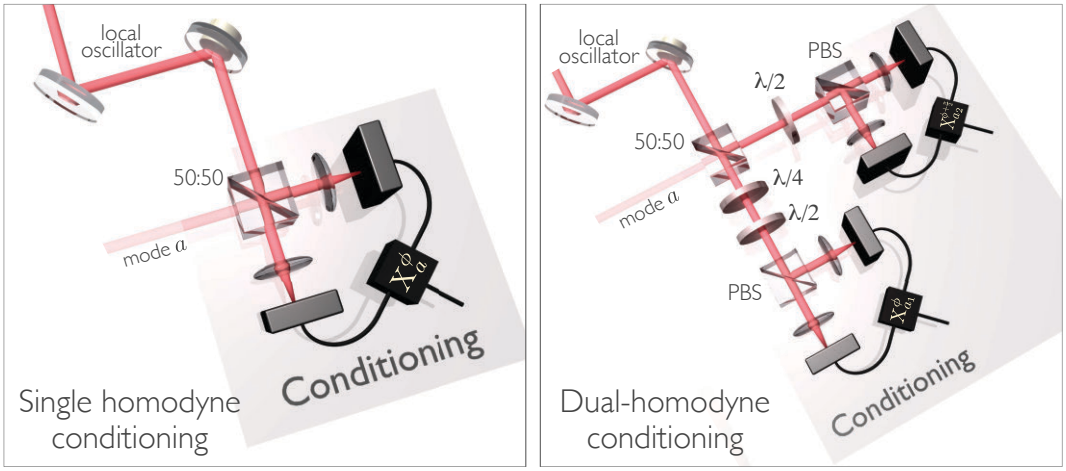

Figure 4.3: The two different conditioning measurement schemes.

detection angle, θ, requires we introduce both classical amplitude and phase sidebands. As we cannot afford the requisite loss from direct modulation of the squeezed state, we instead make use of an auxiliary beam on which we do the encoding.

A small portion of the available laser light was tapped off after the 1064 nm mode cleaner. The light was first focused through a phase electro-optic modulator, and then through an amplitude electro-optic modulator. The polarisation of the light was rotated from vertical to circular after the phase modulator to ensure a linear response of the polarisation modulation to the RF signal. The tomographic locking technique that exploits the phase and amplitude modulation will be discussed in §4.2.2.

The auxiliary mode is then interfered with the squeezed coherent state on a very unbalanced 98:2 beamsplitter. The intensity of the 2% transmitted auxiliary mode is matched to that of the 98% reflected squeezed coherent state, and the relative phase between the auxiliary and squeezed modes is controlled using the 11.25MHz phase modulation available on the squeezed state. The error signal extracted from the small 2% transmitted squeezed light is substantially 'boosted' by the auxiliary mode that is ≈ 2400 times brighter. In addition to introducing the required phase and amplitude modulation on the squeezed coherent state, by choosing the relative phase appropriately, it allows us to displace the coherent squeezed state to a coherent amplitude near to zero. This allows more flexibility in the intensity of the local oscillator. The displacement is primarily limited by the quality of the mode matching between the auxiliary mode and squeezed coherent state. Our fringe visibility was 97.5%.

The squeezed resource state is then split between a *conditioning* measurement and a *characterisation* measurement. Akin to a traditional photon subtraction implementation, a small portion of the light is reflected to the analogous conditioning measurement, whilst the remainder is transmitted for the reconstruction of the conditioned states. The beam-splitter reflectivity defines the fidelity of the "subtraction" operation with an ideal implementation of the annihilation operator. A half-wave plate combined with a polarising beam splitter (PBS) gives us a variable beam-splitter, allowing us to vary the proportion of light used for the conditioning measurement. Typically we reflect 10% of the squeezed

resource for conditioning, the transmitted 90% for characterisation.

4.2.1 Conditioning Measurement

The difference in the conditioning measurement is the only fundamental change between the two experimental demonstrations. Schematics for both conditioning measurements are provided in Figure 4.3. Other changes occurring in the second implementation were largely technical improvements.

Dual-homodyne detection

The first experiment used dual-homodyne detection (or heterodyne detection) as the conditioning measurement. Historically, the idea of replacing the dual-homodyne detection with a single phase-randomised homodyne detection occurred to us after these first measurements were made.

The dual-homodyne measurement was implemented using a polarisation technique outlined in Figure 4.3. The $\sim$ 10% or so of the vertically polarised squeezed field reflected for conditioning is split on a 50:50 non-polarising beam-splitter (NPBS). An orthogonally (horizontally) polarised bright local-oscillator is injected in the unused input port. Of the two output paths, the first (mode a_1) contains a half-wave plate orientated at 45° followed by a PBS that divides the light between two photodetectors, the subtraction of these two photocurrents providing one homodyne measurement. The second output path (mode a_2) also includes a quarter-wave plate orientated at 45°, which introduces a $\pi/2$ phase shift between the two orthogonally polarised local oscillator and squeezed field. The $\pi/2$ phase shift introduced between the local oscillator and the squeezed field ensures the homodyne observable, $\hat{X}_{a_1}^{\phi}$ sampled at mode a_1 will be in quadrature with the observable sampled at mode a_2.

In principle, this technique makes the active control loop of the second homodyne redundant, as the relative phase between the two fields is already appropriately defined. In this instance however, it also exempts us of the need to actively control the conditioning measurement, because, for our estimation of $\hat{n}_a$, measurement of $\hat{X}_{a_1}$ and $\hat{P}_{a_2}$ is equivalent to measurement of $\hat{X}_{a_1}^{\phi}$ and $\hat{X}_{a_2}^{\phi+\frac{\pi}{2}}$. For the dual-homodyne measurement the local oscillator phase is permitted to drift.

The two homodyne stages each utilised two universal photodetectors (Uni-PD). These Uni-PD circuits were designed in-house and named such because they balance desirable high gain and low dark noise properties with reasonable large bandwidth, providing a photo-detector that is nearly 'universally' suitable for many purposes. The detectors were electronically matched to provide a common mode rejection of 45 dB at the subtraction. For the sideband frequencies up to 5 MHz the Uni-PD design provided up to 20 dB of dark noise clearance with a suitably bright local oscillator.

The homodyne efficiency of the entire detection stage was originally estimated at 94% ±2%. This was primarily limited by the *Epitaxx ETX-500* InGaAs photodiodes which have a manufacturer specified quantum efficiency of $95 \pm 2\%$. Later independent measurements made with squeezed light contend this number was more likely to be $\approx 90\%$. Improvements of approximately 2% in the absorbed light could be made by tweaking the angle of incidence, and the total efficiency could have been improved by retro-reflection

of the light onto the diode surface. The secondary limitations to the homodyne efficiency include the accumulated transmission losses and polarisation mis-match introduced by the polarisation optics required for dual-homodyne technique, and mode-matching, with typical fringe visibilities of 99.5 on the dual homodyne stage.

Phase-scanned Homodyne Detection

In the second experiment the conditioning was instead implemented using a single homodyne detection stage. Accurate state reconstruction with this technique relies on phase-randomised homodyne detection with equal representation of all angles. To reduce error resulting in a systematic under or over sampling of quadratures angles, we used a symmetric sawtooth signal to ramp the phase of the homodyne over several π at approximately 100 Hz – significantly faster than the drift of the global phase of the lasers. The encoding of phase and amplitude modulation sidebands to allow control of the homodyne angle X_b^θ for tomographic reconstruction also allows us to verify that for our conditioning measurement, X_a^ϕ, we are equally sampling all quadratures.

The second experiment utilised a new universal photodetector design with a performance comparable to the original design. These new photodetectors used InGaAs diodes custom manufactured by the *Fraunhofer institute* which had a manufacturer specified quantum efficiency of $> 99\%$. The measured fringe visibility of the homodyne detection stage was typically 99.7%, giving an overall homodyne efficiency $> 98\%$.

4.2.2 Characterisation Measurement

The details of the characterisation or reconstruction measurement are largely identical for both implementations. The second experiment benefitted from a much improved quantum efficiency due to the aforementioned improved InGaAs photodiodes.

Tomographic Locking

Optical homodyne tomography requires we stabilise the phase of a balanced homodyne detector to any arbitrary quadrature of our choosing. To do so, we introduced amplitude and phase sidebands onto our squeezed field via a displacement (§4.2). Our technique requires both the frequency, ω, and magnitude, ξ of the phase and amplitude modulation be identical, but the signals be in quadrature (requiring a $\pi/2$ phase shift between them). The modulation of both the amplitude *and* phase quadrature can be described by the coherent addition of the two,

$$\begin{aligned}\hat{a}(t) &= \frac{1}{2}\hat{a}_0(t)\left[1+\xi\cos\omega t\right] + \frac{1}{2}\hat{a}_0(t)\left[1+i\xi\cos\left(\omega t+\tfrac{\pi}{2}\right)\right] \\ &= \hat{a}_0(t)\left[1+\frac{\xi}{2}(\cos\omega t - i\sin(\omega t))\right] = \hat{a}_0(t)\left(1+\frac{\xi}{2}e^{-i\omega t}\right). \end{aligned} \tag{4.1}$$

From the form of (4.1) it is clear we have introduced a single sideband to the signal field. As we are interested in stabilising the relative phase θ between the squeezed field, $\hat{a}$, and the local oscillator, $\hat{a}_{lo}$, we only need to make reference to their classical amplitudes.

Interfering the signal field and the local oscillator on a 50:50 beamsplitter, we obtain the output fields,

$$\alpha_1 = \frac{1}{\sqrt{2}}\left[\alpha(1+\xi e^{-i\omega t}) + \alpha_{lo}e^{i\theta}\right]$$
$$\alpha_2 = \frac{1}{\sqrt{2}}\left[\alpha(1+\xi e^{-i\omega t}) - \alpha_{lo}e^{i\theta}\right]. \tag{4.2}$$

Taking the difference of the corresponding photocurrents i_1 and i_2 (where $i_k \propto \alpha_k^\dagger \alpha_k$) gives the expression

$$\Delta i \propto 2\alpha\,\alpha_{lo}\,\xi\cos(\omega t - \theta). \tag{4.3}$$

Electronically mixing the difference photocurrent with an electronic local oscillator, $\sin(\omega t - \phi)$, and low-pass filtering retrieves the error signal,

$$e \propto \alpha\,\alpha_{lo}\,\xi\,\sin(\theta - \phi), \tag{4.4}$$

where ϕ denotes the phase of the electronic local oscillator. Choosing the demodulation phase, ϕ, equal to the desired homodyne angle θ, produces an error signal with a zero-crossing at θ. This allows stabilisation of the homodyne detection to any arbitrary angle by simply choosing the electronic demodulation phase.

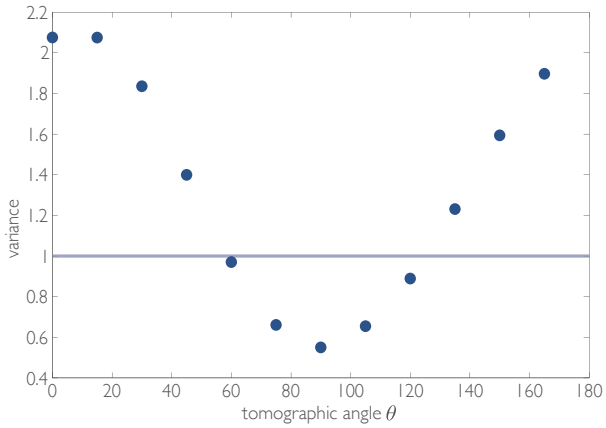

Figure 4.4: Variance of the measured squeezed state at the characterisation stage as a function of the tomographic angle, $\theta = 0°, 15°, \ldots 165°$

Tomographic measurement

The characterisation used a single balanced homodyne detection stage, with the typically $\sim$ 90% of remaining squeezed light split on a 50:50 beam splitter and interfered with a bright local oscillator. The signal field and local oscillator were matched to a fringe visibility of 99.7%. The overall homodyne efficiency for the system was estimated to be $92 \pm 2\%$ for the first experiment, and $> 98\%$ for the second experiment.

The auxiliary mode used to displace the squeezed field introduced phase and amplitude

sidebands at a frequency of 29.0625 MHz (changed to 30 MHz for the second experiment). The homodyne difference photocurrent was split, with half sent for acquisition whilst the remainder was used to control the quadrature angle via the aforementioned method.

4.2.3 Experiment Control & Measurement Acquisition

The experiment utilised a digital control system in *National Instruments LabView* developed by Ben Sparkes, with an overview of the control system published in [101]. A frequency generator (FG - *National Instruments* PXI-5404) provides an 80MHz clock for the system. The 80 MHz clock is split between a clock generator board (CGB - *Analog Devices* AD9959) and a high speed analog-to-digital convertor (ADC - *Analog Devices* AD9460BSVZ-80) . The CGB is controlled via a field programable gate array (FPGA - *National Instruments* PXI-7852R), providing a sine wave of arbitrary frequency for modulation of the light for the purposes of locking. The photocurrent used to to extract an error signal was acquired via the high-speed ADC, and input to the controller algorithm (discussed in detail in [101]). Given our typical control bandwidth, the resulting error signal only required a low speed analog output (AOP), and was subsequently amplified by an analog high voltage amplifier before driving a piezo-electric transducer. At capacity the system required 6 active control loops.

Crucially, the digital system provided additional flexibility for running an experiment over a long period of time. The digital system allowed for the inclusion of logic that allowed for automated re-locking, and also sequential locking, whereby a lock was only obtained when all the dependencies were themselves locked. Significantly, it also allowed for integration of the control system with the acquisition system. The acquisition utilised a pair of 2 channel high-resolution digitisers (*National Instruments* PXI-5124) mounted in the same PXI chassis, providing a sampling rate of 200 MS/s and 12-bit vertical resolution. The acquisition algorithm, developed by Thomas Symul and Ben Sparkes, allowed for automation of the measurement procedure. In addition to automation of the acquisition itself, measurement acquisition could also be conditional on the state of the system controller, ensuring results were only taken when the system was appropriately stabilised. Depending on the intended reconstruction, acquisition times would vary from 10 minutes to several hours.

4.3 Data Analysis and Tomographic Reconstruction

If we first ignore the role of conditioning, the ensemble of homodyne measurements at the tomographic characterisation stage allows construction of the histograms describing the probability distribution of each measured X_b^θ. To do so, we discretise our continuous measurement spectrum by decomposing our X^θ into a finite number of bins, M_x (with our chosen number of tomographic angles, N_θ). For each sample, x_b^θ, we increment the relevant bin, (m, n), by one. Once we have reconstructed histograms describing X_b^θ for several values for θ, we then reconstruct the Wigner function or density matrix. Here we use two different techniques: direct sampling via the pattern functions (§2.5.2) and the MaxEnt principle (§2.5.3).

The extension to 'conditioning' in post-processing is implemented as follows. For

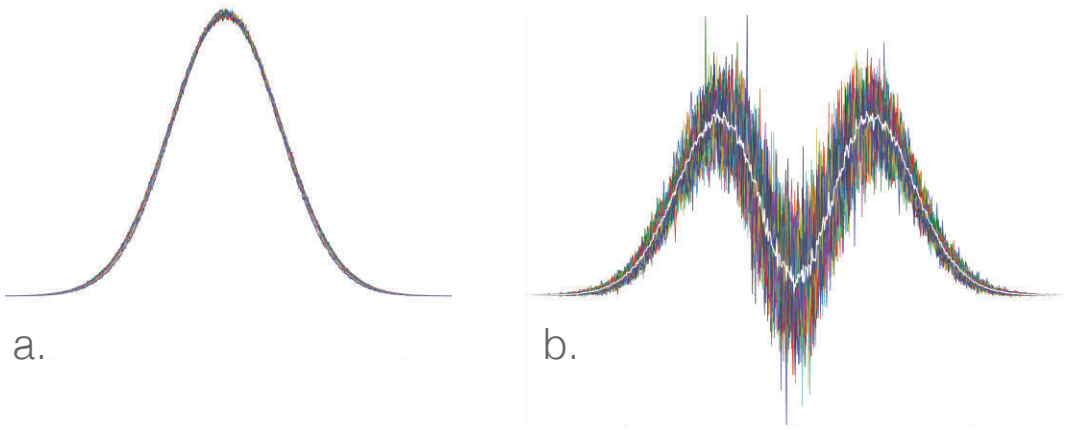

Figure 4.5: A single data set consisting of 10^7 data points is decomposed into 100 sets. (a) Measured probability distribution for the phase quadrature for the 100 sets (b) The reconstructed probability distribution for the $\hat{n}_a$ conditioning for each of the 100 sets. The familiar non-Gaussian shape emerges from the average, provided in white.

each sample x_b^θ we have a corresponding measurement of mode a, x_a^ϕ, which provides the value for the relevant weighting. Instead of incrementing the bin (m, n) corresponding to x_b^θ by one, we instead increment the bin by the outcome of a function of our choosing $\mathcal{P}(\hat{n}_a)$, for which our measured homodyne (or heterodyne) sample at a is the input. We repeat this process until we accumulate sufficient statistics for the non-Gaussian state that corresponds to our 'conditioning' polynomial emerges.

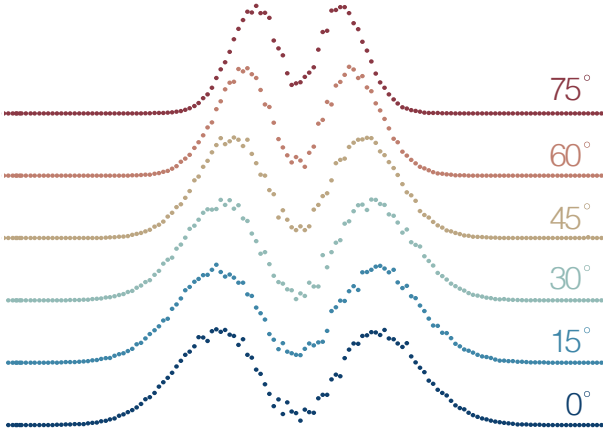

Figure 4.6: Reconstructed probability distributions as a function of the tomographic angle, θ.

4.4 Results & Discussion

4.4.1 Dual-Homodyne Conditioning

We present the results chronologically, beginning with the *dual-homodyne conditioning* experiment. For Figure 4.7 we reconstruct the Wigner functions (and the corresponding

density matrices in Figure 4.8) by directly sampling the density matrix via the pattern functions, using the method described in §2.5.2. We begin by focusing on the reconstruction of the 1-PSSV state. Figure 4.7(a) gives the Wigner function obtained using the simplest conditioning polynomial, $\mathcal{P} = \hat{n}_a$. This conditioning should ideally remove any contribution corresponding to a measurement of the vacuum, $n_a = 0$, in the conditioning mode, a. All other contributions remain, and their contributions are additionally weighted by their corresponding eigenvalues, n_a. In essence we reconstruct a statistical mixture of primarily the 1-PSSV and 2-PSSV states, where their contributions are not solely weighted by the likelihood of successful 'conditioning', but additionally by their corresponding eigenvalues. For instance, the contributions from $n_a = 2$ are weighted at twice that of contributions from $n_a = 1$.

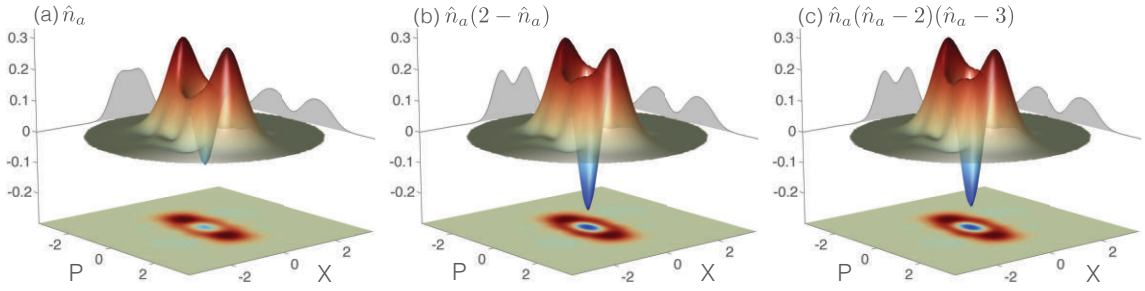

Figure 4.7: Reconstructed Wigner functions for the 1-PSSV state using the pattern function method of sampling the density matrix: The purity and negativity of the reconstructed state improves as we remove contributions from $n_a = 2$ (b.) and $n_a = 3$ (c).

As the ideal squeezed vacuum populates only the even photon number pairs, the ideal subtraction of one photon from squeezed vacuum should produce a superposition of the odd photons numbers (removing any vacuum contribution). An idealised implementation of a photon annihilation corresponds to a beam splitter with reflectivity approaching zero. This permits statistical isolation of a single photon subtraction event from the considerably less likely two photon subtraction event. However, with an experimental implementation, the requirement of a finite tap-off (typically around 10%) inevitably introduces spurious higher order photon subtraction contributions. This is evident when we consider the prominence of the even-photon number terms in the reconstructed density matrix in Figure 4.8(a).

As described §3.2.3, one can instead consider a higher order polynomial in $\hat{n}_a$ that removes potential contributions to the reconstructed state from higher order subtractions that are unwanted and are sufficiently statistically significant to warrant removal. Figure 4.7(c) demonstrates the dramatic improvement in the reconstructed 1-PSSV state by implementing the conditioning polynomial $\mathcal{P} = \hat{n}_a(\hat{n}_a - 2)(\hat{n}_a - 3)$, removing polluting contributions from the 2 and 3 photon subtractions. As the odd-photon number and the even-photon number subtractions approximate theoretical cat states of different parities, contributions from 2 photon 'events' in the reconstructed 1-PSSV state degrade the negativity.

The pattern function method for reconstructing the density matrix, and thus the Wigner function, is limited by its statistical precision in estimating the photon number elements. If the measurement ensemble is too small, the error associated with the obtained

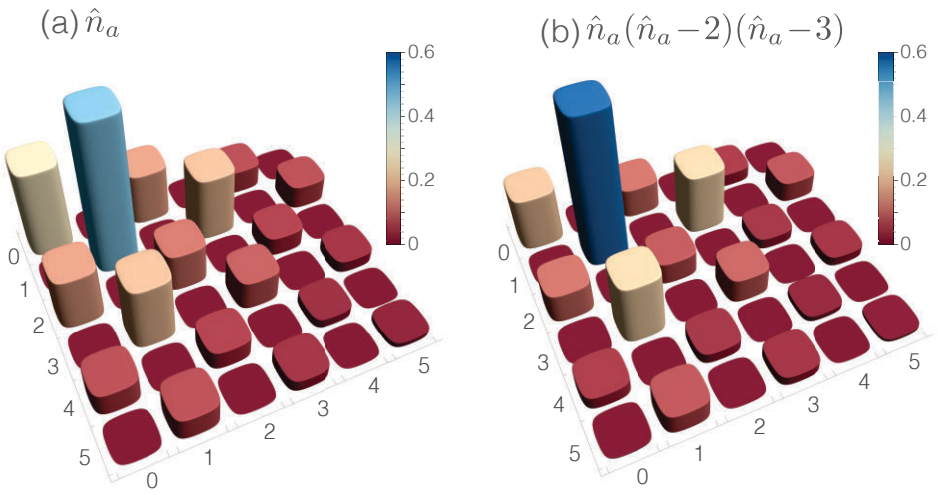

Figure 4.8: Reconstructed density matrices for the 1-PSSV state using the pattern function method of sampling the density matrix: the dominance of the single photon contribution improves as the $n_a = 2$ and $n_a = 3$ contributions are removed.

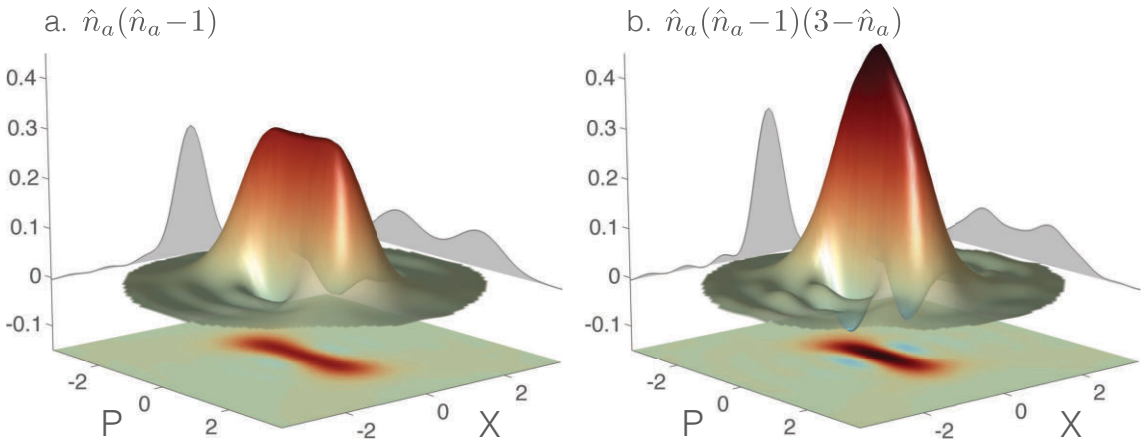

Figure 4.9: Reconstructed Wigner functions for the 2-PSSV state using the pattern function method of sampling the density matrix. The purity and negativity of the initial reconstructed state (a) improves as we remove contributions from $n_a = 3$ (b).

mean value of the photon number element can result in negative contributions or an unphysical density matrix. It might at first seem stange strange to talk quantitatively about the size of the measurement ensemble for the technique presented here, as we cannot make reference to individual, heralded 'events', and the reconstruction itself requires the entire measurement record to succeed. But consider an attempt to reconstruct the 2-PSSV state in the usual manner. Ignoring the effects of squeezed state purity, and given a typically beamsplitter reflectivity of 10%, the likelihood that two photon subtraction event occurs is a tenth as probable as a single photon subtraction event. For a hybrid system it is easy to understand why you need to measure for longer - for every 10 one photon events you

see a single two-photon event. Here, we see a similar effect: to obtain the same statistical precision, you need to acquire an order of magnitude more data.

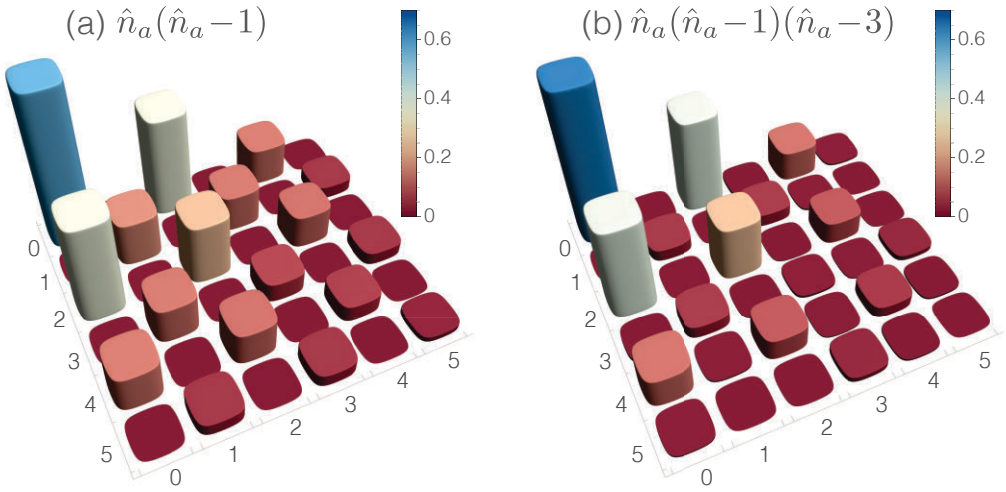

Figure 4.10: Reconstructed density matrices for the 2-PSSV state reconstructed using the pattern functions. Correction for the $n_a = 3$ contributions largely eliminates the odd photon number contributions.

Figure 4.9(a) considers the polynomial, $\mathcal{P}(X_a^{\phi}) = \hat{n}_a(\hat{n}_a - 1)$, removing contributions corresponding to a photon number measurement of $n_a = 0$ and $n_a = 1$. The ideal reconstructed 2-PSSV state has high fidelity with the even kitten state. When we additionally correct for the contributions of the 3-PSSV state there is a clear improvement (Figure 4.9(b)) in the purity of the reconstructed state, evidenced by the increasing isolation of the even-photon number contributions to the density matrix (Figure 4.10(b)).

Dark noise correction

There a handful of subtleties involved in the estimation of the photon statistics with homodyne measurements that we discussed in Chapter 3. Analogies with many of these can be drawn with the usual problems that afflict the analogous real photon counting. This technique relies on correlations shared between modes a and b, and may be degraded by any process that introduces uncorrelated classical or quantum noise. The inherently 'ensemble' nature of this approach means any noise contributions that are uncorrelated with the quantum state may be corrected for, assuming that the noise can be correctly characterised. Unlike experiments with actual photon counters, a correction for dark noise on the conditioning step can integrated directly into the conditioning polynomial. For both experiments we routinely characterised the dark noise of the homodyne detectors for both the conditioning and characterisation measurement. The photodetectors used had a typical electronic noise floor 20 dB smaller than the shot noise variance. With the 20 dB clearance over our measurement band, any improvement provided by the correction was negligible, and consequently, all the results presented here are without dark noise correction.

The role of loss

However, we are still exposed to the effects of loss. Any loss of purity on the initial squeezed vacuum state constrains the non-Gaussian nature of the reconstructed state. The role of loss can be accurately modelled as a beam-splitter with transmissivity, λ, and can be qualitatively understood by drawing analogy to traditional photon counting. Inefficiencies arising from imperfect homodyne detection efficiency or transmission losses scale the rate of success of the homodyne conditioning, analogous to loss on a photon counting measurement. Whilst here we cannot refer to individual events, as this approach succeeds by considering the entire ensemble, we essentially require a larger ensemble to obtain the same conditioned statistics. Additionally, it can also lead to erroneous conditioning, where a loss of photon may see a 3-photon subtraction event contributing as two photon subtraction.

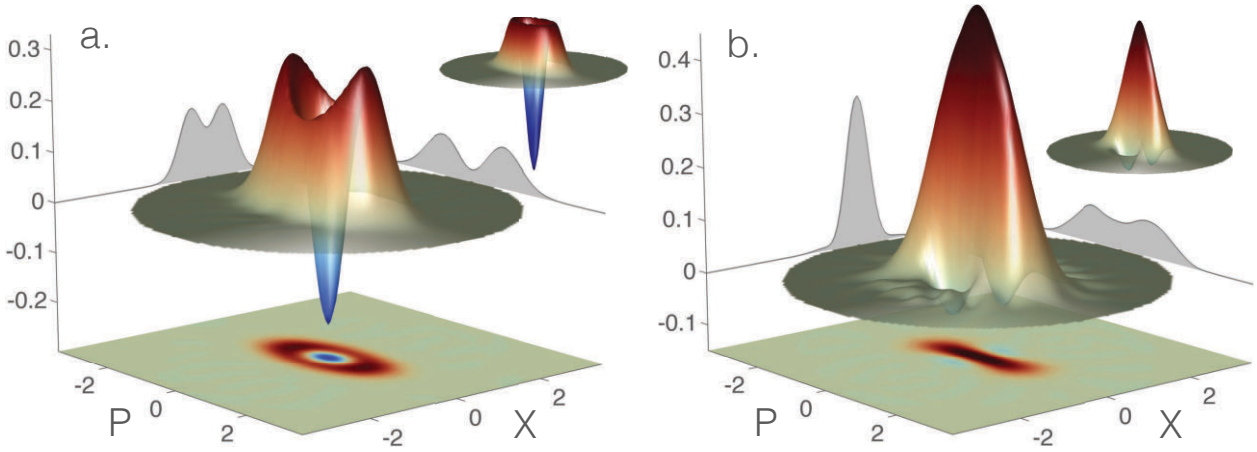

Figure 4.11: Reconstructed Wigner functions for the (a). 1-PSSV and (b). 2-PSSV states using the MaxEnt principle.

4.4.2 Phase Randomised Homodyne Conditioning

The second experiment experiment saw a large improvement in the overall purity of the squeezed resource owing to much improved quantum efficiency of the photodiodes, and improved quality of the non-linear crystal surfaces and AR coatings. We also enclosed the experiment table within a box, which reduced temperature gradients and air currents across the experiment table, minimising loss and noise contributions from beam-pointing. Figures 4.12 and 4.13 demonstrate the considerable improvements in the purity of the reconstructed one and two PSSV states, respectively. A comparison of the purities $\mathrm{Tr}\rho^2$ of the best 1-PSSV states for the two experimental implementations gives 0.58 for Figure 4.7(c) and 0.72 for Figure 4.12(c). The use of the single homodyne conditioning measurement instead of the dual-homodyne also reduced loss contributions on the conditioning mode. The single homodyne conditioning is also compatible with the pattern function conditioning introduced in §2.5.2, which, save for experimental limitations, should amount to the continuous variable analog of *perfect* photon number discrimination. The f_{11} pattern function reconstructs the state with theoretically perfect isolation of the one photon

statistics on the conditioning mode, the analog of a perfect projective measurement on $\tilde{\rho}_a = |1\rangle\langle 1|$.

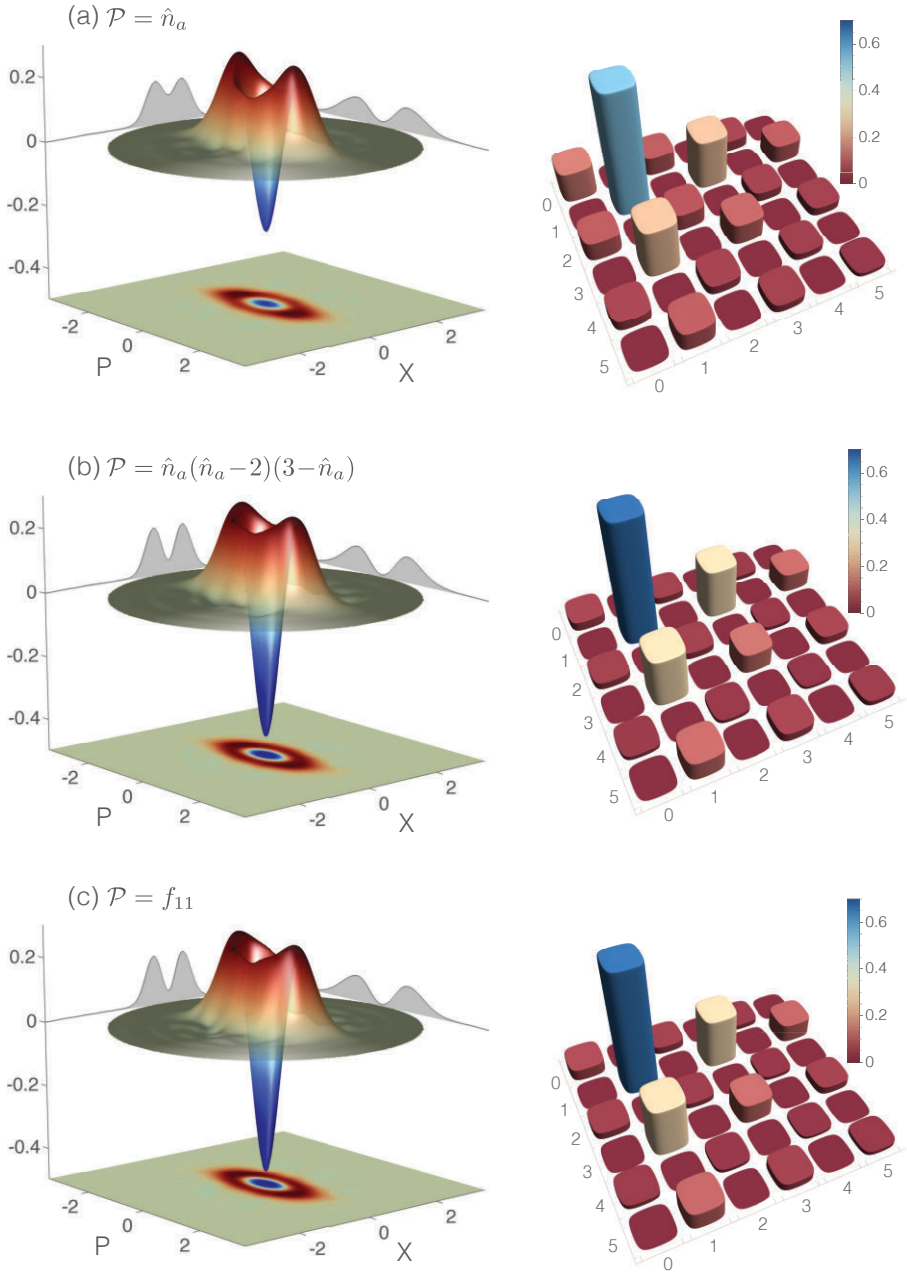

Figure 4.12: Reconstructed Wigner functions and density matrices for the 1-PSSV state using MaxEnt method of reconstructing the density matrix.

Given the homodyne efficiency of typically 98% our primary source of loss in the experiment arises from the impurity of the squeezed vacuum resource, and this is most evident with the reconstruction of the 3-PSSV state (Figure 4.14). In endeavouring to reconstruct the 3-PSSV state, we optimised the experimental parameters to increase the

likelihood of having 3 photons in mode a without sacrificing the quality of the reconstructed state. The likelihood of encountering a 3 photon subtraction event is low. Whilst the probability of subtracting n-photons with a beam splitter of reflectivity η scales as η^n, attempting to measure 3 or 4 photons from mode a also enforces the additional requirement of having at least 4 photons in the original squeezed vacuum mode. As a result, the likelihood of having 3 or more photons in mode a scales poorly. We can improve this predicament by first increasing the percentage of the input mode used for conditioning from 10% to 15%, and second, moving to a stronger squeezed resource, enhancing the population of the higher-order photon pairs. Increasing the squeezing level is often detrimental to the squeezing purity as it introduces noise sources only dominant at high pump power, such as phase noise. In our doubly-resonate system, the requirement of the stronger pump field also has consequences for the long-term stability of the experiment. Obtaining sufficient statistics requires longer acquisition time which concatenates the typical experimental drifts in the measured tomographic angle θ, alignment and squeezing levels over time, reducing the overall purity of the reconstructed state. This problem is further exacerbated by the requirement for loss and noise mechanism that become more dominant at higher pump powers. As a result the reconstructed 3-PSSV state in Figure 4.14(a) has lower reconstructed state purity (evidenced by the smaller observable negativities at the origin) than the reconstructed 1 and 2 PSSV states which require smaller data sets.

If we attempt to reconstruct the 3-PSSV state with an additional correction for the 4 photons events in mode a, the reconstructed state becomes noisier. It is not immediately apparent that removing unwanted contributions should introduce statistical noise into the ensemble, but conditioning on higher photon numbers or the removal of higher order terms essentially requires extraction of finer correlations between modes a and b. For a polynomial $\mathcal{P}(n_a)$ of degree k, we essentially estimate moments of X_b^ϕ up to order $2k$. When coupled with the rapid divergence of the polynomials in X_b^ϕ, sufficient statistics must be acquired to minimise error. This prevents us from implementing a purification of the 3-PSSV state in Figure 4.14(a) with the polynomial approach, even though it is successful with the corresponding f_{33} pattern function (Figure 4.14(b)).

While the pattern functions extract the statistics of ideal photon number discriminating measurement at mode a, limited only by the experimental imperfections, it is worth noting that one can essentially obtain the same outcome by implementing a polynomial weighting to only a few orders. This is despite the fact the polynomials calculated to any $\mathcal{P}(\hat{n_a})$ rapidly diverge for sufficiently large X_b^θ. To emulate a conditioning photon number measurement a low-order implementation of the $\hat{n}$ polynomials is generally sufficient. This point is primarily academic - there is no advantage to choosing the polynomials over the pattern functions which they approximate. To the contrary, when reconstructing with limited statistics, the pattern functions appear at least, if not more efficient than the equivalent polynomial that would approximate the same operation.

What are these results good for? The first point of interest is purely academic. These results build on the several very important works of the mid-to-late 1990's, where the idea of investigating the discrete variables of light by interrogating the continuous variables emerged [46, 50, 54, 102, 51, 103]. On the surface it does appear remarkable that the statistics of a non-Gaussian state can be extracted from only Gaussian measurements. And these results are physically meaningful, corresponding to the state that would be

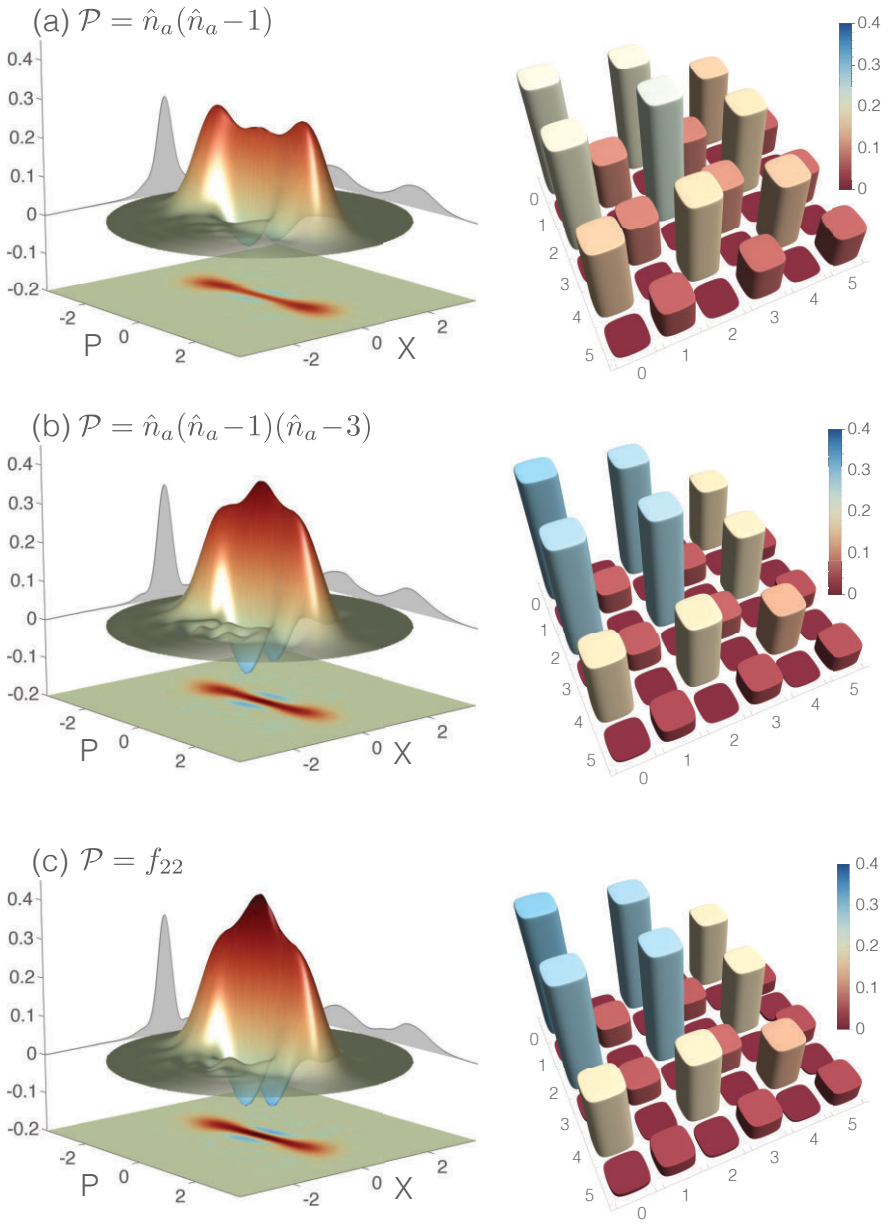

Figure 4.13: Reconstructed Wigner functions and density matrices for the 2-PSSV state using MaxEnt principle.

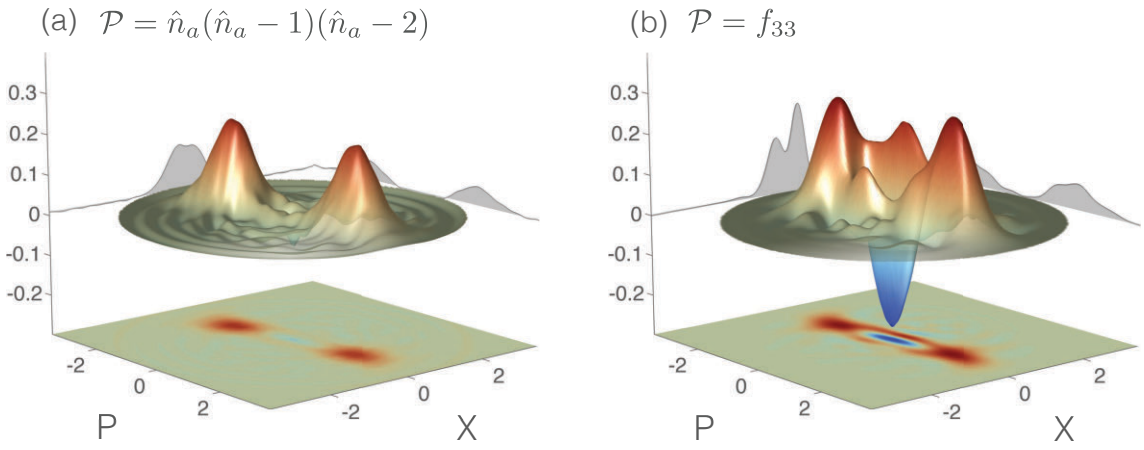

Figure 4.14: Reconstructed Wigner functions for the 3-PSSV state reconstructed using the Max-Ent principle.

prepared given the same physical measurement on mode a. However, if one has access to ρ_{AB} - which can be completely determined by way of regular two-mode tomography - one has access to the same non-Gaussian statistics that we extract with our procedure. In this sense, what I have presented is a variation upon the problem of two-mode tomography.

Even if it is of fundamental interest, the ensemble nature of the approach renders it useless for most (or all) quantum information or communication applications. However, there may be value is this technique in fields where measurements of the continuous variables are favoured because of technical challenges. Opto-mechanical systems, especially those in the microwave regime, are restricted to measurements of the field quadratures as photon counting technology is still undeveloped or experimentally unfeasible. For the moment, techniques similar to those discussed here are indispensable tools for probing quantum effects in such systems [104]. This technique could prove useful to understand and characterise the quantised behaviour of such systems.

4.5 Summary

In this Chapter we have experimentally demonstrated the reconstruction of the photon subtracted squeezed vacuum states using using only measurements of the field quadratures. Previously, extracting such statistics would have required a full tomographic reconstruction of the two-mode Wigner function. These techniques allow for complete characterisation of the outcome of a conditional measurement on a system, and might prove useful in systems where measurements of the DV of the system are limited or unavailable.

Chapter 5

Measurement-Based Noiseless Amplification

5.1 Introduction

The impossibility of determining all properties of a system, as exemplified by Heisenberg's uncertainty principle [105] is a well known signature of quantum mechanics. It results in phase and amplitude fluctuations in the vacuum, enables applications such as quantum key distribution and is at the heart of fundamental results such as the no-cloning theorem [106], quantum limited metrology [107], and the unavoidable addition of noise during amplification [108, 109]. This last constraint means even an ideal quantum amplifier cannot be used for entanglement distillation [110, 111, 112] which is a critical step in the creation of large scale quantum information networks [113, 114].

Distillation protocols, originally conceived for discrete variables [110, 111], proved initially more elusive in the continuous variable (CV) regime. The most experimentally feasible and theoretically well studied class of CV states and operations are the Gaussian states and the operations that preserve them [115], however restricting to this subset has been shown to make distillation impossible [116, 117]. Nevertheless protocols that distill Gaussian states have been discovered [112, 118] involving an initial non-Gaussian operation that increases the entanglement followed by a 'Gaussification' step that iteratively drives the output towards a Gaussian state. However, these protocols are experimentally demanding: Takahashi *et al.* demonstrated distillation through the de-Gaussification of a Gaussian state [90], with Kurochkin *et al.* implementing a quite similar proposal recently [119].

Returning to quantum limited amplifiers one can still avoid the unavoidable by moving to a non-deterministic protocol. This ingenious concept and a linear optics implementation have been proposed [120, 121, 122] and experimentally realised for the case of amplifying coherent states [123, 124, 125, 126], qubits [127, 128, 129], and the concentration of phase information [130]. All of these were extremely challenging experiments, with only Ref.[123] demonstrating entanglement distillation and none directly showing an increase in Einstein-Podolsky-Rosen (EPR) correlations [131]. Moreover the success probability of these experiments was substantially worse than the maximum set by theoretical bounds.

In the context of quantum key distribution (QKD), References [132, 133] proposed the possibility of implementing a non-deterministic measurement-based NLA (MB-NLA) to

improve performance. This represents a significant advantage as the difficulty of sophisticated physical operations can be moved from a hardware implementation, where one must suffer penalties related to source and detector efficiencies, to a software implementation where we are limited primarily by quantum theory and the statistics of our sample. Here we apply this protocol to EPR entanglement and observe improvements in the measured correlations consistent with distillation of the entanglement. We emphasise that this method is only equivalent to entanglement distillation for certain applications. Specifically, it is only situations where the desired distillation operation immediately precedes the measurement of the target mode that the two are indistinguishable.

We first derive some general conditions on the limits to implementing arbitrary quantum operations on an ensemble by conditionally filtering the measurement results. Using this method we experimentally implement an MB-NLA protocol achieving significant distillation with a much improved probability of success. Furthermore we illustrate the critical benefit of distillation in combating decoherence by considering the distribution of EPR entanglement through a lossy channel. We first recover an EPR violating correlations from an otherwise non-EPR violating state degraded by loss. Further, we measure an output level of entanglement that exceeds the maximum achievable without distillation, even if one could use a perfect initial entangled state. Finally, we also examine the role of the MB-NLA in quantum key distribution, providing a proof-of-principle demonstration of secret-key extraction from an otherwise insecure regime.

5.2 Theory

In any quantum information application the final result is always some classical measurement record, drawn from a set of possible outcomes $\{k\}$ and described by some probability distribution $p(k)$. In an application where the proposed distillation would take place immediately prior to measurement, for example in quantum key distribution, one could imagine emulating the operation on an ensemble via post-selective measurements. We first consider the process of emulating arbitrary operations via conditioning on measurements in a general setting before describing the results of References [132, 133] in which an explicit procedure applicable to the NLA was proposed. Our analysis will allow us to clarify some of the previous work as well as showing that the $g^{\hat{n}}$ operator key to the operation of the NLA, is particularly well suited to emulation via post-selection.

Consider an arbitrary quantum map applied to an incoming state ρ which can be written using the Kraus decomposition as [134].

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger \tag{5.1}$$

where $\{E_i\}$ are the Kraus operators. Note that this decomposition is valid for any completely positive operator including those which do not preserve the trace, as is the case with many useful conditional operations in quantum optics including photon addition and subtraction and the NLA. In the latter case the Kraus operators fail to satisfy the usual relation $\sum_i E_i^\dagger E_i = \mathbb{I}$, with the extra information needed to restore conservation of probability being the success probability, P, of the conditional process [135].

If this map is immediately followed by a positive-operator valued measure (POVM) described by operators $\{\pi_k\}$ the corresponding probability distribution is,

$$\begin{aligned} p(k) &= \mathrm{Tr}\left[\pi_k \sum_i E_i \rho E_i^\dagger\right] \\ &= \mathrm{Tr}\left[\tilde{\pi}_k\, \rho\right] \end{aligned} \tag{5.2}$$

where $\tilde{\pi}_k = \sum_i E_i^\dagger \pi_k E_i = \tilde{\mathcal{E}}(\pi_k)$ are a new set of POVM elements obtained by applying the mapping $\tilde{\mathcal{E}}$ to the desired output POVM set. What (5.2) tells us is that we may obtain the statistics of a POVM set $\{\pi_k\}$ upon an arbitrarily transformed state $\mathcal{E}(\rho)$ by conditioning upon measurements made with a transformed set $\{\tilde{\pi}_k\}$ on the original input state.

Although the above procedure is quite general, it is not arbitrary in that it does not allow the reconstruction of any desired POVM set in combination with any desired operation. If one wishes to implement arbitrary operations $\mathcal{E}$, one is restricted to certain final POVM sets $\{\pi_k\}$ and vice versa. Intuitively we expect that in order to correctly reconstruct the statistics of an arbitrary POVM upon an arbitrary state it is necessary to obtain maximum information about that state, i.e. to make measurements capable of complete tomographic reconstruction. This requirement can be derived by considering Equation (5.2) and noting that the POVM set with which one must actually measure, $\{\tilde{\mathcal{E}}(\pi_k)\}$, is not necessarily physical for arbitrary $\mathcal{E}$ and $\{\pi_k\}$. The unphysicality occurs because some of the operations that we wish to emulate are not themselves physical. For example the NLA itself, as will be discussed later, is trace increasing [136]. If one demands access to arbitrary operations, then a sufficient condition on $\{\tilde{\mathcal{E}}(\pi_k)\}$ would be that it maps to physical output states and is capable of uniquely determining an arbitrary CP map. This is precisely the same condition required of a POVM set for it to be classified as informationally complete (IC) [137, 138]. Conversely, if one is only able to experimentally realise a certain POVM set then one is limited in the range of operations that can be faithfully implemented.

5.2.1 Noiseless Amplification

Noiseless amplification is commonly defined as the ability to increase the amplitude of an unknown coherent state without any noise penalty, effecting the transformation

$$|\alpha\rangle \to |g\alpha\rangle \tag{5.3}$$

with $g > 1$. By considering the annihilation and creation operators describing a boson mode it becomes clear that such a transformation would violate the canonical commutation relations $[\hat{a}, \hat{a}^\dagger] = 1$. Equivalently, the impossibility of deterministic noiseless amplification can be recast in the terms of the no-cloning theorem [106]. If quantum mechanics prohibits perfect cloning, it is straightforward to show that it also prohibits noiseless amplification. One could perfectly clone an unknown coherent state by tuning their noiseless amplifier to a gain of $g = \sqrt{2}$, and dividing the output on a 50:50 beamsplitter, affecting the

transformation,

$$|\alpha\rangle|0\rangle \xrightarrow{\text{amplification}} |\sqrt{2}\alpha\rangle|0\rangle \xrightarrow{\text{beam-splitter}} |\alpha\rangle|\alpha\rangle. \tag{5.4}$$

Consistency with quantum mechanics can be restored, however, if one instead forgoes non-determinism in favour of a probabilistic transformation,

$$|\alpha\rangle\langle\alpha| \to P\,|g\alpha\rangle\langle g\alpha| + (1-P)\,|0\rangle\langle 0| \tag{5.5}$$

in which amplification succeeds with probability P, and fails otherwise. This transformation is permitted provided, on average, the distinguishability of the amplified state does not increase.[1] Provided the success is heralded, one may enjoy the benefits of entirely noiseless amplification at least some fraction of the time. As was shown in [120, 122] just such a transformation is performed by the operator $g^{\hat{n}}$ where $\hat{n} = \hat{a}^\dagger\hat{a}$ is the number operator. In the amplification regime ($g > 1$), the operation $g^{\hat{n}}$ is unbounded, and as such could only be implemented exactly with a success probability equal to zero. However, for any particular input state and gain, one can always devise an approximation of $g^{\hat{n}}$ that lies within a suitably truncated Hilbert space and amplifies with a fidelity near-indistinguishable from the perfect NLA. Consider the two-mode squeezed state (or EPR state) written in the number basis as

$$|\chi,\chi\rangle = \sqrt{1-\chi^2}\sum_n \chi^n|n\rangle|n\rangle \tag{5.6}$$

where $\chi \in \{0,1\}$ characterises the entanglement. The application of the amplifier on one mode, when successful, results in the state,

$$g^{\hat{n}}|\chi,\chi\rangle = \sqrt{1-\chi^2}\sum_n (g\chi)^n|n,n\rangle. \tag{5.7}$$

The parameter χ characterising the entanglement scaled by the amplified gain, g. Provided $g > 1$, the entanglement of the two-mode squeezed state is probabilistically increased. For a given Gaussian input state of variance, V, there is a theoretical upper bound on the maximum gain that can be applied, $g_{\max} = \sqrt{(V+1)/(V-1)}$[139][2]. For any pure two-mode squeezed input state $g_{\max}$ corresponds to the gain needed to distill to an 'infinitely squeezed' two-mode squeezed state - a perfect EPR state.

Noiseless Amplification & Loss

Instead consider that one arm of a Gaussian two-mode squeezed state has been distributed through a lossy channel with loss η (Figure 5.1(a)). The application of the NLA on the degraded arm results in an output with a greater degree of initial entanglement that

[1] The condition yields an upper bound on P, though it is unclear if it is tight.

[2] For squeezed states this $g_{\max}$ bound corresponds to amplifying the anti-squeezed variance to infinity, and the squeezed variance to zero

appears to have suffered less loss [120]. The amplified state has an effective entanglement,

$$\chi' = \chi\sqrt{1 + (g^2 - 1)\eta} \tag{5.8}$$

and an effective loss,

$$\eta' = \frac{\eta g^2}{1 + (g^2 - 1)\eta}. \tag{5.9}$$

It is these restorative properties, necessary for the the realisation of a quantum repeater, that provide an immediate practical motivation for a scalable implementation of an NLA. In the presence of only passive loss, application of an NLA with $g = g_{\max}$ will distill to a final finite two-mode squeezed state with a infinite variance. The resulting 'finite squeezing' after distillation corresponds to an infinitely squeezed resource before the lossy channel. Whilst for only passive loss the entanglement improves monotonically with the NLA gain, with the addition of noise to the situation monotonic improvement is no longer guaranteed. Thermal noise is a consequence of entanglement between the system and the environment. As the NLA distills *all* entangled correlations without discrimination, noise contributions can be amplified, and the action of the NLA can actually degrade the entanglement.

We note that although $g^{\hat{n}}$ appears Gaussian in the sense of being quadratic in the annihilation/creation operators, it is in fact non-unitary and unbounded. These properties are also the reason that such an operation falls beyond the purview of the no-go theorem [116, 117, 140] which states that Gaussian entanglement cannot be distilled via purely Gaussian operations. In fact an exact implementation of $g^{\hat{n}}$ would necessitate a success probability of zero. However, when considering a given set of input states one may explicitly construct physical operations which have arbitrarily high fidelity with $g^{\hat{n}}$ while succeeding with a finite probability. The most intuitive version of this method, proposed in[120] and utilised in subsequent experiments, is to use a generalised quantum scissors scheme [141] and truncate in the photon-number basis faithfully amplifying low energy input states that have negligible higher order terms. However these truncated experiments are by no means trivial, with all demonstration limited to the single photon case except for [128] in which two stages were achieved. Thus it would be extremely valuable to devise an easier method of implementing the distillation, albeit for a more restricted set of applications. Here we implement a measurement-based version of this protocol (Figure 5.1(b)) where the original state is first measured using heterodyne detection upon Bob's side. Then a sub-ensemble is post-selected according to a filter function defined by the desired NLA gain.

5.2.2 A Measurement-based Implementation

The exact filter function corresponding to $g^{\hat{n}}$ can be derived following reference [132] by considering the coherent state projection, or the Q-function (§2.3.3), on an arbitrary input state ρ,

$$Q_\rho(\alpha) = \frac{1}{\pi}\langle\alpha|\rho|\alpha\rangle. \tag{5.10}$$

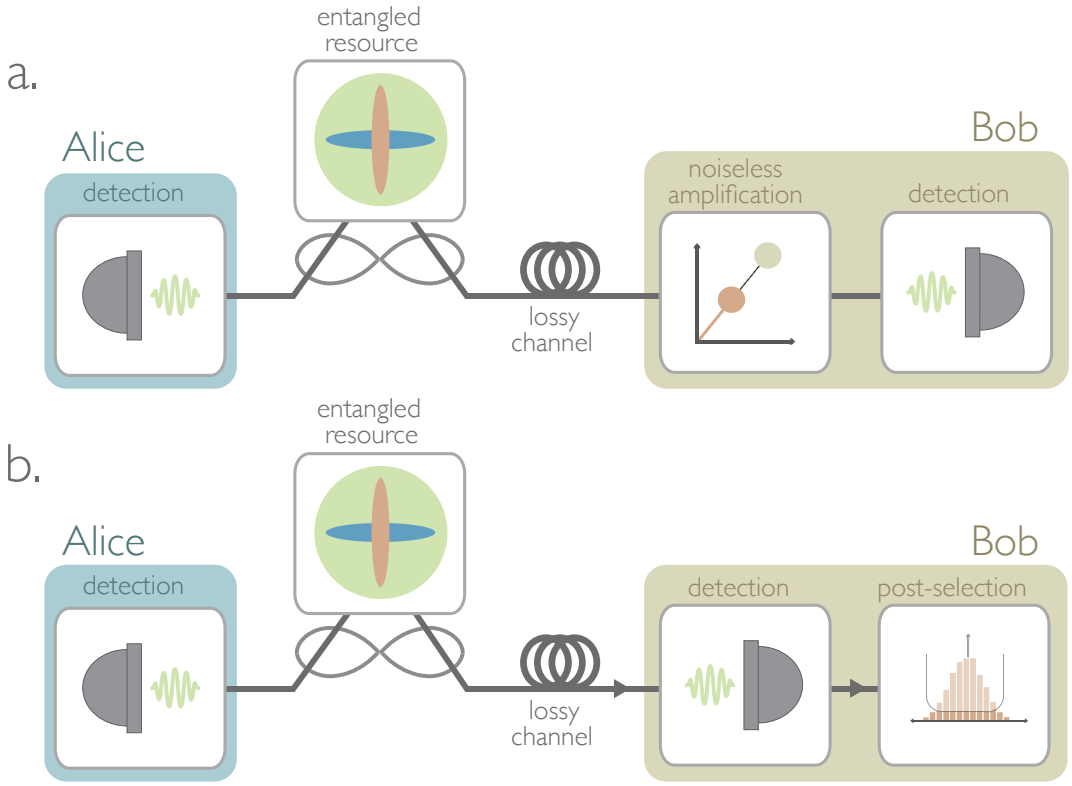

Figure 5.1: Equivalent methods of entanglement distillation with (a) physical and (b) measurement-based noiseless linear amplifiers. Two-mode EPR entanglement is represented by two orthogonally juxtaposed squeezed state. One arm of the EPR entanglement is transmitted through a lossy channel before being noiselessly amplified. In the physical implementation (a) a quantum scissor setup is used to implement the probabilistic amplification before the final measurement. In the measurement-based implementation (b) a post-selective filter is used to keep a remaining fraction of data.

Recalling that the action of the NLA on a coherent state is given by [120],

$$g^{\hat{n}}|\alpha\rangle = e^{\frac{1}{2}(g^2-1)|\alpha|^2}|g\alpha\rangle \tag{5.11}$$

we can write down the Q function of the amplified state ρ',

$$\begin{aligned} Q_{\rho'}(\alpha) &= \langle\alpha|g^{\hat{n}}\rho g^{\hat{n}}|\alpha\rangle \\ &= e^{(g^2-1)|\alpha|^2}\langle g\alpha|\rho|g\alpha\rangle \\ &= e^{(1-1/g^2)|\beta|^2}\langle\beta|\rho|\beta\rangle. \end{aligned} \tag{5.12}$$

where $\beta = g\alpha$. This equation allows us to determine the particular probabilistic filter and rescaling we must apply to the original heterodyne data in order to obtain the same output as a heterodyne measurement applied to the same input state after noiseless amplification with a gain g. Clearly for $g > 1$ the filter defined above does not qualify as a sensible weighting probability as it is always greater than 1. Thus we must renormalise to some cut-off thereby implementing an approximation to the ideal operation. This is analogous to the fact that although the success probability for $g^{\hat{n}}$ has to be zero, one can experimentally achieve a good approximation of an ideal NLA with finite probability. In the measurement-based picture, this corresponds to implementing an approximate operation while keeping a finite fraction of the data after post-selection. In both cases, however, the approximation can be made arbitrarily close to perfect whilst retaining a finite success probability.

The filter function, or acceptance probability, of the $g^{\hat{n}}$ modified post-selection filter with a finite cutoff is given by,

$$P(\alpha) = \begin{cases} e^{\frac{1}{2}(|\alpha|^2-|\alpha_C|^2)(1-g^{-2})}, & \alpha < \alpha_C \\ 1, & \alpha \geq \alpha_C \end{cases} \tag{5.13}$$

where $\alpha = \frac{1}{\sqrt{2}}(x+ip)$ is the coherent state projection for each heterodyne measurement.

Given we have just proposed a 'virtual' implementation of $g^{\hat{n}}$ without extending ourselves beyond our usual Gaussian toolbox, superficially, this work may appear incompatible with the important no-go theorems of [116, 117, 140]. There are a few potentially routes available to resolve this, the most immediate being that our proposal never violates the aforementioned no-go theorem because it never distilled entanglement. As we require a least one of the subsystems be measured, the most we can ever claim is we are distilling correlations between a classical measurement record and an unmeasured quantum state. Of course, the established CV-QKD application (and perhaps others) make no distinction between real distillation and this virtual distillation, because the final success of the protocol only makes reference to the correlations within the classical measurement records. Alternatively, we can look upon the post-selective procedure itself as the requisite source of the non-Gaussianity. Even though our filter inputs are themselves described by Gaussian distributions, and the filter function is itself quadratic in $\hat{a}$ and $\hat{a}^\dagger$, the non-deterministic post-processing of the measurement record could be interpreted as outside the Gaussian toolbox. Nevertheless it is remarkable that, in certain circumstances, we achieve useful results utilising only hardware from the experimentally friendly Gaussian toolbox.

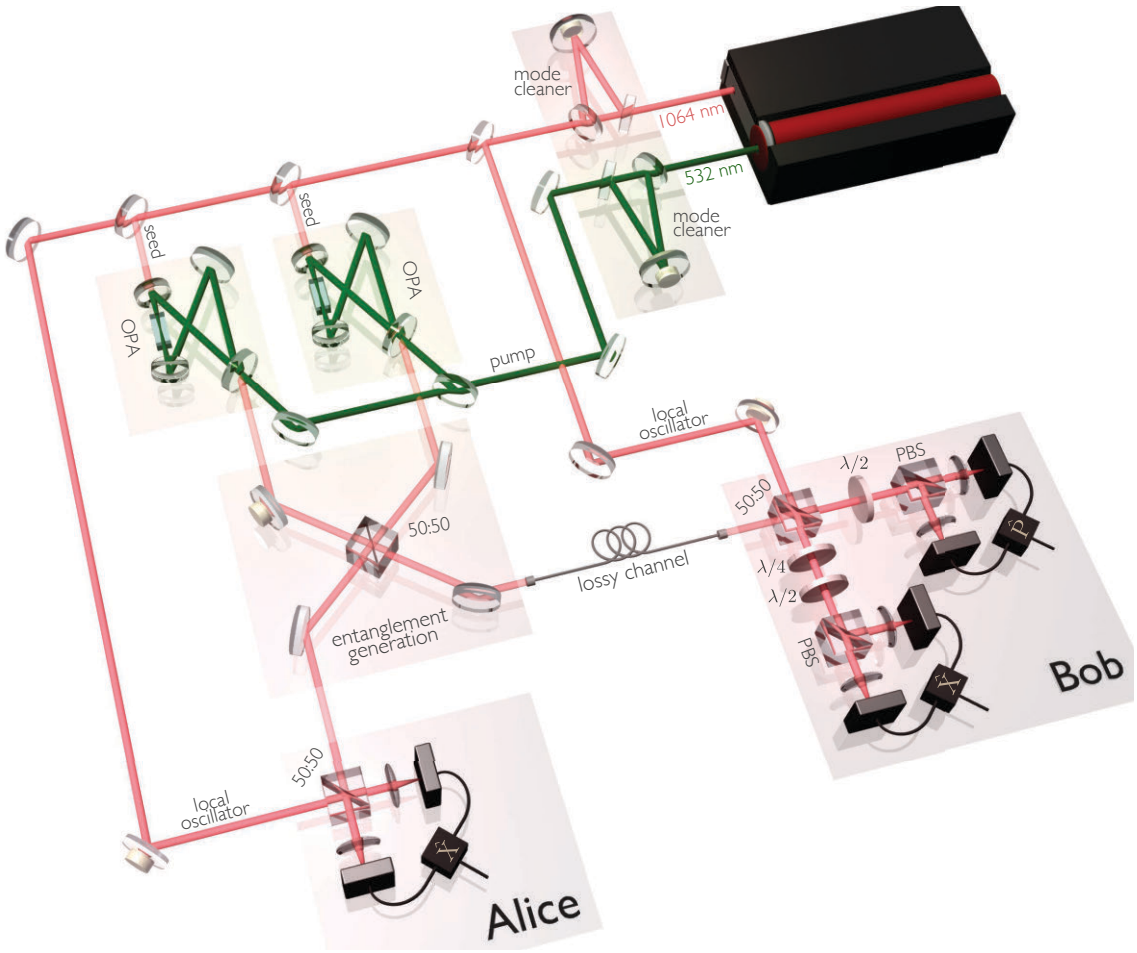

Figure 5.2: Experimental setup of the measurement-based NLA. A laser provides both 1064 nm and 532 nm fields. These fields are spatial and frequency filtered to the quantum noise limit for the sideband detection frequencies between the range of $3 - 4$ MHz. The 1064 nm field is used as the seed and local oscillator fields for two identical degenerate bow-tie optical parametric amplifiers (OPAs), whilst the 532 nm light is used as the pump field. Two amplitude squeezed states are produced and combined on a $50:50$ beam-splitter. With their relative phase locked in quadrature, the beam-splitter produces two-mode EPR state at the output. The entangled beams are sent to Alice locally and through a transmission channel to Bob remotely. Alice performs homodyne detection of her optical states, alternating between conjugate quadratures. Bob on the other hand, performs a heterodyne detection of his state, simultaneously measuring both conjugate quadratures.

5.3 Experiment

Our experimental setup is detailed in Figure 5.2.

5.3.1 Preparation of Seed and Pump Light

The laser source for this experiment was an *Innolight Diablo* Neodymium-doped Yttrium Aluminum Garnet (Nd:YAG) laser producing approximately 400mW of continuous wave single mode light at 1064 nm. The laser head also housed an internal frequency doubler producing approximately 800 mW at 532nm. This laser was identical to the model of laser used for the experiments described in Chapter 4.

As a precaution against optical feedback via unintended backscatter the 1064 nm light was passed through a Faraday isolator. The 1064 nm and 532 nm light was then passed through their respective high-finesse mode cleaners, providing a well-defined TEM-00 spatial mode for the experiment, and additional attenuation of the relaxation oscillation of the laser.

Both the 1064 nm and 532 nm mode cleaners were again of the same design, consisting of a 3-mirror triangular ring resonator, with an optical path length of 800 mm. The 1064 and 532 nm mode-cleaners had respective cavity linewidths of 0.4 MHz and 1.0 MHz. This additional suppression of the remnant relaxation oscillation provides a shot noise limited laser field at frequencies above 4 MHz. Both of the mode cleaners were controlled using Pound-Drever-Hall (PDH) technique using an analog PID system.

5.3.2 Optical Parametric Amplifier

The two OPA's used in this experiment were near identical in design to that of the OPA cavity used in Chapter 4, but were designed to only be resonant for the seed field. The two cavities themselves were physically identical and demonstrated near identical performance.

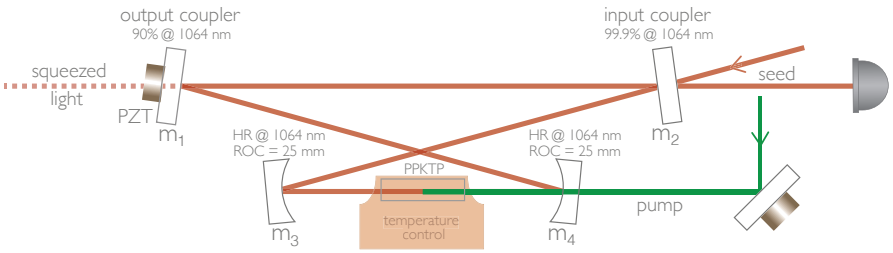

Figure 5.3: Detailed schematic of the optical parametric amplifier cavity.

The two identical OPA cavities were originally constructed by Jiri Janousek. Periodically-poled KTP was again the non-linear material of choice, with a bow-tie cavity resonant at the seed field frequency used to enhance the non-linear interaction. A detailed schematic of the OPA cavity is provided in Figure 5.3. Both travelling-wave cavities comprised of four mirrors in a bow tie geometry: the two inner concave mirrors (m_3 and m_4) with radii of curvature of 25 mm spaced 44 mm apart, and two outer plane mirrors (m_1 and m_2) spaced 90 mm apart. The total round-trip optical path length was 275 mm. The

resulting beam waist of $19\mu m$ centred between the two curved mirrors is almost optimal for the Boyd-Kleinman condition. The output coupler was chosen to be 90% reflective at 1064 nm, producing a cavity linewidth of 19MHz and finesse of 57 for both cavities.

The crystals used were were manufactured by *Raicol* and had identical dimensions of $10 \times 5 \times 1$ mm^3 and poling periods of $\Lambda_p = 9\mu m$. The incident surfaces were polished by *LaserOptik* and the dual-band anti-reflection coating was manufactured by *Advanced Thin Films*. Each crystal was housed within a temperature-stabilised copper oven, with a Peltier element providing temperature control to within $0.1°c$.

In addition to the temperature stabilisation of the non-linear material, two additional active control loops were required for operation of each OPA. The first controlled the cavity length to be resonant with the 1064 nm seed field, and the second stabilised the phase relationship between seed and pump fields, defining direction of the non-linear process and thus, the angle of the squeezed light. A unique phase modulation was introduced onto each of seed fields for the purposes of control; 7.3 MHz for OPA 1 and 16 MHz for OPA 2. Each seed field was coupled into the cavity through the plane input coupler mirror, with an error signal for stabilisation extracted from the detection of the reflected/transmitted light using the PDH technique [96, 97]. This photocurrent also yielded an error signal for the relative phase between the seed and pump light using the method described in §4.1.2.

The pump was initially aligned by production of a small amount of SHG using a bright reverse-propagating seed field. The travelling-wave design of the cavity means the counter-propagating field at the seed frequency will be degenerate with the cavity mode defined by the seed field. This feature is sometimes used for control, and also to provide two squeezed modes from a single OPA. Here, however, it allows generation of a counter-propagating field at the pump frequency that allows us to mode-match the mode of the up-converted light to that of the 532 nm mode cleaner. Examining the SHG conversion efficiency as a function of temperature also provided the optimal crystal temperature. As the periodic polling substantially broadens the effective temperature range over which the seed and pump are phase-matched, and the precision of the crystal temperature was not critical.

5.3.3 Entanglement Generation

The next experimental task was the generation of the two-mode squeezed state or Einstein-Podolsky-Rosen (EPR) state described in §2.2.5. This required we interfere our two amplitude squeezed coherent states on a 50:50 beamsplitter and control the relative phase between the two fields to be $\pi/2$. The stabilisation of a phase between the two squeezed fields is critical to the quality of the produced entanglement. If the two squeezed states are interfered perfectly in quadrature, the resulting EPR criterion (§2.4.5) is somewhat impervious to purity of the component squeezed states, and largely determined by the individual squeezed quadrature variances. However, even small rotations will couple the classical noise associated with the anti-squeezed quadratures into the measured conditional variance, degrading the entanglement.

To control the relative phase between the two squeezed fields we use a familiar DC locking technique. As described, the two amplitude squeezed fields are interfered on a 50:50 NPBS. Using a pellicle AR coated for 1064 nm approximately 1% of each output field is tapped off and detected. The two resulting photocurrents are subtracted to compensate for

their individual DC offsets, and ideally produce a signal insensitive to global fluctuations in the laser power. The result is a sinusoidal signal with a zero-crossing corresponding to $\theta = \frac{\pi}{2}$. The correlation (and anti-correlation) between the two homodyne detection stages was used to further optimise the DC offset of the error signal. Though there are other approaches to this lock, we considered it worthwhile to concede the small amount of additional loss to access the stability this technique affords.

5.3.4 Measurement

Bob's Measurement

Given the operation Bob chooses to implement in post-processing, to do so noiselessly he needs to isolate the compatible POVM. To noiselessly implement $g^{\hat{n}}$, Bob must measure in the coherent state basis, corresponding to a heterodyne detection (or equivalently, a simultaneous homodyne measurement). To do so, we use the same technique described in detail in §4.2. Bob's mode b (vertically polarised) is combined on a 50:50 NPBS with a bright LO (horizontally polarised). The polarisation of the first output mode is then rotated by half-wave plate orientated at 45°, before a PBS that divides the mode between two photodetectors, the subtraction of these two photocurrents providing one homodyne measurement. The second output path additionally encounters a quarter-wave plate orientated at 45°, producing the desired $\pi/2$ phase shift between the signal and local oscillator.

As this technique passively stabilises the phase relationship between Bob's two homodyne detectors, we only require one active control loop to define the correct phase relationship between Bob's mode and the local oscillator. The two different phase modulations (at 7.3 MHz and 16 MHz) used to stabilise the two OPA cavities sit within the cavity linewidth and remain on the two squeezed modes used to generate the entangled state. Ideally, interfering the two squeezed modes in quadrature produces two outputs,

$$\hat{a}_{\pm}(t) = \alpha e^{i\omega t}(1 \pm i + i\xi_1 \cos\omega_1 t + \xi_2 \cos\omega_2 t) \tag{5.14}$$

rotating one of the phase sidebands to the amplitude quadrature. Demodulating our measured homodyne photocurrent at *either* 7.3 MHz or 16 MHz will yield an error signal that (given the passive stabilisation between the two homodyne detections) will allow us to simultaneously sample the phase and amplitude quadrature.

Alice's Measurement

Alice's measurement stage comprised of a single homodyne detection which alternated between sampling the phase and amplitude quadrature. With Bob's dual-homodyne measurement, Alice's measurement of X and P allowed characterisation of the covariance matrix describing the two-mode system. An error signal to stabilise the homodyne detection to *either* the phase or amplitude quadrature was extracted by demodulating at the homodyne photocurrent at either 7.3 MHz or 16 MHz. The quadrature of the relevant modulation was determined by the phase at the entangling beam-splitter.

Detection and Losses

All three homodyne stages both used variations of the Uni-PD circuits described in §4.2 with a combination of the *Epitaxx ETX-500* and the custom-fabricated *Laser Components* InGaAs photodiodes. Additional retro-reflection of the light scattered from the *ETX-500* diode surface produced similar efficiencies to that of the *Laser Components* photodiodes (with a specified quantum efficiency 99%), suggesting the reduced quantum efficiency of the *ETX-500* is largely consequence scattering from the surface of the material, and not the result of electron-hole recombination. The overall detection efficiency of Alice's homodyne measurement was $\sim 98\%$, with a measured fringe visibility of 99.5%. Bob's measurement stage suffered from an additional reduction in the overall detection efficiency due to the requirement for additional surfaces in the optical path, with the optical elements also introducing polarisation mis-match.

5.3.5 Experiment Control & Measurement Acquisition

In its completed state, the experiment required 9 active control loops (not including temperature control). The first four control loops (controlling the 1064 nm and 532 nm mode-cleaners and the cavity length of the two OPAs) used analog Proportional-Integrator (PI) servos manufactured in house. The error signals for the remaining five control loops were extracted via analog demodulation, with the PI control implemented digitally in *Labview* using an algorithm developed by Seiji Armstrong [142]. The acquisition utilised an 8-channel digitiser (*National Instruments* PXI-5105) with a sampling rate of 60 MS/s and 12-bit vertical resolution. .

5.3.6 Filter Implementation

As previously discussed, $g^{\hat{n}}$ is unbounded for $g > 1$, and therefore, whether physically or virtually, it cannot be implemented exactly. However, the virtual implementation, $g^{\hat{n}}$ can be emulated to arbitrary precision by truncating our original post-selection filter at an appropriate amplitude, α_C. The resulting modified post-selection filter is given by

$$P(\alpha) = \begin{cases} e^{\frac{1}{2}(|\alpha|^2 - |\alpha_C|^2)(1-g^{-2})}, & \alpha < \alpha_C \\ 1, & \alpha \geq \alpha_C \end{cases} \tag{5.15}$$

where, α is obtained from the measured heterodyne outcome via $\alpha = \frac{1}{\sqrt{2}}(x + ip)$. Any measurement outcomes falling beyond the cutoff amplitude, α_C, are kept with unit probability.

The procedure for implementing the measurement-based implementation of $g^{\hat{n}}$ is as follows: our experimentally prepared two mode squeezed state is split between two measurement stations, which we identify as Alice (mode a) and Bob (mode b). At a time we denote t_i, Bob performs a heterodyne detection of his mode, obtaining two outcomes x_i and p_i. Bob's measurement-based implementation of $\hat{g}^n$ then amounts to him keeping or rejecting his obtained measurement outcome, α_i (from his heterodyne measurement outcomes, x_i and p_i), with a probability specified by Equation (5.15). If his outcome α_i falls beyond the cutoff α_C, he always keeps it. If we then consider a two-mode scenario

(virtual distillation), Bob also informs Alice to either keep or reject her state - measured or unmeasured. [132]

The Filter cut-off

Care needs to be taken with the choice of α_C to ensure that the truncated approximation emulates $g^{\hat{n}}$ with high fidelity. As suggested by Fiurasek and Cerf [132], indistinguishable fidelity with the $g^{\hat{n}}$ can always be obtained by pushing α_C out to the largest measurement outcome encountered in the measurement ensemble. This approach, however, scales very poorly with increasing ensemble size [132]. A compromise between excellent emulation of $g^{\hat{n}}$ and the post-selection probability can be achieved by implementing an α_C sufficiently large that the purity ($\frac{1}{\sqrt{\det(\sigma)}}$) of the post-selected state decreases consistent with an ideal implementation of $g^{\hat{n}}$. Here, we choose a finite cutoff α_C that optimises post-selection rates whilst preserving high fidelity with the ideal filter. This also ensures that output distributions remain statistically close to a normal distribution, allowing us to only consider the second order moments of the measured distribution to characterise the correlations. There are a few quite subtle points that arise in the implementation of the 'truncation',

As an ideal implementation of $g^{\hat{n}}$ results in a Gaussian mapping of the input state, by here ensuring that our emulation is consistent (within statistical error) with the theoretical mapping of our characterised input state, any non-Gaussianity is necessarily negligible. The Gaussianity of our amplified 'state' itself can be roughly verified by examining the size of the third (skewness) and fourth (kurtosis) order moments. Owing to the symmetry of the filter it should not introduce skewness, but implementation of the filter without a sufficiently large cutoff may introduce a "peakedness" that might be quantified via the fourth moments. One could also consider a more sophisticated approach, like a JarqueBera test.

If we characterise the original state correctly, and accordingly choose α_C sufficiently large to ensure that output distributions remain (close to) normally distributed, we only need consider the first and second order moments of the 'amplified' distribution to characterise the correlations. Consider though, that one erroneously sets the cut-off too small, such that the post-selection produces a 'non-Gaussian' distribution. Given our entanglement witnesses only consider the first and second moments of the 'amplified' distribution it is not immediately clear what outcome setting the cutoff too low has for the distillation operation. One could conceive the resulting non-Gaussianity only punishes your relevant entanglement witnesses? Or the witnesses may be insensitive to it? In reality, setting α_C too small - while offering no real improvement in the final shared correlations - can give the impression of improved distillation for certain witnesses. This is best elucidated by considering what the post selection is physically doing. Each post-selection filter with a gain, $g > 1$ extracts the statistics of a larger two-mode squeezed state from the original measurement ensemble. As the amplifier gain grows, so too does the variance of the desired post-selected state, and accordingly, the relevance of measurement outcomes at the extremities. And as the size of the post-selected state expands, the cut-off needs to also expand to accommodate it. If not, the filter essentially does not act over the entire phase space, and the cut-off has a large effect on the statistics of the post-selected ensemble.

Simply, the NLA ceases to distill. As such, the size of the ensemble usually sets the practical limit on the amount of distillation that can be achieved, as the ensemble itself is usually exhausted well before one reaches the theoretical maximum gain set by the energy of the state [143].

However, even in the afore described situation where additional distillation has ceased being useful, a different issue emerges from setting α_C too small. Whilst the ideal $g^{\hat{n}}$ operation will always decrease in purity of the amplified state, a 'prematurely truncated' post-selection filter can distort the noise floor of Bob's post-selected state. Consequently, his approximated amplifier no longer preserves the vacuum and entanglement witnesses that reference a noise floor are distorted, *i.e.* what we observe as a decrease in the covariance is actually an artificial decrease in Bob's variance. These problems are easily negated by ensuring the purity ($\frac{1}{\sqrt{\det(\sigma)}}$) of the post-selected state decreases consistent with an ideal implementation of $g^{\hat{n}}$.

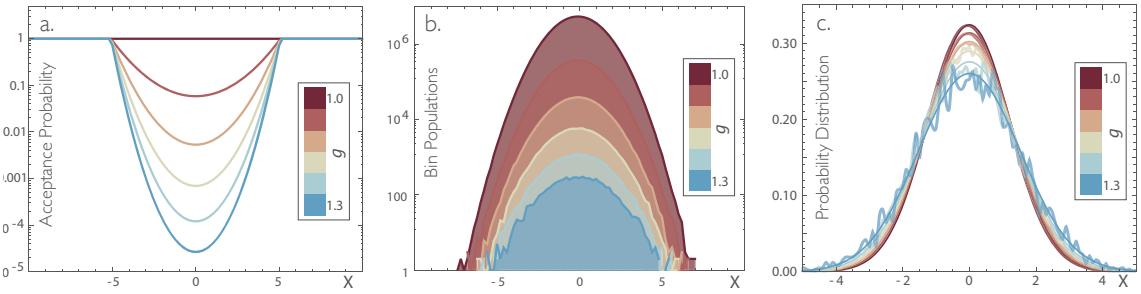

Figure 5.4: Measurement-based NLA performed on the receiver, Bob's, experimental data: (a). Acceptance probability function of the post-selective filters used to obtained, (b). the resulting measurement histograms, and (c). the final normalised probability distributions. The gain, g, is increased by selecting a filter function with increasingly lower acceptance probability. As the gain is increased, the variance of Bob's final distribution increases. This corresponds to a larger, more entangled two-mode squeezed state.

It is also important to note that here that for a given input state, all emulations of $g^{\hat{n}}$ for varying g use the same cut-off α_C. We restrict ourselves to an α_C sufficiently big to accommodate the largest gain applied. One could potentially improve on our presented post-selection rates by considering a gain dependence in the choice of α_C.

For this proof-of-principle demonstration we first characterised the two-mode squeezed state shared between Alice and Bob by reconstructing the covariance matrix given our measurements. Bob's measured variance establishes the size of the cutoff α_C needed to implement $g^{\hat{n}}$. Here, given our ensemble size of 8×10^7, the cut-off of 4.5 standard deviations of Bob's measured state, balanced post selection rates whilst preserving fidelity with the ideal $g^{\hat{n}}$ operation.

5.4 Results & Discussion

Distillation Performance

We begin the results with the simplest scenario: Alice and Bob share a symmetric two-mode EPR state, with no transmission loss between either of their measurement stations. In this scenario, a MB-NLA on Bob's side is indistinguishable from an implementation on Alice's side, and the observed distillation should be symmetric for both parties. Our results use the customary CV entanglement witnesses - the EPR criterion [42] introduced in §2.4.5 and the inseparability criterion introduced in §2.4.3. As the EPR criterion is an inherently directional quantity and we borrow terminology from QKD, and refer to Bob's ability to infer Alice's state as the *direct* inference ($\mathcal{E}_{B\blacktriangleright A} = V_{x_B|x_A} V_{p_B|p_A}$), and the converse as the *reverse* inference ($\mathcal{E}_{A\blacktriangleright B} = V_{x_A|x_B} V_{p_A|p_B}$). Either $\mathcal{E}_{A\blacktriangleright B} < 1$ or $\mathcal{E}_{A\blacktriangleright B} < 1$ is a sufficient but not necessary condition for entanglement. We denote the symmetric inseparability criterion by $\mathcal{I}_{A\blacklozenge B}$. For Gaussian states $\mathcal{I}_{A\blacklozenge B} < 1$ is both a necessary and sufficient condition for entanglement.

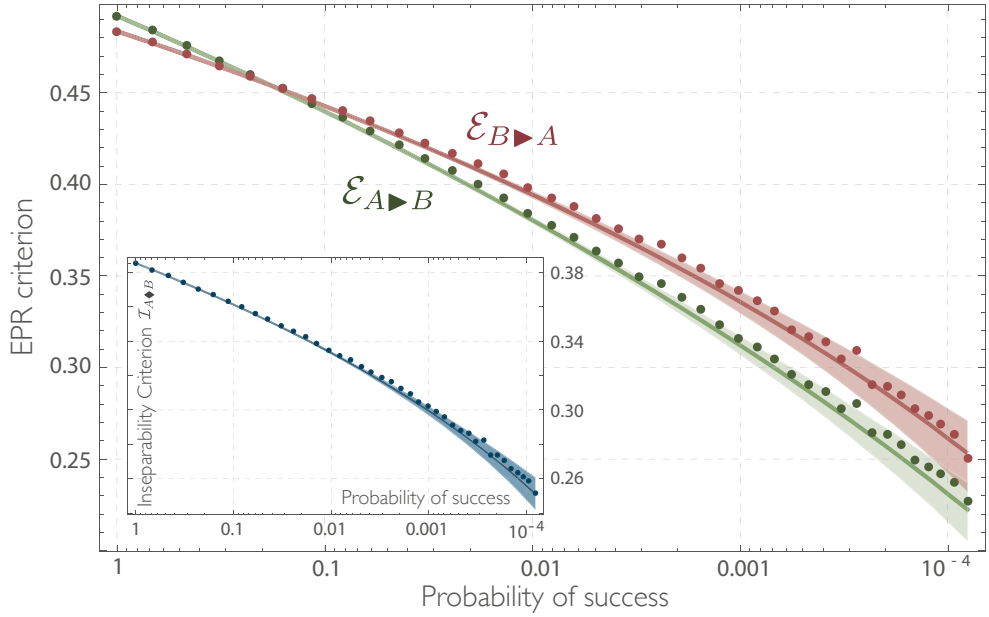

Figure 5.5: EPR criterion as a function of post-selection success probability for the direct ($\mathcal{E}_{A\blacktriangleright B}$, red) and reverse ($\mathcal{E}_{B\blacktriangleright A}$, green) inferences for input state with an initial EPR entanglement strengths of 0.484±0.001 and 0.492±0.001 respectively. Data points presented are the post-selected ensemble average of 10 experimental runs. The solid lines shows the theoretical distillation of an ideal implementation of $g^{\hat{n}}$ given the same input state. Shading represents a 2σ confidence interval on the variance of the implemented filter. Inset shows the effect of the distillation on the inseparability criterion, $\mathcal{I}_{A\blacklozenge B}$, with the same data set.

Our initial entangled resource demonstrates an EPR criterion violation of $\mathcal{E}_{A\blacktriangleright B} = 0.484 \pm 0.001$ and $\mathcal{E}_{B\blacktriangleright A} = 0.492 \pm 0.001$ with an initial ensemble size of 8×10^7 data points. We then apply the post-selection function of (5.15) followed with the required linear rescaling of Bob's post-selected measurement record by $1/g$. A linear increase in

the amplifier gain, g sees an exponential reduction in the probability of success, yielding a smaller, but more correlated subset of the original measurement record. This is equivalent to Bob and Alice sharing a larger initial two-mode squeezed state. Figure 5.5 demonstrates our improvement in the EPR criterion as a function of the success probability. The solid line indicates the behaviour of an ideal implementation of $g^{\hat{n}}$ with the same input state [143], and the shaded area gives a 2σ confidence interval on the theoretical EPR violation. For a post-selection probability of 8×10^{-5} we obtain effective EPR criterions of $\mathcal{E}_{A\blacktriangleright B} = 0.25 \pm 0.02$ and $\mathcal{E}_{B\blacktriangleright A} = 0.23 \pm 0.02$. The asymmetry in the EPR criterions for the direct (green) and reverse (red) inferences arise from variations in the purity of the two-subsystems; Bob's heterodyne measurement introducing additional loss. Figure 5.5 also plots the inseparability criterion as a function of the success probability. We find excellent agreement between theory and experiment.

The declining probability of success as we apply increasingly larger gain to obtain stronger correlations manifests in increased statistical uncertainty. To understand the statistical error we first consider the theoretical probability of success for our truncated approximation of $g^{\hat{n}}$. Given the initial state (characterised by its measured covariance matrix), the size of the initial ensemble, and the choice of cut-off α_C, the theoretical probability of success allows us to ascertain the statistical error associated with the post-selected covariance matrix. We calculate a 2σ confidence interval on the theoretical performance of the NLA, reflecting the role of finite sample size. The 2σ associated with EPR criterion will be larger than that of the inseparability criterion; a consequence of its dependence on the product of the two conditional variances, rather than their sum.

Distillation & Loss

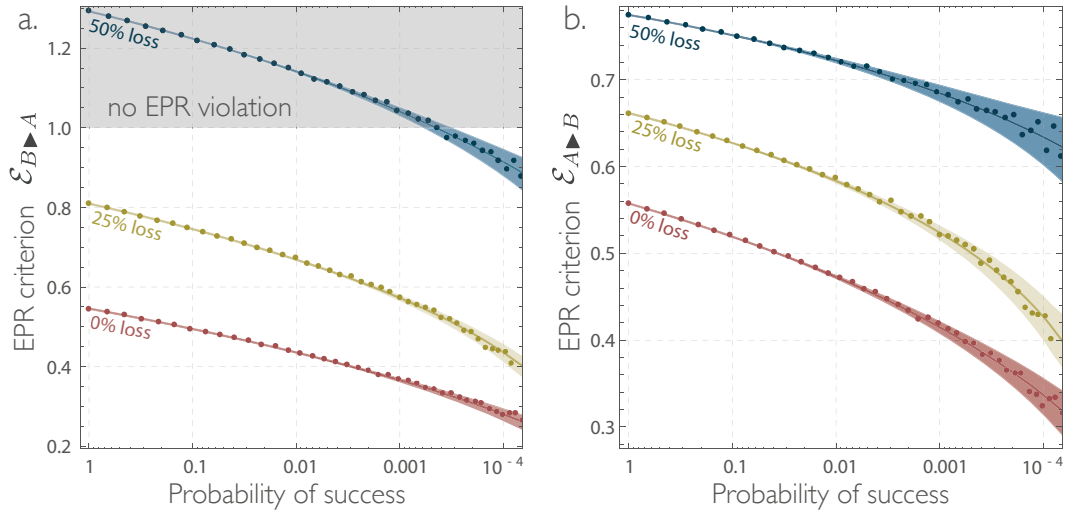

Figure 5.6: Effect of EPR entanglement distillation as a function of probability of success for different losses (0%, 25% and 50%).

Most real world applications of a physical NLA employ it to combat the effects of loss. Here we examine the performance of the MB-NLA in two different loss regimes: that of moderate loss, and that of very high loss. We experimentally introduce loss on Bob's

subsystem, allowing us to model several lossy channels. Bob subsequently implements a MB-NLA. Figure 5.6 demonstrates the performance of the post-selective NLA for a two-mode EPR state with moderate loss on one subsystem. Figure 5.6(a) plots the EPR criterion for the direct inference as a function of post-selection probability for a series of channel transmissions, whilst the reverse inference is plotted in Figure 5.6(b). For 25% loss on Bob's channel, we find that post-selection allows us to fully compensate for any loss incurred by the quantum state and demonstrate a final EPR correlation well beyond that of the original state. Despite the initial asymmetry of their subsystems, sufficient post-selection (a probability of success of $\sim 10^{-4}$) allows Bob to obtain the same EPR violation as Alice.

Even in the limit of a maximally entangled (infinite energy) two-mode squeezed state, Bob's EPR correlations will not survive a perfect passive loss channel of 3dB (50%). And the situation is only worse for a real channel, where excess noise also contributes. For 3dB of passive loss, post-selection allows Bob to recover an EPR criterion < 1 from initial state with a (non-violating) EPR criterion of $\mathcal{E}_{B\blacktriangleright A} \approx 1.3$. Recent works have demonstrated the direct equivalence of EPR violation and the concept of 'steering' for discrete and continuous variable systems[144, 145, 146, 147, 148]. Whilst the concept of steering provides strong operational meaning for EPR correlations, it is also the resource of interest for *semi-device-independent* protocols[149]. Though subtleties likely arise when combining post-selection with the relaxation of the honesty assumption for the parties in semi-device independent communication, it may prove a fruitful research avenue.

Perhaps the most interesting regime for the performance of the MB-NLA occurs at very high loss. In Figure 5.7 we plot the inseparability criterion of the two-mode EPR state as a function of the channel transmission encountered by Bob's subsystem. We consider four different high loss channels (implemented via a $\lambda/2$ and a beam-splitter) with a maximum loss of 99%, equivalent to 100 km of optical fibre.[3] For each channel, we consider examine the improvement that the MB-NLA affords, with a typical maximum gain of $g = 1.6$. Without distillation, the best inseparability one can hope to obtain is is given by the boundary of the shaded area, describing the theoretical inseparability in the theoretical limit of a perfect (infinitely squeezed) EPR state subject to the same channel transmissivity. We find that post-selection allows access to final correlation that – without distillation – proves inaccessible even in the limit of a perfect initial EPR resource.

Entanglement-Based Quantum Key Distribution

Finally, we turn to the application that sparked interest in this protocol, and investigate the performance of entanglement-based CV-QKD protocol that includes an MB-NLA. For the sake of brevity, CV-QKD is not introduced in detail in this thesis. The theses of Raúl García-Patrón [150] and the recent review of Cerf and Grangier [151] provide an excellent introduction to CV-QKD.

The works of [132, 133] provide security proofs for CV-QKD protocols using Gaussian post-selection for arbitrary attacks in the asymptotic limit of large key lengths. These two results essentially emerged from an extension of two previous results: firstly, under collec-

[3] Assuming the usual loss of 0.02 dB per kilometre of optical fibre.

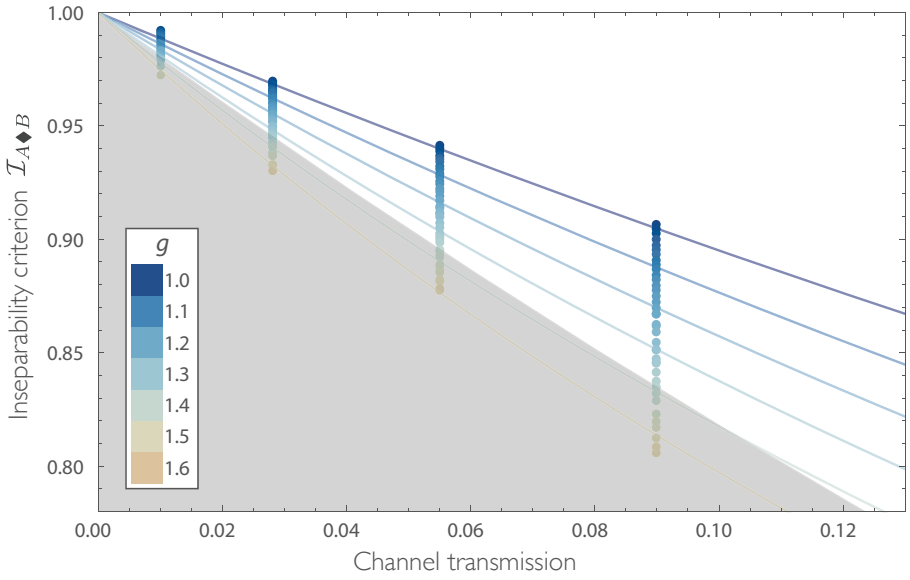

Figure 5.7: Improvement in the inseparability criterion of the two-mode EPR state for a series of lossy channels. For each transmissivity, a series of post-selection corresponding to an NLA gain (specified by the legend) are applied. The boundary of the shaded area describes the theoretical inseparability of a perfect EPR state - infinitely squeezed - subject to the same channel transmissivity. Post-selection allows access to an entangled resource beyond that accessible with even perfect initial resource. The solid lines represent theoretical inseparability of our input state, with an applied post-selection filter of a defined gain ($g = 1, 1.1, \ldots 1.5$) as a function of the channel transmission.

tive attacks[4] the key rate is minimised by assuming the final state is Gaussian[152, 153], and secondly, and secondly, collective attacks are optimal in the asymptotic limit[5][154]. Both conditions hold even if the CV-QKD protocol is not perfectly Gaussian, noting the post-selection procedure could introduce some non-Gaussianity. It is only that the bounds would become very pessimistic were the post-selection to exhibit a strong non-Gaussianity, but as explained above, this is not the case here. This level of analysis is the same as that employed in several CV-QKD experiments [155, 156], including the only previous demonstration of CV-QKD using entangled states [157], but does not comprise all of the finite-size effects [158] and reconciliation and privacy amplification processes of a state of the art CV-QKD demonstration such as [159]. Nevertheless it is sufficient to demonstrate the benefit of the MB-NLA for key distribution.

We conduct a very cautious analysis in which *all* measured imperfections are attributed to the eavesdropper, such that our EPR source is interpreted as a pure EPR source followed by a decohering channel. There are some subtleties involved with using genuine entangled states, as opposed to most CV-QKD experiments, where a theoretical equivalence is established between a prepare and measure scheme (P&M) and one involving real entanglement [160]. P&M schemes can prepare states of near-perfect purity, with a corresponding virtual entangled state considerably purer than those currently feasible experimentally. This can be mitigated by an additional step where the decoherence in the source production is characterised and the purification of that noise is not attributed to the eavesdropper. This method was successfully employed by [157] who showed that the EPR based scheme actually showed improved robustness to channel noise in comparison to coherent state P&M protocols[161]. An extension of this proof to include the MB-NLA would mitigate the effects of impurities within our initial two-mode squeezed state. Regardless, we still demonstrate the value of the MB-NLA while attributing all observed impurities to the eavesdropper.

Our measured covariance matrices are interpreted as coming from a pure EPR source that has been transmitted through a lossy channel with thermal noise. The effective channel has a relatively low loss but high additional noise. In this situation the optimal protocol would be direct reconciliation (DR)[6] with heterodyne detection on both Alice and Bob's side.

The secret key rate for this protocol is given by [152, 153],

$$K^{\blacktriangleright} = \beta I(A:B) - S(A:E) \tag{5.16}$$

where $I(A:B)$ is the classical mutual information between quadrature measurements made by Alice and Bob, $S(A:E)$ is the Holevo quantity between Eve and Alice and $\beta \in [0,1]$ is the reconciliation protocol efficiency. Here we choose an optimistic value of $\beta = .98$, consistent with [157]. Given that the post-selected measurements are still very

[4] *Collective attacks* assume that the eavesdropper acts independently on the quantum systems shared between the two honest parties, Alice and Bob, at each round of the protocol. The eavesdropper can then measure her systems collectively to maximise her information. This is stronger than the *individual attack* strategy, where the eavesdropper attacks *and* measures individually with each round of the protocol.

[5] The *asymptotic limit* is the limit of infinite key length.

[6] In *direct reconciliation* Bob attempts to infer Alice's measurement results. In the converse, *reverse reconciliation*, Alice attempts to guess Bob's measurement results.

Gaussian, Alice and Bob's can be calculated using the formula,

$$I(A:B) = \log \frac{V_A + 1}{V_{A|B} + 1} \tag{5.17}$$

where V_A is the measured homodyne variance on Alice's side and $V_{A|B}$ is the conditional variance of Alice's measurement given Bob's heterodyne detection. Eve's mutual information is given by [152, 153],

$$S(A:E) = S(AB) - S(B|A) \tag{5.18}$$

where $S(AB)$ and $S(B|A)$ are the von Neumann entropies of the inferred state ρ_{AB} and the conditional state following a heterodyne detection by Alice. In general, the von Neumann entropies are bounded by those of a Gaussian state with the same covariance matrix (CM). The Gaussian von Neumann entropy can be calculated straightforwardly from the symplectic eigenvalues of the CM [115], and thus our key rate can be determined directly from the unconditional and conditional CM's of our system. Figure 5.8 plots the secret key rate as a plotted as a function of the post-selection rate. Although we being in an insecure regime, application of an NLA of sufficiently high gain allows us to extract a secure key. As noted in [133, 162] the NLA acts to improve the transmission of the effective channel while actually increasing the noise, but in such a way as to create an information advantage between Alice and Bob. As explained in [143] this can be seen as the amplifier distilling both the Alice-Eve and the Alice-Bob entanglement. Figure 5.8 only considers the effect on the key rate of the post-selected ensemble. When considering the overall key rate, the maximum gain is unlikely to be the optimal gain; rather, the optimal gain recovers a secure key while balancing post-selection rates. The large error bars associated with the experimental results of Figure 5.8 are a consequence of small sample sizes used for our parameter estimation.

These results also have ramifications for applications of the NLA in metrology. Early theoretical and experimental implementations have speculated on the potential usefulness of the NLA for metrological applications. These initial demonstrations had very pessimistic success rates for their $g^{\hat{n}}$ implementations, largely owing to experimental technicalities. Even still, success probabilities for $g^{\hat{n}}$ were not concrete, and it was unclear what limits quantum mechanics actually set. The probabilistic nature of any implementation of $g^{\hat{n}}$ must be carefully accounted for when considering metrological applications, as the value of noiseless amplification must be balanced against the reduced sample size of the measurement[163]. For example, even the success probabilities achieved here would render the operation unsuitable for metrology protocols that scale with the square root of the number of measurements or worse (*i.e.* most of them). By contrast, in many quantum communication protocols it is the final quality of correlations that is of paramount importance.

5.5 Summary

The primary significance of this work is two-fold. Firstly, we experimentally demonstrated the equivalence of the MB-NLA to the implementation of a physical NLA for entangle-

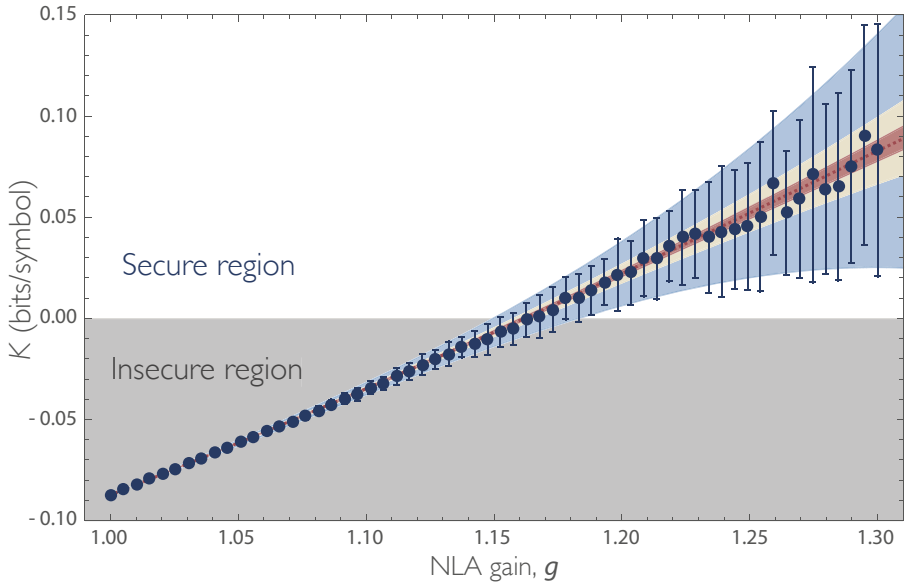

Figure 5.8: Application of MB-NLA to extract positive key rate from otherwise insecure regime in CV-QKD system. Secret key rate as a function of the gain for a direct reconciliation CV-QKD protocol where both parties use heterodyne detection. The application of the MB-NLA allows the recovery of secure key distribution from an initially insecure situation. The dashed line represents the theoretical key rate given the initial state. Error bars and the blue shaded region represent the experimental and theoretical 1σ statistical confidence interval for our initial sample of 8.3×10^7 points, respectively. A larger sample size of 10^9 and 10^{10} would reduce our 1σ confidence interval to the beige and red shaded areas, respectively.

ment distillation when considering scenarios where amplification is directly proceeded by measurement. This equivalence ensures this technique has immediate applications for CV-QKD, where the advantage of an NLA has already been studied[162, 132, 133]. Furthermore, it provides a generalised theoretical explanation of the conditions in which an arbitrary quantum operation could, in principle, be implemented upon an ensemble via post-selective measurements. Secondly, this equivalence is of practical relevance, as the MB-NLA is significantly less demanding than the existing physical implementations of $g^{\hat{n}}$, where inefficiencies in sources and measurement restrict the physical NLA to very small input states[120, 123, 124, 128]. In contrast, the 'software' nature of the MB-NLA, while restricting its applicability, ensures that for compatible applications its performance is superior to its physical predecessor. The inherent flexibility of the software implementation means that the MB-NLA can be used on a wide variety of input states without experimental reconfiguration. By circumventing the requirement for experimental hardware and its a accompanying inefficiencies, it achieves near optimal success probability for an implementation of $g^{\hat{n}}$ of arbitrary precision. Whilst there are clear restrictions on the scenarios where this MB-NLA can be substituted for its physical counterpart, when applicable, it is certainly advantageous to do so. The achievable entanglement distillation is now chiefly limited by the amount of data collected. Here, for feasible sample sizes, we demonstrate distillation of correlations in excellent agreement with close to the theoretical

ideal performance of $g^{\hat{n}}$. For moderate loss channels we showed the recovery of EPR correlations from an entangled state, and applied to high loss channels demonstrated levels of entanglement that are impossible without a distillation process.

Many avenues for further research remain. Beyond the aforementioned applications in CV-QKD, the NLA could find use in other quantum communication protocols including teleportation and remote state preparation. This would be of particular interest as it would enable us to extend these conditioning distillation techniques to improve the quality of a still propagating, albeit unentangled, quantum mode. Furthermore our theory is sufficiently general to allow extensions to other conditional processes. For example using precisely the same setup described here it is also possible to implement the photon addition operation which has been extensively studied [122, 164, 165, 166, 167]. As well as targeting other operations one could also use this formalism to consider conditioning on different POVM sets, opening up many promising candidates for future applications.

Chapter 6

An Operational Interpretation of Quantum Discord

6.1 Introduction

Correlations lie at the heart of our capacity to manipulate information. The fewer the constraints on the correlations we can exploit, the greater our capacity to manipulate information in ways we desire. The rapid development of quantum information science is a testament to this observation. Quantum systems may be so correlated that they are 'entangled', such that each of their subsystems possesses no local reality. Exploitation of such uniquely quantum correlations has led to many remarkable protocols that would otherwise be either impossible or infeasible [168, 14, 169, 15].

However, the absence of entanglement does not eliminate all signatures of quantum behaviour. The most mature quantum information protocol, quantum cryptography, exploits quantum mechanics to provide unconditional security without needing to invoke entanglement [170]. Coherent quantum interactions (i.e., quantum two-body operations) between separable systems that result in negligible entanglement could still lead to exponential speed-ups in computation [171, 172, 173, 174] or the extraction of otherwise inaccessible information [175]. The potential presence of discord [176, 177] within such protocols motivated speculation that discord could prove a better quantifier of the 'quantum resource' that coherent interactions exploit to deliver a 'quantum advantage' [178, 173, 179]. Discord has thus captured a great deal of attention, as evidenced by studies of its role in open dynamics [180], cloning of correlations [181, 182], scaling laws in many-body physics [183], and quantum correlations within continuous variable systems [184, 185].

Here, we demonstrate that under certain measurement constraints, discord between bipartite systems can be consumed to encode information that can only be accessed by coherent quantum interactions. The inability to access this information by any other means allows us to use discord to directly quantify this quantum advantage. We experimentally encode information within the discordant correlations of two separable Gaussian states. The amount of extra information recovered by coherent interaction is quantified and directly linked with the discord consumed during encoding. No entanglement exists at any point of this experiment. Thus we introduce and demonstrate an operational method to use discord as a physical resource.

6.2 Quantum Discord

Early on (§2.7.3), I made reference to subtleties that arise when one considers the quantum generalisations of the classical mutual information. I promised this would be discussed in detail later and that time has now come. But first I should direct readers to the existence of a very comprehensive review article on quantum discord and other similar proposed measures of quantum correlations [186].

Quantum discord emerged in 2001 though the work of Harold Ollivier and Wojciech H. Zurek [176] and, independently, L. Henderson and Vlatko Vedral [177]. In (§2.7.3) we introduced three equivalent expressions for the quantum mutual information of a bipartite system, ρ_{AB}

$$\begin{aligned}\mathcal{I}(\rho_{AB}) &= S(\rho_A) + S(\rho_B) - S(\rho_{AB}) && (6.1)\\ &= S(\rho_A) - S(A|B) && (6.2)\\ &= S(\rho_B) - S(B|A).\end{aligned}$$

The quantum mutual information, $\mathcal{I}(\rho_{AB})$, provides a measure of all the correlations within a bipartite system, whether classical or quantum in nature. Consider the quantum conditional entropies, $S(A|B)$ and $S(B|A)$. One can quickly verify using (6.1) that via state tomography one can infer the quantity $S(A|B)$ perfectly,

$$S(A|B) = S(\rho_{A,B}) - S(\rho_B). \tag{6.3}$$

Equation 6.3 is how the quantum conditional entropy is usually defined within the literature. But consider a direct attempt to measure $S(A|B)$, that is: if we preform a projective measurement on B, how does the entropy of A change? Consider a set of measurements $\{\Pi_i\}$, where $\sum_i \Pi_i = \mathbb{1}$, are made on B. The resulting state of A is given by

$$\rho_{A|i} = \frac{1}{p_i}\mathrm{Tr}_B(\rho_{AB}\,\Pi_i), \text{ where, } p_i = \mathrm{Tr}_{A,B}(\rho_{AB}\,\Pi_i). \tag{6.4}$$

This 'classical-quantum' version of the conditional entropy associated with the post-measurement density matrix, $\rho_{A|i}$, is given by,

$$S_{\Pi_i}(A|B) \equiv \sum_i p_i\, S(\rho_{A|i}). \tag{6.5}$$

This is formulation of the conditional entropy is perhaps a more faithful generalisation of its classical counterpart. Using (6.5) one can also construct a new version of the mutual information of (6.2), commonly referred to as the *one-way classical correlation*,

$$\mathcal{J}(A|\{\Pi_i\}) = S(\rho_A) - S_{\Pi_i}(A|B). \tag{6.6}$$

We interpret $\mathcal{J}(A|\{\Pi_i\})$ as information gained about one subsystem as a result of a measurement on the other. Unsurprisingly, this quantum analog has an explicit measurement dependence; there exist many measurements that may be preformed on partition B, each potentially yielding more or less information regarding system A and thus providing a

better or worse measure for the classically accessible correlations. Henderson and Vedral [177] showed that the total classical correlations within a bipartite state can be obtained by maximising

$$\begin{aligned}\mathcal{J}(A|B) &= \max_{\{\Pi_i\}} \mathcal{J}(A|\{\Pi_i\}) \\ &= S(\rho_A) - \min_{\{\Pi_i\}} S_{\Pi_i}(A|B),\end{aligned} \tag{6.7}$$

over all possible measurements, $\{\Pi_i\}$. This optimisation is necessary to isolate the least disturbing measurement, such that the change of entropy on one subsystem due to measurement on the other quantifies the correlations between the two subsystems.

The respective quantum mechanical generalisations of these two, classically equivalent, forms of the mutual information yield very different quantities: $\mathcal{I}(\rho_{AB})$ provides a measure of the total correlations within a bipartite system, while $\mathcal{J}(A|B)$ captures the classical correlations. The difference between these two expressions, defined as the *quantum discord*

$$\mathcal{D}(A|B) = \mathcal{I}(\rho_{AB}) - \mathcal{J}(A|B). \tag{6.8}$$

provides a measure of all non-classical correlations [177, 176]. As one might suspect from the form of (6.7), the discord is a directional quantity. For a bipartite system, ρ_{AB}, we also need to define the discord on partition B,

$$\mathcal{D}(B|A) = \mathcal{I}(\rho_{AB}) - \mathcal{J}(B|A). \tag{6.9}$$

In general, the discord is asymmetric, such that $\mathcal{D}(A|B) \neq \mathcal{D}(A|B)$. Using (6.1) and (6.7) we can define,

$$\mathcal{D}(A|B) = S(\rho_A) - S(\rho_{AB}) + \min_{\{\Pi_i\}} S_{\Pi_i}(A|B) S_{\Pi_i}(A|B), \tag{6.10}$$

and equivalently for $\mathcal{D}(B|A)$. While it is standard that the optimisation be taken over the set of all POVMs, this is generally an arduous task. A handful of theoretical works [184, 185] introduce *discord-esque* quantities that instead consider a restricted sets of measurements, making the optimisation problem tenable. One such measure of significance for this thesis was introduced independently in [184, 185] and considers a restriction to two-mode Gaussian states and Gaussian measurements.

When quantum discord initially appeared [177, 187] the quantum physics community proved to be quite indifferent. It was a mixed state quantum computing protocol introduced by Knill and Laflamme [171] in the late 90's, referred to as *Deterministic Quantum Computing with One Qubit* (DQC1), that proved critical to discords reemergence from obscurity. DQC1, though certainly not as powerful as a pure-state quantum computing protocol, it is nonetheless capable of providing an exponential speed up on some computational tasks compared to any known classical algorithm. With the resources of a solitary qubit coupled to a completely mixed bath of dimension n, DQC1 allows efficient estimation of the trace of any unitary operation applied to the bath [171]. DQC1 proved to be quite counterintuitive as a quantum computing protocol. Historically, quantum computing emerged from the formalism of pure states, and perhaps unsurprisingly, entanglement

was understood to be the resource responsible for any quantum speed up. In 2008 Datta *et al.* [188] showed that while there was 'some' entanglement between the qubit and the bath, it was bounded by a constant independent of on the size of the bath n, such that any entanglement present becomes vanishingly small with large n. This result conflicted with pure state quantum computing, where to retain exponential speed up entanglement is required to increase with the size of the problem. Datta *et al.* [173] demonstrated that when applied to the problem of DQC1, quantum discord scaled appropriately with the size of the problem. This scaling was subsequently verified experimentally [174]. However, there remained (and still remains) no operational link between mixed state computing and discord, just conjecture regarding whether discord was, or was not, the resource of interest. But this significant result propelled discord from obscurity, and nearly a decade after the original proposal, discord was suddenly in vogue and defining itself as a new research field.

One could conceive of discord as an attempt to generalise the concept of entanglement to mixed states. Historically, much of the theoretical work concerning correlations and quantum states was done in the framework of pure states, and within this framework correlations and entanglement are conceptually identical. This may have formed the sometimes prevailing equivalence within the quantum mechanics community regarding the ideas of entanglement and quantum correlations.

As a measure of the 'quantumness' of correlations, discord has some desirable, and some less desirable qualities. There are two important properties that quantum discord satisfies:

1. When restricted to pure states, the discord coincides with the entropy of entanglement, $\mathcal{D}(A|B) = S(\rho_A) = S(\rho_B)$.[1]

2. The discord is zero for product states, $\rho_{AB} = \rho_A \otimes \rho_B$.

The first satisfies our intuition that if discord is a measure of quantum correlations, as all correlations in pure state are 'entangled' correlations, discord must reduce to our established entanglement measures. The second asserts the discord of any state without *any* correlations must be zero. In addition to the above, the discord:

3. Is invariant under local unitary transformations.

4. Is non-increasing under local operations.

This bring us to perhaps the most distinctive difference between discord and the established notion of entanglement; if in addition to local operations, we permit classical communication (LOCC) between the subsystems, properties 3. and 4. no longer hold. Entanglement, however, is strictly invariant under LOCC. Unlike entanglement, one can manufacture discord within a bipartite system without requiring the two subsystems have ever interacted. As a result, discord is a "broad net" that encompasses correlations of varying degrees of 'quantumness'; from those that violate classical models of 'local realism' to separable, but non-orthogonal states.

[1] The entropy of entanglement is an entanglement measure for pure bipartite states. It is given by Von Neumann entropy of the reduced subsystems, $\mathcal{E}(\rho_{AB}) = S(\rho_A) = S(\rho_B)$.

Non-orthogonality as a criterion for quantumness establishes discord as a ubiquitous quantity. Ferraro *et al.* in 2009 showed that almost every state picked at random would have non-zero discord, and that a generic arbitrarily small perturbation of a state with zero discord will generate discord[189]. Discord's ubiquity in itself is not a problem. It is perhaps unsurprising that most any state described by quantum mechanics could, by some measure, be considered uniquely quantum. But if discord is to be considered useful for quantum information protocols – especially given the association with mixed state quantum computation – the discrepancy between the apparent hardness of implementing quantum information protocols and ease of availability of discord as a resource needs to be addressed.

There has been some progress in establishing an operational meaning for discord. In their initial work [177], Henderson and Zurek related discord to the process of decoherence, where vanishing discord is a condition for the evolution of a quantum state into its 'classical' pointer states. Discord has also found favour as a measure in the quantum thermodynamics community, and is closely related to the *quantum deficit* introduced by Oppenheim *et al.*[190]. Zurek later related discord to difference in efficiency of quantum and classical Maxwells demons [191]. The majority of literature on quantum discord, however, has emerged from the quantum information community. In addition to the speculative link to mixed state quantum computation, discord has been theoretically [192, 193] and experimentally [194] linked to quantum state merging [195]. Links have also been drawn to dense-coding [192]. As non-orthogonality of basis states is a sufficient resource for quantum cryptography, and for separable states the discord is a measure of the non-orthogonality, one could conceive discord should be an excellent candidate to describe cryptographic protocols.

6.3 Theory

Consider our two usual characters from information theory, Alice and Bob. Alice prepares some correlated resource ρ_{AB} on a bipartite quantum system. As discord is, in general, an asymmetric quantity, Alice has two expressions for her discord,

$$\mathcal{D}(A|B) = \mathcal{I}(\rho_{AB}) - \mathcal{J}(A|B) \quad \text{and} \quad \mathcal{D}(B|A) = \mathcal{I}(\rho_{AB}) - \mathcal{J}(B|A). \tag{6.11}$$

She the chooses the labelling for her bipartitions, A and B, such that $\mathcal{D}(A|B) \leq \mathcal{D}(B|A)$ and gives subsystem B to Bob.[2]. Alice possesses a classical random variable $\mathbf{K}$ that takes on the value k with corresponding probability p_k. She privately encodes $\mathbf{K}$ onto her subsystem by application of a corresponding unitary operator U_k. The preparation and encoding scheme is publicly announced. For Alice, with knowledge of her encoding, the discord of ρ_{AB} is unchanged. But to anyone oblivious to which unitary U_k was applied, Alice's encoding results in the state

$$\tilde{\rho}_{AB} = \sum_k p_k U_k \rho_{AB} U_k^\dagger. \tag{6.12}$$

[2]This labelling choice is made for the purposes of the proof of (6.14)

with a corresponding discord $\tilde{\mathcal{D}}(A|B)$. To a party oblivious to Alice's choice of U_k the encoding is non-unitary, and appears as random, uncorrelated noise introduced on partition A. The amount of discord destroyed or *consumed* during the encoding of Alice's partition is simply given by the difference,

$$\Delta\mathcal{D}(A|B) = \mathcal{D}(A|B) - \tilde{\mathcal{D}}(A|B) \tag{6.13}$$

This quantity is always greater or equal to 0. Alice then gives her partition A to Bob and challenges him to estimate $\mathbf{K}$ from $\tilde{\rho}_{AB}$. Bob would need to proceed with some decoding protocol on $\tilde{\rho}_{AB}$ that will output a classical variable $\mathbf{K}_o$ describing his estimate of $\mathbf{K}$. The performance of his chosen decoding protocol is then determined by the classical mutual information between Alice's original encoding and Bob's estimate, given by $I(\mathbf{K}_o, \mathbf{K})$.

We define two different general strategies that Bob can take in estimating $\mathbf{K}$: an *incoherent* strategy, and a *coherent* strategy. The incoherent case restricts Bob to performing individual local measurements on each bipartition and post-processing: *i.e.* Bob can make a local measurement first on A, then B, or vice versa, and use combine his two classical measurement records to estimate $\mathbf{K}$. We denote the upper bound on Bob's performance for when restricted to the incoherent strategy as I_c.

If we extend Bob capabilities to also permit arbitrary coherent operations between A and B, allowing him to optimisation over all possible measurements of the joint system AB, we obtain - in general - a different upper bound on Bob's information, which we label I_q.

We then ask the question, *when are coherent interactions advantageous for Bob?* That is, when is $I_q > I_c$?

In Appendix B.1 we prove that

$$\Delta\mathcal{D}(A|B) - \tilde{\mathcal{J}}(A|B) \leq \Delta I \leq \Delta\mathcal{D}(A|B), \tag{6.14}$$

where $\tilde{\mathcal{J}}(A|B)$ represents the classical correlations remaining *after* encoding (i.e., in $\tilde{\rho}_{AB}$). Equation (6.14) indicates that the additional information accessible to Bob regarding Alice's original encoding, $\mathbf{K}$, when he is allowed to coherently interact his bipartions (that is, $\Delta I = I_Q - I_C$) is related to the discord of Alice's original state, ρ_{AB}. Furthermore, we demonstrate the coherent strategy (*i.e.* a joint measurement) is advantageous if and only if ρ_{AB} contains discord. And that the amount of discord Alice consumes during encoding bounds exactly this advantage. Should Bob and Alice share no discord ($\Delta\mathcal{D}(A|B) = 0$), the coherent and incoherent strategies are indistinguishable in their performance.

The lower bound $\Delta\mathcal{D}(A|B) - \tilde{\mathcal{J}}(A|B) \leq \Delta I$ indicates that an advantage is available for any encoding such that the discord consumed is strictly greater the classical correlations after encoding. This is possible for any discorded ρ_{AB}, since there exists since there exists maximal encodings, such that $\tilde{\mathcal{J}} = 0$ for any ρ_{AB} (see Appendix B.1). In this scenario, all available discord initially available is consumed and the advantage coincides exactly with the discord of the initial state,

$$\Delta I = \Delta\mathcal{D}(A|B) = \mathcal{D}(A|B). \tag{6.15}$$

Discord therefore quantifies exactly a resource that coherent interactions can exploit.

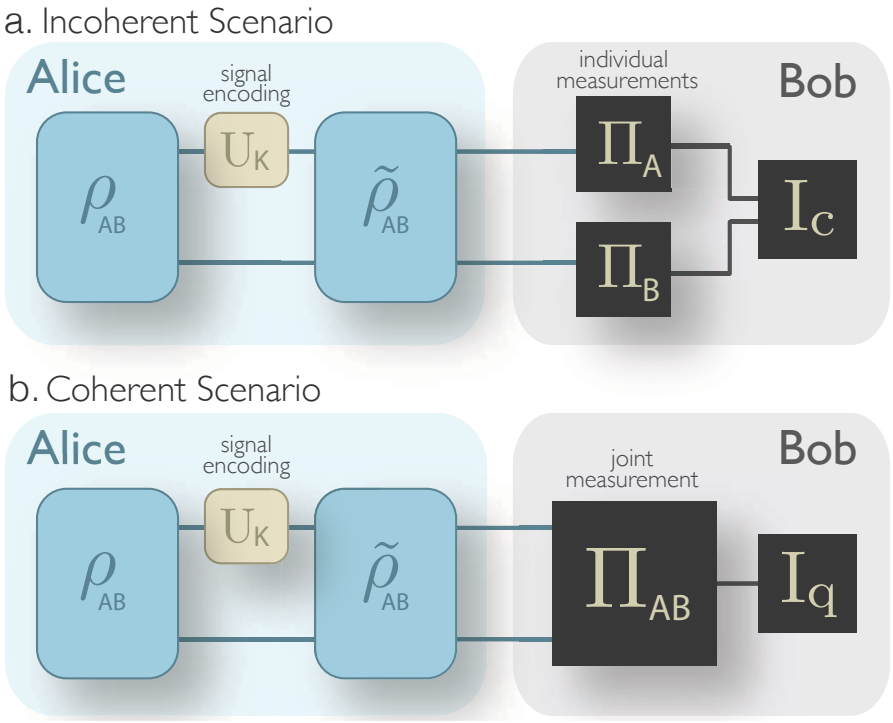

Figure 6.1: Alice begins with a bi-partite resource state, ρ_{AB}. She then encodes a signal **K** via a unitary operation, U_k on subsystem A. She then sends the entire system ρ_{AB} to Bob, who has two avaliable scenarios to try to ascertain **K**: *incoherent* scenario, and an *coherent scenario*. In the incoherent scenario, Bob is permitted an individual measurement on each of his subsystems and post-processing to arrive at his best estimate of **K**. In the coherent scenario, Bob is allowed to optimise over all possible measurements of the joint system, AB.

An example of maximal encoding on two qubits are the Pauli operators $\{I, \sigma_x, \sigma_z, \sigma_x\sigma_z\}$ chosen with equal probability. The special case where this encoding is applied to a singlet state coincides with dense coding [168]. Coherent interactions allow Bob to extract one extra bit of knowledge about which of the four unitary transformations was applied by Alice. This equals the discord consumed when we encode onto the singlet state.

The operational significance of discord beyond entanglement is highlighted when we repeat the above protocol on a separable discordant resource. For example, take $\rho_{AB} = \sum_{i=\{x,y,z\}}(|0\rangle_i|0\rangle_i\langle 0|_i\langle 0|_i + |1\rangle_i|1\rangle_i\langle 1|_i\langle 1|_i)$, where $|0\rangle_i$ and $|1\rangle_i$ represent the computational basis states with respect to σ_i. This resource is clearly separable, and yet possesses a discord of $\frac{1}{3}$. Therefore coherent processing can harness the discord within this resource to extract $\frac{1}{3}$ extra bits of information despite the absence of entanglement.

So far, we have assumed in our definition of $\mathcal{J}(A|B)$ that Bob may choose any POVM to gain information about Alice's encoding. It was mentioned briefly in §6.2 that there exist discord 'variants' where $\mathcal{J}(A|B)$ is optimised over a restricted class of measurements, such as projective measurements [176]; and in the case of continuous variables, Gaussian measurements [184, 185] . The results can be adapted to such variants, where ΔI now bounds the extra advantage gained by Bob if he can implement arbitrary coherent interactions within the restricted class of measurements. The experimental demonstration presented in the next chapter considers the restriction to Gaussian states and Gaussian measurements.

6.4 Gaussian Discord

The optimisation over *all* measurements that the discord requires makes it, in general, onerous to calculate. As such, analytical formulas for the discord of general quantum systems are not ordinarily forthcoming, with most of the existing success occurring in very restricted Hilbert spaces, such as a system of two qubits [196, 197]. The infinite dimensionality of the Hilbert space for continuous-variable systems ensures the calculation of the discord for even a specific state is an arduous task. To reduce this problem to a more tractable one, Giorda and Paris [184], and Adesso and Datta [185] introduced a discord variant that - under the restriction to Gaussian measurements - gave an analytical result for two-mode Gaussian states.[3] This is the so-called 'Gaussian discord'.

As both $\mathcal{D}$ and $\mathcal{J}$ are invariant under local unitary operations, we can exploit the Standard Form of a general two-mode Gaussian state (discussed briefly in §2.4.1) to simplify the analytical result. Recall that any two mode Gaussian state is fully specified by its covariance matrix and mean vector. By the means of a series of unitary (symplectic) operations, any covariance matrix can be transformed into *Standard Form I* with diagonal sub-blocks,

$$\sigma_{AB} = \begin{pmatrix} A & \Gamma \\ \Gamma^{\mathrm{T}} & B \end{pmatrix}. \tag{6.16}$$

[3] Giorda and Paris [184] gave an analytical result for the (admittedly, very general) squeezed-thermal states. Adesso and Datta [185] produced the general result for all two-mode Gaussian states presented here.

For ease of calculation we define the symplectic invariants as $\alpha = \det A$, $\beta = \det B$, and $\gamma = \det \Gamma$. The covariance matrix corresponds to a physical state if and only if $\alpha, \beta \geq 1$ and its symplectic eigenvalues, $\lambda_\pm \geq 1$, where

$$\lambda_\pm^2 = \tfrac{1}{2}(\Delta + \sqrt{\Delta^2 - 4 \det \sigma_{AB}}\,) \tag{6.17}$$

and $\Delta = \alpha + \beta + 2\gamma$. The Gaussian discord for two-mode Gaussian state described a covariance matrix σ_{AB} is given by,

$$\mathcal{D}_G(A|B) = \Phi(\sqrt{\beta}) - \Phi(\lambda_-) - \Phi(\lambda_+) + \min_{\sigma_\Pi} \Phi(\sqrt{\det \sigma_{A|B}}) \tag{6.18}$$

where $\Phi(x) = x_+ \log_2 x_+ - x_- \log_2 x_-$, and $x_\pm = \frac{1}{2}(x \pm 1)$ [185]. We can also define the Gaussian discord in the opposite direction,

$$\mathcal{D}_G(B|A) = \Phi(\sqrt{\det \alpha}) - \Phi(\lambda_-) - \Phi(\lambda_+) + \min_{\sigma_\Pi} \Phi(\sqrt{\det \sigma_{B|A}}). \tag{6.19}$$

Here, $\sigma_{A|B}$ corresponds to the covariance matrix of the conditional state $\rho_{A|b}$ given a measurement Π_i on B. We require a minimisation of $\det \sigma_{A|B}$ over all possible Gaussian measurements σ_Π (projections onto pure single-mode Gaussian states). Adesso and Datta provided an analytical result for this minimisation [185],

$$\text{For} \quad (\det \sigma_{AB} - \alpha\beta)^2 \leq (1+\beta)\,\gamma^2\,(\alpha + \det \sigma_{AB})\,, \tag{6.20}$$

$$\min_{\sigma_\Pi} \det \sigma_{A|B} = \frac{2\gamma^2 + (-1+\beta)\,(-\alpha + \det \sigma_{AB}) + 2|\gamma|\sqrt{\gamma^2 + (-1+\beta)\,(-\alpha + \det \sigma_{AB})}}{(-1+\beta)^2}.$$

Otherwise,

$$\min_{\sigma_\Pi} \det \sigma_{A|B} = \frac{\alpha\beta - \gamma^2 + \det \sigma_{AB} - \sqrt{\gamma^4 + (-\alpha\beta + \det \sigma_{AB})^2 - 2\gamma^2\,(\alpha\beta + \det \sigma_{AB})}}{2\beta}.$$

Whilst the result above is in a form that renders it quite unilluminating, it has quite interesting consequences for the way we consider Gaussian states. The results of [185, 184] showed that the only Gaussian states that do not possess Gaussian discord are product states; that is to say, all bi-partite states with any correlations whatsoever possess discord. A large subset of these states would have been broadly identified as 'effectively' classical systems, consisting only of statistical mixtures coherent states.

However, the restriction to a class of measurements is undesirable in that it ensures that you can only ever overestimate the 'quantumness' of the state. As such, the Gaussian discord only ever provides an upper bound on the actual discord of the state. Recent work has shown that for Gaussian states, the presence of Gaussian discord necessary condition for genuine discord [198]. There has also been some numerical evidence for the equivalence of Gaussian discord and discord proper for Gaussian states, arguing for the optimality of Gaussian measurements for Gaussian states [199]. However this work only considered a handful of very restrictive non-Gaussian measurements. Gaussian measurements are understood to be, in general, not optimal for Gaussian states, and non-Gaussian

measurements are usually required to achieve the Holevo information[200].[4]

6.5 A Continuous Variables Implementation

We now turn our attention applying the theory of §6.3 to a continuous variable physical system suitable for our demonstration. Consider A and B are continuous variables modes, with respective quadrature operators $\hat{X}_A, \hat{P}_A$ and $\hat{X}_B, \hat{P}_B$ that obey the commutation relations $[X_j, P_k] = 2i\delta_{jk}$.

To emphasise that discord is the quantity of interest in this proof of principle demonstration we designed Alice's resource state to be separable, but discordant. As any bipartite Gaussian state that is not a product state has non-zero Gaussian discord, all that we require is the introduction of correlations between partitions A and B. Alice prepares her resource state ρ_{AB} by random, and correlated displacement of two vacuum states. The resulting resource is described by the covariance matrix $\sigma(\rho_{AB})$

$$\sigma(\rho_{AB}) = \begin{pmatrix} V+1 & 0 & V & 0 \\ 0 & V+1 & 0 & -V \\ V & 0 & V+1 & 0 \\ 0 & -V & 0 & V+1 \end{pmatrix} \tag{6.21}$$

where V is the variance of the correlated noise. As the role of this correlated noise is to introduce discord between partitions A and B, we will herein reluctantly refer to it as *discording noise*. Using (6.20), the Gaussian discord, $\mathcal{D}_G(A|B)$ of Alice's resource, ρ_{AB}

$$\mathcal{D}_G(A|B) = \Phi\left(V+1\right) - 2\Phi\left(\sqrt{2V+1}\right) + \Phi\left(1 + \frac{2V}{2+V}\right). \tag{6.22}$$

where $\Phi(x) = x_+ \log_2 x_+ - x_- \log_2 x_-$, and $x_\pm = \frac{1}{2}(x \pm 1)$. Alice then encodes separate signals x_s and p_s governed respectively by Gaussian distributed random variables $\mathbf{X}_s$ and $\mathbf{P}_s$ of variance V_s in the quadratures of her mode by application of

$$\mathcal{E}_A(x_s, y_s) = \exp(-\tfrac{1}{2} i x_s X_A) \exp(-\tfrac{1}{2} i p_s P_A) \tag{6.23}$$

This results in an encoded state $\tilde{\rho}_{AB} = \int \mathcal{E}_A \rho_{AB} \mathcal{E}_A^\dagger dx_s dp_s$ with the covariance matrix,

$$\sigma(\tilde{\rho}_{AB}) = \begin{pmatrix} V+1 & 0 & V & 0 \\ 0 & V+1 & 0 & -V \\ V & 0 & V+V_s+1 & 0 \\ 0 & -V & 0 & V+V_s+1 \end{pmatrix}. \tag{6.24}$$

Alice then gives her encoded state $\tilde{\rho}_{AB}$ to Bob, and tasks him with estimating the encoded signal $(\mathbf{X}_s, \mathbf{P}_s)$. In §6.3 we identified two different strategies undertaken by Bob in estimating $(\mathbf{X}_s, \mathbf{P}_s)$: which we dubbed the *incoherent*, and *coherent* strategy. The incoherent

[4] After the submission this thesis, a proof of the bosonic minimum output entropy conjecture appeared [201]. This result leads to the optimality of Gaussian discord - that is, for Gaussian states that the Gaussian discord is the real discord [202].

strategy restricts Bob to an individual measurement on each mode and post-processing. An upper bound on Bob's accessible information for this scenario (with the additional restriction to Gaussian measurements) is given by the Holevo information of Bob's state after he has preformed his optimal Gaussian measurement on one of his partitions. Application of (B.7) gives

$$I_c = \Phi\left(1 + \frac{2V}{V+2} + V_s\right) - \Phi\left(1 + \frac{2V}{2+V}\right). \tag{6.25}$$

In contrast, the coherent strategy allows Bob to optimise over all possible measurements of the joint system, $\tilde{\rho}_{AB}$. The upper bound on his accessible information I_q is simply the Holevo bound,

$$I_q = S(\tilde{\rho}_{AB}) - S(\rho_{AB}) = \Phi(\lambda_+) + \Phi(\lambda_-) - 2\Phi(\sqrt{2V+1}), \tag{6.26}$$

where

$$\lambda_\pm = \sqrt{2V + 1 + \frac{V_s}{2}(V_s + 2V + 2 \pm \sqrt{(V_s+2)(4V+V_s+2)})}.$$

In the limit that Alice encodes maximally ($V_S \to \infty$) and consequently consumes all the discord in the encoded state, $\tilde{\rho}_{AB}$, then $\lim_{V_s\to\infty}(I_q - I_c) = \mathcal{D}(\rho_{AB})$ and the additional information available to Bob through coherent interactions is equal to the discord of Alice's initial state, ρ_{AB}.

The Holevo bound dictated by (6.26) corresponds to a theoretical bound that in theory can *always* be saturated. Practically, however, we rarely know the experimental strategy that achieves it. For the purposes of this demonstration, however, we only need to isolate a strategy for which I_q^{exp} outperforms the ideal incoherent strategy, I_c. And the identified strategy is notably simple: Bob coherently interacts his bipartite system on a 50-50 beamsplitter and the resulting modes are then measured independently via homodyne detection in orthogonal quadrature basis. If Bob does this perfectly he can achieve an information,

$$I_q^{\text{exp}} = \log\left(1 + \frac{V_s}{2}\right). \tag{6.27}$$

Note that the result of (6.32) is independent of the discording correlations, V. Interfering the two subsystems in phase on a 50:50 beamsplitter results in the covariance matrix,

$$\sigma(\tilde{\rho}_{AB}) = \begin{pmatrix} 2V + \frac{V_s}{2} + 1 & 0 & \frac{V_s}{2} & 0 \\ 0 & \frac{V_s}{2} + 1 & 0 & -\frac{V_s}{2} \\ \frac{V_s}{2} & 0 & \frac{V_s}{2} + 1 & 0 \\ 0 & -\frac{V_s}{2} & 0 & 2V + \frac{V_s}{2} + 1 \end{pmatrix}. \tag{6.28}$$

In this idealised scenario Bob can estimate Alice's encoding without penalty from the initial encoded noise, his measurement equivalent to a heterodyne detection of Alice's encoding. One can check that as $V \to \infty$, $I_q^{\text{exp}} \to I_q$ (Figure 6.2.a). Provided the discord in the initial resource is very large, the protocol defined above is *almost* optimal. Experimentally however, 'infinite' noise proves quite the pain to realise, but we only require 'discording noise' variance comparable to the variance of the signal encoding to see a clear discepency between I_q^{exp} and I_c.

We experimentally prepare the aforementioned resource state ρ_{AB}, and encode within it the signals $(\mathbf{X}_s, \mathbf{Y}_s)$. We then take on the role of Bob, and attempt to measure some observable pairs $(\mathbf{X}_s^{\text{exp}}, \mathbf{Y}_s^{\text{exp}})$ such that $I(\mathbf{X}_s, \mathbf{Y}_s; \mathbf{X}_s^{\text{exp}}, \mathbf{Y}_s^{\text{exp}})$ is maximized. Theory dictates that when limited to a single local measurement on each subsystem, $I(\mathbf{X}_s, \mathbf{Y}_s; \mathbf{X}_s^{\text{exp}}, \mathbf{Y}_s^{\text{exp}}) \leq I_c$. Experimental violation of the above inequality will demonstrate coherent processing can harness uniquely quantum correlations. The magnitude with which we can violate this inequality

$$\Delta I^{\text{exp}} = I(\mathbf{X}_s, \mathbf{Y}_s; \mathbf{X}_s^{\text{exp}}, \mathbf{Y}_s^{\text{exp}}) - I_c(\rho_{AB}) \tag{6.29}$$

then defines the observed quantum advantage.

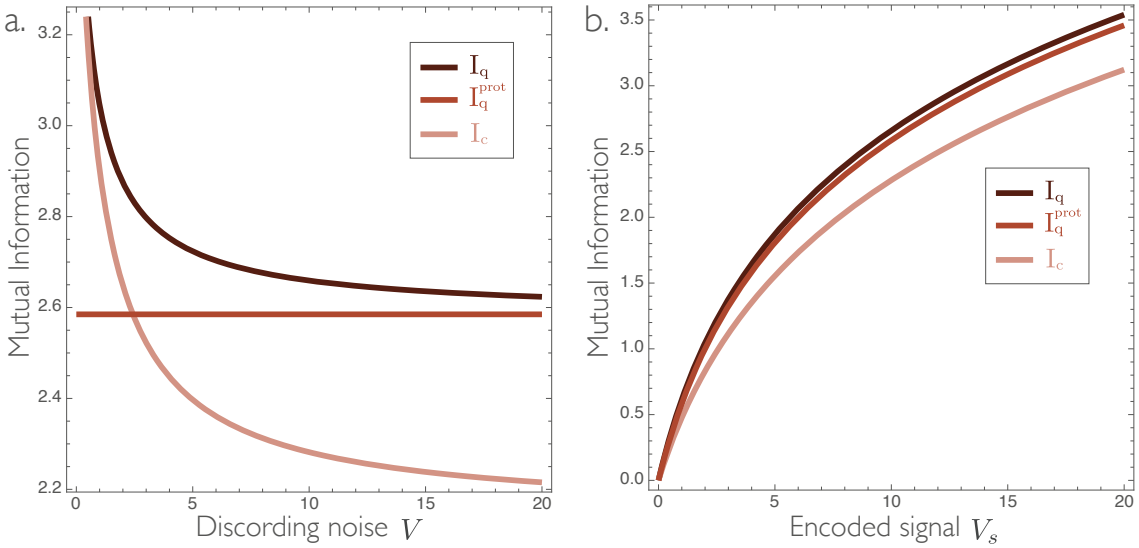

Figure 6.2: The theoretical mutual information between Alice's encoding ($\mathbf{X}_s$, $\mathbf{P}_s$) and Bob's estimates ($\mathbf{X}_o$ and $\mathbf{P}_o$) for three scenarios: I_c, Bob's best performance when he is restricted to individual measurements; I_q, when Bob is also able to make joint measurements; and I_q^{prot}, the non-optimal joint measurement protocol specified in §6.5. Figure *a.* plots the mutual information as a function of the discording noise for a constant signal strength, $V_s = 10$, whereas *b.* considers a constant discording noise ($V = 10$) and a varying signal strength.

We can also compare this rate to the optimal known decoding scheme when limited to a single local measurement of A and B [203, 204]. In this scheme, Bob makes simultaneous quadrature measurement of the two modes. The information Bob can extract from a state having covariance matrix of the form (6.24) is given by

$$I_c^{prot} = \log\left(1 + \frac{1+V}{1+2V} V_s\right). \tag{6.30}$$

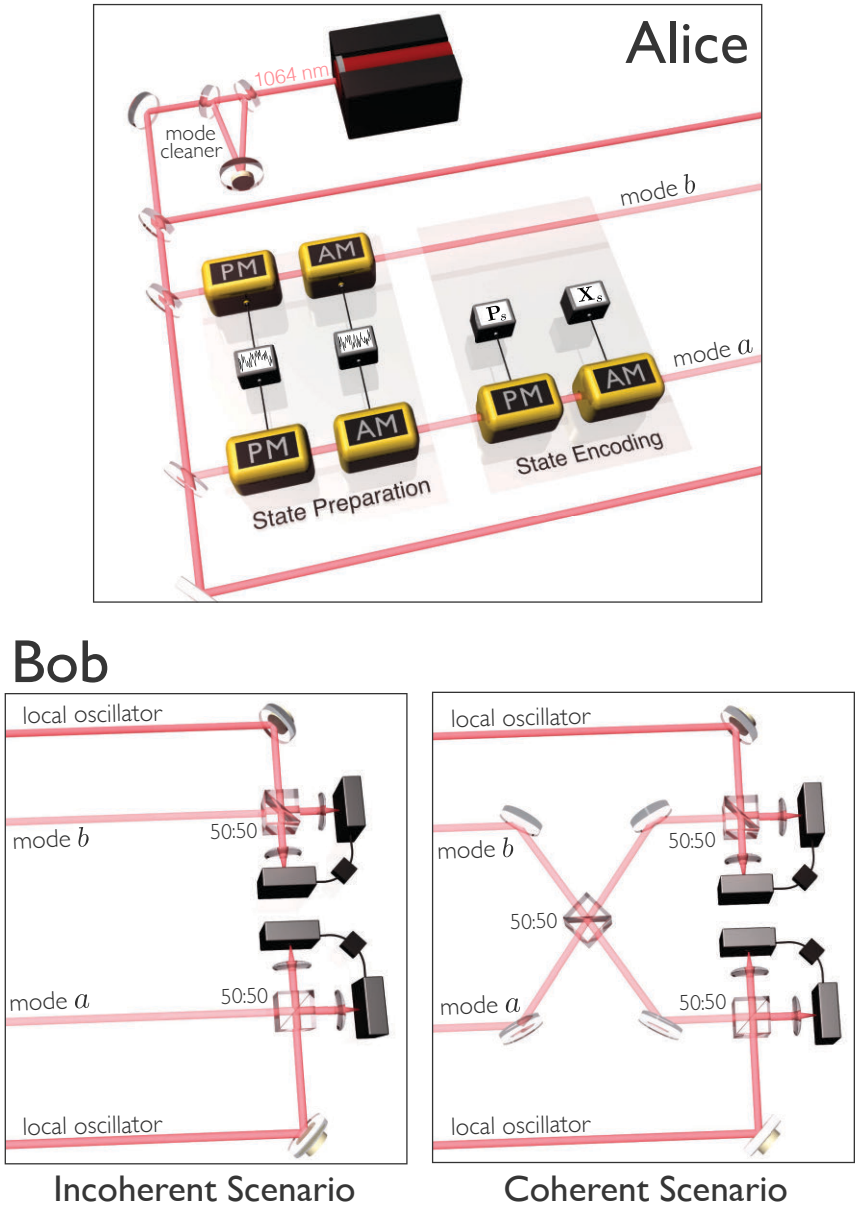

Figure 6.3: A laser provides coherent light that is encoded using modulation of the sideband frequencies. Alice prepares her discordant bi-partite state ρ_{AB} by correlated (anti-correlated) displacement of two coherent vacuum states in the amplitude (phase) quadrature with Gaussian distributed noise using electro-optic modulation. Alice then encodes independent signals $\mathbf{X}_s$ and $\mathbf{Y}_s$ on the phase and amplitude quadrature of her subsystem using EOM and subsequently transmits her state to Bob. We compare Bob's capacity to extract information in two different scenarios. The theoretical limit to Bob's performance when Bob makes individual measurements of A and B, and the experimental observed performance when Bob uses a particular protocol involving coherent interference to enhance his knowledge of Alice's encoding.

6.6 The Experiment

In the previous section we have devised a recipe for implementation of the theory of §6.3, now we turn to the details of the implementation. This experiment is notably more straightforward than those discussed in the previous chapters, requiring only electro-optic modulation and homodyne detection.

6.6.1 Light source

The experiment was constructed on an actively damped table using entirely free space optics. The source of laser light for this demonstration was an *Innolight Mephisto* Neodymium-doped Yttrium Aluminum Garnet (Nd:YAG) laser producing up to 2.1W of single mode continuous wave light at 1064nm. The laser FWHW linewidth was specified to by the manufacturer to be ≈ 1 kHz. The ND:Yag crystals non-planar ring geometry has a natural relaxation oscillation at ~ 750 kHz, which is attenuated considerably by the inclusion of a 'noise-eater' option. Despite the noise-eater, the roll off of the relaxation oscillation was evident up to 4 MHz.

A Faraday isolator was introduced directly after the laser as a precaution against unintended optical feedback. To further improve the suppression of the relaxation oscillation, and also ensure a well-defined TEM 00 spatial mode for our experiment, the laser field was passed through a high-finesse optical mode-cleaner. The triangular ring geometry mode cleaner had a linewidth of approximately 2 MHz and an observed finesse of 160. The mode cleaner was controlled via Pound-Drever-Hall locking, with a phase-modulator introduced immediately after the Faraday isolator for this purpose.

6.6.2 State preparation

We now assume the role of Alice. A small portion of the filtered laser light was partitioned into two beams of equal power, which we labelled partitions A and B. The remainder is reserved to provide a bright local oscillator for measurement. Consider the ideal form of the covariance matrix for the initial resource state,

$$\sigma(\rho_{AB}) = \begin{pmatrix} V+1 & 0 & V & 0 \\ 0 & V+1 & 0 & -V \\ V & 0 & V+V_s+1 & 0 \\ 0 & -V & 0 & V+V_s+1 \end{pmatrix} \tag{6.31}$$

Our first requirement is to introduce discord between subsystems A and B. We do so via correlated and anti-correlated random displacements of amplitude and phase quadratures of initial vacuum state. The displacement employed a series of electro-optic modulators from *New Focus* that were all broadband coated for the near to far infrared (900-1600 nm). Each beam passed through one phase and one amplitude electro-optic modulator. The light was polarised vertically to the crystals propagation axis as it entered the phase modulator. We opted for a quarter wave plate before the amplitude modulator to ensure a linear response to the modulation signal. Function generators provided two Gaussian-distributed white noise sources, which we label $\mathbf{P}_d$ and $\mathbf{X}_d$.

To correlate the amplitude quadratures of the partitions A and B, our classical signal $\mathbf{X}_d$ is split to drive both amplitude modulators. Equivalently, the phase quadratures were anti-correlated by dividing our classical signal $\mathbf{P}_d$ between both phase modulators, with one output phase shifted by 180°. As this demonstration was concerned with both the successfully encoding and detection of correlations, the amplitude and phase response of the system is critical. The requirement for passive electronics like delay, attenuators and splitters, in addition to the detection circuits, makes ensuring a close to identical frequency response for both system A and B difficult across the entire available detection band of 3-10 MHz difficult. As such, we narrowed our focus to a sideband frequency measurement window between 3.2-4 MHz, which would later be fine-tuned further for data acquisition.

To ensure the quadratures of modes A and B were optimally correlated we made use of the two homodyne detection stages required for the measurement. Choosing a frequency band we would focus on for our measurements, we examined the correlations of both quadrature across a 1 MHz frequency window around our central sideband frequency of 3.6 MHz. Delay was accordingly introduced on mode B to synchronise it to mode A. The magnitudes for the white noise were also matched to ensure the closest realisation to the ideal symmetric form of the covariance matrix of (6.31).

Alice is also required to encode a signal on mode A. We denote Alice's signal encoding in the amplitude and phase quadratures of mode A by the Gaussian distributed random variables, $\mathbf{P}_s$ and $\mathbf{X}_s$, experimentally provided by two independent function generators. The signal encoding made use of the existing phase and amplitude modulators on mode A, electronically adding the 'signal' noise to the 'discording' noise. A copy of the encoded signal is recorded for the purposes of characterising success of the protocols. The size of the encoded signals in both quadratures was balanced $V_{s_x} \approx V_{s_p}$ to ensure the encoded state was close to the form of (6.31).

6.6.3 State measurement

We now consider the implementation of the two measurement scenarios: the *incoherent* scenario and the *coherent* scenario. As we want to establish an upper bound for Bob's information when restricted to individual measurements and post-processing, we first characterised Bob's system by reconstructing the covariance matrix describing the encoded state, $\tilde{\rho}_{AB}$. The covariance matrix allowed direct inference of the theoretical incoherent limit, I_c from the Holevo bound. The coherent scenario, however, required we implement our chosen protocol: interaction of the two subsystems on a beamsplitter, followed by a direct measurement via homodyne detection. Both scenarios require two homodyne detection stages. As such, experiment was designed such that that we could transition between the coherent and incoherent scenarios by rotating flip mirror mounts in and out of the beam paths, while still maintaining the same path lengths for mode matching purposes.

Homodyne detection

Both the coherent and incoherent scenarios made use of the same two balanced homodyne detection stages (labelled 1 and 2). Each homodyne detection stage utilised two Uni-PD circuits. The homodyne efficiency of the entire detection stage was originally estimated at $91 \pm 2\%$. This was primarily limited by the *Epitaxx ETX-500* InGaAs photodiodes which

have a manufacturer specified quantum efficiency of $95 \pm 2\%$. Improvements of approximately 1-2% in the absorbed light could be made by tweaking the angle of incidence, and the total efficiency could have been improved by retro-reflection of the light onto the diode surface. The secondary limitation to the homodyne efficiency was the mode-matching, with typical fringe visibilities of 98%. This was limited by wavefront distortions introduced by the electro-optic modulators.

For control of the homodyne detection angle of mode A and mode B we introduced additional sideband modulations at 21.25 MHz and 33.125 MHz respectively. Some care had to be taken with choice of these frequencies to ensure a clear detection band. Depending on our choice of quadrature measurement, an analog switch was used to route the control frequencies to either their respective amplitude or phase modulators. An error signal for controlling the homodyne detection was extracted a straightforward demodulating the photocurrent at the modulation frequency. This approach allowed us to easily switch between a lock of either the phase or amplitude quadrature without requiring any changes to the electronic hardware.

Incoherent Scenario

We reconstruct the two-mode covariance matrix of the encoded state ρ_{AB} by individual measurements of modes A and B on two homodyne detection stages. Full reconstruction of the covariance matrix required we sample all possible combinations of quadrature measurements of phase and amplitude on modes A and B. This allows us to characterise not only the matrix elements corresponding to the ideal standard form, but also the cross-terms that arise due to imperfect modulation and control.

Coherent Scenario

The coherent protocol we identified in §6.5, while not optimal, is sufficient to demonstrate the improvement afforded by joint measurements. The protocol requires we interfere the two input systems on a 50:50 beamsplitter, with the relative phase, ϕ, of the two beams controlled to be zero. The two resulting beams are then measured via the same balanced homodyne detection stages described above, one stage sampling the amplitude quadrature, whilst the other measures the phase quadrature.

The interference beamsplitter uses a non-polarising beamsplitter cube that deviated from a perfect 50:50 beamsplitter with an observed splitting ratio of 48:52. This ratio has consequences for the protocol as the ideal implementation of the protocol identified requires perfect destructive interference of the 'discording' noise to realise the information of

$$I_q^{\exp} = \log\left(1 + \frac{V_s}{2}\right). \tag{6.32}$$

Any deviation from a balanced beamsplitter will introduce a noise term that will rapidly degrade Bob's estimation.

Control of the relative phase, ϕ, between the two beams at the beamsplitter required we encode an additional phase modulation on subsystem A. Consider we block subsystem B before the beamsplitter, such that we are effectively splitting mode A between two

homodyne detection stages. We then lock the homodyne detection angle to the sample the amplitude quadrature of our signal mode ($\theta = \pi/2$). Unblocking mode B will produce some rotation of the phase modulation, resulting in an amplitude modulation contribution in our measured homodyne photocurrent proportional to the relative phase between A and B. By demodulating the homodyne photocurrent at this new frequency we can extract an error signal that is proportional to the relative phase, ϕ. In combination with our existing error signal for the homodyne angle, we can control the three mode system as desired. This is not the most elegant approach as the systems are coupled to each other, but it works well for accommodating small perturbations around $\phi = 0$ and $\theta = \pi/2$ and requires we introduce no additional loss for the purposes of control.

Otherwise, the technical details of the control of the two modes for homodyne measurement are identical to the incoherent scenario. But as we do not need to fully characterise Bob's state, but directly extract his estimates $\mathbf{X}_o$ and $\mathbf{P}_o$, we only require one set of measurements. For this the amplitude quadrature of the bright output and the phase quadrature of the dark output are sampled using the two available homodyne detection stations.

6.6.4 Acquisition and analysis

For each choice of V and V_s the characterisation of Bob's covariance matrix for the coherent strategy required four measurements, corresponding to the permutations of X and P at the two detection stations. The incoherent strategy required only required the single measurement of X and P. For each measurement we also record Alice's encoded signal. For each separate homodyne detection 5×10^6 data points are sampled at 20 MHz using a digital acquisition system. All measurements were digitally filtered to 3.6-3.8 MHz and then re-sampled.

To obtain I_q^{exp}, we directly calculated the classical mutual information between our record of Alice's encodings ($\mathbf{X}_s$,$\mathbf{P}_s$), and Bob's measurement strings ($\mathbf{X}_o$,$\mathbf{P}_o$). Our upper bound I_c requires we first reconstruct the CM of Bob's two mode state ρ_{AB}. We then symmetrised the reconstructed CM before compensating for the presence of loss and noise. This compensation allowed us to obtain a strong upper bound on the Holevo information for Bob's state given the limitations of our experimental implementation. The direct reconstruction of the CM also allowed us to model both the coherent and incoherent scenarios accurately, with some additional characterisation of the experimental parameters.

6.6.5 Alignment and optimisation

A successful demonstration of the theory presented in §6.5 required we address a few technical subtleties that arose. Experimentally, the electro-optic modulation of one quadrature will always introduce a parasitic modulation contribution in the orthogonal quadrature. This effect is largely attributed to parasitic etalon effects that occur within the crystal that induce a small rotation of the modulation sideband, producing a contribution in the orthogonal quadrature. This effect is routinely worse for amplitude modulators which require two crystals and thus more interfaces. The broadband coating of the modulators is the likely culprit, with small changes in alignment and temperature producing large variation in the contribution from parasitic modulation. When encoding in one quadrature,

a successful demonstration of the theory required typical suppression of the orthogonal quadrature to be 25 dB or greater.

6.7 Model

Whilst ideally, we strive to achieve the state described by the covariance matrix of (6.24), in practise the actual state is never such. To refine the experimental model, we included effects of imperfect correlations, passive losses, excess noise and unbalanced beam splitter ratio. The covariance matrix of the bipartite state is a function of the input signal, the input noise and the quantum noise. It can be expressed as $C_0 = \hat{v}^\dagger \hat{v}$ where

$$\hat{v} = \left(\vec{X}_A, \vec{P}_A, \vec{X}_B, \vec{P}_B\right) \tag{6.33}$$

and $\vec{X}_{A(B)}$ and $\vec{P}_{A(B)}$ represent the modulation on the amplitude and phase quadratures of mode $A(B)$ written as a linear combination of eight independent inputs: the input signals for $X(P)$, $\sigma_{sx(p)}$, the input classical noise for $X(P)$, $\sigma_{nx(p)}$ and the vacuum noises in $X(P)$ in mode A and beam B, σ_v. We write $\hat{v}$ as

$$\hat{v} = \begin{pmatrix} \eta^A_{xx}\sigma_{sx} & \eta^A_{px}\sigma_{sx} & 0 & 0 \\ \eta^A_{xp}\sigma_{sp} & \eta^A_{py}\sigma_{sp} & 0 & 0 \\ \beta^A_{xx}\sigma_{nx} & \beta^A_{px}\sigma_{nx} & \beta^B_{xx}\sigma_{nx} & \beta^B_{px}\sigma_{nx} \\ \beta^A_{xp}\sigma_{np} & \beta^A_{pp}\sigma_{np} & \beta^B_{xp}\sigma_{np} & -\beta^B_{pp}\sigma_{np} \\ \xi^A_x\sigma_v & 0 & 0 & 0 \\ 0 & \xi^A_p\sigma_v & 0 & 0 \\ 0 & 0 & \xi^B_x\sigma_v & 0 \\ 0 & 0 & 0 & \xi^B_p\sigma_v \end{pmatrix}. \tag{6.34}$$

The coefficients η and β characterise the linear correlations between the quadrature modulations and the applied signal and noise voltages. The terms η_{xp} and η_{px} capture the parasitic cross correlations that arise due to imperfectly orthogonal measurements and imperfect modulation of the quadratures. A non-zero correlation will degrade the mutual information, with this degradation proving more pronounced for the coherent scenario owing to the restricted nature of the decoding scheme. The flexibility for additional post-processing in the incoherent scenario moderates its effect. The terms η_{xx} and η_{yy} are the correlations between the signal and the quadrature modulation. Imperfect correlation will again degrade both the resulting information for both the coherent and incoherent scenarios. The coefficients ξ characterise excess noise in the quadratures.

The identified decoding strategy for the coherent scenario is quite artificially restricted, and our strategy is only *almost* optimal in the limit of a perfect implementation concerning *very* large encoding and *very* large noise. To maximise the observed quantum advantage the 'discording' noise should be as close to equal in magnitude on both modes, allowing perfect cancellation. An unequal magnitude of the signal encoding on phase and amplitude quadrature also punishes the coherent scenario where the measurement cannot be biased appropriately. For the coherent case, we also include a small nonlinear loss that increases with the signal variance around the order of $\eta_{loss} = 0.0001\sigma^2_{sx(sp)} + 0.00003\sigma^4_{sx(sp)}$ just

before the beam splitter. This is attributed to the nonlinear response of the electro-optic modulators and gives rise to the observed plateauing of the quantum advantage in Figure 6.5. The loss is simulated by propagating the covariance matrix C_0 through a beam splitter and tracing over the output of the vacuum port to get the new covariance matrix:

$$C_1^A = \mathrm{Tr}_v\{BS(\eta_{loss}) \cdot C_0^A \oplus C_v \cdot BS(\eta_{loss})^\dagger\} \tag{6.35}$$

where

$$C_v = \begin{pmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix} \tag{6.36}$$

is the covariance matrix for the vacuum input and

$$B(\eta) = \begin{pmatrix} \sqrt{\eta} & 0 & -\sqrt{1-\eta} & 0 \\ 0 & \sqrt{\eta} & 0 & -\sqrt{1-\eta} \\ \sqrt{1-\eta} & 0 & \sqrt{\eta} & 0 \\ 0 & \sqrt{1-\eta} & 0 & \sqrt{\eta} \end{pmatrix} \tag{6.37}$$

is the beam splitter transformation with transmission η. $C_0^A = \mathrm{Tr}_B\{C_0\}$ is the covariance matrix for beam A. For the coherent scenario, the modes A and B are then propagated through an interference beam splitter with transmission coefficient $\eta_i = 0.48$ and the relative phase between the two beams fixed at $\phi_A - \phi_B = 0$. The output covariance matrix is then

$$C_2 = B(\eta_i)P(\phi_A, \phi_B) \cdot C_1 \cdot P(\phi_A, \phi_B)^\dagger B(\eta_i)^\dagger \tag{6.38}$$

where $P(\phi_A, \phi_B) = P(\phi_A) \oplus P(\phi_B)$ shifts the phases of beam $A(B)$ by $\phi_{A(B)}$ with

$$PS(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} . \tag{6.39}$$

The homodyne efficiencies are modelled as a vacuum noise contaminating the signal. Moreover, we take into account an imperfect locking angle between the local oscillator and the signal, modelled as a rotation of the beam quadrature before the measurement

$$C_3^A = \mathrm{Tr}_v\{B(\eta_{lo}^A)P(\phi_{lo}^A) \cdot C_2^A \cdot P(\phi_{lo}^A)^\dagger B(\eta_{lo}^A)^\dagger\} \tag{6.40}$$

and a similar expression for C_3^B with $\phi_{lo}^A = 0$ and $\phi_{lo}^B = \pi/2$. Finally, tracing over the phase quadrature gives the measured output of the detectors in the coherent interaction setup $SX_{measured} = \mathrm{Tr}_P\left[C_3^A\right]$ and $SP_{measured} = \mathrm{Tr}_X\left[C_3^B\right]$. In the incoherent interaction case, the covariance matrix C_0 is directly propagated through to the homodyne detection to sequentially measure both the $\hat{X}$ and $\hat{P}$ quadratures of both beams. The information rate for the incoherent scenario is then calculated using the full covariance matrix.

6.8 Results & Discussion

The experimental results for this work are summarised in only two figures. The first, Figure 6.4 compares the coherent and incoherent scenarios as a function of the discording noise. Here we fix the signal variance (normalised to the standard quantum limit)

at 9.10 ± 0.05. The blue data points represent the directly observed values of I_q^{exp}. Our identified protocol, I_q^{prot} should show no 'advantage' for very low discording noise, and should prove *almost* optimal for very large discording noise. Experimentally, we find that

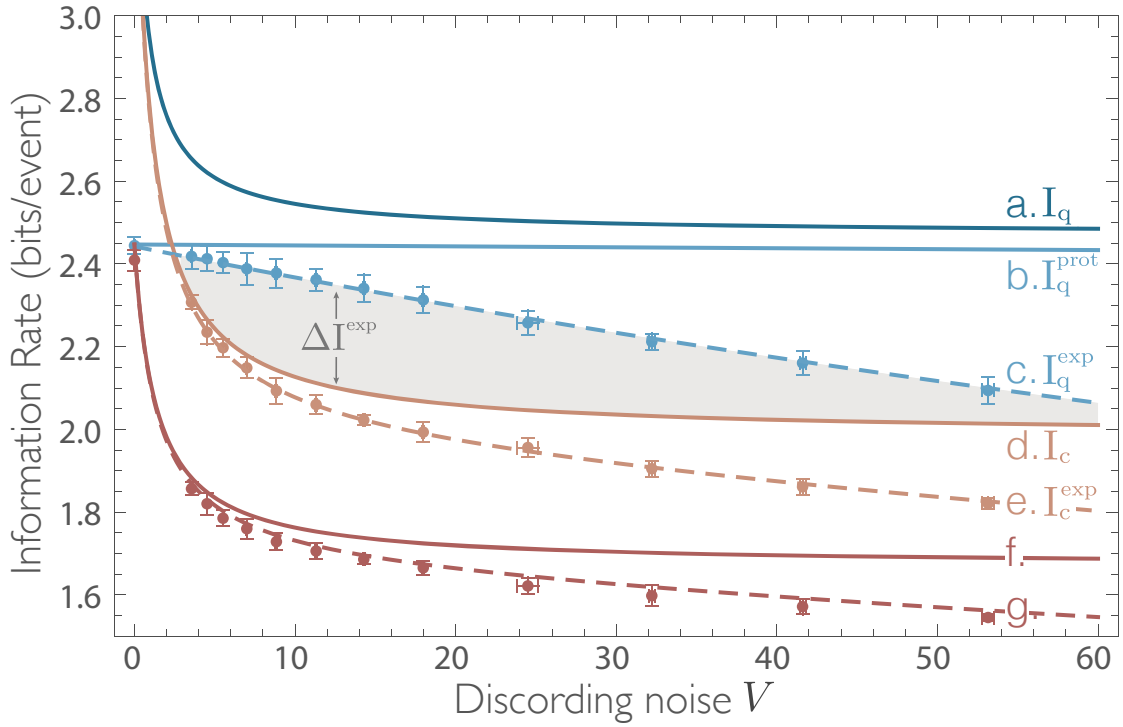

Figure 6.4: Plot of Bob's knowledge of the encoded signal for bipartite resource states with varying discording noise and fixed encoding variance V_s. (a) represents the amount of information Bob can theoretical gain should he be capable of coherent interactions. For our proposed implementation, this maximum is reduced to (b). Experimentally, Bob's knowledge about the encoded signal is represented by the green data points. The line (c) models these observations by taking experimental imperfections into account. Despite these imperfections, Bob is still able to gain more information than the incoherent limit (d). The shaded region highlights this quantum advantage. This advantage is more apparent if we compare Bob's performance to the reduced incoherent limit when experimental imperfections are accounted for (e). We can also compare these rates to a practical decoding scheme for Bob when limited to a single measurement on each optimal mode (f) and its imperfect experimental realisation (g). The error bars represent a statistical confidence interval of 3σ.

provided the discording noise is sufficiently large (such that the original resource has a significant amount of discord) I_q^{exp} clearly exceeds the ideal incoherent limit I_c (Figure 6.4.d.). There is a considerable deviation between the amount of information we experimentally extract (Figure 6.4.c.), and the theoretical prediction I_q^{prot} of the idealised protocol (Figure 6.4.b.). This discrepancy is largely due to loss on Bob's measurement. Because we cannot ensure identical loss contributions for both the coherent and incoherent strategies, when computing our theoretical bound of Bob's incoherent information, I_c, we correct for loss on Bob's reconstructed co-variance matrix. This correction lifts Bob's information from

the Holevo information of the measured co-variance matrix, I_c^{exp} (Figure 6.4.e.) to the idealised version, I_c. While our measured I_q^{exp} is monotonically decreasing with increasing discording noise, the same trend is evident in the 'uncorrected' incoherent performance, and is therefore largely attributed to loss.

Other experimental imperfections that contribute to the discrepancy between I_q^{prot} and I_q^{exp} include asymmetric modulation variance in the phase and amplitude quadratures of both the encoded signal and the discording noise, limited suppression of the parasitic phase and amplitude modulations, and crucially, asymmetry of the '50:50' beam splitter. Ideally, the beamsplitter interaction eliminates any effect of the discording noise in Bob's estimate of Alice's encoding. If the beamsplitter deviates from 50:50, Bob's estimate will be contaminated by a contribution from the 'uncancelled' discording noise. As our chosen coherent protocol is inherently 'symmetric', any deviation from symmetric encodings of the quadratures punishes I_q^{exp} much faster than I_c. These imperfections are well captured by our model, with theory and experiment showing excellent agreement.

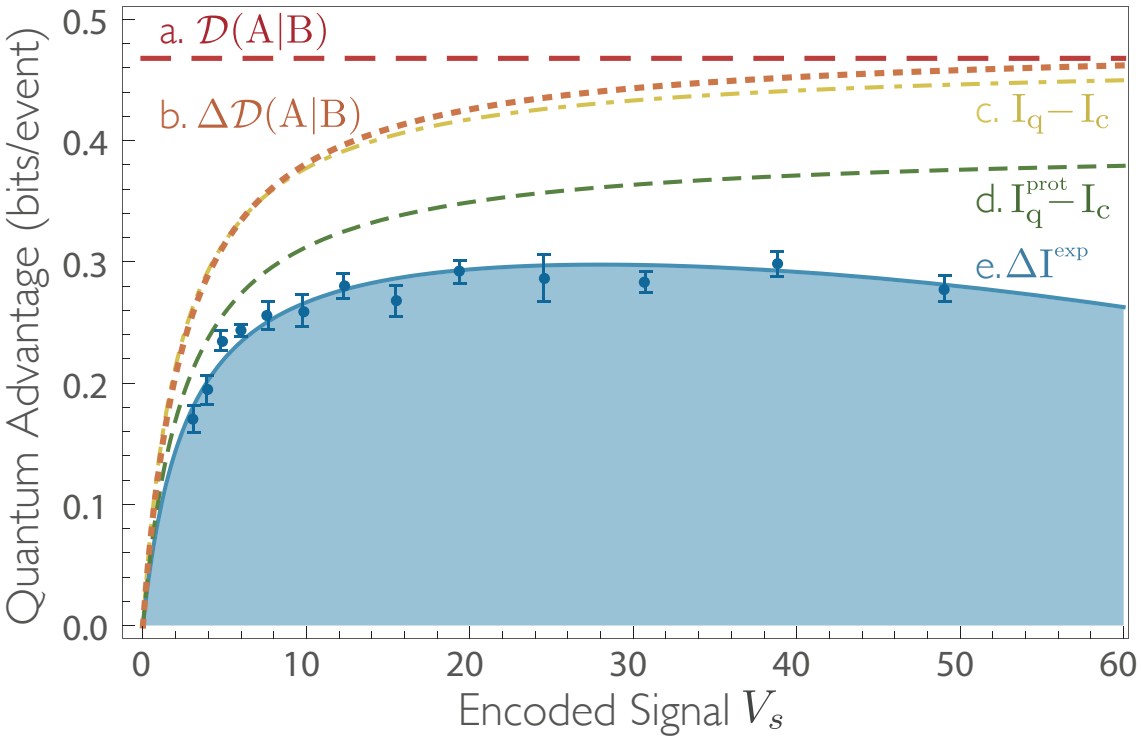

Figure 6.5: Plot of quantum advantage for a fixed resource state (with $V = 10.0 \pm 0.1$) with varying strength of the the encoded signal, V_s. (a) represents the maximum available amount of discord in the original resource ρ_{AB}, of which we progressively consume more of as we increase V_s (b). This bounds the maximum possible quantum advantage, assuming Bob can perform an ideal decoding protocol that saturates the Holevo limit (c). In the limit of large V_s, the encoding becomes maximal, and this tends to the discord of the original resource (a). The actual advantage that can be harnessed by our proposed protocol is represented by (d). In practice, experimental imperfections reduce the experimentally measured advantage to (e). The error bars denote a statistical confidence interval of 99%.

Figure 6.5 gives the experimentally observed 'advantage' for ρ_{AB} with varying strength of the encoded signal, V_s, with the discording noise is fixed at $V = 10.0 \pm 0.1$ (normalised to shot noise). The 'advantage' is defined as the difference between Bob's experimentally measured coherent performance, I_q^{exp}, and Bob's incoherent limit, I_c. The amount of Gaussian discord within the initial resource is fixed at $\mathcal{D}(A|B)$ (Figure 6.5.a). quantifies the quantum correlations shared by Alice and Bob, proving the resource that can be potentially harnessed to demonstrate an advantage. As Alice increases the strength of the encoded signal, progressively more of this initial resource is consumed (Figure 6.5.b), the amount of discord consumed bounding the potential quantum advantage (Figure 6.5.c). If Alice encodes maximally ($V_s \to \infty$) the potential quantum advantage is exactly equal to the discord in the initial state. For our identified decoding protocol (Figure 6.5.d) we observed a clear advantage of coherent protocol over the incoherent protocol over all encoded signals. In an ideal version of the our identified decoding protocol, the advantage would increase monotonically with the signal strength (Figure 6.5.d). With our imperfect experimental setup, there is initially an increase in the observed quantum advantage for increasing signal. However, there exists a saturation point around $V_s \sim 20$, beyond which the extra theoretical gain from increased signal strength is offset by the extra experimental imperfections in encoding. This is attributed to the nonlinear response of the electro-optic modulators and could also be explained by the limited dynamic range of the photodetectors. When we include these imperfections within our theoretical model, observations and theory agree (Figure 6.5.d).

More intuitively, why do coherent interactions help Bob in this protocol? In short, Bob's enhanced performance for the coherent scenario arises from the non-orthogonality of $\hat{X}$ and $\hat{P}$. This non-orthogonality ensures that when Alice encodes in both $\hat{X}$ and $\hat{P}$, the requirement for Bob to estimate both quadratures requires he always incurs a noise penalty. Bob is burdened with that penalty in both scenarios. It is the additional introduction of correlated 'discording' noise between the two partitions makes his 'incoherent' challenge harder. When Bob is restricted to individual measurements, his estimate of the discording noise is also limited by the precision to which he can estimate both quadratures. There is an inherent quantumness in the joint measurement of $\hat{X}$ and $\hat{P}$ that manifests as uniquely quantum correlations, inaccessible via local operations. The freedom to do a joint measurement, however, means he is no longer required to precisely measure and compensate for the correlated noise that degrades his estimate of Alice's encoded signal. Instead, the coherent interaction allows for (ideally) perfect cancellation of the unwanted noise, and retrieves a shot-noise limited measurement of $\hat{X}$ and $\hat{P}$.

In this Chapter, we have introduced and experimentally verified a protocol - albeit a bit artificial - that gives a clear operational interpretation to discord. We argue that correlations between two subsystems form a resource that allows one bi-partition to gain information about the other. This narrative naturally divides correlations into classical and quantum components. The former can be harnessed by an LOCC alone. The latter, quantified by discord, are inaccessible without the additional requirement of coherent interactions. Our demonstration within a separable quantum system, broadly thought of as *classical* [5], emphasises that discord is the resource of interest and that systems need

[5] As correlated, but separable, statistical mixture of coherent states

not be entangled benefit from joint measurements. As the capacity to coherently interact quantum systems is essential to quantum information, our results provide evidence that some of the advantages quantum information pertains over its classical counterpart can be attributed to its potential to harness discord.

The relation between the advantage of coherent interactions and non-classical correlations has also been studied within related paradigms. The thermodynamic variant of discord, for example, characterises the advantage of coherent interactions in energy extraction from a given quantum state [190, 191, 205]. Our protocol gives similar interpretation for the standard notion of discord in terms of information extraction.

This protocol leads to a direct practical application: By challenging untrusted parties to perform such tasks, Alice is capable of harnessing discord for the purpose of 'quantum processing authentication'. Alice can convince herself that an untrusted device is capable coherently interacting two spatially separated quantum systems. Unlike entangled states, Bell's inequalities, or tomography-based tests for coherent processing, our challenge does not require Alice to perform any quantum measurements or interact the systems herself. Our results demonstrate such test remain possible, even when we relax the initial resource to the point where it contains no entanglement.

The real significance of this work does not lie in the applicability of the protocol itself, but rather in its implications for established quantum information applications. The operational link provides between discord and coherent interactions allows potential reinterpretation of many existing protocols. If we regard our proposal as an attempt for Alice to communicate the contents of $\mathbf{K}$ to Bob via a pre-shared resource ρ_{AB}, the protocol resembles a quantum one-time pad with a generic resource [206]. Discord now plays a role in measuring the amount of extra information coherent interactions can unlock. In the special case where A and B are entangled, this protocol corresponds to dense coding, where the additional gain in communication rates is made possible by coherent interactions that decode information within the discordant correlations. Meanwhile, if we regard the task of trace estimation in DQC1 [171] as Bob's attempt to extract information Alice has encoded within the trace of a given unitary, our protocol may shed light on where the power in DQC1 originates. These connections are worth further investigation, and may not only lead to additional insight on the role of discord within a diverse range of applications, but also indicate whether we are already harnessing discord in many existing proposals without realising it.

Chapter 7

Summary and Future Outlooks

In Chapters 3 & 4 we discussed the problem of mimicking a conditional photon number measurement with homodyne and heterodyne detection. We showed that given an entangled bi-partite state, ρ_{ab}, with an informationally complete measurement of the field quadratures at a, one can reconstruct the state at b that corresponds to any projective measurement in the Fock basis at a. We illustrated this with the characterisation of the non-Gaussian photon subtracted squeezed vacuum states. We believe this is best understood as a variant upon two-mode tomography, where one can isolate the desired measurement outcome at a and reconstruct the corresponding conditional state at b, without needing to reconstruct the complete two-mode density operator. These techniques allow for complete characterisation of the outcome of a conditional measurement on a system, and might prove useful in systems where measurements of the DV of the system are limited or unavailable.

In Chapter 5 we instead found superior avenue for our desire to 'meddle' with Gaussian measurement records. The results of [132, 133] showed that, under certain constraints, a noiseless linear amplifier was equivalent to an appropriate post selection upon the measurement record. This result came with a clear application, itself born out of its relevance to QKD. In Chapter 5 we first obtain some general conditions on the limits of implementing arbitrary quantum operations on an ensemble via a conditional filtering of the measurement outcomes. We examined the performance of the MB-NLA in several different regimes, most notably the performance of the MB-NLA in restoring correlations degraded by a lossy channel. The MB-NLA allowed recovery of EPR violating correlations from a resource state, that after a lossy channel, no longer was no longer EPR correlated. We demonstrate that even in situations of exceptionally high loss (up to 99%) the MB-NLA could recover correlations that exceed the maximum achievable with a perfect squeezed state and an identical lossy channel. We also provided a proof-of-principle demonstration of the MB-NLA for QKD.

The MB-NLA is a promising tool for quantum communication applications. Whilst there are clear limitations on its applicability - crucially that the physical amplification must directly precede the measurement stage for the equivalence to hold - when applicable, it offers considerable advantages over a physical amplification. The results of Chapter 5 provide two promising avenues of future research. The first is the application of the MB-NLA to existing quantum communication and information protocols. Numerous CV-QKD implementations could be improved by its inclusion - whether they be prepare & measure

or entanglement-based implementations. Quantum teleportation and its variant, remote state preparation, look to be clear candidates for improved performance via post-selection, where post-selection could be integrated with the Bell measurement. The generality of the theory presented in Chapter 5 also provides a second avenue of research, the application of this recipe to the emulation of other desirable operations.

In Chapter 6 we examined the recently popularised measure of all quantum correlations, quantum discord. We introduced a simple protocol that provides a clear operational interpretation for discord: that it describes uniquely quantum information only accessible via coherent interactions. We demonstrate that under certain measurement constraints, discord between bipartite systems can be consumed to encode information that can only be accessed by coherent quantum interactions. We experimentally encoded information within the discordant correlations of two separable Gaussian states. The amount of extra information recovered by joint measurements, when compared to invidual measuremenets. is quantified and directly linked with the discord consumed during encoding. No entanglement exists at any point of this experiment. Thus we introduce and demonstrate an operational method to use discord as a physical resource.

There is still much work to be done on understanding the role of discord in quantum information and quantum communications. While discord certainly is a measure of quantum correlations, it is perhaps unsurprising that links between existing protocols and discord have not been forthcoming. While discord itself is defined in terms of established quantum information quantities, neither the quantum mutual information nor the one way classical correlation are constructive, in the sense of giving a concrete set of measurements for *both* parties to carry out in order to saturate the appropriate entropic quantities. Therefore if one takes a real world implementation of a quantum information protocol where the measurements involved are constrained by experimental technology, it becomes very difficult to relate the performance of such a protocol directly to the quantum discord. By constructing a specific protocol we were able to identify as well defined limit in which the improved performance could be related to the consumption of discord. What would be more desirable however, would to show whether it is discord or a discord-variant that this really is the driving quantum advantage of this and other protocols as they are implemented in the laboratory.

Appendix A

Conditioning Polynomials

In this Appendix, we demonstrate how the sampling polynomials can be obtained for arbitrary functions of $\hat{n}$. We provide two equivalent methods for doing this.

The first method involves writing the polynomial functions of the phase randomised quadrature operators $\bar{X}$ in terms of $\hat{n}$ via the creation and annihilation operators. These functions can then be inverted to solve for functions of $\hat{n}$ in term of $\bar{X}$.

The second method reproduce the same polynomials via measuring the moment of the Fock state by integration of Hermite polynomials.

Method 1

For an arbitrary function of $f(\hat{n})$, the analogue of equation (3.11) that we want to estimate using a phase randomised homodyne measurement would be

$$f\left(X_b^\theta\right) = \mathrm{Tr}\left[\hat{\rho}_{ab} f(\hat{n}) \otimes \left|X_b^\theta\right\rangle\!\left\langle X_b^\theta\right|\right] \tag{A.1}$$

$$= \mathrm{pr}(X_b^\theta)\mathrm{Tr}_a\left[\hat{\rho}_a(X_b^\theta) f(\hat{n})\right] \ , \tag{A.2}$$

where $\rho_a(X_b^\theta)$ is the state at a after tracing out b. Our goal is to find a function $F(\bar{X})$ corresponding to $f(\hat{n})$ such that

$$\mathrm{Tr}\left[\hat{\rho} f(\hat{n})\right] = \mathrm{Tr}\left[\hat{\rho} F(\bar{X})\right] \ , \tag{A.3}$$

where

$$F(\bar{X}) = \frac{1}{2\pi}\int_0^{2\pi} d\theta F(\hat{a}_\phi + \hat{a}_\phi^\dagger) \tag{A.4}$$

and $\hat{a}_\phi = \hat{a}\exp(-i\phi)$. Let us consider polynomial functions of $\bar{X}$ for which the monomials $\bar{X}^m$ for $m = 0, 1, \ldots$ forms a basis.

For all odd values of m, $\bar{X}^m$ vanish since the exponential terms $\exp(-i\phi)$ integrate to zero. For even m, the only terms in the expansion of $(\hat{a}_\phi + \hat{a}_\phi^\dagger)^m$ that are not a function of ϕ are those having equal numbers of $\hat{a}_\phi$ and $\hat{a}_\phi^\dagger$. These are the only terms that are non-zero after performing the integral in equation (A.4). They can be expressed as a function of $\hat{n}$ using the identity $\hat{a}^\dagger\hat{a} = \hat{n}$ and the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$.

We provide an example for the case of $m = 4$:

$$\bar{X}^4 = \frac{1}{2\pi}\int_0^{2\pi} d\phi \left(\hat{a}_\phi + \hat{a}_\phi^\dagger\right)^4 \tag{A.5}$$
$$= \hat{a}\hat{a}\hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a}^\dagger\hat{a}\hat{a}^\dagger + \hat{a}\hat{a}^\dagger\hat{a}^\dagger\hat{a} \tag{A.6}$$
$$+ \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{a} + \hat{a}^\dagger\hat{a}\hat{a}\hat{a}^\dagger \tag{A.7}$$
$$= 6\hat{n}^2 + 6\hat{n} + 3\,. \tag{A.8}$$

Results for various powers of $\bar{X}$ are tabulated below.

$$\bar{X}^0 = 1 \tag{A.9}$$
$$\bar{X}^2 = 1 + 2\hat{n} \tag{A.10}$$
$$\bar{X}^4 = 3 + 6\hat{n} + 6\hat{n}^2 \tag{A.11}$$
$$\bar{X}^6 = 15 + 40\hat{n} + 30\hat{n}^2 + 20\hat{n}^3 \tag{A.12}$$
$$\bar{X}^8 = 105 + 280\hat{n} + 350\hat{n}^2 + 140\hat{n}^3 + 70\hat{n}^4 \tag{A.13}$$
$$\bar{X}^{10} = 945 + 2898\hat{n} + 3150\hat{n}^2 + 2520\hat{n}^3 + 630\hat{n}^4 + 252\hat{n}^5 \tag{A.14}$$

Method 2

As an alternative method, we note that equation (A.3) must hold for arbitrary inputs $\hat{\rho}$. In particular, when $\hat{\rho} = |n\rangle\langle n|$ we get

$$\mathrm{Tr}\left[F(\bar{X})\,|n\rangle\langle n|\right] = f(n) \tag{A.15}$$
$$\iint dx d\tilde{x}\,\phi_n(x)\phi_n^\star(\tilde{x})F(x)\delta(x-\tilde{x}) = f(n) \tag{A.16}$$
$$\int dx\;|\phi_n(x)|^2\,F(x) = f(n) \tag{A.17}$$

where $\phi_n(x) = \langle n|x\rangle$ are the eigenstates of the harmonic oscillators. For $F(\bar{X}) = \bar{X}^m$, the associated functions of n would correspond to the m-th moment of the eigenstates.

While this integration can be performed directly using the Hermite polynomials, it turns out that it is more convenient to express $\bar{X}$ in terms of the annihilation and creation operators instead. As an example, we evaluate $f(n)$ when $F(\bar{X}) = \bar{X}^4$:

$$\langle n|\bar{X}^4|n\rangle = \frac{1}{2\pi}\int_0^{2\pi} d\phi\langle n|\left(\hat{a}_\phi + \hat{a}_\phi^\dagger\right)^4|n\rangle \tag{A.18}$$
$$= \langle n|\hat{a}\hat{a}\hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a}^\dagger\hat{a}\hat{a}^\dagger + \hat{a}\hat{a}^\dagger\hat{a}^\dagger\hat{a} \tag{A.19}$$
$$+ \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{a} + \hat{a}^\dagger\hat{a}\hat{a}\hat{a}^\dagger|n\rangle \tag{A.20}$$
$$= 6n^2 + 6n + 3 \tag{A.21}$$

which is the same result as Equation (A.8) as to be expected.

Appendix B

Proof of discord relations

B.1 Proof that Discord is a quantifier of quantum advantage

In this section, we explicitly prove that

$$\Delta\mathcal{D}(A|B) - \tilde{\mathcal{J}}(A,B) \leq \Delta I \leq \Delta\mathcal{D}(A|B), \tag{B.1}$$

which is equivalent to the statement $\mathcal{D}(A|B) - \tilde{I}(A,B) \leq \Delta I \leq \mathcal{D}(A|B) - \tilde{\mathcal{D}}(A|B)$, where $\tilde{I}(A,B)$ denotes the mutual information of $\tilde{\rho}_{AB}$. To do this, we make use of the Holevo information. Let $\mathbf{K}$ be a random variable that takes on value k with probability p_k. If each k is encoded in a quantum state with density operator ρ_k, then the maximum amount of information that may later be extracted about $\mathbf{K}$ is given by

$$S\left(\sum_k p_k\rho_k\right) - \sum_k p_k S(\rho_k). \tag{B.2}$$

when there are no constraints on what quantum operations are allowed.

To evaluate ΔI, we first introduce an additional scenario where Bob has no access to system B, and attempts to the find the best estimate of $\mathbf{K}$ using only measurements on system A. Let I_0 be Bob's maximum performance in this scenario. Recall that after encoding, the bipartite state between Alice and Bob is given by

$$\tilde{\rho}_{AB} = \sum_k p_k U_k \rho_{AB} U_k^\dagger \tag{B.3}$$

Since Bob has no access to B, we can trace over system B. Noting that U_k acts only on system A and is thus preserved under the partial trace, this results in codewords $U_k\rho_A U_k^\dagger$, which give Bob

$$I_0 = S(\tilde{\rho}_A) - S(\rho_A) \tag{B.4}$$

bits of accessible information by application of Eqn. (B.2). Here, $\tilde{\rho}_A = \mathrm{Tr}_B(\sum_k p_k U_k \rho_{AB} U_k^\dagger) = \sum_k p_k U_k \rho_A U_k^\dagger$. This case can be considered the control, i.e., the amount of information accessible to Bob when he cannot access any of the correlations between A and B.

We now compute I_q, the maximum extra information available to Bob when he can

implement arbitrary interaction between A and B. In this case, we have codewords $\rho_k = U_k \rho_{AB} U_k^\dagger$, such that $S(\rho_k) = S(\rho_{AB})$. This results in a Holevo information of $I_q = S(\tilde{\rho}_{AB}) - S(\rho_{AB})$. Therefore, the extra information Bob gains over the control case is

$$\Delta_q \equiv I_q - I_0 = S(\tilde{\rho}_{AB}) - S(\rho_{AB}) - S(\tilde{\rho}_A) + S(\rho_A) = I(A,B) - \tilde{I}(A,B). \tag{B.5}$$

I and $\tilde{I}$ respectively represent the total correlations between A and B before and after encoding. Thus, the advantage of being able to implement arbitrary two-body interactions over having no way to make use of system B coincides with the total amount of correlations consumed during the encoding process.

Similarly, we compute I_c, the maximum amount of information available to Bob by a single local measurement on each bipartition. We note that this constraint is equivalent to local operations and one-way communication, since multiple rounds of two way communication does not help Bob if does not measure a single partition more than once. In this scenario, the best Bob can do is to first measure either A or B in some basis $\{\Pi_b\}$, and make use of the classical output to improve his estimate of $\mathbf{K}$.

Consider first a measurement on B. Let Bob's resulting performance be $\overleftarrow{I}_c$. The state after measurement is $\rho_{A|b} = \mathrm{Tr}_B(\rho_{AB}\Pi_b)/q_b$ with probability q_b, where $q_b = \mathrm{Tr}(\rho_{AB}\Pi_b)$. Thus, Alice has effectively encoded $\mathbf{K}$ onto codewords $U_k \rho_{A|b} U_k^\dagger$. This results in $S(\sum_k p_k U_k \rho_{A|b} U_k^\dagger) - S(\rho_{A|b})$ bits of information accessible about $\mathbf{K}$ with probability q_b. To obtain the upper bound on how much information accessible to Bob, we maximize the expected value of the above subject to all possible measurements Bob could have made, thus

$$\overleftarrow{I}_c = \sup_{\{\Pi_b\}} \left(\sum_b q_b S(\tilde{\rho}_{A|b}) - \sum_b q_b S(\rho_{A|b}) \right). \tag{B.6}$$

Here, we have used the fact that Alice's application of U_k on system A, and Bob's measurement of system B act on different Hilbert spaces, and thus commute.

Now consider the case where Bob first measures A. We partition to total amount of information Bob can gain, $\overrightarrow{I}_c$ into two components; the component $\overrightarrow{I}_c^{(A)}$, that he gains directly from his measurement of system A; and $\overrightarrow{I}_c^{(B)}$, that he gains from the resulting collapsed quantum state on system B. Clearly $\overrightarrow{I}_c^{(A)} = I_0$, since Bob has not yet measured A.

To bound $\overrightarrow{I}_c^{(B)}$, note that measurement of $\rho_k = U_k \rho_{AB} U_k^\dagger$ on system A in a basis $\{\Pi_a\}$ is equivalent to measurement of ρ_{AB} in a rotated basis $\{U_k^\dagger \Pi_a U_k\}$. Thus, for each possible encoding k, the entropy of B after measurement is bounded below by $\sum_a \inf_{\{\Pi_a\}} S(\rho_{B|a})$. Therefore, the Holevo bound gives $\overrightarrow{I}_c^{(B)} \leq S(\rho_B) - \inf_{\{\Pi_a\}} \sum_a S(\rho_{B|a})$, and thus

$$\overrightarrow{I}_c \leq \overrightarrow{I}_c^{(A)} + \overrightarrow{I}_c^{(B)} \leq I_0 + S(\rho_B) - \inf_{\{\Pi_a\}} \sum_a S(\rho_{B|a}). \tag{B.7}$$

The optimal amount of information Bob can extract without coherent interactions is thus the maximal of $\overleftarrow{I}_c$ and $\overrightarrow{I}_c$, i.e., $I_c = \max\{\overrightarrow{I}_c, \overleftarrow{I}_c\}$. Noting that, $\Delta_c = I_c - I_0 = \max\{\overrightarrow{\Delta}_c, \overleftarrow{\Delta}_c\}$, where $\overrightarrow{\Delta}_c = \overrightarrow{I}_c - I_0$ and $\overleftarrow{\Delta}_c = \overleftarrow{I}_c - I_0$, we first evaluate $\overrightarrow{\Delta}_c$ and $\overleftarrow{\Delta}_c$ separately.

Substraction of Eq. (B.4) from Eq. (B.6) gives

$$\overleftarrow{\Delta}_c = S(\rho_A) + \sup_{\{\Pi_b\}} \left(\sum_b q_b S(\tilde{\rho}_{A|b}) - S(\tilde{\rho}_A) - \sum_b q_b S(\rho_{A|b}) \right). \tag{B.8}$$

Noting that $S[\tilde{\rho}_{A|b}] \leq S[\tilde{\rho}_A]$ since entropy can never increase under conditioning, we immediately find

$$\overleftarrow{\Delta}_c \leq S(\rho_A) - \inf_{\{\Pi_b\}} \sum_b q_b S(\rho_{A|b}) = J(A|B). \tag{B.9}$$

Also, rearranging Eq. B.8 gives

$$\begin{aligned}
\overleftarrow{\Delta}_c =& \sup_{\{\Pi_b\}} \left[S(\rho_A) - \sum_b q_b S(\rho_{A|b}) - \left(S(\tilde{\rho}_A) - \sum_b q_b S(\tilde{\rho}_{A|b}) \right) \right], \\
\geq& \sup_{\{\Pi_b\}} \left[S(\rho_A) - \sum_b q_b S(\rho_{A|b}) \right] - \sup_{\{\Pi_b\}} \left[S(\tilde{\rho}_A) - \sum_b q_b S(\tilde{\rho}_{A|b}) \right], \\
=& S(\rho_A) - \inf_{\{\Pi_b\}} \sum_b q_b S(\rho_{A|b}) - \left(S(\tilde{\rho}_A) - \inf_{\{\Pi_b\}} \sum_b q_b S(\tilde{\rho}_{A|b}) \right) \\
=& J(A|B) - \tilde{J}(A|B),
\end{aligned} \tag{B.10}$$

where $\tilde{\mathcal{J}}(A|B) = S(\tilde{\rho}_A) - \inf_{\{\Pi_b\}} \sum_b S(\tilde{\rho}_{A|b})$ denote the classical correlations of $\tilde{\rho}_{AB}$.

Therefore

$$\mathcal{J}(A|B) - \tilde{\mathcal{J}}(A|B) \leq \overleftarrow{\Delta}_c \leq \mathcal{J}(A|B), \tag{B.11}$$

Meanwhile, subtraction (B.4) from (B.7) gives $\overrightarrow{\Delta}_c = \overrightarrow{I}_c - I_0 \leq \overrightarrow{\mathcal{J}}$, therefore

$$\mathcal{J}(A|B) - \tilde{\mathcal{J}}(A|B) \leq \Delta_c \leq \max\{\mathcal{J}(A|B), \mathcal{J}(B|A)\}. \tag{B.12}$$

Subtraction of this equation from (B.5) immediately bounds the extra performance of coherent processing over its incoherent counterpart.

$$\min\{\mathcal{D}(A|B), \mathcal{D}(B|A)\} - \tilde{I}(A,B) \leq \Delta_q - \Delta_c \leq \Delta\mathcal{D}(A|B). \tag{B.13}$$

Applying our assumption that $\mathcal{D}(A|B) \leq \mathcal{D}(B|A)$, and the observation that $\Delta_q - \Delta_c = I_q - I_c = \Delta I$, results in

$$\Delta\mathcal{D}(A|B) - \tilde{\mathcal{J}}(A,B) \leq \Delta I \leq \Delta\mathcal{D}(A|B), \tag{B.14}$$

as required.

B.2 Example of Maximal Encodings

In this section, we prove the assertion made in the paper that there always exists maximal encodings. Recall that we may define maximal encodings as follows:

Definition 1 (Maximal Encoding) *Consider a bipartite quantum system with subsys-*

tems A and B that is described by density operator ρ_{AB}. The encoding of a random variable **K** *that takes on values k with probability p_k, by application of unitaries U_k is a maximal encoding if and only if $I\left(\sum_k p_k U_k\ \rho_{AB}\ U_k^\dagger\right) = 0$ for any ρ_{AB}, where $I(\rho_{AB})$ denotes the mutual information of ρ_{AB}.*

In particular, we prove the following:

Theorem 1 *Suppose Alice's bipartition has dimension d, then U_k is a maximum encoding whenever $\tilde{\rho}_{AB}$ is locally a maximally mixed state for any input state ρ_A on Alice's bipartition.*

Proof: To prove the result, it suffices to show that $\tilde{\rho}_{AB}$ is a product state. Consider an arbitrary projective measurement of the B subsystem in some basis $\{\Pi_b\}$ on $\tilde{\rho}_{AB}$. Since these measurements commute with U_k, it follows that $\mathrm{Tr}_B(\Pi_b \tilde{\rho}_{AB}) = \mathbf{I}/d$ for all j. Thus $\tilde{\rho}_{AB}$ must be a product state and the result follows. ■

Therefore, any encoding that looks like a maximally mixing channel is a maximal encoding. One example, on a system of qubits, for example, is application of the set of unitary transformations $\{I, \sigma_x, \sigma_z, \sigma_x\sigma_z\}$. In an continuous variable mode with annihilation operator a, application of an operation selected uniformly from the set of displacement operators $D(\alpha) = \exp\left(\alpha\hat{a}^\dagger + \alpha^*\hat{a}\right)$ is also a maximal encoding.

Bibliography

[1] M. Planck, "On the law of distribution of energy in the normal spectrum," Annalen der Physik **4**, 1 (1901).

[2] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," Annalen der Physik **322**, 132 (1905).

[3] R. H. Brown and R. Q. Twiss, "Correlation between Photons in two Coherent Beams of Light," Nature **177**, 27 (1956).

[4] L. M. Wolf and Emil, Optical Coherence and Quantum Optics (Cambridge University Press, 1995).

[5] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," J. of Math **58**, 345 (1936).

[6] C. E. Shannon, "A Mathematical Theory of Communication," The Bell System Technical Journal **379-423**, 1 (1948).

[7] R. P. Feynman, "Simulating physics with computers," International Journal of Theoretical Physics **21**, 467 (1982).

[8] S. Wiesner, "Conjugate coding," ACM Sigact News **15**, 78 (1983).

[9] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," Proceedings of International Conference on Computers, Systems and Signal Processing, Bangalore, India 1984.

[10] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," Nature **409**, 46 (2001).

[11] D. Browne and T. Rudolph, "Resource-Efficient Linear Optical Quantum Computation," Physical Review Letters **95**, 010501 (2005).

[12] M. Nielsen, "Optical Quantum Computation Using Cluster States," Physical Review Letters **93**, 040503 (2004).

[13] N. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. Ralph, and M. Nielsen, "Universal Quantum Computation with Continuous-Variable Cluster States," Physical Review Letters **97**, 110501 (2006).

[14] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM journal on computing **26**, 1484 (1997).

[15] L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," Physical Review Letters **79**, 325 (1997).

[16] G. Brassard, "Brief history of quantum cryptography: A personal perspective," p. 19 (2005).

[17] Z. Y. Ou, S. F. Pereira, and H. J. Kimble, "Realization of the Einstein-Podolsky-Rosen paradox for continuous variables in nondegenerate parametric amplification," Applied Physics B **55**, 265 (1992).

[18] S. L. Braunstein and H. J. Kimble, "Dense coding for continuous variables," Physical Review A **61**, 042302 (2000).

[19] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, "Quantum key distribution using gaussian-modulated coherent states," Nature **421**, 238 (2003).

[20] S. Braunstein and H. Kimble, "Teleportation of continuous quantum variables," Phys. Rev. Lett. **80**, 869 (1998).

[21] A. Furusawa, "Unconditional Quantum Teleportation," Science **282**, 706 (1998).

[22] J. Eisert, S. Scheel, and M. B. Plenio, "Distilling Gaussian States with Gaussian Operations is Impossible," Physical Review Letters **89**, 137903 (2002).

[23] J. Fiurasek, "Gaussian Transformations and Distillation of Entangled Gaussian States," Physical Review Letters **89**, 137904 (2002).

[24] G. Giedke and J. Ignacio Cirac, "Characterization of Gaussian operations and distillation of Gaussian states," Physical Review A **66**, 032316 (2002).

[25] J. Niset, J. Fiurasek, and N. J. Cerf, "No-Go Theorem for Gaussian Quantum Error Correction," Physical Review Letters **102**, 120501 (2009).

[26] U. Leonhardt, Measuring the Quantum State of Light (Cambridge Studies in Modern Optics) (Cambridge University Press, 2005).

[27] D. F. Walls and G. J. Milburn, Quantum Optics, Springer ed. (Springer, 2010).

[28] R. Loudon, The Quantum Theory of Light, Oxford university press, usa ed. (Oxford University Press, USA, 2000).

[29] C. C. Gerry and P. L. Knight, Introductory Quantum Optics (Cambridge Univ. Press, Cambridge, 2004).

[30] D. Stoler, "Equivalence classes of minimum uncertainty packets," Physical Review D **1**, 3217 (1970).

[31] H. P. Yuen, "Two-photon coherent states of the radiation field," Physical Review A **13**, 2226 (1976).

[32] H. P. Yuen and J. H. Shapiro, "Generation and detection of two-photon coherent states in degenerate four-wave mixing," Optics letters **4**, 334 (1979).

[33] R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz, and J. F. Valley, "Observation of Squeezed States Generated by Four-Wave Mixing in an Optical Cavity," Physical Review Letters **56**, 788 (1986).

[34] L. A. Wu, M. Xiao, and H. J. Kimble, "Squeezed states of light from an optical parametric oscillator," J. Opt. Soc. Am. B **4**, 1465 (1987).

[35] P. D. Drummond and Z. Ficek, Quantum Squeezing (Springer, 2004).

[36] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," Physical review **47**, 777 (1935).

[37] R. J. Glauber, "Coherent and incoherent states of the radiation field," Physical review **131**, 2766 (1963).

[38] E. Sudarshan, "Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams," Physical Review Letters **52**, 1962 (1963).

[39] K. Banaszek, "Operational theory of homodyne detection," Physical Review A (1997).

[40] L. M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, "Inseparability criterion for continuous variable systems," Physical Review Letters **84**, 2722 (2000).

[41] M. D. Reid and P. D. Drummond, "Quantum correlations of phase in nondegenerate parametric oscillation," Physical Review Letters **60**, 2731 (1988).

[42] M. D. Reid, "Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification," Physical Review A **40**, 913 (1989).

[43] K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," Physical Review A **40**, 2847 (1989).

[44] A. I. Lvovsky, "Continuous-variable optical quantum-state tomography," Reviews of Modern Physics **81**, 299 (2009).

[45] M. Paris and J. Řeháček, Quantum State Estimation (Springer, 2004).

[46] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum," Physical Review Letters **70**, 1244 (1993).

[47] H. M. Chrzanowski, J. Bernu, B. M. Sparkes, B. Hage, A. P. Lund, T. C. Ralph, P. K. Lam, and T. Symul, "Photon-number discrimination without a photon counter and its application to reconstructing non-Gaussian states," Physical Review A **84**, 050302 (2011).

[48] G. M. d'Ariano, C. Macchiavello, and M. Paris, "Detection of the density matrix through optical homodyne tomography without filtered back projection," Physical Review A **50**, 4298 (1994).

[49] H. Kühn, D. G. Welsch, and W. Vogel, "Determination of density matrices from field distributions and quasiprobabilities," Journal of Modern Optics **41**, 1607 (1994).

[50] U. Leonhardt, H. Paul, and G. M. d'Ariano, "Tomographic reconstruction of the density matrix via pattern functions," Physical Review A **52**, 4899 (1995).

[51] U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. Raymer, "Sampling of photon statistics and density matrix using homodyne detection," Optics communications **127**, 144 (1996).

[52] T. Richter, "Pattern functions used in tomographic reconstruction of photon statistics revisited," Physics Letters A **211**, 327 (1996).

[53] T. Richter, "Determination of field correlation functions from measured quadrature component distributions," Physical Review A **53**, 1197 (1996).

[54] G. M. d'Ariano, U. Leonhardt, and H. Paul, "Homodyne detection of the density matrix of the radiation field," Physical Review A **52**, R1801 (1995).

[55] V. Buzek and G. Drobny, "Quantum tomography via the MaxEnt principle," Journal of Modern Optics **47**, 2823 (2000).

[56] A. I. Lvovsky, "Iterative maximum-likelihood reconstruction in quantum homodyne tomography," Journal of Optics B: Quantum and Semiclassical Optics **6**, S556 (2004).

[57] N. J. Cerf and C. Adami, "Negative entropy and information in quantum mechanics," Physical Review Letters **79**, 5194 (1997).

[58] M. Horodecki, J. Oppenheim, and A. Winter, "Partial quantum information," Nature **436**, 673 (2005).

[59] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," Problemy Peredachi Informatsii **9**, 3 (1973).

[60] R. Loudon and P. Knight, "Squeezed light," Journal of Modern Optics **34**, 709 (1987).

[61] C. M. Caves and B. L. Schumaker, "New formalism for two-photon quantum optics. I. Quadrature phases and squeezed states," Physical Review A **31**, 3068 (1985).

[62] B. L. Schumaker and C. M. Caves, "New formalism for two-photon quantum optics. II. Mathematical foundation and compact notation," Physical Review A **31**, 3093 (1985).

[63] B. Yurke, "Use of cavities in squeezed-state generation," Physical Review A **29**, 408 (1984).

[64] H. Yuen and J. Shapiro, "Optical communication with two-photon coherent states–Part III: quantum measurements realizable with photoemissive detectors," Information Theory, IEEE Transactions on **26**, 78 (1980).

[65] H. P. Yuen and V. W. Chan, "Noise in homodyne and heterodyne detection," Optics letters **8**, 177 (1983).

[66] B. L. Schumaker, "Noise in homodyne detection," Optics letters **9**, 189 (1984).

[67] J. Shapiro, "Quantum noise and excess noise in optical homodyne and heterodyne receivers," Quantum Electronics, IEEE Journal of **21**, 237 (1985).

[68] T. Richter, "Determination of photon statistics and density matrix from double homodyne detection measurements," Journal of Modern Optics **45**, 1735 (1998).

[69] T. C. Ralph, W. J. Munro, and R. Polkinghorne, "Proposal for the Measurement of Bell-Type Correlations from Continuous Variables," Physical Review Letters **85**, 2035 (2000).

[70] T. C. Ralph, E. H. Huntington, and T. Symul, "Single-photon side bands," Physical Review A **77**, 1 (2008).

[71] K. Banaszek, "Maximum likelihood estimation of photon number distribution from homodyne statistics," arXiv.org (1997) 3901993119491712828related:PGdSixGqJjYJ.

[72] M. Vasilyev, S. Choi, P. Kumar, and G. D'Ariano, "Tomographic measurement of joint photon statistics of the twin-beam quantum state," Physical Review Letters **84**, 2354 (2000).

[73] J. G. Webb, T. C. Ralph, and E. H. Huntington, "Homodyne measurement of the average photon number," Physical Review A **73**, 1 (2006).

[74] N. B. Grosse, T. Symul, M. Stobińska, T. C. Ralph, and P. K. Lam, "Measuring Photon Antibunching from Continuous Variable Sideband Squeezing," Physical Review Letters **98**, 1 (2007).

[75] P. Van Loock, "Optical hybrid approaches to quantum information," Laser & Photonics Reviews **5**, 167 (2011).

[76] S. Deléglise, I. Dotsenko, C. Sayrin, J. Bernu, M. Brune, J.-M. Raimond, and S. Haroche, "Reconstruction of non-classical cavity field states with snapshots of their decoherence," Nature **455**, 510 (2008).

[77] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, "Quantum homodyne tomography of a two-photon Fock state," Physical Review Letters **96**, 213601 (2006).

[78] J. Neergaard-Nielsen, B. Nielsen, C. Hettich, K. Mølmer, and E. Polzik, "Generation of a superposition of odd photon number states for quantum information networks," Physical Review Letters **97**, 83604 (2006).

[79] K. Wakui, H. Takahashi, A. Furusawa, and M. Sasaki, "Photon subtracted squeezed states generated with periodically poled KTiOPO4," Opt. Express **15**, 3568 (2007).

[80] T. Gerrits, S. Glancy, T. Clement, and B. Calkins, "Generation of optical coherent-state superpositions by number-resolved photon subtraction from the squeezed vacuum," Physical Review A (2010).

[81] B. Yurke and D. Stoler, "Generating quantum mechanical superpositions of macroscopically distinguishable states via amplitude dispersion," Physical Review Letters **57**, 13 (1986).

[82] M. Dakna, T. Anhut, T. Opatrny, L. Knöll, and D. Welsch, "Generating Schrödinger-cat-like states by means of conditional measurements on a beam splitter," Physical Review A **55**, 3184 (1997).

[83] A. Lund, H. Jeong, T. Ralph, and M. Kim, "Conditional production of superpositions of coherent states with inefficient photon detection," Physical Review A **70**, 020101 (2004).

[84] M. Sasaki and S. Suzuki, "Multimode theory of measurement-induced non-Gaussian operation on wideband squeezed light: Analytical formula," Physical Review A **73**, 043807 (2006).

[85] K. Mølmer, "Non-Gaussian states from continuous-wave Gaussian light sources," Physical Review A **73**, 063804 (2006).

[86] A. Ourjoumtsev, F. Ferreyrol, R. Tualle-Brouri, and P. Grangier, "Preparation of non-local superpositions of quasi-classical light states," Nature Physics **5**, 1 (2009).

[87] J. Neergaard-Nielsen, M. Takeuchi, K. Wakui, H. Takahashi, K. Hayasaka, M. Takeoka, and M. Sasaki, "Optical continuous-variable qubit," Physical Review Letters **105**, 53602 (2010).

[88] N. Lee, H. Benichi, Y. Takeno, S. Takeda, J. Webb, E. Huntington, and A. Furusawa, "Teleportation of Nonclassical Wave Packets of Light," Science **332**, 330 (2011).

[89] O. Morin, C. Fabre, and J. Laurat, "Experimentally Accessing the Optimal Temporal Mode of Traveling Quantum Light States," Physical Review Letters **111**, 213602 (2013).

[90] H. Takahashi, J. S. Neergaard-Nielsen, M. Takeuchi, M. Takeoka, K. Hayasaka, A. Furusawa, and M. Sasaki, "Entanglement distillation from Gaussian input states," Nature Photonics **4**, 178 (2010).

[91] J. S. Neergaard-Nielsen, Y. Eto, C.-W. Lee, H. Jeong, and M. Sasaki, "Quantum tele-amplification with a continuous-variable superposition state," p. 1 (2013).

[92] S. Takeda, T. Mizuta, M. Fuwa, P. van Loock, and A. Furusawa, "Deterministic quantum teleportation of photonic quantum bits by a hybrid technique," Nature **500**, 315 (2013).

[93] U. Leonhardt, Measuring the quantum state of light (Cambridge Univ Pr, 1997).

[94] D. Sych, J. Řeháček, Z. Hradil, G. Leuchs, and L. L. Sánchez-Soto, "Informational completeness of continuous-variable measurements," Physical Review A **86**, 052123 (2012).

[95] M. G. A. Paris, "On density matrix reconstruction from measured distributions," Optics communications **124**, 277 (2003).

[96] R. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," Applied Physics B **31**, 97 (1983).

[97] E. D. Black, "An introduction to Pound–Drever–Hall laser frequency stabilization," American Journal of Physics **69**, 79 (2001).

[98] N. B. Grosse, Harmonic Entanglement and Photon Anti-Bunching, PhD thesis 2009.

[99] B. Hage, Purification and Distillation of Continuous Variable Entanglement, PhD thesis 2010.

[100] Y. Takeno, M. Yukawa, H. Yonezawa, and A. Furusawa, "Observation of-9 dB quadrature squeezing with improvement of phase stability in homodyne measurement," Optics Express **15**, 4321 (2007).

[101] B. M. Sparkes, H. M. Chrzanowski, D. P. Parrain, B. C. Buchler, P. K. Lam, and T. Symul, "A scalable, self-analyzing digital locking system for use on quantum optics experiments," Review of Scientific Instruments **82**, 075113 (2011).

[102] M. Munroe, D. Boggavarapu, M. Anderson, and M. Raymer, "Photon-number statistics from the phase-averaged quadrature-field distribution: Theory and ultrafast measurement," Physical Review A **52**, 924 (1995).

[103] D. F. McAlister and M. G. Raymer, "Correlation and joint density matrix of two spatial–temporal modes from balanced-homodyne sampling," Journal of Modern Optics **44**, 2359 (1997).

[104] D. Bozyigit, C. Lang, L. Steffen, J. M. Fink, C. Eichler, M. Baur, R. Bianchetti, P. J. Leek, S. Filipp, M. P. da Silva, A. Blais, and A. Wallraff, "Antibunching of microwave-frequency photonsobserved in correlation measurements usinglinear detectors," Nature Physics **7**, 154 (2010).

[105] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Z. Physik **43**, 172 (1927).

[106] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," Nature **299**, 802 (1982).

[107] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," Nature Photonics **5**, 222 (2011).

[108] C. M. Caves, "Quantum limits on noise in linear amplifiers," Phys. Rev. D **26**, 1817 (1982).

[109] C. M. Caves, J. Combes, Z. Jiang, and S. Pandey, "Quantum limits on phase-preserving linear amplifiers," Physical Review A **86**, 063802 (2012).

[110] C. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. Wootters, "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels," Phys. Rev. Lett. **76**, 722 (1996).

[111] M. Horodecki, P. Horodecki, and R. Horodecki, "Inseparable Two Spin- 12 Density Matrices Can Be Distilled to a Singlet Form," Phys. Rev. Lett. **78**, 574 (1997).

[112] D. Browne, J. Eisert, S. Scheel, and M. Plenio, "Driving non-Gaussian to Gaussian states with linear optics," Phys. Rev. A **67**, 062320 (2003).

[113] L. M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, "Long-distance quantum communication with atomic ensembles and linear optics," Nature **414**, 413 (2001).

[114] H. J. Kimble, "The quantum internet," Nature **453**, 1023 (2008).

[115] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. Shapiro, and S. Lloyd, "Gaussian quantum information," Rev. Mod. Phys. **84**, 621 (2012).

[116] J. Eisert, S. Scheel, and M. Plenio, "Distilling Gaussian states with Gaussian operations is impossible," Phys. Rev. Lett. **89**, 137903 (2002).

[117] J. Fiurášek, "Gaussian Transformations and Distillation of Entangled Gaussian States," Phys. Rev. Lett. **89**, 137904 (2002).

[118] J. Eisert, D. Browne, S. Scheel, and M. Plenio, "Distillation of continuous-variable entanglement with optical means," Annals of Physics **311**, 431 (2004).

[119] Y. Kurochkin, A. S. Prasad, and A. I. Lvovsky, "Distillation of The Two-Mode Squeezed State," Physical Review Letters **112**, 070402 (2014).

[120] T. C. Ralph and A. P. Lund, Quantum Communication Measurement and Computing Proceedings of 9th International Conference , 155 (2009).

[121] P. Marek and R. Filip, "Coherent-state phase concentration by quantum probabilistic amplification," Physical Review A **81**, 022302 (2010).

[122] J. Fiurášek, "Engineering quantum operations on traveling light beams by multiple photon addition and subtraction," Physical Review A **80**, 053822 (2009).

[123] G. Y. Xiang, T. C. Ralph, A. P. Lund, N. Walk, and G. J. Pryde, "Heralded noiseless linear amplification and distillation of entanglement," Nature Photonics **4**, 316 (2010).

[124] F. Ferreyrol, M. Barbieri, R. Blandino, S. Fossier, R. Tualle-Brouri, and P. Grangier, "Implementation of a nondeterministic optical noiseless amplifier," Phys. Rev. Lett. **104**, 123603 (2010).

[125] F. Ferreyrol, R. Blandino, M. Barbieri, R. Tualle-Brouri, and P. Grangier, "Experimental realization of a nondeterministic optical noiseless amplifier," Phys. Rev. A **83**, 063801 (2011).

[126] A. Zavatta, J. Fiurášek, and M. Bellini, "A high-fidelity noiseless amplifier for quantum light states," Nature Photonics **5**, 52 (2010).

[127] C. I. Osorio, N. Bruno, N. Sangouard, H. Zbinden, N. Gisin, and R. T. Thew, "Heralded photon amplification for quantum communication," Physical Review A **86**, 023815 (2012).

[128] S. Kocsis, G. Y. Xiang, T. C. Ralph, and G. J. Pryde, "Heralded noiseless amplification of a photon polarization qubit," Nat Phys **9**, 23 (2012).

[129] M. Mičuda, I. Straka, M. Miková, M. Dušek, N. J. Cerf, J. Fiurášek, and M. Ježek, "Noiseless Loss Suppression in Quantum Optical Communication," Physical Review Letters **109**, 180503 (2012).

[130] M. A. Usuga, C. R. Müller, C. Wittmann, P. Marek, R. Filip, C. Marquardt, G. Leuchs, and U. L. Andersen, "Noise-powered probabilistic concentration of phase information," Nature Physics **6**, 767 (2010).

[131] M. D. Reid, P. D. Drummond, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, "Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications," Reviews of Modern Physics **81**, 1727 (2009).

[132] J. Fiurasek and N. J. Cerf, "Gaussian postselection and virtual noiseless amplification in continuous-variable quantum key distribution," Physical Review A **86**, 060302 (2012).

[133] N. Walk, T. C. Ralph, T. Symul, and P. K. Lam, "Security of continuous-variable quantum cryptography with Gaussian postselection," Physical Review A **87**, 020303 (2013).

[134] K. Hellwig and K. Kraus, "Operations and measurements. II," Commun.Math. Phys. **16**, 142 (1970).

[135] F. Ferreyrol, N. Spagnolo, R. Blandino, M. Barbieri, and R. Tualle-Brouri, "Heralded processes on continuous-variable spaces as quantum maps," Physical Review A **86**, 062327 (2012).

[136] Supplementary Material.

[137] E. Prugovečki, "Information-theoretical aspects of quantum measurement," International Journal of Theoretical Physics **16**, 321 (1977).

[138] P. Busch and P. J. Lahti, "The determination of the past and the future of a physical system in quantum mechanics," Found Phys **19**, 633 (1989).

[139] N. Walk, A. P. Lund, and T. C. Ralph, "Nondeterministic noiseless amplification via non-symplectic phase space transformations," New Journal of Physics **15**, 073014 (2013).

[140] G. Giedke and J. I. Cirac, "Characterization of Gaussian operations and distillation of Gaussian states," Phys. Rev. A **66**, 032316 (2002).

[141] D. Pegg, L. Phillips, and S. Barnett, "Optical state truncation by projection synthesis," Phys. Rev. Lett. **81**, 1604 (1998).

[142] S. C. Armstrong, Experiments in Quantum Optics: Scalable Entangled States and Quantum Computation with Cluster States, PhD thesis 2014.

[143] N. Walk, A. P. Lund, and T. C. Ralph, "Nondeterministic noiseless amplification via non-symplectic phase space transformations," New J. Phys. **15**, 073014 (2013).

[144] H. M. Wiseman, S. J. Jones, and A. C. Doherty, "Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox," Physical Review Letters **98**, 140402 (2007).

[145] E. Cavalcanti, S. Jones, H. Wiseman, and M. Reid, "Experimental criteria for steering and the Einstein-Podolsky-Rosen paradox," Physical Review A **80**, 032112 (2009).

[146] D. H. Smith, G. Gillett, M. P. de Almeida, C. Branciard, A. Fedrizzi, T. J. Weinhold, A. Lita, B. Calkins, T. Gerrits, H. M. Wiseman, S. W. Nam, and A. G. White, "Conclusive quantum steering with superconducting transition-edge sensors," Nature Communications **3**, 625 (2011).

[147] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. M. Wiseman, R. Ursin, and A. Zeilinger, "Loophole-free Einstein–Podolsky–Rosen experiment via quantum steering," New Journal of Physics **14**, 053030 (2012).

[148] A. J. Bennet, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, "Arbitrarily Loss-Tolerant Einstein-Podolsky-Rosen Steering Allowing a Demonstration over 1 km of Optical Fiber with No Detection Loophole," Physical Review X **2**, 031003 (2012).

[149] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, "One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering," Physical Review A **85**, 010301 (2012).

[150] R. García-Patrón, "Quantum Information with Optical Continuous Variables: from Bell Tests to Key Distribution," Universit e Libre de Bruxelles (2007).

[151] N. J. Cerf and P. Grangier, "From quantum cloning to quantum key distribution with continuous variables: a review (Invited)," J. Opt. Soc. Am. B **24**, 324 (2007).

[152] M. Navascués, F. Grosshans, and A. Acín, "Optimality of Gaussian Attacks in Continuous-Variable Quantum Cryptography," Phys. Rev. Lett. **97**, 190502 (2006).

[153] R. García-Patrón and N. Cerf, "Unconditional Optimality of Gaussian Attacks against Continuous-Variable Quantum Key Distribution," Phys. Rev. Lett. **97**, 190503 (2006).

[154] R. Renner and J. Cirac, "de Finetti Representation Theorem for Infinite-Dimensional Quantum Systems and Applications to Quantum Cryptography," Phys. Rev. Lett. **102**, 110504 (2009).

[155] S. Fossier, E. Diamanti, T. Debuisschert, A. Villing, R. Tualle-Brouri, and P. Grangier, "Field test of a continuous-variable quantum key distribution prototype," New Journal of Physics **11**, 045023 (2009).

[156] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. Cerf, R. Tualle-Brouri, and S. McLaughlin, "Quantum key distribution over 25km with an all-fiber continuous-variable system," Phys. Rev. A **76**, 042305 (2007).

[157] L. Madsen, V. Usenko, M. Lassen, R. Filip, and U. Andersen, "Continuous variable quantum key distribution with modulated entangled states," Nat Comms **3**, 1083 (2012).

[158] A. Leverrier, F. Grosshans, and P. Grangier, "Finite-size analysis of a continuous-variable quantum key distribution," Phys. Rev. A **81**, 062343 (2010).

[159] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, "Experimental demonstration of long-distance continuous-variable quantum key distribution," Nature Photon **7**, 378 (2013).

[160] F. Grosshans, N. Cerf, P. Grangier, J. Wenger, and R. Tualle-Brouri, "Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables," Quantum Inf. Comput. **3**, 535 (2003).

[161] R. García-Patrón and N. Cerf, "Continuous-Variable Quantum Key Distribution Protocols Over Noisy Channels," Physical Review Letters **102**, 130501 (2009).

[162] R. Blandino, A. Leverrier, M. Barbieri, J. Etesse, P. Grangier, and R. Tualle-Brouri, "Improving the maximum transmission distance of continuous-variable quantum key distribution using a noiseless amplifier," Phys. Rev. A **86**, 012327 (2012).

[163] S. Pandey, Z. Jiang, J. Combes, and C. M. Caves, "Quantum limits on probabilistic amplifiers," Physical Review A **88**, 033852 (2013).

[164] S.-Y. Lee, S.-W. Ji, H.-J. Kim, and H. Nha, "Enhancing quantum entanglement for continuous variables by a coherent superposition of photon subtraction and addition," Phys. Rev. A **84**, 012302 (2011).

[165] H.-J. Kim, S.-Y. Lee, S.-W. Ji, and H. Nha, "Quantum linear amplifier enhanced by photon subtraction and addition," Phys. Rev. A **85**, 013839 (2012).

[166] M. Barbieri, N. Spagnolo, M. G. Genoni, F. Ferreyrol, R. Blandino, M. G. A. Paris, P. Grangier, and R. Tualle-Brouri, "Non-Gaussianity of quantum states: An experimental test on single-photon-added coherent states," Physical Review A **82**, 063833 (2010).

[167] A. Zavatta, S. Viciani, and M. Bellini, "Quantum-to-Classical Transition with Single-Photon-Added Coherent States of Light," Science **306**, 660 (2004).

[168] C. H. Bennett and S. J. Wiesner, "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states," Physical Review Letters **69**, 2881 (1992).

[169] A. K. Ekert, "Quantum cryptography based on Bell's theorem," Physical Review Letters **67**, 661 (1991).

[170] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," **175**, 8 (1984).

[171] E. Knill and R. Laflamme, "Power of One Bit of Quantum Information," Physical Review Letters **81**, 5672 (1998).

[172] A. Datta and G. Vidal, "Role of entanglement and correlations in mixed-state quantum computation," Physical Review A **75**, 042310 (2007).

[173] A. Datta, A. Shaji, and C. M. Caves, "Quantum Discord and the Power of One Qubit," Physical Review Letters **100**, 050502 (2008).

[174] B. P. Lanyon, M. Barbieri, M. P. Almeida, and A. G. White, "Experimental Quantum Computing without Entanglement," Physical Review Letters **101**, 200501 (2008).

[175] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, "Quantum nonlocality without entanglement," Physical Review A **59**, 1070 (1999).

[176] H. Ollivier and W. Zurek, "Quantum discord: A measure of the quantumness of correlations," Physical Review Letters **88**, 17901 (2001).

[177] L. Henderson and V. Vedral, "Classical, quantum and total correlations," Journal of Physics A: Mathematical and General **34**, 6899 (2001).

[178] R. Laflamme, D. Cory, C. Negrevergne, and L. Viola, "NMR Quantum Information Processing and Entanglement," Quantum Information & Computation **2**, 166 (2002).

[179] V. Vedral, "The Elusive Source of Quantum Speedup," Foundations of Physics **40**, 1141 (2010).

[180] C. A. Rodríguez-Rosario, K. Modi, A.-m. Kuah, A. Shaji, and E. C. G. Sudarshan, "Completely positive maps and classical correlations," Journal of Physics A: Mathematical and Theoretical **41**, 205301 (2008).

[181] M. Piani, P. Horodecki, and R. Horodecki, "No-Local-Broadcasting Theorem for Multipartite Quantum Correlations," Physical Review Letters **100**, 090502 (2008).

[182] S. Luo, "On Quantum No-Broadcasting," Letters in Mathematical Physics **92**, 143 (2010).

[183] B. Tomasello, D. Rossini, A. Hamma, and L. Amico, "Ground-state factorization and correlations with broken symmetry," EPL (Europhysics Letters) **96**, 27002 (2011).

[184] P. Giorda and M. Paris, "Gaussian quantum discord," Phys. Rev. Lett. **105**, 20503 (2010).

[185] G. Adesso and A. Datta, "Quantum versus classical correlations in Gaussian states," Physical Review Letters **105**, 030501 (2010).

[186] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, "The classical-quantum boundary for correlations: discord and related measures," Reviews of Modern Physics **84**, 1655 (2012).

[187] H. Ollivier and W. H. Zurek, "Quantum Discord: A Measure of the Quantumness of Correlations," Phys. Rev. Lett. **88**, 017901 (2001).

[188] A. Datta, S. Flammia, and C. Caves, "Entanglement and the power of one qubit," Physical Review A **72**, 042316 (2005).

[189] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, and A. Acin, "Almost all quantum states have nonclassical correlations," Physical Review A **81**, 052318 (2010).

[190] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, "Thermodynamical approach to quantifying quantum correlations," Physical Review Letters **89**, 180402 (2002).

[191] W. H. Zurek, "Quantum discord and Maxwell's demons," Physical Review A **67**, 012320 (2003).

[192] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, and M. Piani, "Operational interpretations of quantum discord," Physical Review A (2011).

[193] V. Madhok and A. Datta, "Quantum Discord as a Resource in Quantum Communication," International Journal of Modern Physics B **27**, 1345041 (2013).

[194] B. Dakić, Y. O. Lipp, X. Ma, M. Ringbauer, S. Kropatschek, S. Barz, T. Paterek, V. Vedral, A. Zeilinger, Č. Brukner, and P. Walther, "Quantum discord as resource for remote state preparation," Nature Physics **8**, 666 (2012).

[195] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," Communications in mathematical physics **269**, 107 (2007).

[196] S. Luo, "Quantum discord for two-qubit systems," Physical Review A **77**, 042303 (2008).

[197] D. Girolami and G. Adesso, "Quantum discord for general two-qubit states: Analytical progress," Physical Review A **83**, 052108 (2011).

[198] S. Rahimi-Keshari, C. M. Caves, and T. C. Ralph, "Measurement-based method for verifying quantum discord," Physical Review A **87**, 012119 (2013).

[199] P. Giorda, M. Allegra, and M. G. A. Paris, "Quantum discord for Gaussian states with non-Gaussian measurements," Physical Review A **86**, 052328 (2012).

[200] M. M. Wilde, S. Guha, S.-H. Tan, and S. Lloyd, "Explicit capacity-achieving receivers for optical communication and quantum reading," p. 551 (2012).

[201] A. S. Holevo, A. Mari, and V. Giovannetti, "Quantum state majorization at the output of bosonic Gaussian channels," Nature Communications **5**, 1 (2014).

[202] S. Pirandola, G. Spedalieri, S. L. Braunstein, N. J. Cerf, and S. Lloyd, "Optimality of Gaussian Discord," Physical Review Letters **113**, 140405 (2014).

[203] J. Lodewyck and P. Grangier, "Tight bound on the coherent-state quantum key distribution with heterodyne detection," Physical Review A **76**, 022332 (2007).

[204] J. Sudjana, L. Magnin, R. García-Patrón, and N. Cerf, "Tight bounds on the eavesdropping of a continuous-variable quantum cryptographic protocol with no basis switching," Physical Review A **76**, 052301 (2007).

[205] A. Brodutch and D. R. Terno, "Quantum discord, local operations, and Maxwell's demons," Physical Review A **81**, 062103 (2010).

[206] B. Schumacher and M. Westmoreland, "Quantum mutual information and the one-time pad," Physical Review A **74**, 042305 (2006).