



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Quantum Enhanced Metrology

*Quantum Mechanical Correlations
and Uncertainty Relations*

Anthony Hayes

Submitted for the degree of Doctor of Philosophy

University of Sussex

April 2018

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Anthony Hayes

UNIVERSITY OF SUSSEX

ANTHONY HAYES, DOCTOR OF PHILOSOPHY

QUANTUM ENHANCED METROLOGYQUANTUM MECHANICAL CORRELATIONS AND UNCERTAINTY RELATIONSABSTRACT

The foundational theory of quantum enhanced metrology for parameter estimation is of fundamental importance to the progression of science and technology as the scientific method is built upon empirical evidence, the acquisition of which is entirely reliant on measurement. Quantum mechanical properties can be exploited to yield measurement results to a greater precision (lesser uncertainty) than that which is permitted by classical methods. This has been mathematically demonstrated by the derivation of theoretical bounds which place a fundamental limit on the uncertainty of a measurement. Furthermore, quantum metrology is of immediate interest in the application of quantum technologies since measurement plays a central role.

This thesis focuses on the role of quantum correlations and uncertainty relations which govern the precision bounds. We show how correlations can be distributed amongst limited resources in realistic scenarios, as permitted by current experimental capabilities, to achieve higher precision measurements than current approaches. This is extended to the setting of multiparameter estimation in which we demonstrate a more technologically feasible method of correlation distribution than those previously posited which perform as well as, or worse than, our scheme.

Furthermore, a quantum metrology protocol is typically comprised of three stages: probe state preparation, sensing and then readout, where the time required for the first and last stages is usually neglected. We consider the more realistic sensing scenario of time being a limited resource which is divided amongst the three stages and demonstrate the most efficient use of this resource.

Additionally, we take an information theoretic approach to quantum mechanical uncertainty relations and derive a one-parameter class of uncertainty relations which supplies more information about the quantum mechanical system of interest than conventional uncertainty relations. Finally, we demonstrate how we can use this class of uncertainty relations to reconstruct information of the state of the quantum mechanical system.

Acknowledgements

This thesis is dedicated to all those; teachers, colleagues, friends and family that have taught, guided, encouraged and facilitated my glimpse of an understanding of the sublime structure that describes our universe. To those, I cannot begin to express my gratitude - this thesis is as much yours as it is my own.

Contents

List of Figures	xiv
1 Introduction	1
1.1 Thesis Outline	1
1.2 Quantum Mechanics	2
1.2.1 Entanglement	2
1.2.2 Dirac Notation	3
1.2.3 The Postulates of Quantum Mechanics	5
1.2.4 Mixed states and Density Matrices	7
1.2.5 The Schrödinger, Heisenberg and Interaction Pictures	9
1.3 Metrology: An Overview of the Science of Measurement	10
2 Bosonic Systems	13
2.1 The Quantum Harmonic Oscillator	13
2.1.1 Field Quadrature Operators	13
2.1.2 The Number Operator	15
2.2 Coherent States	16
2.2.1 The Wigner Function	19
2.3 Squeezed States	20
2.3.1 Quadrature Squeezing	20
2.3.2 Practical Generation of Squeezed Light	24
2.4 Gaussian and Non-Gaussian States	25
2.4.1 Gaussian Probability Distributions	25
2.4.2 Non-Gaussian States	27
3 Fermionic Systems	32
3.1 Two-Level System	32
3.2 N Spin-1/2 System	34

3.3	Coherent Spin States and Rotations	37
3.4	Spin Squeezing	39
3.4.1	One-Axis Twisting	41
3.4.2	Two-Axis Twisting	43
3.5	Bosonic limit	44
4	Quantum Enhanced Metrology and Parameter Estimation	46
4.1	Measurements	46
4.1.1	Classical and Quantum Measurements	46
4.1.2	Projective Measurements	47
4.1.3	Entanglement	49
4.1.4	von Neumann Measurement Schemes	50
4.2	Uncertainty Measures	52
4.2.1	Helstrom-Holevo Lower bound	52
4.2.2	The Fisher Information	54
4.3	Quantum Fisher Information	56
4.4	Mach-Zehnder Interferometer	58
4.4.1	Independent Photons	58
4.4.2	Coherent State	60
4.4.3	Entangled Photons	62
4.4.4	QFI Example	63
5	Quantum Correlations in Quantum Enhanced Metrology	66
5.1	Quantum Correlations	66
5.1.1	QFI for Path-Symmetric Pure States	68
5.2	Squeezed-Entangled State	70
5.3	Squeezed-Cat State	72
5.4	Multiparameter Estimation	74
5.5	Local vs. Global Strategies	75
5.5.1	Parallel Interferometers	76
5.5.2	Multimode Quantum Enhanced Imaging	79
5.5.3	General Procedure	80
6	Quantum Metrology with Time as a Limited Resource: Concurrent State Preparation and Readout	82
6.1	Magnetic Field Sensing and Two-Axis Twisting	82

6.1.1	Scheme A	84
6.1.2	Scheme B	84
6.1.3	Scheme C	88
6.2	Magnetic Field Sensing and One-Axis Twisting	90
6.2.1	Scheme B'	91
6.2.2	Scheme C'	95
7	Entropy in Information Theory	97
7.1	Hartley and Shannon Entropy	97
7.1.1	Hartley Entropy	97
7.1.2	Shannon Entropy	98
7.2	Differential Entropy	101
7.3	Rényi Entropy	104
7.3.1	Discrete Rényi Entropy	104
7.3.2	Geometric Interpretation of the Discrete Rényi Entropy	106
7.3.3	Continuous Rényi Entropy	108
8	A One-Parameter Class of Uncertainty Relations	110
8.1	Uncertainty Relations	110
8.1.1	Variance-Based Uncertainty Relations	110
8.1.2	Entropic Uncertainty Relations	112
8.1.3	Entropy Power Uncertainty Relations	112
8.2	REPUR for Gaussian States	114
8.3	REPUR for Non-Gaussian States	122
8.3.1	Cat States	122
8.3.2	Superstition State of Vacuum and Squeezed Vacuum	125
8.3.3	Uninformative VUR	128
8.3.4	Discussion	130
9	Information Scan and Reconstruction of Quantum States	132
9.1	Variance is a Deceptive Measure of Uncertainty	132
9.1.1	Example 1	132
9.1.2	Example 2	134
9.1.3	Entropies of Heuristic Examples	136
9.2	The Information Scan	137
9.3	Reconstruction of Quantum States	139

9.3.1	Relating REPs to Cumulants	142
9.4	BCS (Equimeasurable to Gaussian) Example	144
9.5	Analytic Expressions for the Cumulants of the UCS	146
9.5.1	Reconstruction	147
10	Conclusion	149
	Bibliography	152
A	Accurate Numerical Higher Order Derivatives	165
B	Optical System	168
B.1	Lossless	168
B.1.1	Physical Setup	171
B.2	Lossy	173
B.2.1	Sequential Scheme	173
B.2.2	Simultaneous	180

List of Figures

2.1	Plots of Wigner functions for vacuum state (<i>left</i>) coherent state with $\alpha = 2$ (<i>centre</i>) and coherent state with $\alpha = \sqrt{2} + i\sqrt{2}$ (<i>right</i>). Taking $\alpha \in \mathbb{R}$ displaces the vacuum state along the position axis whereas taking $\alpha \in \mathbb{C}$ displaces the vacuum in the plane of phase space. Note that the vacuum state is equivalent to a coherent state with $\alpha = 0$	18
2.2	Plots of Wigner functions for the squeezed vacuum state with $\phi = 0$ (<i>left</i>) $\phi = \pi$ (<i>centre</i>) and $\phi = \pi/2$ (<i>right</i>) and $r = 1/2$ for all. This corresponds to squeezing in the position, momentum and general quadrature axis respectively.	22
2.3	Plots of Wigner functions for the vacuum state (<i>left</i>), the squeezed vacuum state with $r = \frac{1}{2}$ and $\phi = 0$ (<i>centre</i>) and the squeezed-displaced vacuum state with $r = \frac{1}{2}$, $\phi = 0$ and $\alpha = 2$ (<i>right</i>).	24
2.4	Plots of position and momentum probability distributions for a coherent state with $\alpha = 2$ (<i>left</i>) and for a squeezed vacuum state with $r = 1/2$ and $\phi = 0$ (<i>right</i>). All are Gaussian PDFs.	26
2.5	Plots of Wigner functions for the Fock states $ n\rangle$ with $n = 1$ (<i>left</i>) $n = 2$ (<i>centre</i>) and $n = 4$ (<i>right</i>).	27
2.6	Probability distribution functions for the Fock states $ n\rangle$ with $n = 1$ (<i>left</i>) $n = 2$ (<i>centre</i>) and $n = 4$ (<i>right</i>). All PDFs display an overall non-Gaussian structure.	28
2.7	Plots of Wigner functions for odd cat states with $\phi = 0$ and $\tilde{\phi} = -\pi$ (<i>left</i>) and with $\phi = \pi/2$ and $\tilde{\phi} = -\pi/2$ (<i>right</i>).	29
2.8	Probability distribution functions for odd cat states with $\phi = 0$ and $\tilde{\phi} = -\pi$ (<i>left</i>) and with $\phi = \pi/2$ and $\tilde{\phi} = -\pi/2$ (<i>right</i>). Both PDFs display an overall non-Gaussian structure.	31

3.1	The Bloch sphere representation of a two-level system such as a spin-1/2 particle. The choice of axis that represent the $ 0\rangle$ and $ 1\rangle$ state are arbitrary but conventionally taken to be along the positive and negative z axis. . . .	34
3.2	Representation of a CSS on the unit sphere.	38
3.3	Schematic representation of Spin Squeezing of a N spin-1/2 system which constitutes a coherent spin state (CSS). The squeezing process introduces quantum mechanical correlations into the system.	39
3.4	Plots of the spin Q-function for the Dicke state $ N/2, m\rangle$ with $N = 20$ and $m = 0$ (<i>left</i>), the CSS $ \downarrow\rangle^{\otimes N}$ with $N = 10$ (<i>centre</i>) and the squeezed spin state (<i>right</i>) where the OAT Hamiltonian (3.54) has been applied to the CSS.	43
3.5	Plots of the spin Q-function for the CSS $ \downarrow\rangle^{\otimes N}$ with $N = 10$ (<i>left</i>) and the squeezed spin state (<i>right</i>) where the TAT Hamiltonian (3.59) has been applied to the CSS.	44
4.1	A Mach-Zehnder interferometer (MZI) consisting of two inputs states of light, a linear crystal referred to as a beam splitter (BS) which allows the states of light to interact, a phase shift ϕ on one arm (relative to the other arm), a second beam splitter which allows for an interference interaction between the now phase-shifted light and finally two photodetectors output of the MZI.	58
4.2	An MZI with a coherent state input in mode a and a vacuum state in mode b	60
5.1	The gravitational wave detection scheme as proposed by C. Caves [74] using a Michelson interferometer. A squeezed vacuum state $ \zeta\rangle$ is injected into the previously unused port thus introducing intramode correlations into the scheme.	67
5.2	The general interferometry scheme used in the following investigation. Path symmetric states are used to exploit the natural symmetries of the MZI.	68
5.3	[79] Plots of the quantum Fisher information for the NOON state, the Squeezed Entangled State (SES), the Separable Squeezed Vacuum (SSV), the Caves' Squeezed-Vacuum-Coherent-State (SVCS) and the Squeezed Cat State (SCS).	71

- 5.4 Plots of the Wigner functions for the Cat State (left), the Squeezed Cat State (centre) and the phase-shifted Squeezed Cat State (right). It is evident that squeezing the cat state reduces the overlap of the fringes of the states before and after a phase shift. 73
- 5.5 The general setup of a quantum optical multiparameter estimation scheme in which the “preparation” stage involves the creation of an M -mode probe state which is subject the $\vec{\theta} = (\theta_1, \dots, \theta_M)$ linear phase shifts and subsequently the multiple parameters ϕ_i for $i \in \{1, \dots, d\}$, which are functions of $\vec{\theta}$, are read out in the “measurement” stage. 74
- 5.6 A network of quantum sensors modelled by a multimode collection of parallel Mach-Zehnder interferometers, each of which yields a phase difference to be measured. 76
- 6.1 In scheme A , the magnetic field is applied over the entire time τ , by the operation $\hat{D}_\omega(\tau)$. The spins remain in a separable state throughout. In scheme B , the two-axis twisting operation $\hat{S}_\eta(t')$ generates a sensitive entangled state before exposure to the magnetic field through $\hat{D}_\omega(t)$. In scheme C the spins are subject to the operation $\hat{U}_{\omega,\eta}(t')$ (as defined in section 6.1.3) which exposes them to the magnetic field during the twisting operation. Each scheme ends with a measurement of the final state $|\psi_i\rangle$ ($i \in \{A, B, C\}$), which we assume can be done in a negligible time. For a fair comparison, between the three schemes, each is constrained by the time τ 83
- 6.2 Bloch Sphere representation (with inverted z axis for ease of representation) of a) a rotation operation around the y axis, as given by equation (6.4), applied to the CSS $|\downarrow\rangle^{\otimes N}$ and b) a TAT operation as given by equation (6.6). 86
- 6.3 These plots show that for a sufficiently small value of $\eta\tau$ (e.g. $\eta\tau = 0.4$ in the upper plots), scheme B gives no improvement over scheme A . For a sufficiently large value of $\eta\tau$ (e.g. $\eta\tau = 4$ in the lower plots), both scheme B and scheme C can give a better sensitivity than scheme A (i.e the two-axis twisting state preparation is worthwhile), if the sensing time t/τ is optimised. 87

- 6.4 The upper plots show the optimised sensitivity bound $\max_{t/\tau}(\sqrt{F_i}/\tau)$ as a function of $\eta\tau$, and the lower plots show the corresponding optimal sensing times, $(t/\tau)_{\text{opt}}$. Comparison of schemes reveals that scheme B gives no advantage over scheme A for $\eta\tau \lesssim 0.5$. Scheme C , however, does better than scheme A for all values of $\eta\tau$, although the advantage vanishes as $\eta\tau \rightarrow 0$. 87
- 6.5 In scheme B' , the one-axis twisting operation $\hat{T}_\chi(t')$ generates a spin squeezed state before exposure to the magnetic field through $\hat{D}_\omega(t)$. The “echo” (anti-squeezing) operation $\hat{T}_\chi^\dagger(t') = \hat{T}_{-\chi}(t')$ (as introduced in section 6.2.1) is applied before the final measurement. In scheme C' the spins are exposed to the magnetic field during the OAT and echo operations. For a fair comparison, each protocol is constrained by the time τ 90
- 6.6 Bloch Sphere representation (with inverted z axis for ease of representation) of a OAT operation, as given by equation (6.24), on the CSS $|\downarrow\rangle^{\otimes N}$. In contrast to TAT, the resulting reduction in quantum fluctuations is not in line with any of the primary axes. 92
- 6.7 These plots show that for a sufficiently small value of $\chi\tau$ (e.g. $\chi\tau = 4$ and $N = 10$), scheme A gives a better sensitivity than scheme B' and scheme C' . For a sufficiently large value of $\chi\tau$ (e.g. $\chi\tau = 50$ and $N = 10$ or $N = 100$), both scheme B' and scheme C' give a better sensitivity than scheme A , i.e the spin squeezing is worthwhile. 94
- 6.8 The upper plots show the optimised sensitivity bound $\max_{t/\tau} [(\sqrt{\nu}\tau\delta\omega)^{-1}]$ as a function of $\chi\tau$. These plots are optimised over time, but not over measurements, in contrast to Fig.6.3 and Fig.6.4 which are optimised over both. We choose the measurement in line with the scheme of Davis *et.al* [107]. Comparison of schemes reveals that when $N = 10$ scheme A outperforms scheme B' for $\chi\tau \lesssim 11.5$ and scheme C' for $\chi\tau \lesssim 5$. These threshold values decrease for larger N . For very large $\chi\tau$, the sensitivities of schemes B' and C' converge. The lower plots show the optimal sensing time $(t/\tau)_{\text{opt}}$ as a function of $\chi\tau$ 94
- 7.1 [116] The “entropy Venn diagram”, with the discrete Shannon entropy of the random variable X denoted by $H(X)$, allows the heuristic deduction of many of the entropy inequalities but is not a completely reliable guide. . . . 101
- 7.2 [117] Quantisation of a continuous random variable. 103

7.3	[120] The 2- and 3-dimensional representations of the probability simplex with the entropy α -norm of arbitrary point (or PMF) p	107
7.4	[120] The 3-dimensional probability simplex isoentropy contours (contours of equal Rényi entropy) for various values of α	108
8.1	Plots of position and momentum probability distributions for a coherent state with $\beta = 2$. Both are Gaussian PDFs.	117
8.2	Plots of position and momentum probability distributions for a squeezed vacuum state with $z = 1/2$ and $\phi = 0$. Both are Gaussian PDFs.	120
8.3	Plot of REPUR against the Rényi parameter r for the Cat State. It is evident that the bound is saturated for $r = -1/2$, begins to diverge as r increases then plateaus for large r	125
8.4	Plots of the position and momentum PDFs for the superposition state consisting of the vacuum and squeezed vacuum with $z = 2$	127
8.5	Plots of the REPUR (in units of \hbar) for the superposition state consisting of the vacuum and squeezed vacuum against $\log_{10}(1 + r)$ for squeezing parameters $\zeta \in \{1, 2, 3\}$. The REPUR bound is saturated in both $N_{\infty}(x)N_{1/2}(p)$ and $N_{1/2}(x)N_{\infty}(p)$ cases. Deviation from this bound saturation is observed for all other values of r with maximal deviation at $r = 0$ corresponding to the Shannon EP.	127
8.6	Plot of the variance based uncertainty relation for the superposition state of a vacuum and squeezed vacuum. The bound is saturated for $z = 0$ as this corresponds to the Gaussian vacuum state but blows up rapidly as z increases.	130
9.1	A heuristic example depicting the failings of variance as a measure of uncertainty in which a particle in state A can be found over length L and in state B it can be found over a length $L/2$. Clearly we know more about position in state B but the variance suggests otherwise.	133
9.2	Plot of the PDF of the Balanced Cat state in the x -quadrature with $\alpha = 4$. This is a physical analogue of State B.	134
9.3	An alternative example depicting the failings of variance as a measure of uncertainty.	134
9.4	Plot of the PDF for the unbalanced cat state in the x -quadrature with $\alpha = 1$ and $\nu = 0.2$. This is a physical analogue of State C.	135

- 9.5 Information scan of $\mathcal{F}(y)$ and the equimeasurably rearranged PDF $\tilde{\mathcal{F}}(y)$. Cumulative distribution function $f(x)$ measures the area of $\mathcal{F}(y)$ within the limits dictated by the intercept line $\mathcal{F}(y) = 2^{-x}$ with $\mathcal{F}(y)$ (shaded area). For the entropy measured in nats $2^{-x} \rightarrow e^{-x}$. The information PDF $g(x)$ represents the rate of change of the area of the cumulative distribution $f(x)$. Note that $f(x)$ and $g(x)$ are identical for the equimeasurably rearranged PDFs $\mathcal{F}(y)$ and $\tilde{\mathcal{F}}(y)$. The 3-peak structure of $g(x)$ is one of the invariant characteristics of the equimeasurable family of PDFs. 138
- 9.6 (a) Cumulative distribution functions $f_{\zeta}(x)$ and information PDFs $g_{\zeta}(x)$ for the superposition state of a vacuum with a squeezed vacuum. (b) The logarithmic scaling depicting the tail behaviour (corresponding to $x > 1$) of $\mathcal{F}(y) = |\psi_V(y)|^2$. From (a) it is apparent that the larger the value of ζ , the higher the peak of $\mathcal{F}(y)$, and the peaks have heights $2^{-a_{\zeta}^+}$. For $\zeta = 2$ and $\zeta = 3$ we see a second peak near $x = 3$ which is due to a sharp change in the overall shape of the PDF at height $\mathcal{F}(y) = 2^{-x}$. From (b) the tail behaviour is displayed and the best-fit analysis reveals the tails to be Gaussian. . . . 140
- 9.7 Tail convergence of the reconstructed information distribution of a Gaussian for higher order terms of the exponential expansion denoted by m . Only the first cumulant contributes to the reconstruction of a Gaussian PDF. . . . 146
- 9.8 Convergence of the reconstructed information distribution of an Unbalanced Cat State (UCS) with $\nu = 0.97$ and $\alpha = 10$. The value a_2^+ corresponds to the value of x at the point of intersection with the second (lower) peak of the $\mathcal{F}(y)$ for the UCS. The Edgeworth expansion has been used here to order $n^{-3/2}$ requiring control of the first 5 cumulants. 148

Chapter 1

Introduction

1.1 Thesis Outline

In classical mechanics a system can be modelled as a collection of particles, each with a definite position and momentum. In contrast to this, it is a fundamental property of the quantum mechanical model that the state of the position and momentum of a particle has some inherent uncertainty. Quantum mechanics utilizes the mathematical framework provided by linear algebra in order to produce predictions, with unprecedented accuracy, on the probabilistic nature of measurement outcomes. Furthermore, this mathematical framework predicts natural behaviour which cannot be reproduced within a classical framework. These non-classical phenomena are quantum mechanical properties which can be harnessed and exploited in modern technologies. A fundamental use of such quantum technologies is the bootstrapping of the quantum theory, describing how to perform enhanced precision measurements, with the technological progress on which theory is reliant. This particular area of research is collectively known as quantum enhanced metrology and is the focus of this thesis.

In chapter 2 and 3 we provide essential background on the mechanics of bosonic and fermionic systems respectively. In chapter 4 we review quantum enhanced metrology and parameter estimation while introducing the theoretical tools required to understand the results that follow. In chapter 5 we explore the role of quantum correlations in optical quantum metrology and moreover, we present schemes for both single and multiparameter estimation strategies which are better suited to practical implementation than the current proposed strategies and can return equivalent, or higher, precisions. Chapter 6 begins by addressing the traditional three stage quantum metrology protocol which consists of

preparation, sensing and readout. Regarding time as a limited resource in such protocols motivates the proposal of concurrent sensing during the probe preparation and readout stages which we show outperforms its sequential preparation-sensing-readout counterpart. In chapter 7 we give an overview of entropic measures in information theory. In chapter 8 we present a generalised uncertainty relation based on a specific entropy measure known as the Rényi entropy and apply this bound to a system state of particular interest in quantum metrology. This yields an infinite class of uncertainty relations and as such, provides far more information on the underlying physical system than the generic uncertainty relation can supply. This notion is furthered in chapter 9 where we establish an “information scan” using the aforementioned Rényi entropies and apply this technique to system states of particular interest in quantum metrology. In chapter 10 we give concluding remarks.

In the subsections that immediately follow, we review some of the underlying physical and mathematical principals that form the building blocks of quantum mechanics and give a brief overview of the development of metrology, the more experienced reader may wish to advance to chapter 2.

1.2 Quantum Mechanics

1.2.1 Entanglement

Entanglement is fundamental phenomenon from which some of the most peculiar aspects of quantum mechanics emerge such as nonlocality which is the mechanical interaction of objects that are spatially separated or as Einstein famously disparaged it as “spooky action at a distance”. This was brought to attention in 1935 through the EPR paradox [1] which was put forward to highlight the incompleteness of the Copenhagen interpretation of quantum mechanics (this interpretation takes the wavefunction to be a mathematical object that predicts the probability of specific outcomes of an experiment), however in the same year Erwin Schrödinger introduced the counter-intuitive notion of “Verschränkung” roughly meaning “entanglement” [2] (alongside his famous thought experiment of a cat that has curiously entered a state of being simultaneously dead and alive based on quantum mechanical principles). Schrödinger emphasises the importance of this concept in stating “I would not call [entanglement] *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought”. This recognition of the fundamental importance that entanglement manifests eventually facilitated

the resolution of the paradox and now has experimental evidence based upon violations of the so called “Bell inequalities” [3, 4]. In particular, violation of the Bell inequalities challenges the assumption of locality and shows it to be mathematically inconsistent with the underlying theory of quantum mechanics [5]. Even so, whether evidence of nonlocality is demonstrated by the evidence supporting the violation of the Bell inequalities depends on the interpretation of quantum mechanics employed and this is still an area of highly active debate which strays into the all-encompassing territory of philosophy. Our interests lie in the utilisation of entanglement as a resource to be exploited, some particularly fruitful applications of entanglement include quantum cryptography including key distribution [6, 7, 8], quantum teleportation [9, 10], the development of quantum computers [11] and quantum enhanced metrology [12, 13] which is the area of investigation for this thesis. Though entanglement has many useful features it is not without its drawbacks; it is extremely fragile and susceptible to unwanted interaction from its environment and it is not possible to impart entanglement on spatially separated objects i.e one cannot “increase” entanglement between objects that are not directly in contact (but entanglement can persist over spatially separated regions once established).

1.2.2 Dirac Notation

Here we introduce the notation used throughout this thesis, it is referred to as Dirac notation or “bra-ket” notation and is used to describe a quantum mechanical system’s physical state and evolution. This notation is not only compact and concise, it is designed to describe quantum mechanical systems that cannot be described by position and momentum alone, examples of this include the two-level atom and the single quanta of light; the photon.

A general state of a quantum mechanical system is denoted by the “ket” $|\psi\rangle$, which is a vector built of component basis vectors $|i\rangle$

$$|\psi\rangle = \sum_i^N c_i |i\rangle \quad (1.1)$$

where $c_i \in \mathbb{C}$ are the normalising coefficients and the basis states $|i\rangle$ are an orthonormal basis of the N -dimensional vector space which describes the physical system and thus the inner product $\langle i|j\rangle = \delta_{i,j}$ which is the Kronecker delta given by

$$\delta_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (1.2)$$

and it is now apparent that

$$\langle \psi | \psi \rangle = \sum_i^N |c_i|^2 = 1 \quad (1.3)$$

where this condition is known as “normalisation”. Furthermore, given a vector

$$\langle \phi | = \sum_j^N b_j^* \langle j | \quad (1.4)$$

the inner product or “overlap” of this and (ket) vector (1.1) is

$$\begin{aligned} \langle \phi | \psi \rangle &= \left(\sum_j^N b_j^* \langle j | \right) \left(\sum_i^N c_i | i \rangle \right) \\ &= \sum_i^N b_i^* c_i \end{aligned} \quad (1.5)$$

Note that since the states are represented by vectors, the terms are interchangeable. The $\langle \cdot |$ is known as a “bra vector” and as we’ve already seen the $|\cdot\rangle$ is referred to as a “ket vector”, this constitutes the “bra-ket” notation more commonly referred to as Dirac notation. As is now evident, Dirac notation is simply another way of writing vectors but done so in a way that distils the minimum amount of information required to perform quantum mechanical calculations. It is useful to highlight this by introducing the notion of operators; given a quantum mechanical operator \hat{A} , it is always possible to find states $|\lambda_i\rangle$ such that

$$\hat{A} |\lambda_i\rangle = \lambda_i |\lambda_i\rangle \quad (1.6)$$

hence, $|\lambda_i\rangle$ are referred to as eigenstates of \hat{A} with eigenvalues λ_i . As a consequence of the spectral theorem, $\lambda_i \in \mathbb{R}$ and the eigenstates form an orthonormal basis i.e. $\langle \lambda_i | \lambda_j \rangle = \delta_{i,j}$ so the operator can be written

$$\hat{A} = \sum_i^N \lambda_i |\lambda_i\rangle \langle \lambda_i|. \quad (1.7)$$

The notion of using operators to model physical observables is covered more formally in the following subsection. For now, in order to see how Dirac notation translates, consider the more general operator written in matrix representation

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \dots & A_{NN} \end{bmatrix}$$

which acts on the vector

$$\psi = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}$$

in the same way that the operator

$$\begin{aligned} \hat{A} = & A_{11} |1\rangle \langle 1| + A_{12} |1\rangle \langle 2| + A_{13} |1\rangle \langle 3| \dots A_{1N} |1\rangle \langle N| \\ & + A_{21} |2\rangle \langle 1| + A_{22} |2\rangle \langle 2| + A_{23} |2\rangle \langle 3| \dots A_{2N} |2\rangle \langle N| \\ & \vdots \\ & + A_{N1} |N\rangle \langle 1| + A_{N2} |N\rangle \langle 2| + A_{N3} |N\rangle \langle 3| \dots A_{NN} |N\rangle \langle N| \end{aligned} \quad (1.8)$$

acts on the vector

$$|\psi\rangle = c_1 |1\rangle + c_2 |2\rangle + c_3 |3\rangle + \dots + c_N |N\rangle \quad (1.9)$$

then it is clear that operator in its eigenbasis is one in which all off diagonal terms are zero i.e $A_{ij} = 0, \forall i \neq j$; an operator of this form is referred to as diagonalised. A special case of such operator is known as the Identity matrix and is of the following form

$$\mathbb{I} = \sum_i^N |i\rangle \langle i| \quad (1.10)$$

which, when applied to a general state of the same dimension, has the following effect

$$\mathbb{I} |\psi\rangle = \left(\sum_i^N |i\rangle \langle i| \right) \left(\sum_j^N c_j |j\rangle \right) = \sum_i^N c_i |i\rangle = |\psi\rangle \quad (1.11)$$

hence, the Identity operator returns the state it acts upon. Note that up until now, only finite dimensional systems have been considered i.e vectors and operator of size N however, there are many systems that require the limiting case of $N \rightarrow \infty$. Although the mathematics required is different and more involved, the same underlying physical principles outlined here still apply [14].

1.2.3 The Postulates of Quantum Mechanics

The physical laws presented here underpin the entirety of quantum theory as a formalism for predicting the probabilistic outcomes of measurements on a system. In particular, given an ensemble of identically prepared systems, the aim of quantum theory is to accurately predict the underlying probability distributions of measurement outcomes. With that, the postulates of quantum mechanics are as follows

1. The state of a quantum mechanical system is represented by a *state vector* $|\psi\rangle$, taken to be of unit length, which exists in a state space known formally as a *Hilbert space* \mathcal{H} . This implies that if two vectors $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ then it is possible for the superposition of these to yield a vector $|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle$ with $|\psi\rangle \in \mathcal{H}$ where $c_1, c_2 \in \mathbb{C}$. Additionally, the Hilbert space associates a complex number to any two state vectors defined within the Hilbert space via the *inner product* i.e for any $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, $\exists \langle\psi|\phi\rangle = (\langle\phi|\psi\rangle)^* \in \mathbb{C}$. Strictly speaking, the “bra” belongs to the *dual space* of \mathcal{H} and defines a *linear map* from the Hilbert space to the complex numbers $L_\psi : \mathcal{H} \rightarrow \mathbb{C}$. For a state vector to be of unit length it must satisfy the *normalisation* condition $|\langle\psi|\psi\rangle|^2 = 1$.
2. All *quantum mechanical observables* are represented by *Hermitian operators* of the form

$$\hat{A} = \sum_{i=1} \lambda_i |\lambda_i\rangle \langle\lambda_i| \quad (1.12)$$

where $|\lambda_i\rangle$ are the *eigenstates* satisfying $\hat{A} |\lambda_i\rangle = \lambda_i |\lambda_i\rangle$ with the *eigenvalues* λ_i . In general, the action of the operator on the state vector alters the state itself however, in the special case where the system exists in an eigenstate of the operator the system remains unchanged by measurement.

3. The eigenvalues of the observable \hat{A} represent the only possible outcomes upon measurement of \hat{A} . Since \hat{A} is a Hermitian operator this implies $\lambda_i \in \mathbb{R}$ as expected of measured results. These quantized results of measurement are the very manifestation of quantum theory. Moreover, the complete set of eigenstates $S = \{|\lambda_1\rangle, \dots, |\lambda_N\rangle\}$ form an *orthonormal basis* with $\langle\lambda_i|\lambda_j\rangle = \delta_{i,j} \forall i, j$; this set forms the *linear span* of the Hilbert space denoted $[S] = \mathcal{H}$.
4. For a general state $|\psi\rangle = \sum_i c_i |\lambda_i\rangle$, a measurement of observable \hat{A} will return the eigenvalue λ_j given by the conditional probability $P(A = \lambda_j; |\psi\rangle) = |\langle\lambda_j|\psi\rangle|^2$, the overlap $\langle\lambda_j|\psi\rangle$ is referred to as the *probability amplitude*. The expectation value (in the usual statistical sense) of observable \hat{A} for the state $|\psi\rangle$ is given by

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} = \langle\psi|\hat{A}|\psi\rangle = \sum_{i=1} \lambda_i |c_i|^2. \quad (1.13)$$

In practise, this number is obtained by performing a large number of measurements on an ensemble of identically prepared systems.

5. Once a measurement of a general observable \hat{A} has submitted an eigenvalue λ_j , all subsequent measurements of observable \hat{A} on the system will yield the same value λ_j . This is interpreted as the “collapse of the wave function” onto the eigenstate λ_j .
6. The state vector of the system evolves in time in accordance with the time-dependent Shrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (1.14)$$

where $|\psi(t)\rangle$ is a time-dependent state vector that moves along a trajectory through the Hilbert space. As a consequence of the conservation of probability, this evolution preserves the normalisation of the states and is thus referred to as *unitary evolution*. An operator \hat{U} is defined to be unitary iff $\hat{U}^\dagger \hat{U} = \mathbb{I}$. Equation (1.14) gives the unitary time evolution operator $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$ so a system initially in state $|\psi(0)\rangle$ evolves to $|\psi(t)\rangle = \hat{U}(t) |\psi(0)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle$ thus preserving normalisation.

We note that the number of postulates given can vary depending on convention, here we have given a relatively comprehensive account.

1.2.4 Mixed states and Density Matrices

So far, we have considered quantum systems in known states denoted by the vector $|\psi\rangle$, we now formally refer to this as a *pure state*. Here we consider the case where the state of a quantum system is not completely known i.e the system can be in state $|\psi_i\rangle$ with probability p_i , we refer to this as an *ensemble of pure states* - from which we can construct a mathematical object for the mixed state given by the so called *density matrix*

$$\hat{\rho} = \sum_{i=1} p_i |\psi_i\rangle \langle \psi_i| \quad (1.15)$$

here we shall explore some of the properties of the density matrix. As such, let \hat{A} be a general observable and let $\{|e_1\rangle, |e_2\rangle, \dots, |e_N\rangle\}$ be an orthonormal basis on the Hilbert space, the *trace* of \hat{A} is defined by [14]

$$\text{Tr}[\hat{A}] = \sum_{i=1}^N \langle e_i | \hat{A} | e_i \rangle \quad (1.16)$$

then taking a general state vector defined on the Hilbert space, we find

$$\begin{aligned} \langle \psi | \hat{A} | \psi \rangle &= \sum_{i=1}^N \langle \psi | \hat{A} | e_i \rangle \langle e_i | \psi \rangle = \sum_{i=1}^N \langle e_i | \psi \rangle \langle \psi | \hat{A} | e_i \rangle \\ &= \sum_{i=1}^N \langle e_i | \hat{P}_\psi \hat{A} | e_i \rangle = \text{Tr}[\hat{P}_\psi \hat{A}] \end{aligned} \quad (1.17)$$

where $\hat{P}_\psi = |\psi\rangle\langle\psi|$ is known as a projection operator; these are explored in detail in section 4.1.2. It is clear that the conditional probability $P(A = \lambda_j; |\psi\rangle) = \text{Tr}[\hat{P}_\psi \hat{P}_j]$ and from the linearity of the trace function we consequently have

$$\langle\psi| \hat{A} |\psi\rangle = \sum_{i=1} p_i \text{Tr}[\hat{P}_{\psi_i} \hat{A}] = \text{Tr} \left[\sum_{i=1} p_i \hat{P}_{\psi_i} \hat{A} \right] = \text{Tr}[\hat{\rho} \hat{A}] \quad (1.18)$$

and similarly $P(A = \lambda_j; \hat{\rho}) = \text{Tr}[\hat{\rho} \hat{P}_j]$. In essence, the density matrix embodies an extension of the type of mathematical object that constitutes a probability measure. This can be seen more clearly via the following properties of the operator:

1. $\hat{\rho} = \hat{\rho}^\dagger$
2. $\langle\psi| \hat{\rho} |\psi\rangle \geq 0, \forall |\psi\rangle$
3. $\text{Tr}[\hat{\rho}] = 1$

and in general, any operator satisfying these properties is a density matrix. It is useful to clarify the language here, density operator and density matrix are used equivalently, a quantum state of which we have full knowledge is known as a pure state $|\psi\rangle$ and has an equivalent density matrix representation $\hat{\rho} = |\psi\rangle\langle\psi|$, a quantum state of which we have partial knowledge is known as a mixed state i.e a mixture of different pure states from the ensemble. A notable mathematical difference is that for pure state $\text{Tr}[\rho^2] = 1$ and for mixed states $\text{Tr}[\hat{\rho}^2] < 1$. So far we have rephrased some of the basic notions of quantum mechanics in terms of density matrices, a final example of this is to consider unitary evolution of the system that is in the initial state $|\psi_i\rangle$ with probability p_i and thus enters the final state $\hat{U} |\psi_i\rangle$ with probability p_i . Then it follows that the evolution of the density matrix is given by

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \rightarrow \sum_i p_i \hat{U} |\psi_i\rangle\langle\psi_i| \hat{U}^\dagger = \hat{U} \hat{\rho} \hat{U}^\dagger. \quad (1.19)$$

We can then determine the time evolution of a system in a mixed state (taking the states $|\psi_i\rangle$ to be time dependent)

$$\frac{\partial \rho}{\partial t} = \sum_i p_i \left[\left(\frac{\partial}{\partial t} |\psi_i\rangle \right) \langle\psi_i| + |\psi_i\rangle \left(\frac{\partial}{\partial t} \langle\psi_i| \right) \right] \quad (1.20)$$

then using the Schrödinger equation and noting that for the bra

$$-i\hbar \frac{\partial}{\partial t} \langle\psi_i| = \langle\psi_i| \hat{H} \quad (1.21)$$

it follows that

$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \frac{1}{i\hbar} \sum_i p_i \left[\left(\hat{H} |\psi_i\rangle \right) \langle \psi_i| - |\psi_i\rangle \left(\langle \psi_i| \hat{H} \right) \right] \\
&= \frac{1}{i\hbar} (\hat{H} \hat{\rho} - \hat{\rho} \hat{H}) \\
&= \frac{1}{i\hbar} [\hat{H}, \hat{\rho}]
\end{aligned} \tag{1.22}$$

It is possible to give an intrinsic characterisation of the density matrix [15] which is achieved by using the three mathematical properties listed above, in this way a reformulation of quantum mechanics in terms of density matrices is possible.

1.2.5 The Schrödinger, Heisenberg and Interaction Pictures

In the Schrödinger picture, the quantum states evolve in time (governed by the Schrödinger equation) and the operators are time independent, this is given by (1.14). In the Heisenberg picture, the time dependency exists in the operators and the states are time independent. Using the subscripts S and H to distinguish between the Schrödinger and Heisenberg pictures respectively, we have

$$|\psi(t)\rangle_S = \hat{U}(t - t_i) |\psi(t_i)\rangle_S = e^{-i\hat{H}(t-t_i)/\hbar} |\psi(t_i)\rangle_S \tag{1.23}$$

and we define

$$|\psi(t)\rangle_H \equiv \hat{U}^\dagger(t - t_i) |\psi(t)\rangle_S = |\psi(t_i)\rangle_S. \tag{1.24}$$

It is clear that the expectation value of an arbitrary operator should be the same regardless of the picture used, thus

$$\begin{aligned}
\langle \psi(t) |_S \hat{A}_S |\psi(t)\rangle_S &= \langle \psi(t) |_S \hat{U}(t - t_i) \hat{U}^\dagger(t - t_i) \hat{A}_S \hat{U}(t - t_i) \hat{U}^\dagger(t - t_i) |\psi(t)\rangle_S \\
&= \langle \psi(t) |_H \hat{A}_H |\psi(t)\rangle_H
\end{aligned} \tag{1.25}$$

where we have defined

$$\hat{A}_H \equiv \hat{U}^\dagger(t - t_i) \hat{A}_S \hat{U}(t - t_i) \tag{1.26}$$

as the Heisenberg operator. An importance consequence of this is the evolution equation

$$i\hbar \frac{\partial}{\partial t} \hat{A}_H = [\hat{A}_H, \hat{H}] \tag{1.27}$$

which is known as Heisenberg's equation of motion. Note that if an operator commutes with \hat{H} we can conclude that the associated physical observable is conserved over time.

Another widely used and extremely helpful picture is the Interaction picture. In order to understand this picture it is first useful to introduce the following Schrödinger picture Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{V} \quad (1.28)$$

where, \hat{H}_0 is the “main” Hamiltonian and \hat{V} is an “interaction” or “perturbation” term. The utility of this decomposition is that we can use \hat{H}_0 to put the time dependence into operators and use \hat{V} to give the evolution of states. If we let

$$|\psi(t)\rangle_I = e^{i\hat{H}_0 t/\hbar} |\psi(t)\rangle_S \quad (1.29)$$

then we have

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle_I = -\hat{H}_0 e^{i\hat{H}_0 t/\hbar} |\psi(t)\rangle_S + e^{i\hat{H}_0 t/\hbar} i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle_S \quad (1.30)$$

then defining the Interaction picture operator to be $\hat{V}_I \equiv e^{i\hat{H}_0 t/\hbar} \hat{V} e^{-i\hat{H}_0 t/\hbar}$ we have

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle_I = \hat{V}_I |\psi(t)\rangle_I \quad (1.31)$$

and finally, for an arbitrary Interaction picture operator $\hat{A}_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{A}_S e^{-i\hat{H}_0 t/\hbar}$ we find the following equation of motion

$$i\hbar \frac{\partial}{\partial t} \hat{A}_I = [\hat{A}_I, \hat{H}_0] \quad (1.32)$$

note the subtle difference in signs with the density matrix evolution operator (1.22).

1.3 Metrology: An Overview of the Science of Measurement

“To measure is to know” - Lord Kelvin

Empirical investigation is the acquisition of evidence, through sensory experience, that verifies or falsifies claims of reality. The practise of empirical investigation can be traced back to classical antiquity, these investigations have been described by the natural philosophers of ancient Greece such as Thales and Aristotle [16]. The scientific method is built upon empirical evidence, the acquisition of which is entirely reliant on measurement. Indeed the scientific method consists of question, hypothesis, prediction, testing and analysis - it is the testing aspect that employs empirical investigation. The outlined scientific method has been employed since the middle ages (notable pioneers include Galileo Galilei

and Johannes Kepler) and has since been refined and bolstered by the additional steps of replication, peer review and data sharing.

Our ability to measure the natural world thus imposes a fundamental limitation on scientific progress. This immediately raises the fundamental question of what defines evidence however, a shift in intuition neatly bypasses such ambiguity through demanding that the nuance must lie in the question asked; measured evidence is that which answers empirical questions (hypothesis) i.e one must ask the right questions. How to do this is far beyond the scope of this thesis, instead the focus of the following work is on how to improve upon probing the universe for precise information. Crucially, measurement cannot prove a scientific hypothesis, it can either disprove, or act as statistical support, to a current working hypothesis. Metrology encompasses both the experimental and theoretical aspect of the measurement itself. Significant developments in classical metrology include the development of telescopes, clocks, the refutation of the ether by Michelson and Morley (1887) and the discovery of the atomic nucleus by Rutherford (1911) to name a few. Modern metrology primarily concerns itself with the measurement of particles, the discrete nature of which results in Poisson noise or “Shot noise” as we shall refer to it. This is easily conceptualised by counting electrons, or indeed by counting photon numbers. A specific understanding of the origins of shot noise can be exemplified by a simple coin tossing experiment; after a great many number of coin tosses $N \gg 1$, the number of occurrences of heads and tails will be approximately equal $n_h/n_t \approx 1$ however for relatively small N there can be a significant difference in the number of occurrences of heads and tails $n_h/n_t \gg 1$ or $n_h/n_t \ll 1$. Nonetheless, upon repeating such an experiment of small N the difference in outcomes will fluctuate significantly and it can be proven that this fluctuation reduces as $1/\sqrt{\nu N}$ where ν is the number of experimental repeats. Hence in order to reduce the noise (increase the precision) of a measurement we can simply increase the number of resources used. However this may not always be possible, for example it may be desired to probe a fragile living sample (e.g DNA combing [17]) which would be denatured if too many photons are incident upon it, another such example includes trying to illuminate a distant object where only a few photons are reflected thus reducing the number of measurements [18]. To overcome this limitation of precision we turn to quantum mechanics.

Quantum enhanced metrology utilizes purely quantum mechanical correlations in order to enhance the precision of measurements. Measurement is at the core of quantum mechanics

as is evident from the postulates 1.2.3 and the language used which deals with observables, outcomes and expectation values. A natural question can arise: what distinguishes measurement from any other kind of interaction in quantum mechanics, an interesting question but beyond the scope of this thesis which follows the aforementioned postulates. Regardless of interpretation, quantum metrology has been successfully used in applications such as gravitational wave detection [19], biological physics and medicine [20]. A key theoretical result is the use of entanglement to obtain a precision that scales with the number of particles as the Heisenberg limit $1/N$ which is a \sqrt{N} improvement over the shot noise limit.

Chapter 2

Bosonic Systems

In this chapter an overview is given of bosonic systems corresponding to the quantised electromagnetic field. With this we introduce some of the more prolific states in quantum optics and investigate the properties of particular interest to quantum metrology.

2.1 The Quantum Harmonic Oscillator

2.1.1 Field Quadrature Operators

Quantum optics treats the electric and magnetic fields as Hermitian operators, the ubiquitous method used to arrive at this treatment is to decompose the fields into modes and to treat each mode as a quantum harmonic oscillator [21, 22]. To achieve this, it is useful and concise to define the non-commutative, non-Hermitian field operators a^\dagger and a which satisfy the following commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1 \quad (2.1)$$

One method that facilitates the emergence of this relation is to simply use the correspondence principle for classical position and momentum i.e replace the variables with their corresponding Hermitian operators, this gives the commutation relation

$$[\hat{q}, \hat{p}] = i\hbar. \quad (2.2)$$

The classical Hamiltonian that describes the single mode confined to a cavity of volume V is given by

$$H = \frac{1}{2} \int dV \left[\epsilon_0 E_x^2(z, t) + \frac{1}{\mu_0} B_y^2(z, t) \right] \quad (2.3)$$

where ϵ_0 and μ_0 are the vacuum permittivity and permeability respectively, x, y and z are orthogonal spatial dimensions, t is the temporal dimension and the electric and magnetic

fields are given by

$$E_x(z, t) = \left(\frac{2\omega^2}{V\epsilon_0} \right)^{1/2} q(t) \sin(kz) \quad (2.4)$$

$$B_y(z, t) = \left(\frac{\mu_0\epsilon_0}{k} \right) \left(\frac{2\omega^2}{V\epsilon_0} \right)^{1/2} \dot{q}(t) \cos(kz) \quad (2.5)$$

where ω is the frequency of the mode. Hamiltonian (2.3) can then be neatly rewritten, using the correspondence rule (as postulated and demonstrated by Bohr, Heisenberg and Jordan (1926) [23]), in the following form

$$\hat{H} = \frac{1}{2}(\hat{p}^2 + \omega^2 \hat{q}^2) \quad (2.6)$$

which is the Hamiltonian for the 1-dimensional quantum harmonic oscillator (QHO). At this point it is convenient to define the creation and annihilation operators in terms of the position and momentum operators

$$\hat{a} = (2\hbar\omega)^{-1/2}(\omega\hat{q} + i\hat{p}) \quad (2.7)$$

$$\hat{a}^\dagger = (2\hbar\omega)^{-1/2}(\omega\hat{q} - i\hat{p}) \quad (2.8)$$

which reproduces eqn. (2.2) and reveals the following form of the Hamiltonian of the QHO

$$\hat{H}_Q = \hbar\omega \left(\hat{a}\hat{a}^\dagger + \frac{1}{2} \right). \quad (2.9)$$

Furthermore, we are now able to define a specific type of operator, in terms of the creation and annihilation operators, that act like the position and momentum operators but are scaled to be dimensionless; these are known as quadrature operators

$$\hat{X} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger) \quad (2.10)$$

$$\hat{P} = \frac{1}{i\sqrt{2}}(\hat{a} - \hat{a}^\dagger) \quad (2.11)$$

noting that the factor $1/\sqrt{2}$ is chosen by convention and is sometimes taken to be $1/2$, thus one should take particular care to keep track of factors of 2 when using these operators. With this it is simple to find

$$[\hat{X}, \hat{P}] = i. \quad (2.12)$$

This non-commutivity leads to an important consequence that distinguishes quantum optics from the classical formalism. To see this, it is instructive to first recall [24] that for two operators satisfying $[\hat{A}, \hat{B}] = \hat{C}$, it follows that

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}|\langle\hat{C}\rangle| \quad (2.13)$$

where $\Delta\hat{A}$ and $\Delta\hat{B}$ are the standard deviations of the respective operators and $\langle\hat{C}\rangle$ is the expectation value (in the usual statistical sense) of the operator \hat{C} , hence

$$\Delta^2\hat{X}\Delta^2\hat{P} \geq \frac{1}{4} \quad (2.14)$$

Since variance quantitatively captures the spread among the outcomes as deviation around the mean, it can be used as a measure of uncertainty (or equivalently, certainty) of an expectation value (or equivalently, the average measurement outcome) thus we deduce that *simultaneous* knowledge of a quantum mechanical system's position and momentum is fundamentally limited. This is in stark contrast to classical mechanics with which one can, in principle, attain certainty about a system's position and momentum simultaneously. A more general statement can be made by defining the general quadrature operator

$$\hat{X}_\lambda = \frac{1}{\sqrt{2}}(\hat{a}e^{-i\lambda} + \hat{a}^\dagger e^{i\lambda}) \quad (2.15)$$

then we have that for any two quadrature operators \hat{X}_λ and $\hat{X}_{\lambda+\pi/2}$ (which for $\lambda = 0$ corresponds to (2.10) and (2.11) respectively) i.e any two operators associated with field amplitudes that are oscillating with a phase difference of $\pi/2$, the following commutation relation holds

$$[\hat{X}_\lambda, \hat{X}_{\lambda+\pi/2}] = i \quad (2.16)$$

and pairs of such operators are referred to as canonically conjugate. From this, we obtain the uncertainty relation for general conjugate pairs of quadrature operators

$$\Delta^2\hat{X}_\lambda\Delta^2\hat{X}_{\lambda+\pi/2} \geq \frac{1}{4}. \quad (2.17)$$

2.1.2 The Number Operator

Another important Hermitian operator that can be constructed from the creation and annihilation operators is the number operator, this is given by

$$\hat{n} = \hat{a}^\dagger\hat{a}. \quad (2.18)$$

In order to gain some intuitive understanding of this operator, it is useful to introduce the number states $|n\rangle$. These can be defined as the eigenstates of the QHO Hamiltonian

$$\hat{H}_Q |n\rangle = \hbar\omega \left(n + \frac{1}{2}\right) |n\rangle \quad (2.19)$$

or equivalently, the eigenstates of the number operator

$$\hat{n} |n\rangle = n |n\rangle \quad (2.20)$$

which thus form an orthonormal set, consequently

$$\sum_{n=0}^{\infty} |n\rangle \langle n| = \mathbb{I} \quad (2.21)$$

and $\langle n|m\rangle = \delta_{n,m}$. The individual actions of the creation and annihilation operators upon the number states are given by

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \quad (2.22)$$

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad (2.23)$$

i.e the creation and annihilation operators raise and lower the photon number by one respectively. With this it is evident that the number operator serves to count the number of photons that exist in a given state and the vacuum state is represented by $|0\rangle$, a single photon is given by $|1\rangle$, two photons by $|2\rangle$ etc. In fact, the general number state can be given in terms of the vacuum state and the creation operator as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle. \quad (2.24)$$

2.2 Coherent States

Following the introduction of the creation and annihilation operators along with the number states, a natural question arises; do eigenstates of the creation and annihilation operators themselves exist and if so, what are they? With hindsight, this question is answered by introducing the (Glauber) coherent state [25] which, in the number state basis, is given by

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (2.25)$$

where $\alpha = |\alpha|e^{i\theta}$ with $\theta \in [0, 2\pi]$ so that in general, $\alpha \in \mathbb{C}$. From this we have

$$\begin{aligned} \hat{a} |\alpha\rangle &= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \hat{a} |n\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \sqrt{n} |n-1\rangle \\ &= e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{\alpha^{m+1}}{\sqrt{m!}} |m\rangle \\ &= \alpha |\alpha\rangle \end{aligned} \quad (2.26)$$

thus $|\alpha\rangle$ is the eigenstate of \hat{a} however it is essential to note that $|\alpha\rangle$ is clearly not a *right* eigenstate of \hat{a}^\dagger . Nonetheless, it does follow that

$$\langle \alpha | \hat{a}^\dagger = (\hat{a} |\alpha\rangle)^\dagger = (\alpha |\alpha\rangle)^\dagger = \alpha^* \langle \alpha | \quad (2.27)$$

so $|\alpha\rangle$ is a *left* eigenstate of \hat{a}^\dagger . The coherent state can be defined by the condition of being an eigenstate of the annihilation operator and its coefficients in the number state basis can be deduced from this [21]. Evidently, the coherent state is a superposition of the infinity of number states. This constitutes the state of the QHO which behaves “most classically” while implicitly being a quantum mechanical state, furthermore this state accurately describes the monochromatic light produced by lasers. With this we review some of the properties of the coherent state. The expectation value of the general quadrature variable (2.15) with respect to the coherent state is given by

$$\begin{aligned}\langle\alpha|\hat{X}_\lambda|\alpha\rangle &= \langle\alpha|\frac{1}{\sqrt{2}}(\hat{a}e^{-i\lambda} + \hat{a}^\dagger e^{i\lambda})|\alpha\rangle \\ &= \frac{1}{\sqrt{2}}(\alpha e^{-i\lambda} + \alpha^* e^{i\lambda})\end{aligned}\quad (2.28)$$

and for \hat{X}_λ^2 , the expectation value is found to be

$$\begin{aligned}\langle\alpha|\hat{X}_\lambda^2|\alpha\rangle &= \langle\alpha|\frac{1}{2}\left[\hat{a}^2 e^{-2i\lambda} + (\hat{a}^\dagger)^2 e^{2i\lambda} + 2\hat{a}^\dagger \hat{a} + 1\right]|\alpha\rangle \\ &= \frac{1}{2}\left[\alpha^2 e^{-2i\lambda} + (\alpha^*)^2 e^{2i\lambda} + 2|\alpha|^2 + 1\right]\end{aligned}\quad (2.29)$$

thus the variance of the general quadrature variable for all coherent states is

$$\Delta\hat{X}_\lambda^2 = \langle\alpha|\hat{X}_\lambda^2|\alpha\rangle - \left(\langle\alpha|\hat{X}_\lambda|\alpha\rangle\right)^2 = \frac{1}{2}\quad (2.30)$$

and furthermore, for any conjugate pair of quadrature observables \hat{X}_λ and $\hat{X}_{\lambda+\pi/2}$ the following uncertainty relation holds for all coherent states

$$\Delta\hat{X}_\lambda^2 \Delta\hat{X}_{\lambda+\pi/2}^2 = \frac{1}{4}\quad (2.31)$$

which saturates the bound given by (2.17). In other words, coherent states minimise the product of uncertainties of conjugate observables, it is for this reason that coherent states are also referred to as minimum uncertainty states [26]. It’s helpful to visualise the coherent state via phase space plots and to do so we first note that from (2.28) the expectation values in the dimensionless position and momentum quadratures are revealed to be $\langle\hat{X}\rangle = (\alpha + \alpha^*)/\sqrt{2} = \sqrt{2}\text{Re}(\alpha)$ and $\langle\hat{P}\rangle = (\alpha - \alpha^*)/i\sqrt{2} = \sqrt{2}\text{Im}(\alpha)$ hence the complex plane of α models phase space (that has been scaled to be dimensionless). Fig.2.1 displays such plots of the quasi probability distributions known as Wigner functions [27] for the vacuum state and a coherent state. We shall explore the Wigner functions in more detail in the following subsection, for now we are only interested in the phase space plane it affords. By inspection of the plots given in Fig.2.1, the coherent state is revealed to be

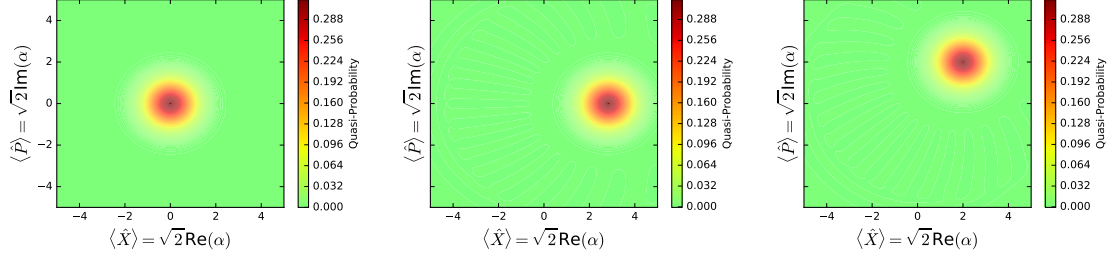


Figure 2.1: Plots of Wigner functions for vacuum state (*left*) coherent state with $\alpha = 2$ (*centre*) and coherent state with $\alpha = \sqrt{2} + i\sqrt{2}$ (*right*). Taking $\alpha \in \mathbb{R}$ displaces the vacuum state along the position axis whereas taking $\alpha \in \mathbb{C}$ displaces the vacuum in the plane of phase space. Note that the vacuum state is equivalent to a coherent state with $\alpha = 0$.

the vacuum state displaced in phase space and can indeed be defined in such a manner, as we shall see by the introduction of the so called displacement operator

$$\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}}. \quad (2.32)$$

It is immediately apparent that $\hat{D}^\dagger(\alpha) = e^{-\alpha \hat{a}^\dagger + \alpha^* \hat{a}} = D(-\alpha) \Rightarrow \hat{D}^\dagger(\alpha) \hat{D}(\alpha) = \mathbb{I}$, hence the displacement operator is unitary. To see how the displacement operator defines a coherent state, it is useful to first recall the Baker-Campbell-Hausdorff formula (also known as the disentangling theorem)

$$e^{\hat{A} + \hat{B}} = e^{\hat{A}} e^{\hat{B}} e^{-\frac{1}{2}[\hat{A}, \hat{B}]} = e^{\hat{B}} e^{\hat{A}} e^{\frac{1}{2}[\hat{A}, \hat{B}]} \quad (2.33)$$

which holds iff $[\hat{A}, [\hat{A}, \hat{B}]] = [\hat{B}, [\hat{A}, \hat{B}]] = 0$. With this, it follows that

$$\begin{aligned} \hat{D}(\alpha) |0\rangle &= e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}} |0\rangle \\ &= e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}} e^{-\frac{1}{2}|\alpha|^2 [\hat{a}^\dagger, -\hat{a}]} |0\rangle \\ &= e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}} e^{-\frac{1}{2}|\alpha|^2} |0\rangle \end{aligned} \quad (2.34)$$

then since $e^{-\alpha^* \hat{a}} |0\rangle = |0\rangle$, we have

$$\hat{D}(\alpha) |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n (\hat{a}^\dagger)^n}{n!} |0\rangle = |\alpha\rangle \quad (2.35)$$

where we have used (2.24). This is perhaps the most revealing definition of the coherent state since it is now straight forward to give the parameter $\alpha = |\alpha|e^{i\theta}$ physical meaning; $|\alpha|$ is the magnitude of displacement from the vacuum and θ dictates the direction of the displacement. Moreover, many useful properties are easily unveiled by the displacement operator such as

$$\langle \alpha | \alpha \rangle = \langle 0 | \hat{D}^\dagger(\alpha) \hat{D}(\alpha) | 0 \rangle = \langle 0 | 0 \rangle = 1 \quad (2.36)$$

hence the coherent state is normalised. Another useful operator theorem which states

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!} [\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!} [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots \quad (2.37)$$

enables us to inspect how the displacement operator transforms the annihilation operator

$$\hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha) = \hat{a} + [-\alpha\hat{a}^\dagger + \alpha^*\hat{a}, \hat{a}] + \dots = \hat{a} + \alpha \quad (2.38)$$

from which it immediately follows that $\hat{D}^\dagger(\alpha)\hat{a}^\dagger\hat{D}(\alpha) = \hat{a}^\dagger + \alpha^*$ (using the potentially subtle fact that $(\hat{A}\hat{B}\hat{C})^\dagger = \hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger$). Consequently, the expectation value of the number operator is found to be

$$\begin{aligned} \langle \alpha | \hat{n} | \alpha \rangle &= \langle 0 | \hat{D}^\dagger(\alpha) \hat{n} \hat{D}(\alpha) | 0 \rangle \\ &= \langle 0 | \hat{D}^\dagger(\alpha) \hat{a}^\dagger \hat{D}(\alpha) \hat{D}^\dagger(\alpha) \hat{a} \hat{D}(\alpha) | 0 \rangle \\ &= \langle 0 | (\hat{a}^\dagger + \alpha^*)(\hat{a} + \alpha) | 0 \rangle = |\alpha|^2. \end{aligned} \quad (2.39)$$

Additionally, the for two distinct displacement operators we find

$$\begin{aligned} \hat{D}(\alpha)\hat{D}(\beta) &= e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} e^{\beta\hat{a}^\dagger - \beta^*\hat{a}} \\ &= e^{i\text{Im}(\alpha\beta^*)} e^{(\alpha+\beta)\hat{a}^\dagger - (\alpha^*+\beta^*)\hat{a}} \\ &= e^{i\text{Im}(\alpha\beta^*)} \hat{D}(\alpha + \beta) \end{aligned} \quad (2.40)$$

where we have used $[\alpha\hat{a}^\dagger - \alpha^*\hat{a}, \beta\hat{a}^\dagger - \beta^*\hat{a}] = \alpha\beta^* - \alpha^*\beta = 2i\text{Im}(\alpha\beta^*)$. It is then simple to find that the overlap of two coherent states is given by

$$\begin{aligned} \langle \beta | \alpha \rangle &= \langle 0 | \hat{D}^\dagger(\beta) \hat{D}(\alpha) | 0 \rangle \\ &= e^{-\frac{1}{2}(|\beta|^2 + |\alpha|^2 - 2\beta^*\alpha)} \end{aligned} \quad (2.41)$$

2.2.1 The Wigner Function

In Fig. 2.1 we introduced the Wigner function in which enabled a phase space representation of the coherent state to be given, following the work of [21] the Wigner function itself is detailed here. For pure states, the Wigner function is defined as

$$W(x, p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \psi^* \left(x - \frac{q}{2} \right) \psi \left(x + \frac{q}{2} \right) e^{ipq/\hbar} dq \quad (2.42)$$

where $|x \pm q/2\rangle$ are the eigenstates of the position operator and $\psi(x + q/2) = \langle x + q/2 | \psi \rangle$ etc. Importantly, the Wigner function is a quasi-probability distribution since (2.42) can take negative values which arise for some non-classical states, as demonstrated in subsection 2.4. Indeed, negativity of the Wigner function is a good indicator of quantum

mechanical interference however, not all quantum mechanical states necessarily display negative regions of the Wigner function. Integrating (2.42) over momentum we find

$$\begin{aligned}
\int_{-\infty}^{\infty} W(x, p) dp &= \int_{-\infty}^{\infty} \psi^* \left(x - \frac{q}{2} \right) \psi \left(x + \frac{q}{2} \right) dq \int_{-\infty}^{\infty} e^{ipq/\hbar} dp dq \\
&= \int_{-\infty}^{\infty} \psi^* \left(x - \frac{q}{2} \right) \psi \left(x + \frac{q}{2} \right) \delta(q) dq \\
&= |\psi(x)|^2
\end{aligned} \tag{2.43}$$

and similarly by integrating (2.42) over position we find $\int_{-\infty}^{\infty} W(q, p) dx = |\varphi(p)|^2$ where $\varphi(p)$ is the probability amplitude of the momentum space related to the position space probability amplitude via a Fourier transform. In other words, the integral of the Wigner function over position or momentum gives the respective conjugate probability distribution. This can be generalised further [22] in that the integral of the Wigner function over any quadrature variable X_λ must be positive semi-definite. Hence, the Wigner function gives a useful phase space representation of optical states.

2.3 Squeezed States

2.3.1 Quadrature Squeezing

In the previous section we reviewed the “most classical” state of the QHO system, here we will look at one of many non-classical states. The squeezed state is defined to be a state whose variance in the general quadrature variable is less than that of the vacuum (or equivalently less than that of the coherent state). The mathematical statement is then

$$\Delta^2 \hat{X}_\lambda \leq \frac{1}{2} \tag{2.44}$$

and from the uncertainty relation (2.17) it is clear that for a conjugate pair of quadrature operators, if we have (2.44) then this imposes the condition

$$\Delta^2 \hat{X}_{\lambda+\pi/2} \geq \frac{1}{2} \tag{2.45}$$

and of course the reverse is true; $\Delta^2 \hat{X}_{\lambda+\pi/2} \leq \frac{1}{2} \Rightarrow \Delta^2 \hat{X}_\lambda \geq \frac{1}{2}$. Clearly the variance in any given quadrature has become dependent on λ , the state is said to be squeezed in the quadrature of smaller variance. Squeezed states are generated by the so-called squeezing operator

$$\hat{S}(\zeta) = \exp \left(\frac{1}{2} [\zeta^* a^2 - \zeta (a^\dagger)^2] \right) \tag{2.46}$$

where $\zeta = re^{i\phi} \in \mathbb{C}$ with $r \in [0, \infty)$ and $\phi \in [0, 2\pi]$. Note that much like the displacement operator, the squeezing operator is unitary since

$$\hat{S}^\dagger(\zeta) = \exp\left(-\frac{1}{2}[\zeta^* a^2 - \zeta(a^\dagger)^2]\right) = \hat{S}(-\zeta) \quad (2.47)$$

thus $\hat{S}^\dagger(\zeta)\hat{S}(\zeta) = \mathbb{I}$. To reveal how the squeezing operator results in reduced quadrature uncertainties consider the action of the squeezing operator on the general state $|\psi\rangle$ which gives a squeezed state $|\psi(\zeta)\rangle = \hat{S}(\zeta)|\psi\rangle$. For now, we shall aim to inspect the variances of the dimensionless position and momentum quadratures (2.10) and (2.11) respectively. In order to accomplish this, it is useful to know how the squeezing operators transform the annihilation operator (from which the action on the creation operator immediately follows) which, through use of (2.37), is found to be

$$\hat{S}^\dagger(\zeta)\hat{a}\hat{S}(\zeta) = \hat{a} \cosh(r) - \hat{a}^\dagger e^{i\phi} \sinh(r) \quad (2.48)$$

$$\hat{S}^\dagger(\zeta)\hat{a}^\dagger\hat{S}(\zeta) = \hat{a}^\dagger \cosh(r) - \hat{a} e^{-i\phi} \sinh(r) \quad (2.49)$$

hence

$$\langle\psi(\zeta)|\hat{a}|\psi(\zeta)\rangle = \langle\psi|\left(\hat{a} \cosh(r) - \hat{a}^\dagger e^{i\phi} \sinh(r)\right)|\psi\rangle \quad (2.50)$$

and further

$$\langle\psi(\zeta)|\hat{a}^2|\psi(\zeta)\rangle = \langle\psi|\left(\hat{a}^2 \cosh^2(r) - (2\hat{a}^\dagger\hat{a} + 1)e^{i\phi} \cosh(r) \sinh(r) + (\hat{a}^\dagger)^2 e^{2i\phi} \sinh^2(r)\right)|\psi\rangle. \quad (2.51)$$

For a specific example, the simplest squeezed state to investigate is the squeezed vacuum $|\zeta\rangle = \hat{S}(\zeta)|0\rangle$ for which it is immediately clear that $\langle\hat{X}\rangle = \langle\hat{P}\rangle = 0$. Then using (2.51) we find

$$\begin{aligned} \Delta^2 \hat{X} &= \frac{1}{2} (\cosh^2(r) + \sinh^2(r) - 2 \sinh(r) \cosh(r) \cos(\phi)) \\ \Delta^2 \hat{P} &= \frac{1}{2} (\cosh^2(r) + \sinh^2(r) + 2 \sinh(r) \cosh(r) \cos(\phi)) \end{aligned} \quad (2.52)$$

Setting $\phi = 0$ results in squeezing in the position axis

$$\Delta^2 \hat{X} = \frac{1}{2} e^{-2r}, \quad \Delta^2 \hat{P} = \frac{1}{2} e^{2r} \quad (2.53)$$

i.e, since $r \in [0, \infty)$, uncertainty in the position quadrature decreases with increased squeezing while uncertainty in momentum increases hence this is referred to as position quadrature squeezing. Setting $\phi = \pi$ gives

$$\Delta^2 \hat{X} = \frac{1}{2} e^{2r}, \quad \Delta^2 \hat{P} = \frac{1}{2} e^{-2r} \quad (2.54)$$

which evidently describes momentum quadrature squeezing. This can be furthered to encompass the general quadrature operator for which the variance is found to be

$$\Delta^2 \hat{X}_\lambda = \frac{1}{2} [e^{2r} \sin^2(\lambda - \phi/2) + e^{-2r} \cos^2(\lambda - \phi/2)]. \quad (2.55)$$

Squeezing of the vacuum state in various quadratures is demonstrated in Fig. 2.2, as previously noted the Wigner function is negative for some non-classical states, the squeezed states are a notable exception in being non-classical states that have entirely positive Wigner functions.

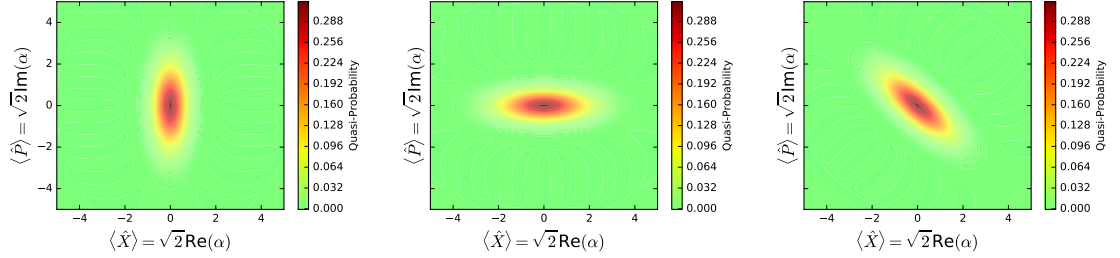


Figure 2.2: Plots of Wigner functions for the squeezed vacuum state with $\phi = 0$ (left) $\phi = \pi$ (centre) and $\phi = \pi/2$ (right) and $r = 1/2$ for all. This corresponds to squeezing in the position, momentum and general quadrature axis respectively.

It is useful to inspect some of the overlaps of squeezed states with other common optical states. Firstly, the squeezed vacuum state takes the following form in the number state basis

$$|\zeta\rangle = \frac{1}{\sqrt{\cosh(r)}} \sum_{m=0}^{\infty} (-1)^m \frac{\sqrt{(2m)!}}{2^m m!} e^{im\theta} (\tanh(r))^m |2m\rangle \quad (2.56)$$

the overlap of two squeezed vacuum states is then found to be

$$\langle \zeta | \zeta' \rangle = \left(\frac{\text{sech}(r) \text{sech}(r')}{1 - e^{i(\phi' - \phi)} \tanh(r) \tanh(r')} \right)^{1/2} \quad (2.57)$$

where $\zeta' = r' e^{i\phi'}$. The overlap of a squeezed vacuum with a coherent state is given by

$$\langle \alpha | \zeta \rangle = \sqrt{\text{sech}(r)} \exp \left(-\frac{1}{2} (\alpha^*)^2 e^{i\phi} \tanh(r) \right) e^{-|\alpha|^2/2}. \quad (2.58)$$

Finally, we note that the expectation value of the number operator with respect to the squeezed vacuum is given by

$$\begin{aligned} \langle \zeta | \hat{n} | \zeta \rangle &= \langle 0 | \hat{S}(\zeta)^\dagger \hat{a}^\dagger \hat{S}(\zeta) \hat{S}^\dagger(\zeta) \hat{a} \hat{S}(\zeta) | 0 \rangle \\ &= \langle 0 | [\hat{a}^\dagger \cosh(r) - \hat{a} e^{-i\phi} \sinh(r)] [\hat{a} \cosh(r) - \hat{a}^\dagger e^{i\phi} \sinh(r)] | 0 \rangle \\ &= \sinh^2(r) \end{aligned} \quad (2.59)$$

hence as squeezing is increased, the photon number increases or in other words, squeezing the vacuum state requires energy input and results in a non-zero mean photon number in the resultant state. As we have seen, the vacuum is a special case of the coherent state and similarly the squeezed vacuum is a special case of the more general squeezed-displaced vacuum states which are given by

$$|\alpha, \zeta\rangle = \hat{D}(\alpha)\hat{S}(\zeta)|0\rangle. \quad (2.60)$$

It is crucial to note that the squeezing and displacement operators do not commute $[\hat{S}(\zeta), \hat{D}(\alpha)] \neq 0$. In fact, by using (2.49), we have

$$\begin{aligned} \hat{S}(\zeta)\hat{D}(\alpha)|0\rangle &= \hat{S}(\zeta)\hat{D}(\alpha)\hat{S}(-\zeta)\hat{S}(\zeta)|0\rangle \\ &= \exp\left[\alpha(\hat{a}^\dagger \cosh(r) + \hat{a}e^{-i\phi} \sinh(r)) - \alpha^*(\hat{a} \cosh(r) + \hat{a}^\dagger e^{i\phi} \sinh(r))\hat{S}(\zeta)\right]|0\rangle \\ &= \hat{D}(\alpha \cosh(r) - \alpha^* e^{i\phi} \sinh(r))\hat{S}(\zeta)|0\rangle \\ &= |\gamma, \zeta\rangle \end{aligned} \quad (2.61)$$

where $\gamma = \alpha \cosh(r) - \alpha^* e^{i\phi} \sinh(r)$. Thus,

$$|\alpha, \zeta\rangle = \hat{D}(\alpha)\hat{S}(\zeta)|0\rangle = \hat{S}(\zeta)\hat{D}(\gamma)|0\rangle. \quad (2.62)$$

The translation of the annihilation operator under the joint action of the squeezing and displacement operators is given by

$$\begin{aligned} \hat{S}^\dagger(\zeta)\hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha)\hat{S}(\zeta) &= \hat{S}^\dagger(\zeta)(\hat{a} + \alpha)\hat{S}(\zeta) \\ &= \hat{a} \cosh(r) - \hat{a}^\dagger e^{i\phi} \sinh(r) + \alpha \end{aligned} \quad (2.63)$$

and similarly for the creation operator

$$\hat{S}^\dagger(\zeta)\hat{D}^\dagger(\alpha)\hat{a}^\dagger\hat{D}(\alpha)\hat{S}(\zeta) = \hat{a}^\dagger \sinh(r) - \hat{a}e^{-i\phi} \sinh(r) + \alpha^* \quad (2.64)$$

then it is straight forward to find that

$$\begin{aligned} \langle\alpha, \zeta|\hat{a}|\alpha, \zeta\rangle &= \hat{S}^\dagger(\zeta)\hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha)\hat{S}(\zeta) \\ &= \langle 0|(\hat{a} \cosh(r) - \hat{a}^\dagger e^{i\phi} \sinh(r) + \alpha)|0\rangle \\ &= \alpha \end{aligned} \quad (2.65)$$

and similarly $\langle\alpha, \zeta|\hat{a}|\alpha, \zeta\rangle = \alpha^*$. With this the expectation value of the general quadrature operator is revealed to be non-zero

$$\langle\alpha, \zeta|\hat{X}_\lambda|\alpha, \zeta\rangle = \frac{1}{\sqrt{2}}(\alpha e^{-i\lambda} + \alpha^* e^{i\lambda}) \quad (2.66)$$

and furthermore,

$$\langle \alpha, \zeta | \hat{X}_\lambda^2 | \alpha, \zeta \rangle = \frac{1}{2} (e^{2r} \sin^2(\lambda - \phi/2) + e^{-2r} \cos^2(\lambda - \phi/2)) + \frac{1}{2} \left(\alpha e^{-i\lambda} + \alpha^* e^{i\lambda} \right)^2 \quad (2.67)$$

so that the variance of the general quadrature operator with respect to the displaced squeezed state is found to be

$$\Delta^2 \hat{X}_\lambda = \frac{1}{2} (e^{2r} \sin^2(\lambda - \phi/2) + e^{-2r} \cos^2(\lambda - \phi/2)) \quad (2.68)$$

which is precisely the expression of the variance for the squeezed vacuum state (2.55) thus the expectation value is dependent only on the displacement while the variance is dependent only on the squeezing. The Wigner function for the squeezed-displaced vacuum is plotted in Fig.2.3

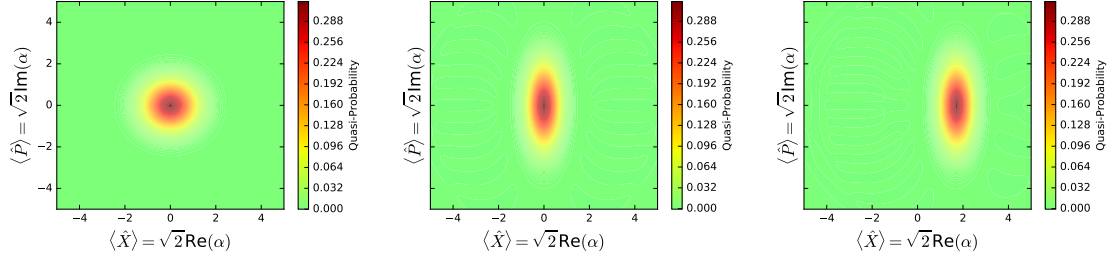


Figure 2.3: Plots of Wigner functions for the vacuum state (*left*), the squeezed vacuum state with $r = \frac{1}{2}$ and $\phi = 0$ (*centre*) and the squeezed-displaced vacuum state with $r = \frac{1}{2}$, $\phi = 0$ and $\alpha = 2$ (*right*).

2.3.2 Practical Generation of Squeezed Light

A widely practised method of generating quadrature-squeezed light uses a device known as a “parametric down converter” in which a nonlinear medium is pumped by a field of frequency ω_p . From this pump field a signal field is generated via the conversion of some of the pump photons into pairs of identical (signal) photons each of frequency $\omega = \omega_p/2$. Following the work of [21] we give an account of the process here. The Hamiltonian governing the process is given by

$$H = \hbar\omega a^\dagger a + \hbar\omega_p b^\dagger b + i\hbar\chi^{(2)}[a^2 b^\dagger - (a^\dagger)^2 b] \quad (2.69)$$

where, a and b are the annihilation operators (for brevity we have dropped the “hat” notation for these operators) for signal and pump modes respectively and $\chi^{(2)}$ is the second order nonlinear susceptibility (of the medium). Now by making the “parametric

approximation” i.e by assuming the pump field is a coherent classical field which loses no photons over the duration of the process, denoted by time t , the field state is given by $|\beta e^{-i\omega_p t}\rangle$ where $\beta = |\beta|e^{i\theta}$ with $\theta \in [0, 2\pi]$ and the operators b and b^\dagger are approximated by $\beta e^{-i\omega_p t}$ and $\beta^* e^{i\omega_p t}$ respectively. The Hamiltonian then becomes

$$H = \hbar\omega a^\dagger a + i\hbar[\eta^* a^2 e^{i\omega_p t} - \eta(a^\dagger)^2 e^{-i\omega_p t}] \quad (2.70)$$

where, $\eta = \chi^{(2)}\beta$ and the irrelevant constant term in the Hamiltonian has been dropped. Transforming to the interaction picture yields the Hamiltonian

$$H_I(t) = i\hbar \left[\eta^* a^2 e^{i(\omega_p - 2\omega)t} - \eta(a^\dagger)^2 e^{-i(\omega_p - 2\omega)t} \right] \quad (2.71)$$

whose time dependency is overcome by choosing the classical pump field frequency to be $\omega_p = 2\omega$ thus defining the relationship between pump and signal photons. The final Hamiltonian is the given by

$$H_I = i\hbar[\eta^* a^2 - \eta(a^\dagger)^2] \quad (2.72)$$

and the associated unitary time evolution operator is given by

$$\begin{aligned} U_I(t) &= \exp \left(\eta^* t a^2 - \eta t (a^\dagger)^2 \right) \\ &= \exp \left((\chi^{(2)}|\beta|e^{i\theta})^* t a^2 - \chi^{(2)}|\beta|e^{i\theta} t (a^\dagger)^2 \right) \end{aligned} \quad (2.73)$$

which we compare to the familiar form of the squeezing operator (2.46) and discern the relationship between the physical parameters involved in the parametric down conversion and the compact form of the squeezing parameter used in (2.46)

$$\begin{aligned} r &= 2\chi^{(2)}|\beta|t \\ &= 2\eta t. \end{aligned} \quad (2.74)$$

2.4 Gaussian and Non-Gaussian States

2.4.1 Gaussian Probability Distributions

So far, we have encountered the “most classical” and quadrature squeezed states of the QHO. If we are to inspect the probability distribution function (PDF) of each of these states in the position and momentum quadratures we find that each is a Gaussian distribution as depicted in Fig.2.4. To see this more concretely, we first find the overlap of a general coherent state with $\alpha \in \mathbb{R}$

$$|e^{i\theta}\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{(e^{i\theta}\alpha)^n}{\sqrt{n!}} |n\rangle \quad (2.75)$$

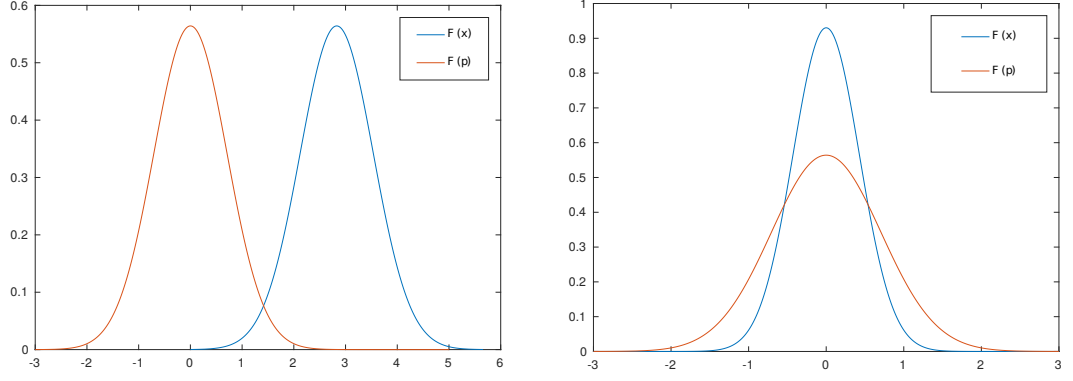


Figure 2.4: Plots of position and momentum probability distributions for a coherent state with $\alpha = 2$ (*left*) and for a squeezed vacuum state with $r = 1/2$ and $\phi = 0$ (*right*). All are Gaussian PDFs.

with the general quadrature state (the eigenstate of \hat{X}_λ as given by (2.15)) in the number state representation

$$|X_\lambda\rangle = \sum_{n=0}^{\infty} (2^n n!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-X_\lambda^2/2} e^{-in\lambda} H_n(X_\lambda) |n\rangle. \quad (2.76)$$

where, $H_n(X_\lambda)$ are the Hermite polynomials of degree n . Hence, for the overlap we find

$$\begin{aligned} \langle X_\lambda | e^{i\theta} \alpha \rangle &= \sum_{m=0}^{\infty} (2^m m!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-X_\lambda^2/2} e^{im\lambda} H_m(X_\lambda) \cdot e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{(e^{i\theta} \alpha)^n}{\sqrt{n!}} \langle m | n \rangle \\ &= \pi^{-\frac{1}{4}} \exp\left(-\frac{1}{2}(X_\lambda^2 + \alpha^2)\right) \sum_{n=0}^{\infty} \frac{1}{n!} H_n(X_\lambda) (2^{-\frac{1}{2}} e^{i\lambda} e^{i\theta} \alpha)^n \end{aligned} \quad (2.77)$$

noting that $\alpha \in \mathbb{R}$, hence $|\alpha|^2 = \alpha^2$. Then using the following relation

$$\sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) t^n = e^{-t^2 + 2tx} \quad (2.78)$$

we find

$$\begin{aligned} \langle X_\lambda | e^{i\theta} \alpha \rangle &= \pi^{-\frac{1}{4}} \exp\left(-\frac{1}{2}(X_\lambda^2 + \alpha^2)\right) \exp\left(-2^{-1} e^{2i(\lambda+\theta)} \alpha^2 + 2(2^{-\frac{1}{2}} e^{i(\lambda+\theta)} \alpha) X_\lambda\right) \\ &= \pi^{-\frac{1}{4}} \exp\left(-\frac{1}{2}(X_\lambda^2 + \alpha^2)\right) \exp\left(-\frac{1}{2} e^{2i(\lambda+\theta)} \alpha^2 + \sqrt{2} e^{i(\lambda+\theta)} \alpha X_\lambda\right). \end{aligned} \quad (2.79)$$

thus the PDFs of a coherent state with $\theta = 0$ for the conjugate quadrature variables X_0 and $X_{\pi/2}$ are given by

$$\mathcal{F}(X_0) = |\langle X_0 | \alpha \rangle|^2 = \pi^{-\frac{1}{2}} e^{-(X_0 - \sqrt{2}\alpha)^2} \quad (2.80)$$

$$\mathcal{F}(X_{\pi/2}) = |\langle X_{\pi/2} | \alpha \rangle|^2 = \pi^{-\frac{1}{2}} e^{-X_{\pi/2}^2}. \quad (2.81)$$

Comparing this to the general form of a Gaussian distribution

$$f(x) = a \exp\left(\frac{-(x-b)^2}{2c^2}\right) \quad (2.82)$$

it is clear that both (2.80) and (2.81) are (normalised) Gaussian PDFs. Similarly for the squeezed vacuum state, the overlap is given by [22]

$$\langle X_\lambda | \zeta \rangle = (2\pi\Delta^2 X_\lambda)^{-1/4} \exp\left(\frac{X_\lambda^2}{4\Delta^2 X_\lambda} [1 - i \sin(2\lambda - \phi) \sinh(2r)]\right) \quad (2.83)$$

where the variance $\Delta^2 X_\lambda$ is given by (2.55) hence, taking $\phi = 0$, the PDFs of the squeezed vacuum state for the conjugate quadrature variables X_0 and $X_{\pi/2}$ are found to be

$$\mathcal{F}(X_0) = \pi^{-1/2} \exp(-X_0^2 e^{2r} + r) \quad (2.84)$$

$$\mathcal{F}(X_{\pi/2}) = \pi^{-1/2} e^{-X_{\pi/2}^2} \quad (2.85)$$

and again, in comparison to (2.82) the PDFs are revealed to be Gaussian even though the states are non-classical. At this point we note that the displacement operator governs the first statistical moment (the mean) of the Gaussian PDF and the squeezing operator governs the second moment (the variance). Furthermore, the probability amplitudes given by $\psi(X_\lambda) = \langle X_\lambda | \psi \rangle$ and $\varphi(X_{\lambda+\pi/2}) = \langle X_{\lambda+\pi/2} | \psi \rangle$ are related by the Fourier transform

$$\psi(\hat{X}_\lambda) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} e^{\frac{i}{\hbar} X_\lambda X_{\lambda+\pi/2}} \varphi(X_{\lambda+\pi/2}) dX_{\lambda+\pi/2} \quad (2.86)$$

2.4.2 Non-Gaussian States

As a first example of a state with a non-Gaussian PDF in phase space, we inspect the number states $|n\rangle$ (as given by (2.24)) also referred to as Fock states [28]. In Fig.2.5 we see the Wigner functions take on negative values implying the quantum (non-classical) nature of the Fock states and clearly the nature of the position and momentum PDFs are more exotic than Gaussian. Indeed, inspection of the overlap with the general quadrature

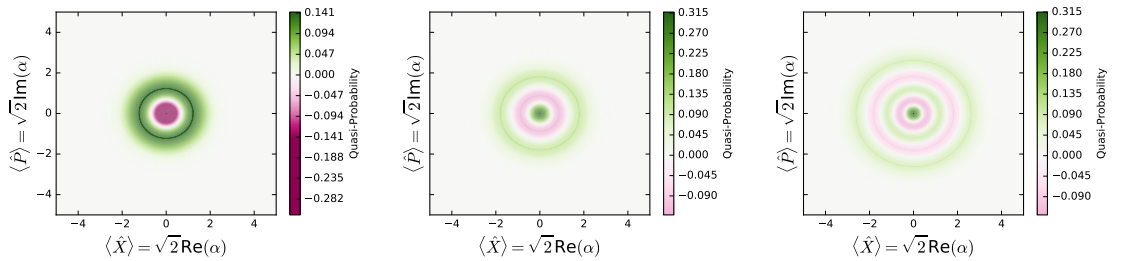


Figure 2.5: Plots of Wigner functions for the Fock states $|n\rangle$ with $n = 1$ (left) $n = 2$ (centre) and $n = 4$ (right).

variable reveals

$$\begin{aligned}\langle n|X_\lambda\rangle &= \sum_{m=0}^{\infty} (2^m m!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-X_\lambda^2/2} e^{-im\lambda} H_m(X_\lambda) \langle n|m\rangle \\ &= (2^n n!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-X_\lambda^2/2} e^{-in\lambda} H_n(X_\lambda)\end{aligned}\quad (2.87)$$

thus the overlaps for the Fock states with the quadrature states for $X = X_0$ and $P = X_{\pi/2}$ only differ by a phase factor and thus $\mathcal{F}(x) = \mathcal{F}(p)$ which is demonstrated in Fig.2.5. Furthermore, Fig.2.6 shows the overall non-Gaussian structure of the PDFs for various Fock states.

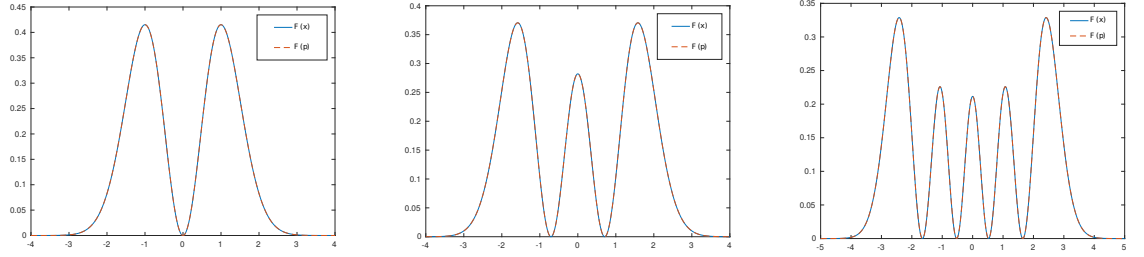


Figure 2.6: Probability distribution functions for the Fock states $|n\rangle$ with $n = 1$ (*left*) $n = 2$ (*centre*) and $n = 4$ (*right*). All PDFs display an overall non-Gaussian structure.

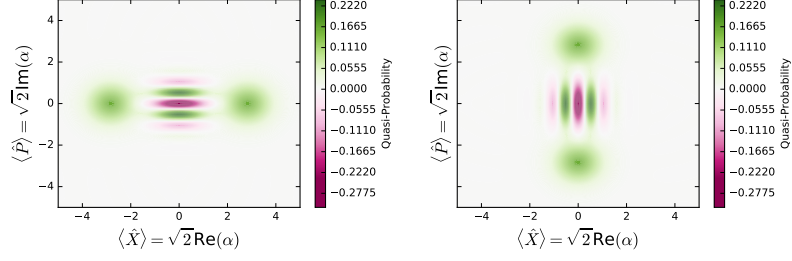


Figure 2.7: Plots of Wigner functions for odd cat states with $\phi = 0$ and $\tilde{\phi} = -\pi$ (left) and with $\phi = \pi/2$ and $\tilde{\phi} = -\pi/2$ (right).

We now turn our attention to the following state

$$|\psi_{CS}^{\pm}(\phi, \tilde{\phi})\rangle = \mathcal{N}_{\alpha}^{\pm}(\phi, \tilde{\phi})(|e^{i\phi}\alpha\rangle \pm |e^{i\tilde{\phi}}\alpha\rangle) \quad (2.88)$$

where the normalising factor is given by

$$\mathcal{N}_{\alpha}^{\pm}(\phi, \tilde{\phi}) = \left[2 \pm e^{-\alpha^2} \left(\exp(\alpha^2 e^{i(\tilde{\phi}-\phi)}) + \exp(\alpha^2 e^{i(\phi-\tilde{\phi})}) \right) \right]^{-\frac{1}{2}}. \quad (2.89)$$

This state is a superposition of two macroscopic states and as such, this type of state is known as a “cat state” in reflection of Shrödinger’s thought experiment. When the constituent superposition states are $\phi = \pi$ out of phase the state is referred to as an “even cat state” for the “+” case and an “odd cat state” for the “-” case, this is because the even and odd states contain only even or odd numbers of photons, this can be seen from the simple example of $\phi = 0$ and $\tilde{\phi} = \pi$

$$|\psi_{CS}^{+}(0, \pi)\rangle = 2\mathcal{N}_{\alpha}^{+}(0, \pi)e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^{2n}}{(2n)!} |2n\rangle \quad (2.90)$$

$$|\psi_{CS}^{-}(0, \pi)\rangle = 2\mathcal{N}_{\alpha}^{-}(0, \pi)e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^{2n+1}}{(2n+1)!} |2n+1\rangle. \quad (2.91)$$

The Wigner functions for two examples of odd cat states are plotted in Fig.2.7 where interference fringes - a manifestation of the quantum superposition principle - are clearly visible. The probability distribution of the general quadrature variable for state (2.88) is given by

$$\begin{aligned} \mathcal{F}^{\pm}(X_{\lambda}, \phi, \tilde{\phi}) &= |\langle X_{\lambda} | \psi_{CS}^{\pm}(\phi, \tilde{\phi}) \rangle|^2 \\ &= \langle X_{\lambda} | \psi_{CS}^{\pm}(\phi, \tilde{\phi}) \rangle \langle \psi_{CS}^{\pm}(\phi, \tilde{\phi}) | X_{\lambda} \rangle \\ &= \langle X_{\lambda} | \psi_{CS}^{\pm}(\phi, \tilde{\phi}) \rangle \left(\langle X_{\lambda} | \psi_{CS}^{\pm}(\phi, \tilde{\phi}) \rangle \right)^{\dagger} \\ &= (N_{\alpha}^{\pm})^2 \left(\langle X_{\lambda} | e^{i\phi}\alpha \rangle \pm \langle X_{\lambda} | e^{i\tilde{\phi}}\alpha \rangle \right) \left[\left(\langle X_{\lambda} | e^{i\phi}\alpha \rangle \pm \langle X_{\lambda} | e^{i\tilde{\phi}}\alpha \rangle \right) \right]^{\dagger}. \end{aligned} \quad (2.92)$$

Everything we need here is given by (2.79), we find

$$\begin{aligned} \langle X_\lambda | e^{i\phi} \alpha \rangle + \langle X_\lambda | e^{i\tilde{\phi}} \alpha \rangle = \pi^{-\frac{1}{4}} \exp \left(-\frac{1}{2} (X_\lambda^2 + \alpha^2) \right) \left[\exp \left(-\frac{1}{2} e^{2i(\lambda+\phi)} \alpha^2 + \sqrt{2} e^{i(\lambda+\phi)} \alpha X_\lambda \right) \right. \\ \left. + \exp \left(-\frac{1}{2} e^{2i(\lambda+\tilde{\phi})} \alpha^2 + \sqrt{2} e^{i(\lambda+\tilde{\phi})} \alpha X_\lambda \right) \right] \end{aligned} \quad (2.93)$$

thus the general probability density function is given by

$$\begin{aligned} \mathcal{F}^\pm(X_\lambda, \phi, \tilde{\phi}) = \pi^{-\frac{1}{2}} (N_\alpha^\pm)^2 \left| \exp \left(-\frac{1}{2} (X_\lambda^2 + \alpha^2) \right) \left[\exp \left(-\frac{1}{2} e^{2i(\lambda+\phi)} \alpha^2 + \sqrt{2} e^{i(\lambda+\phi)} \alpha X_\lambda \right) \right. \right. \\ \left. \left. \pm \exp \left(-\frac{1}{2} e^{2i(\lambda+\tilde{\phi})} \alpha^2 + \sqrt{2} e^{i(\lambda+\tilde{\phi})} \alpha X_\lambda \right) \right] \right|^2. \end{aligned} \quad (2.94)$$

Finding a specific state-variable probability distribution function is then simply a matter of inputting the desired phase values λ, ϕ and $\tilde{\phi}$. As an example we shall inspect the usual probability distribution functions of orthogonal quadrature variables x_0 and $x_{\pi/2}$ for the even cat state

$$|\psi_{CS}^+(\phi = \pi/2, \tilde{\phi} = -\pi/2)\rangle = \mathcal{N}_\alpha^+ (|i\alpha\rangle + |-i\alpha\rangle) \quad (2.95)$$

which gives the following

$$\begin{aligned} \mathcal{F}^+(x_0, \pi/2, -\pi/2) = \pi^{-\frac{1}{2}} (N_\beta^+)^2 \left| \exp \left(-\frac{1}{2} (x_0^2 + \beta^2) \right) \left[\exp \left(-\frac{1}{2} (-1)\beta^2 + \sqrt{2}(i)\beta x_0 \right) \right. \right. \\ \left. \left. + \exp \left(-\frac{1}{2} (-1)\beta^2 + \sqrt{2}(-i)\beta x_0 \right) \right] \right|^2 \end{aligned} \quad (2.96)$$

thus,

$$\begin{aligned} \mathcal{F}^+(x_0, \pi/2, -\pi/2) &= \pi^{-\frac{1}{2}} (N_\beta^+)^2 \left| \exp \left(-\frac{x_0^2}{2} \right) \left(e^{i\sqrt{2}\beta x_0} + e^{-i\sqrt{2}\beta x_0} \right) \right|^2 \\ &= \pi^{-\frac{1}{2}} (N_\beta^+)^2 \left| \exp \left(-\frac{x_0^2}{2} \right) 2 \cos(\sqrt{2}\beta x_0) \right|^2 \\ &= 4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_0^2} \cos^2(\sqrt{2}\beta x_0). \end{aligned} \quad (2.97)$$

where we have used $\cos(x) = (e^{ix} + e^{-ix})/2$. Through very similar working we find for the conjugate variable (i.e for $\lambda = \pi/2$)

$$\mathcal{F}^+(x_{\pi/2}, \pi/2, -\pi/2) = 4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_{\pi/2}^2 - 2\beta^2} \cosh^2(\sqrt{2}\beta x_{\pi/2}). \quad (2.98)$$

Noting further that $\sin(x) = (e^{ix} - e^{-ix})/2i$ and $\sinh(x) = (e^x - e^{-x})/2$ we can immediately conclude that the probability distributions of the conjugate quadrature variables for the odd cat state are given by

$$\mathcal{F}^-(x_0, \pi/2, -\pi/2) = 4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_0^2} \sin^2(\sqrt{2}\beta x_0) \quad (2.99)$$

and

$$\mathcal{F}^-(x_{\pi/2}, \pi/2, -\pi/2) = 4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_{\pi/2}^2 - 2\beta^2} \sinh^2(\sqrt{2}\beta x_{\pi/2}) \quad (2.100)$$

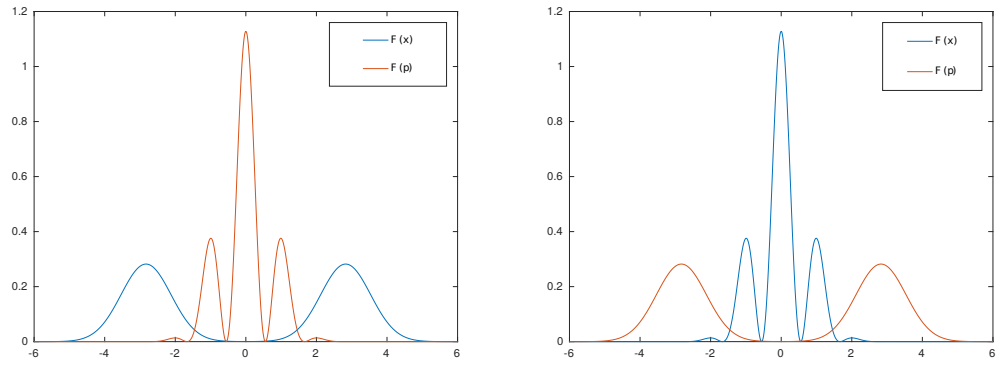


Figure 2.8: Probability distribution functions for odd cat states with $\phi = 0$ and $\tilde{\phi} = -\pi$ (*left*) and with $\phi = \pi/2$ and $\tilde{\phi} = -\pi/2$ (*right*). Both PDFs display an overall non-Gaussian structure.

Chapter 3

Fermionic Systems

3.1 Two-Level System

All elementary fermions are spin-1/2 particles [29] meaning that observation of the magnetic quantum number along any one axis yields the eigenvalues $\pm\hbar/2$ only, this is predicted through solving the Shrödinger equation for angular momentum. This phenomenon was first measured in the famous Stern-Gerlach experiment [30] which was designed to test the “intrinsic angular momentum” (spin) of silver atoms, the values of which were expected to be integer valued which would have required the observation of three fringes. Contrary to the expected results, only two fringes were observed and thus the conclusion was that the silver atoms had intrinsic angular momentum of 1/2. As these spin-1/2 particles can only be observed in one of two states they are often referred to as two-level systems; these are among the simplest of quantum mechanical systems. The spin operators (S_x, S_y, S_z) can be described in terms of the Pauli matrices

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z \quad (3.1)$$

where the Pauli matrices are given by

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (3.2)$$

which satisfy the following commutation relations

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (3.3)$$

The eigenvectors of the Pauli matrices are denoted by

$$\begin{aligned}\sigma_x |\rightarrow\rangle &= |\rightarrow\rangle, & \sigma_x |\leftarrow\rangle &= -|\leftarrow\rangle \\ \sigma_y |\boxtimes\rangle &= |\boxtimes\rangle, & \sigma_y |\odot\rangle &= -|\odot\rangle \\ \sigma_z |\uparrow\rangle &= |\uparrow\rangle, & \sigma_z |\downarrow\rangle &= -|\downarrow\rangle\end{aligned}\tag{3.4}$$

where,

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad |\boxtimes\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad |\odot\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad |\uparrow\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\downarrow\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\tag{3.5}$$

and from this it is evident that

$$\begin{aligned}|\rightarrow\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle), & |\leftarrow\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) \\ |\boxtimes\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle), & |\odot\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle - i|\downarrow\rangle)\end{aligned}\tag{3.6}$$

which reveals that any pure state of a two-level quantum can be written as a superposition of the basis vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ thus, taking $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$, a general pure state vector of a two-level system can be written as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |\uparrow\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |\downarrow\rangle\tag{3.7}$$

where we have taken the coefficient of $|\uparrow\rangle$ to be real and positive since it is only the relative phase of the basis states that has physical consequence. It is then possible to use parameters θ and ϕ to specify a vector in spherical coordinates

$$\vec{r} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)\tag{3.8}$$

which submits a useful representation of the general two-level system i.e any state of a spin-1/2 particle can be represented by the vector \vec{r} ; this is known as the Bloch sphere representation and is depicted in Fig.3.1. Furthermore, the Pauli matrices, along with the 2×2 identity matrix \mathbb{I} (sometimes denoted σ_0), form a basis for the 2×2 vector space of Hermitian matrices. This means that any operator on the spin-1/2 Hilbert space ($\mathcal{H} = \mathbb{C}^2$) can be represented by a linear combination of the Pauli matrices and thus the general density matrix of a two-level system can be given by

$$\rho = \frac{1}{2} (\mathbb{I} + \vec{r} \cdot \vec{\sigma})\tag{3.9}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. Furthermore, pure states have $|\vec{r}| = 1$ and exist on the surface of the Bloch sphere while mixed states have $|\vec{r}| < 1$ and thus exists within the sphere.

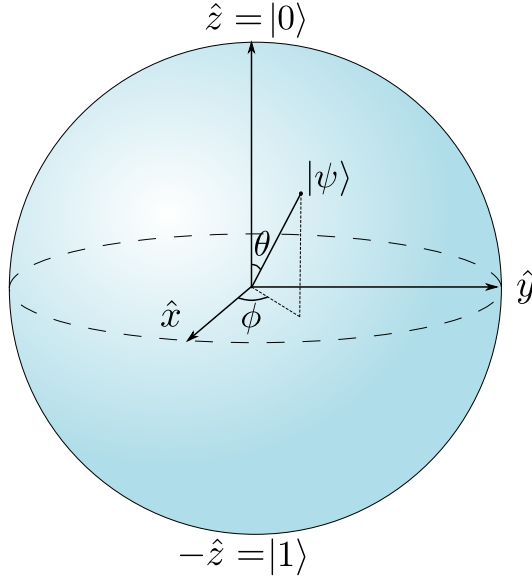


Figure 3.1: The Bloch sphere representation of a two-level system such as a spin-1/2 particle. The choice of axis that represent the $|0\rangle$ and $|1\rangle$ state are arbitrary but conventionally taken to be along the positive and negative z axis.

3.2 N Spin-1/2 System

We now extend the two level system to that of an N spin-1/2 system. We begin by considering a system of two spin-1/2 particles which exists in a Hilbert space described by the tensor produce of the individual spaces of the two particles $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$, here the basis states can have the property of being either symmetric

$$|\uparrow, \uparrow\rangle, \quad |\downarrow, \downarrow\rangle, \quad \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) \quad (3.10)$$

or antisymmetric

$$\frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle) \quad (3.11)$$

which is determined by how the state changes when any two spins are swapped - a property bosons do not have. For a system of N spin-1/2 particles, the Hilbert space is given by $\mathcal{H}_N = (\mathbb{C}^2)^{\otimes N}$ and the associated spin operators are sometimes referred to as “big-spin operators” or “collective spin operators”, given by

$$J_\alpha = \frac{1}{2} \sum_{i=1}^N \sigma_\alpha^{(i)}, \quad \text{for } \alpha \in \{x, y, z\} \quad (3.12)$$

which satisfy the commutation relation

$$[J_\mu, J_\nu] = i\epsilon^{\mu\nu\rho} J_\rho \quad (3.13)$$

where $\epsilon^{\mu\nu\rho}$ is the antisymmetric tensor (such that $\epsilon^{xyz} = 1$). This reveals the uncertainty relation

$$\Delta^2 J_z \Delta^2 J_y \geq \frac{1}{4} \langle J_x^2 \rangle. \quad (3.14)$$

The dimension of the Hilbert space for this N spin-1/2 system is $\text{Dim}(\mathcal{H}_N) = 2^N$ i.e it grows exponentially with the number of particles. We now introduce the so-called ‘‘Dicke states’’ $|j, m\rangle_N$. These states must satisfy two conditions, the first is that

$$J_z |j, m\rangle_N = m |j, m\rangle_N \quad (3.15)$$

where $j \in \{0, 1, \dots, \frac{N}{2}\}$ for N even, $j \in \{\frac{1}{2}, \frac{3}{2}, \dots, \frac{N}{2}\}$ for N odd, $m \in \{-j, -j+1, \dots, j\}$ and the subscript of $|\cdot\rangle_N$ denotes a state of the N spin-1/2 system. For the second requirement we must introduce the ‘‘total spin’’ operator

$$J^2 = J_x^2 + J_y^2 + J_z^2 \quad (3.16)$$

then the state $|j, m\rangle_N$ is a Dicke state if it satisfies both (3.15) and

$$J^2 |j, m\rangle_N = j(j+1) |j, m\rangle_N \quad (3.17)$$

or in other words, a Dicke state is defined to be the *simultaneous* eigenstate of J_z and J^2 (hence these two operators commute). Akin to the creation and annihilation operators (2.23) and (2.22), we are able to define the raising and lowering operators

$$J_{\pm} |j, m\rangle_N = \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle_N \quad (3.18)$$

which satisfy the commutation relations

$$[J_-, J_+] = -2J_z, \quad [J_z, J_{\pm}] = \pm J_{\pm}, \quad [J^2, J_{\pm}] = 0. \quad (3.19)$$

The $j = N/2$ subspace is one of particular interest, the Dicke states are completely symmetric in this subspace [31] and $m \in \{-N/2, \dots, N/2\}$, thus there are $N+1$ of these states. In this subspace it is sometimes helpful to relabel the Dicke states as $|N/2, m\rangle_N = |N/2, n - N/2\rangle_N \equiv |n\rangle_N$. There is an underlying connection between the model of quantised angular momentum and the model of two uncorrelated QHO's which we will explore here. It is first useful to define

$$N_+ \equiv a_{\uparrow}^{\dagger} a_{\uparrow} \equiv \frac{N}{2} + J_z \quad \text{and} \quad N_- \equiv a_{\downarrow}^{\dagger} a_{\downarrow} \equiv \frac{N}{2} - J_z \quad (3.20)$$

which, upon acting on a Dicke state in the $j = N/2$ subspace, returns the number of spin-up and spin-down particles as the eigenvalue

$$a_{\uparrow}^{\dagger} a_{\uparrow} |N/2, m\rangle_N = \left(\frac{N}{2} + m\right) |N/2, m\rangle_N \quad (3.21)$$

$$a_{\downarrow}^{\dagger} a_{\downarrow} |N/2, m\rangle_N = \left(\frac{N}{2} - m\right) |N/2, m\rangle_N. \quad (3.22)$$

If we rewrite this as $N_+ |n_+, n_-\rangle = n_+ |n_+, n_-\rangle$ and $N_- |n_+, n_-\rangle = n_- |n_+, n_-\rangle$, then in complete analogy with the creation and annihilation operators (2.23) and (2.22) we have

$$a_{\uparrow}^{\dagger} |n_+, n_-\rangle = \sqrt{n_+ + 1} |n_+ + 1, n_-\rangle, \quad a_{\uparrow} |n_+, n_-\rangle = \sqrt{n_+} |n_+ - 1, n_-\rangle \quad (3.23)$$

$$a_{\downarrow}^{\dagger} |n_+, n_-\rangle = \sqrt{n_- + 1} |n_+, n_- + 1\rangle, \quad a_{\downarrow} |n_+, n_-\rangle = \sqrt{n_-} |n_+, n_- - 1\rangle \quad (3.24)$$

and as such we are able to obtain any general eigenstate of N_+ and N_- by repeatedly applying a_{\uparrow}^{\dagger} and a_{\downarrow}^{\dagger} to the vacuum state but this requires redefining the notion of the vacuum state since we have the ground state $|N/2, -N/2\rangle_N = |\downarrow\rangle^{\otimes N}$ (all spins down) and the “roof state” $|N/2, N/2\rangle_N = |\uparrow\rangle^{\otimes N}$ (all spins up), thus we define the vacuum state to be $|0, 0\rangle$ so that

$$a_{\uparrow} |0, 0\rangle_N = 0 = a_{\downarrow} |0, 0\rangle_N \quad (3.25)$$

and

$$|n_+, n_-\rangle_N = \frac{(a_{\uparrow}^{\dagger})^{n_+} (a_{\downarrow}^{\dagger})^{n_-}}{\sqrt{n_+!} \sqrt{n_-!}} |0, 0\rangle_N. \quad (3.26)$$

With this, it is straightforward to see that

$$a_{\uparrow}^{\dagger} a_{\downarrow} |n_+, n_-\rangle = \sqrt{n_- (n_+ + 1)} |n_+ + 1, n_- - 1\rangle \quad (3.27)$$

$$a_{\downarrow}^{\dagger} a_{\uparrow} |n_+, n_-\rangle = \sqrt{n_+ (n_- + 1)} |n_+ - 1, n_- + 1\rangle \quad (3.28)$$

then substituting $n_+ \rightarrow j + m$ and $n_- \rightarrow j - m$ we find the eigenvalues to be exactly that of (3.18). Furthermore, it is now natural to use $j \equiv (n_+ + n_-)/2$ and $m \equiv (n_+ - n_-)/2$ which reveals that the raising and lowering operators (3.18) are composed of

$$J_+ = a_{\uparrow}^{\dagger} a_{\downarrow}, \quad J_- = a_{\downarrow}^{\dagger} a_{\uparrow} \quad (3.29)$$

and additionally, the general Dicke state can be written

$$|j, m\rangle_N = \frac{(a_{\uparrow}^{\dagger})^{j+m} (a_{\downarrow}^{\dagger})^{j-m}}{\sqrt{(j+m)! (j-m)!}} |0, 0\rangle_N \quad (3.30)$$

3.3 Coherent Spin States and Rotations

There are multiple ways to define the Coherent Spin State (CSS), we begin with a specific example and establish the notion of a CSS in the $j = N/2$ subspace in which we can utilize the symmetry properties. In this setting the CSS can be regarded as the “most classical” state of N spin-1/2 particles or indeed of N two-mode bosons [32, 33] and are realised by placing all N particles in the same (arbitrary) superposition state of the two modes

$$|N/2, (\theta, \phi)\rangle_N = \frac{1}{\sqrt{N!}} \left[\cos\left(\frac{\theta}{2}\right) a_{\downarrow}^{\dagger} + e^{-i\phi} \sin\left(\frac{\theta}{2}\right) a_{\uparrow}^{\dagger} \right]^{\otimes N} |0, 0\rangle_N \quad (3.31)$$

and hence are parametrised by θ and ϕ . It is simple to find that for a single spin-1/2 particle oriented along the x -axis is $\Delta^2 \sigma_z \Delta^2 \sigma_y = \frac{1}{4} \cdot \frac{1}{4}$ and since the CSS is a separable state of N spin-1/2 particles it immediately follows that

$$\Delta^2 J_z = \Delta^2 J_y = \frac{N}{4} \quad (3.32)$$

thus the CSS is characterised by having equal variance in the directions orthogonal to the axis of spin orientations (parametrised by θ and ϕ). A general coherent state, i.e in the j th subspace, can be given as a superposition of Dicke states

$$|j, (\theta, \phi)\rangle_N = \sum_{m=-j}^j c_m(\theta) e^{-i(j+m)\phi} |j, m\rangle_N \quad (3.33)$$

where,

$$c_m(\theta) = \binom{2j}{j+m}^{1/2} \cos(\theta/2)^{j-m} \sin(\theta/2)^{j+m} \quad (3.34)$$

hence the PDF of a CSS is a binomial distribution

$$|\langle j, (\theta, \phi) | j, m \rangle|^2 = \binom{2j}{j+m} p^{j+m} (1-p)^{j-m}. \quad (3.35)$$

An alternative representation of the CSS is given by the stereographic projection of the spherical coordinates θ and ϕ to $\varsigma \in \mathbb{C}$ via

$$\varsigma = e^{-i\phi} \tan\left(\frac{\theta}{2}\right) \quad (3.36)$$

thus the CSS becomes

$$|j, \varsigma\rangle_N = \sum_{m=-j}^j \binom{2j}{j+m}^{1/2} \frac{\varsigma^{j+m}}{(1+|\varsigma|^2)^j} |j, m\rangle_N \quad (3.37)$$

A particularly useful definition of the CSS is that it must be simultaneous eigenstate of the \hat{J}^2 and $\vec{r} \cdot \vec{J}$ operators where $\vec{J} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ and the associated eigenvalues are $j(j+1)$

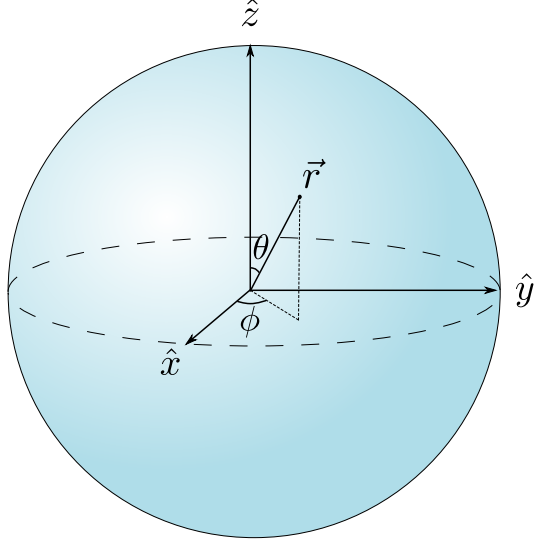


Figure 3.2: Representation of a CSS on the unit sphere.

and j respectively. This allows the visualisation of the CSS as given in Fig.3.2 where it is characterised as a point on the unit sphere specified by θ and ϕ . The CSS can then be regarded as a displacement of a “reference CSS” where the displacement is modelled by a rotation operator

$$R(\theta, \phi) = \exp(-i\theta \vec{J} \cdot \vec{n}) \quad (3.38)$$

which acts to rotate the reference CSS about the unit vector \vec{n} by the angle θ . Indeed, taking the reference CSS to be the Dicke state $|j, -j\rangle_N$ and the unit vector to be $\vec{n} = (\sin \phi, -\cos \phi, 0)$ in the xy -plane, then we find

$$\begin{aligned} |j, (\theta, \phi)\rangle &= R(\theta, \phi) |j, -j\rangle_N \\ &= \exp[-i\theta(J_x \sin \phi - J_y \cos \phi)] |j, -j\rangle_N \\ &= \left[\cos\left(\frac{\theta}{2}\right) + (e^{-i\phi}\sigma_+ - e^{i\phi}\sigma_-) \sin\left(\frac{\theta}{2}\right) \right]^{\otimes N} \end{aligned} \quad (3.39)$$

then considering the symmetric $j = N/2$ subspace, the reference state becomes $|N/2, -N/2\rangle_N = |\downarrow\rangle^{\otimes N}$, and using (3.4) we find

$$|N/2, (\theta, \phi)\rangle_N = \left[\cos\left(\frac{\theta}{2}\right) |\downarrow\rangle + e^{-i\phi} \sin\left(\frac{\theta}{2}\right) |\uparrow\rangle \right]^{\otimes N} \quad (3.40)$$

which is precisely the CSS (3.31) which was introduced as the state of all (independent) N spin-1/2 particles in the same (arbitrary) superposition state. A physical example of this is the application of a magnetic field to the reference state $|\downarrow\rangle^{\otimes N}$. Taking the free Hamiltonian of the N spin-1/2 system to be $H_0 = \omega \hat{J}_z$ and an external classical magnetic

field to be governed by the Hamiltonian $H_B = -\gamma \vec{B} \cdot \vec{J}$ where γ is the gyromagnetic ratio of the spins, the total Hamiltonian is $H = H_0 + H_B = \omega \hat{J}_z - \gamma \vec{B} \cdot \vec{J}$ and from this we can prepare the general CSS given by (3.40) via one of the two following methods: applying the magnetic field

$$\vec{B} = \begin{pmatrix} -B \sin \phi \\ B \cos \phi \\ \omega/\gamma \end{pmatrix} \quad (3.41)$$

yielding the Hamiltonian $H = B \sin \phi \hat{J}_x - B \cos \phi \hat{J}_y$. Evolving our reference state within this classical external magnetic field for time $t = \theta/B$ returns the CSS given by (3.40).

3.4 Spin Squeezing

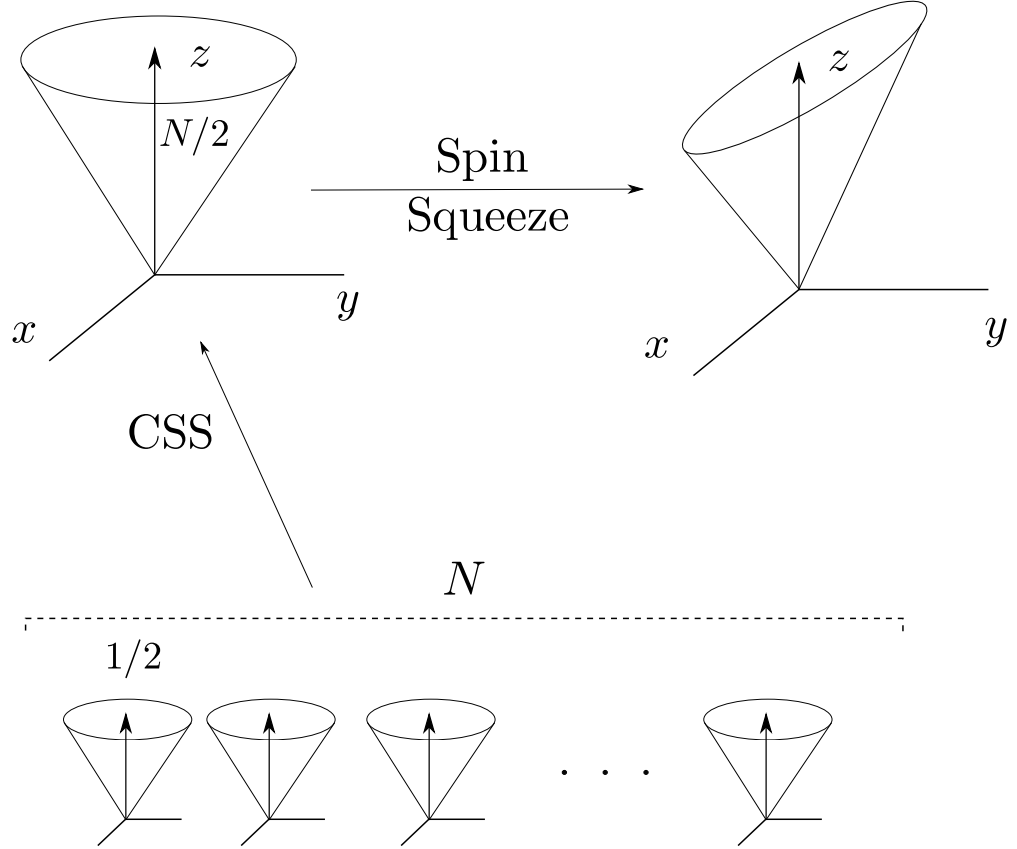


Figure 3.3: Schematic representation of Spin Squeezing of a N spin-1/2 system which constitutes a coherent spin state (CSS). The squeezing process introduces quantum mechanical correlations into the system.

In analogy to the notion of squeezing in the QHO system (section 2.3) we can define spin-squeezing by conditions on the variance for complementary observables i.e since we

have $[J_\alpha, J_\beta] = i\epsilon^{\alpha\beta\gamma} J_\gamma$ and thus

$$\Delta^2 J_\alpha \Delta^2 J_\beta \geq \frac{|\langle J_\gamma \rangle|^2}{4} \quad (3.42)$$

we can quantify spin-squeezing by the condition

$$\Delta^2 J_\alpha \leq \frac{|\langle J_\gamma \rangle|}{2} \quad \text{for } \alpha \neq \gamma \quad (3.43)$$

which would signify spin-squeezing in the α direction. Then for two general orthonormal vectors \vec{n}_1 and \vec{n}_2 , we can define the squeezing parameter

$$\xi_H = \frac{2\Delta^2 J_{\vec{n}_1}}{|\langle J_{\vec{n}_2} \rangle|} \quad (3.44)$$

where the subscript H is used to highlight the underlying connection to the Heisenberg uncertainty relations, if $\xi_H^2 < 1$ then the state is squeezed. Considering the CSS (3.40) with n_1 in the x -direction and introducing n_0 to be in the z -direction it can be shown [34] that

$$\xi_H^2 = \frac{1 - (\vec{n}_0 \cdot \vec{n}_1)^2}{|\vec{n}_0 \cdot \vec{n}_2|} = |\sin(\theta)| \quad (3.45)$$

which implies that by varying θ the CSS can be squeezed but as we have seen, this is clearly not the case and as such we conclude that ξ_H^2 is not a reliable quantifier of spin squeezing.

An alternative spin squeezing operator has been proposed by Kitagawa and Ueda which relies on the idea of “mean spin direction” (MSD). We have previously established that the variance of the coherent state in the bosonic system is equal in all directions whereas the variance of the CSS is dependent on \vec{n} and has a prior direction; the MSD, which is given by

$$\vec{n}_0 = \frac{\langle \vec{J} \rangle}{|\langle \vec{J} \rangle|}. \quad (3.46)$$

with this we define \vec{n}_\perp to be the unit vector in the direction perpendicular to that of the MSD. The spin-squeezing parameter is then given to be

$$\xi_S^2 = \frac{\min(\Delta^2 J_{\vec{n}_\perp})}{j/2} = \frac{4 \min(\Delta^2 J_{\vec{n}_\perp})}{N} \quad (3.47)$$

where $j = N/2$ and the minimisation is over all directions \vec{n}_\perp . It can be shown that for a CSS $\Delta^2 J_{\vec{n}_\perp} = j/2$ and thus we have for a CSS $\Delta^2 J_{\vec{n}_\perp} = N/4$ and as such, $\xi_S = 1$. A spin state is said to be squeezed if $\xi_S^2 < 1$.

Another squeezing parameter of particular significance in quantum metrology is that proposed by Wineland *et. al* [35] in the context of Ramsey spectroscopy. The squeezing parameter ξ_S is the analogue of the bosonic squeezing parameter, the parameter introduced here exhibits a deep connection to sensitivities provided by rotations of angular momentum states. This is a consequence of using the CSS as a noise-reference state. To portray this idea, we consider a spin state $|\psi\rangle$ and take the MSD to be in the z -axis (so $\langle J_x \rangle = 0 = \langle J_y \rangle$). Using the error propagation formula

$$\Delta x = \frac{\Delta f(x)}{\left| \frac{\partial \langle f(x) \rangle}{\partial x} \right|} \quad (3.48)$$

it can be shown that

$$\Delta \phi = \frac{\Delta J_y}{|\langle J_z \rangle|} \quad (3.49)$$

and indeed for general MSD we find

$$\Delta \phi = \frac{\Delta J_{\vec{n}_\perp}}{|\langle \vec{J} \rangle|}. \quad (3.50)$$

For the CSS, the phase sensitivity is found to be

$$(\Delta \phi)_{\text{CSS}} = \frac{1}{\sqrt{N}}. \quad (3.51)$$

The squeezing parameter is then given to be

$$\xi_R^2 = \frac{\Delta^2 \phi}{(\Delta^2 \phi)_{\text{CSS}}} = \frac{N \Delta^2 J_{\vec{n}_\perp}}{|\langle \vec{J} \rangle|^2} \quad (3.52)$$

which is in fact related to parameter ξ_S by

$$\xi_R^2 = \left(\frac{j}{|\langle \vec{J} \rangle|} \right)^2 \xi_S^2 \quad (3.53)$$

and since $j = N/2 \geq |\langle \vec{J} \rangle|$, we have that $\xi_S^2 \leq \xi_R^2$.

3.4.1 One-Axis Twisting

There are multiple ways of producing squeezed spin systems as quantified by the above measures [35, 36, 37], here we investigate a particular method of spin squeezing known as “One-Axis Twisting” (OAT) which uses the following Hamiltonian (in units of \hbar) for a system of N spin-1/2 particles

$$H_{\text{OAT}} = \chi J_x^2 \quad (3.54)$$

so that the unitary time evolution operator that constitutes the OAT operation is given by $U_\chi(t) = \exp[-itH_{OAT}]$. Note that the twisting is often modelled along the z -axis and is applied to a CSS with all spins aligned with the x -axis, here we have reversed the roles of the respective axes. Hence, a spin-squeezed state can be written in the form

$$|\psi_{OAT}\rangle = \exp(-it\chi J_x^2) |j, -j\rangle = \exp(-i\theta J_x^2/2) |j, -j\rangle \quad (3.55)$$

where $\theta = 2\chi t$ is the angle through which the CSS is twisted. It is well known [34, 36] that for sufficiently large N and sufficiently small $|\theta|$ such that $N|\theta|^2 \ll 1$, the optimum value of the spin-squeezing parameter (the value which minimises variance) scales as $\xi_R^2 \sim N^{-2/3}$. Two notable differences between spin squeezing and bosonic squeezing (2.46) are that while (3.54) influences the quantum fluctuations, it also gives rise to a small rotation and for longer interaction times (3.54) yields non-Gaussian states such as oversqueezed states and maximally entangled GHZ states [38]. The resulting spin squeezed state is highly sensitive to rotation operations (3.38) about the y -axis. Practical implementation of spin squeezing via OAT has been demonstrated to be very viable and has been carried out with BECs [39, 40, 41, 42] as well as with atomic ensembles [43, 44, 45, 46].

Visual Representations of Spin Phase Space

Akin to the Wigner function introduced in section 2.2.1 which is implemented as a visual representation of phase space for the QHO, here we introduce phase space plots of finite spin systems. Indeed the so-called spin-Wigner function [47, 48] gives the analogous quasi-probability distribution for the finite spin system. However, we choose to utilize an alternative representation known as the spin Q-function. Before doing so it is useful to note that since the CSSs form an overcomplete basis

$$\int d\Omega |j, (\theta, \phi)\rangle_N \langle j, (\theta, \phi)|_N = \frac{4\pi}{2j+1} \quad (3.56)$$

an arbitrary state of a finite spin system (for fixed j) can be expressed in the CSS basis

$$\rho = \int d\Omega P(\theta, \phi) |j, (\theta, \phi)\rangle_N \langle j, (\theta, \phi)|_N \quad (3.57)$$

where the function $P(\theta, \phi)$ is generally chosen to be a smooth function [32] and is not unique to the particular state ρ . In spherical coordinates, for fixed j , the spin Q-function is defined as

$$Q(\theta, \phi) = \frac{2j+1}{4\pi} \langle j, (\theta, \phi) | \rho | j, (\theta, \phi) \rangle \quad (3.58)$$

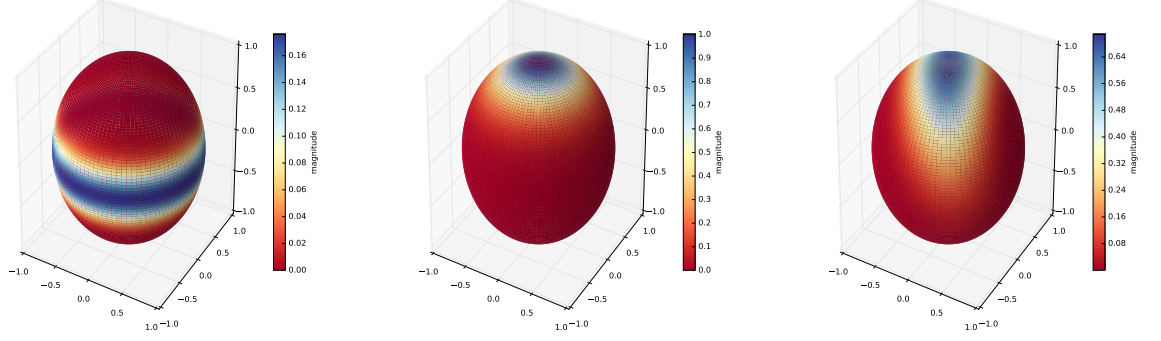


Figure 3.4: Plots of the spin Q-function for the Dicke state $|N/2, m\rangle$ with $N = 20$ and $m = 0$ (*left*), the CSS $|\downarrow\rangle^{\otimes N}$ with $N = 10$ (*centre*) and the squeezed spin state (*right*) where the OAT Hamiltonian (3.54) has been applied to the CSS.

From this it is possible to discern that $Q(\theta, \phi) \geq 0$, $\forall \theta, \phi$ and the spin Q-function is normalised $\int d\Omega Q(\theta, \phi) = 1$. Plots of the Spin Q-function are depicted in Fig. 3.4. Although it may not be immediately clear from Fig. 3.4, the spin squeezed state (*right*) deviates from the geodesic of the sphere; this is a consequence of the aforementioned rotation (the twisting) that the OAT evolution imparts on the CSS during the squeezing of quantum fluctuations.

3.4.2 Two-Axis Twisting

Another method of spin-squeezed state creation is that of “two-axis twisting” (sometimes referred to as “two-axis countertwisting”) [36] which involves the simultaneous twisting, in both the clockwise and anti-clockwise, around two orthogonal axis which lie in the plane normal to the MSD of the initial state $|j, -j\rangle$, this has the effect of cancelling out the rotation (the twisting) that occurs in the OAT case. This is depicted in Fig. 3.5. The associated Hamiltonian is composed of the raising and lowering operators (3.18)

$$H_{TAT} = i\eta(J_-^2 - J_+^2) \quad (3.59)$$

where η dictates the magnitude of squeezing. This method of spin squeezing is in fact analogous to the bosonic squeezing (2.46), indeed we will show in the following section that in the $N \rightarrow \infty$ limit, we recover the bosonic squeezing operator. The disadvantage of TAT in comparison to that of OAT are the complexities involved in the practical implementation, though some viable approaches have been demonstrated [49, 50, 51].

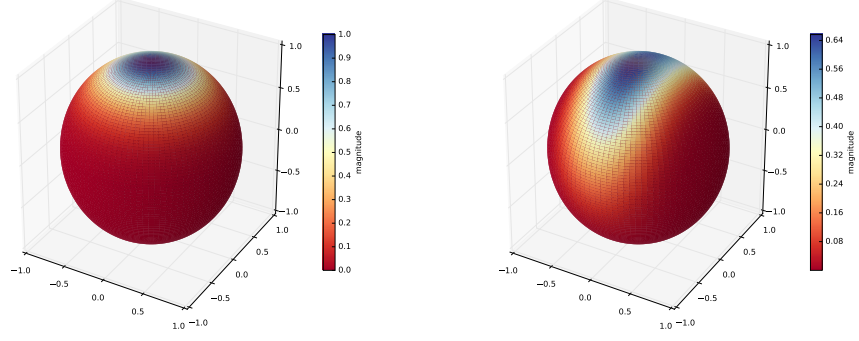


Figure 3.5: Plots of the spin Q-function for the CSS $|\downarrow\rangle^{\otimes N}$ with $N = 10$ (*left*) and the squeezed spin state (*right*) where the TAT Hamiltonian (3.59) has been applied to the CSS.

3.5 Bosonic limit

Here we review how j -subspace spin systems with finite dimensional state spaces correspond to bosonic QHO systems with infinite dimensional state spaces in the so-called “bosonic limit”. Considering the $j = \frac{N}{2}$ symmetric subspace of the N spin-1/2 system, the raising and lowering operators (3.18) take the form

$$\hat{J}_+ = \sum_{n=0}^N \sqrt{(n+1)(N-n)} |n+1\rangle_N \langle n|_N \quad (3.60)$$

$$\hat{J}_- = \sum_{n=0}^N \sqrt{n(N-n+1)} |n-1\rangle_N \langle n|_N \quad (3.61)$$

where the Dicke state $|\frac{N}{2}, m\rangle_N = |\frac{N}{2}, n - \frac{N}{2}\rangle_N \equiv |n\rangle_N$. Additionally, the operators (3.20) can be written in the form $a_{\uparrow} = \sum_{n=0}^N \sqrt{n} |n-1\rangle_N \langle n|_N$ and $a_{\uparrow}^{\dagger} = \sum_{n=0}^{N-1} \sqrt{n+1} |n+1\rangle_N \langle n|_N$ with which, we arrive at the Holstein-Primakoff transformations [52]

$$\frac{\hat{J}_+}{\sqrt{N}} = a_{\uparrow}^{\dagger} \sqrt{1 - \frac{a_{\uparrow}^{\dagger} a_{\uparrow}}{N}} \quad (3.62)$$

$$\frac{\hat{J}_-}{\sqrt{N}} = \sqrt{1 - \frac{a_{\uparrow}^{\dagger} a_{\uparrow}}{N}} a_{\uparrow} \quad (3.63)$$

then taking the limit $N \rightarrow \infty$ reveals

$$\lim_{N \rightarrow \infty} \frac{\hat{J}_+}{\sqrt{N}} = \lim_{N \rightarrow \infty} a_{\uparrow}^{\dagger} = a^{\dagger} \quad (3.64)$$

$$\lim_{N \rightarrow \infty} \frac{\hat{J}_-}{\sqrt{N}} = \lim_{N \rightarrow \infty} a_{\uparrow} = a \quad (3.65)$$

where a, a^\dagger are the QHO creation and annihilation operators which obey the bosonic commutation relation $[a, a^\dagger] = 1$. Furthermore, for finite N , we have

$$\left[\frac{\hat{J}_-}{\sqrt{N}}, \frac{J_+}{\sqrt{N}} \right] = \mathbb{I} - \frac{2a_\uparrow^\dagger a_\uparrow}{N}. \quad (3.66)$$

thus $\left[\frac{\hat{J}_-}{\sqrt{N}}, \frac{J_+}{\sqrt{N}} \right] \approx \mathbb{I} \Rightarrow \frac{a_\uparrow^\dagger a_\uparrow}{N} \approx 0$. The operator $a_\uparrow^\dagger a_\uparrow$ counts the number of spins in the $|\uparrow\rangle$ state, thus if the number of spins in the excited state is small compared to N the spin system can be treated approximately as an optical QHO system. Noting that we are free to scale Hamiltonian (3.59) by $1/N$ (which physically corresponds to scaling of the magnitude of squeezing) it is clear that

$$\begin{aligned} \lim_{N \rightarrow \infty} H_{TAT} &= \lim_{N \rightarrow \infty} i\eta \left(\frac{J_-^2}{N} - \frac{J_+^2}{N} \right) \\ &= i\eta \left(a^2 - (a^\dagger)^2 \right) \end{aligned} \quad (3.67)$$

which is exactly the Hamiltonian that governs squeezing of bosonic systems (noting that (3.59) can be generalized such that $\eta \in \mathbb{C}$). Additionally, it is known [32, 33] that in the limit $N \rightarrow \infty$ the spin coherent state converges to the coherent state of the QHO $|\zeta\rangle_N \rightarrow |\zeta\rangle$ and moreover, in the limit $\zeta \rightarrow 0$, we retrieve the bosonic vacuum state which corresponds to the limit

$$\lim_{N \rightarrow \infty} |\downarrow\rangle^{\otimes N} = |0\rangle. \quad (3.68)$$

When N is finite, the spin coherent $|\zeta/\sqrt{N}\rangle_N$ state is well approximated by the coherent state $|\zeta\rangle$ under the condition $|\zeta| \ll \sqrt{N}$. It is now clear that a natural transition exists between the N spin-1/2 system and the bosonic QHO system.

Chapter 4

Quantum Enhanced Metrology and Parameter Estimation

4.1 Measurements

4.1.1 Classical and Quantum Measurements

Consider a system of N particles, classically this can be described the \mathbb{R}^{6N} phase space and the state of the system is given by a probability distribution on this phase space. A subtle underlying question of this description arises about what the information given by the probability distribution represents. This is an interpretational question, the debate of which is beyond the scope of the present work, however it is important to clarify which interpretation we use here; we directly utilize the interpretation that probability distributions are subjective by nature [53] so the state of a system represents an observer's knowledge of system variables. Classically, a perfect measurement would yield the value of the variable of interest and the system enters a state of complete knowledge. An imperfect measurement would leave the system in a state of incomplete knowledge thus introducing *uncertainty*. Such classical measurements are performed through system-apparatus measurements in which a variable of the apparatus is coupled to the system variable of interest (e.g electrical voltage coupled to the momentum of a voltmeter needle which ultimately varies the position on a dial). Inference of the system variable comes through the probability distribution of the apparatus variable. Imperfections may exist between the coupling of these variables or may be introduced independent of the system in the form of noise affecting the apparatus.

Measurement of a quantum mechanical system is fundamentally different to a classical

system in that an observer may have the maximal amount of knowledge about the system yet the measurement outcome can still not be determined. In other words, maximal knowledge of a quantum mechanical system does not imply complete knowledge and thus, in general, uncertainty remains. As such, measurement outcomes themselves are described by probability distributions. This naturally opens up an interpretational debate of knowledge e.g. attributing the incomplete knowledge to hidden variables [54], again we circumvent such a discussion and take an operational approach by using quantum mechanics as a tool to calculate predictions of experiment. From the postulates of quantum mechanics (see section 1.2.3) it is evident that the notion of knowledge of a quantum state can only refer to predictions of future measurement outcomes since measurement fundamentally changes the state. This implicit uncertainty associated with measurements of quantum mechanical systems is the main focus of the following work.

4.1.2 Projective Measurements

From the 2nd postulate of quantum mechanics (see section 1.2.3) we have that any quantum mechanical observable is given by

$$\hat{A} = \sum_{i=1} \lambda_i |\lambda_i\rangle \langle \lambda_i| \quad (4.1)$$

where $|\lambda_i\rangle$ are the eigenstates satisfying $\hat{A} |\lambda_i\rangle = \lambda_i |\lambda_i\rangle$ with the eigenvalues λ_i . The “butterfly operator” $|\lambda_i\rangle \langle \lambda_i| = \hat{\pi}_\lambda$ is a projector operator which projects onto the eigenstates of \hat{A} with eigenvalue λ_i . This can be generalised to include degenerate eigenvalues [53]; if the eigenvalues of \hat{A} are N_λ -fold degenerate, the butterfly operator is replaced by the projector $\hat{\Pi}_\lambda = \sum_{j=1}^{N_\lambda} |\lambda_i, j\rangle \langle \lambda_i, j|$ and the observable is given by

$$\hat{A} = \sum_{i=1} \lambda_i \hat{\Pi}_\lambda. \quad (4.2)$$

The projectors are orthonormal, that is $\hat{\Pi}_\lambda \hat{\Pi}_{\lambda'} = \delta_{\lambda, \lambda'} \hat{\Pi}_\lambda$ which is equivalent to the statement that the subspaces onto which they project are orthonormal i.e. if the set $\{|\lambda_1\rangle, |\lambda_2\rangle, \dots\}$ constitutes an orthonormal basis for the Hilbert space of the system, the projection operators $|\lambda_i\rangle \langle \lambda_i|$ and $|\lambda_j\rangle \langle \lambda_j|$ are orthonormal. Note that in general, observables that commute represent observables that can be measured simultaneously. The algebraic definition of projection operators is given by [14]

$$\hat{\Pi}_\lambda^2 = \hat{\Pi}_\lambda, \quad \hat{\Pi}_\lambda^\dagger = \hat{\Pi}_\lambda \quad (4.3)$$

that is to say, any operator that satisfies conditions (4.3) is a projection operator. From this we find the sum of projectors is indeed a projector

$$(\hat{\Pi}_\lambda + \hat{\Pi}_{\lambda'})^2 = \hat{\Pi}_\lambda^2 + \hat{\Pi}_{\lambda'}^2 + \hat{\Pi}_\lambda \hat{\Pi}_{\lambda'} + \hat{\Pi}_{\lambda'} \hat{\Pi}_\lambda = \hat{\Pi}_\lambda + \hat{\Pi}_{\lambda'} \quad (4.4)$$

for $\lambda \neq \lambda'$. Performing a measurement of a system variable at time t that then takes a duration of time T , results in the following probability of measurement outcome λ $\Pr[A(t) = \lambda] = \langle \psi(t) | \hat{\Pi}_\lambda | \psi(t) \rangle$, and as a consequence of wave-function collapse (see section 1.2.3) the (conditional) state vector is given by

$$|\psi(t+T)\rangle = \frac{\hat{\Pi}_\lambda |\psi(t)\rangle}{\sqrt{\langle \psi(t) | \hat{\Pi}_\lambda | \psi(t) \rangle}}. \quad (4.5)$$

It is important to note that the projection operator is in general non-unitary, that is unless one keeps track of all measurement results the projection operation is an entropy increasing process and does not preserve normalisation. These results are easily generalised from the discrete case to the continuous case [53] by taking the projection operator to a projection density operator that satisfies the orthonormality condition $\hat{\Pi}(x)\hat{\Pi}(x') = \delta(x-x')\hat{\Pi}(x)$ where $\langle x|x' \rangle = \delta(x-x')$ is the delta function. For a measurement of observable \hat{X} we have

$$\hat{X} = \int_{-\infty}^{\infty} x \hat{\Pi}(x) dx \quad (4.6)$$

and the associated conditional probability of obtaining outcome in the infinitesimal interval x and $x+dx$ is $\Pr[\hat{X} \in (x, x+dx)] = |\langle x | \psi(t) \rangle|^2 dx$.

Modelling quantum mechanical measurements via projective observables is inherently limited, indeed upon considering a photodetector which must absorb quanta of light for detection to be made possible, it is clear that the state of the photons after detection is not an eigenstate of the number operator since the photons have been destroyed. More fundamentally, a system is never measured directly in experiment; a system is coupled to an apparatus which is directly coupled to the variable of interest and an observer subsequently measures the changes in the apparatus much like in the classical case described in the preceding subsection. This of course can involve many steps and couplings; as an example again considering a photodetector, if we wish to measure an atom this can be coupled to a mode of the electromagnetic field which can generate a current which can modify a display which emits more photons etc. this is known as a von Neumann chain. This motivates the following review of system-apparatus measurements which first require a formal introduction of quantum entanglement.

4.1.3 Entanglement

The states we have looked at so far have all been single mode states, we now turn our attention to the effects of including multiple modes. We begin with a two-mode system consisting of modes a and b , the composite state of the two-mode system can be encapsulated by the tensor product of the two modes

$$|\Psi\rangle = |\psi\rangle_a \otimes |\varphi\rangle_b \quad (4.7)$$

and when such a representation is possible the state is said to be separable, however forms of a two-mode system exist where such a state cannot be factorised for example

$$|\Psi\rangle = \mathcal{N} (|\psi\rangle_a |\varphi\rangle_b + |\varphi\rangle_b |\psi\rangle_a) \quad (4.8)$$

where, \mathcal{N} is a normalising factor. Note that the notation of the tensor product varies in the literature with equivalence between $|\psi\rangle_a \otimes |\varphi\rangle_b \equiv |\psi\rangle_a |\varphi\rangle_b \equiv |\psi, \varphi\rangle_{a,b}$ and subscripts are often omitted. When a composite state cannot be written as a tensor product i.e when it is not separable it is said to be entangled, indeed this is the definition of entanglement. Entanglement is a pivotal resource in quantum enhanced metrology, it is important to note that mode entanglement is just a single form the general phenomenon which is dependent on the division of the Hilbert space, which describes the system of interest, into subsystems [55]. As a concrete example, we shall inspect the two-mode Fock state $|1\rangle_a |1\rangle_b$. It is instructive to recall the “first quantisation” of the electromagnetic field (which offers an equivalent yet alternative description of photons) where, although photons are bosonic, the notation used is that of distinguishable particles; for N photons we have

$$|\mathbf{n}\rangle = |n_1\rangle_1 \otimes |n_2\rangle_2 \otimes \cdots \otimes |n_N\rangle_N \quad (4.9)$$

where mode m_i contains the i th photon and the total number of photons is given by $N = n_1 + n_2 + \dots + n_m$. The subscripts on the kets have been included to discern the representation of these distinct particles from the representation of various modes which the kets usually denote. If we are then to rewrite the two-mode Fock state in this way, we must take into account the indistinguishability of the two photons which demands that all possible permutations of particle arrangements must be included in order to maintain symmetry, hence

$$|1\rangle_a |1\rangle_b = \frac{1}{\sqrt{2}} (|a\rangle_1 |b\rangle_2 + |b\rangle_1 |a\rangle_2) \quad (4.10)$$

so although the two-mode Fock state is clearly mode separable, the division of the Hilbert space into “particle” subsystems reveals entanglement. This subtle point is of particular importance when considering squeezed states as we shall see in the following chapter.

4.1.4 von Neumann Measurement Schemes

A system-apparatus measurement scheme is known as a von Neumann measurement scheme, this can be broken down into two steps: i) coupling of the system variable of interest with an apparatus variable ii) a projective measurement of the apparatus variable. If we are to consider the apparatus device to be a needle on a dial, the Hamiltonian governing the measurement interaction is given by

$$\hat{H} = g(t)\hat{A} \otimes \hat{P}_d \quad (4.11)$$

where, \hat{A} is the observable of interest and \hat{P}_d is the conjugate momentum of the canonical position variable of the apparatus device \hat{q}_d such that $[\hat{q}_d, \hat{P}_d] = i\hbar$ and $g(t)$ as a coupling impulse function (corresponding to a non-zero value for only a very short time) and is normalised so that

$$\int g(t)dt = 1. \quad (4.12)$$

Following the calculation presented in Ref.[56], a brief analysis of the coupling process is given here. The position of the needle is determined by the position operator \hat{q}_d such that $\hat{q}_d |x\rangle = x |x\rangle$ where x and $|x\rangle$ are the respective position eigenvalues and eigenvectors of the measurement device's needle. There is a continuum of possible eigenstates of the position operator on our device's Hilbert space so a general state vector of the device can be decompose to

$$|\phi(x)\rangle = \int_x \phi(x) |x\rangle dx \quad (4.13)$$

where, $\phi(x)$ is a functional probability amplitude which is reasonably assumed to be given by

$$\phi(x) = (2\pi\Delta)^{-\frac{1}{4}} e^{-x^2/4\Delta}, \quad (4.14)$$

a normalised Gaussian distribution, of zero mean and variance $\Delta (= \Delta\hat{q}_d)$. Turning our attention to the system that we wish to measure, a general state vector can be decomposed as

$$|\psi\rangle = \sum_i^N \alpha_i |a_i\rangle \quad (4.15)$$

where $\alpha_i \in \mathbb{C}$ is the probability amplitude of obtaining the i^{th} eigenvalue of \hat{A} upon measurement. Since we are interested in the interaction between the system and the apparatus device, we must consider the tensor product of the measurement device's Hilbert

space and the quantum system's (that we wish to measure) Hilbert space $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_d$, thus we will be considering the direct product of the vector states

$$|\psi\rangle \otimes |\phi(x)\rangle = \sum_i^N \alpha_i \int_x \phi(x) |a_i\rangle \otimes |x\rangle dx. \quad (4.16)$$

To inspect how this full system state vector evolves in time, we apply a time evolution operator (a solution of the Shrödinger equation) which is given by

$$\hat{U}(t) = \exp(-i\hat{H}t/\hbar) \quad (4.17)$$

Applying this to the system-device product space state vector (4.16) in conjunction with the (time-dependent) Hamiltonian (4.11), the evolution of the system over time is thus given by

$$\begin{aligned} U(t) |\psi\rangle \otimes |\phi(x)\rangle &= \exp\left(-\frac{i}{\hbar} \int \hat{H} dt\right) |\psi\rangle \otimes |\phi(x)\rangle \\ &= \exp(-i\hat{A} \otimes \hat{P}_d/\hbar) |\psi\rangle \otimes |\phi(x)\rangle \end{aligned} \quad (4.18)$$

from which we gain meaningful insight by inspecting the operator \hat{q}_d , over some interaction time interval T . Using the fundamental theorem of calculus in combination with the Heisenberg equation of motion, we find

$$\begin{aligned} \hat{q}_d(T) - \hat{q}_d(0) &= \int_0^T dt \frac{\partial \hat{q}_d}{\partial t} \\ &= \int_0^T \frac{i}{\hbar} [\hat{H}, \hat{q}_d] dt \\ &= \int_0^T \frac{i}{\hbar} \hat{A} [\hat{P}_d, \hat{q}_d] dt \\ &= \int_0^T \frac{i}{\hbar} (-i\hbar) \hat{A} dt \\ &= a_i \end{aligned} \quad (4.19)$$

where the final equality is true for eigenstates of \hat{A} only. It is then evident that $\hat{q}_d(T) = \hat{q}_d(0) + a_i$ so we infer that the system evolves each state vector by taking \hat{q}_d to $\hat{q}_d + a_i$. Decomposing the general state vectors in terms of the eigenbasis allows us to write

$$\begin{aligned} U(t) |\psi\rangle \otimes |\phi(x)\rangle &= \sum_i^N \alpha_i |a_i\rangle \otimes |\phi(x - a_i)\rangle \\ &= \sum_i^N \alpha_i \int_x (2\pi\Delta)^{-\frac{1}{4}} e^{-(x-a_i)^2/4\Delta} |a_i\rangle \otimes |x\rangle dx. \end{aligned} \quad (4.20)$$

thus the system and apparatus device become entangled. The probability density of measuring the device's pointer in state x' , given by a projective measurement on the

apparatus, is found to be

$$\begin{aligned}
p(x') &= \langle \psi, \phi(x) | U^\dagger(t) | x' \rangle \langle x' | U(t) | \psi, \phi(x) \rangle = | \langle x' | U(t) | \psi, \phi(x) \rangle |^2 \\
&= \left| \sum_j \alpha_j^* \langle a_j | \langle x' | \int_x \sum_i \alpha_i | a_i \rangle (2\pi\Delta)^{-\frac{1}{4}} e^{-(x-a_i)^2/4\Delta} | x \rangle dx \right|^2 \\
&= (2\pi\Delta)^{-\frac{1}{2}} \left| \int_x \sum_{i,j} \alpha_j^* \alpha_i \langle a_j | a_i \rangle e^{-(x-a_i)^2/4\Delta} \langle x' | x \rangle dx \right|^2 \\
&= (2\pi\Delta)^{-\frac{1}{2}} \left| \int_x \sum_{i,j} \alpha_j^* \alpha_i \delta_{ij} e^{-(x-a_i)^2/4\Delta} \delta(x-x') dx \right|^2 \\
&= (2\pi\Delta)^{-\frac{1}{2}} \sum_i |\alpha_i|^2 e^{-(x'-a_i)^2/2\Delta}
\end{aligned}$$

a multinormal distribution with many modes. It is then apparent that the process has shifted the initial Gaussian distribution to another multivariate Gaussian distribution dependent on the eigenvalues of the system observable. Note that this projective measurement disentangles the system and apparatus; this can be seen through inspection of the (normalised) final state

$$|\Psi\rangle_f = \frac{|x'\rangle \langle x' | U(t) | \psi, \phi(x) \rangle}{\sqrt{p(x')}} = \frac{|x'\rangle \hat{M}_{x'} |\psi\rangle}{\sqrt{p(x')}} \quad (4.21)$$

where $\hat{M}_{x'} = \langle x' | \hat{U}(t) | \phi(x) \rangle$ is an operator that acts only on the system Hilbert space \mathcal{H}_s .

4.2 Uncertainty Measures

Quantum parameter estimation concerns itself with the ability to encode and decode information onto, and from, quantum mechanical states of a system of interest. Since, in general, only partial knowledge is attainable even with total knowledge of the preparation of a system state, an ensemble of identically prepared states is necessary. Given this statistical nature, a natural question arises concerning how to quantify the fundamental uncertainties involved in these processes. These uncertainties establish the fundamental limits of encoding and decoding information, after doing so we can then ask how to *optimally* encode and decode information on quantum mechanical systems.

4.2.1 Helstrom-Holevo Lower bound

The first ultimate quantum limit for precision in quantum parameter estimation was independently established by both Holevo [57] and Helstrom [58] in the context of quantum communication. Following the work of Milburn and Wiseman [53] we review the so called Helstrom-Holevo lower bound and further the notion of uncertainty measure to that of the

Fisher information. The underlying method to address optimality is to minimise a cost function associated with the error of the estimated parameter of interest. Firstly, to get a feel for precision bounds we review the derivation of the Helstrom-Holevo lower bound by considering the probe state (given as a density matrix) ρ_0 , the encoding of information onto the probe is given by

$$\rho_0 \rightarrow \rho_X = e^{-iX\hat{G}}\rho_0 e^{iX\hat{G}} \quad (4.22)$$

where \hat{G} is a Hermitian “generator” operator and X is the parameter of interest to be estimated via the measured parameter X_e (the decoders best estimate of X) provided by observable \hat{X}_e . The disparity between X and X_e comes about from the possibility of a systematic bias $b(X) = \langle X_e \rangle_X - X$ and the score function is given by the mean-square error

$$\langle (X_e - X)^2 \rangle_X = \Delta^2 X_e + [b(X)]^2 \quad (4.23)$$

where the variance of the estimator (in ρ_X) is given by

$$\Delta^2 X_e = \text{Tr}[(\hat{X}_e - \langle X_e \rangle_X)^2 \rho_X]. \quad (4.24)$$

We note

$$\frac{d \langle X_e \rangle_X}{dX} = -i \text{Tr} [\hat{X}_e, \hat{G}] \rho_X \quad (4.25)$$

and from the general Heisenberg uncertainty relation

$$\Delta^2 X_e \Delta^2 \hat{G} \geq \frac{1}{4} \left| \text{Tr} [\hat{X}_e, \hat{G}] \rho_X \right|^2 \quad (4.26)$$

which reveals the lower bound

$$\langle (X_e - X)^2 \rangle_X \geq \frac{[1 + b'(X)]^2}{4 \langle (\Delta \hat{G})^2 \rangle_X} + b^2(X). \quad (4.27)$$

Considering the case of no systematic bias $b(X) = 0$ which implies $\langle X_e \rangle_X = X$, thus

$$\Delta^2 X_e \geq \frac{1}{4 \langle \Delta^2 \hat{G} \rangle_0} \quad (4.28)$$

where we have set $X = 0$ by acknowledging the commutativity of \hat{G} with the unitary parameter transformation. Note that for canonically conjugate observables $[\hat{X}, \hat{G}] = i$, it is clear that this result follows immediately from the Heisenberg uncertainty principle.

4.2.2 The Fisher Information

To address the notion of optimality and to introduce an uncertainty measure that will be used extensively in the following work, we present a derivation of the so-called ‘‘Fisher Information’’ based on that presented in Ref.[59]. To begin with, we explicitly define the likelihood function which is the probability of obtaining data s when the true value of the parameter we wish to estimate is θ , thus for a given probability function f_θ , the likelihood is given by $L(\theta|s) = f_\theta(s)$ and the *likelihood function* is given by $L(\cdot|s)$. So if we have $f_{\theta_1}(s) > f_{\theta_2}(s)$, we infer that θ_1 is a more accurate value (closer to the true value θ) than θ_2 . Note that $L(\theta|s)$ is not the probability of θ given that we have observed s . Generally, we are interested in a point estimate of θ so a value that maximises $L(\theta|s)$ is desirable hence we define $\hat{\theta}(s)$ such that $L(\hat{\theta}(s)|s) \geq L(\theta|s)$, $\forall \theta$ and refer $\hat{\theta}(s)$ as the maximum likelihood estimate (MLE). Calculating the MLE often requires using optimisation methods of calculus and as such we require $f_\theta(s)$ to be a continuously differentiable function of θ . To this end, we define the log-likelihood function

$$l(\cdot|s) = \ln [L(\cdot|s)] \quad (4.29)$$

noting that $L(\hat{\theta}(s)|s) \geq L(\theta|s) \Rightarrow l(\hat{\theta}(s)|s) \geq l(\theta|s)$. Furthermore for a sample (s_1, \dots, s_n) , the likelihood function is given by

$$L(\theta|s_1, \dots, s_n) = \prod_{i=1}^n f_\theta(s_i) \quad (4.30)$$

which gives the log-likelihood as

$$l(\theta|s_1, \dots, s_n) = \sum_{i=1}^n \ln[f_\theta(s_i)] \quad (4.31)$$

and since it is generally a simpler task to differentiate a sum than it is a product, the advantage of the log-likelihood is revealed. In fact, this property is so useful that it’s derivative is defined as the score function

$$S(\theta|s) = \frac{\partial l(\theta|s)}{\partial \theta} \quad (4.32)$$

and then it is clear that the MLE is found by solving the equation

$$S(\theta|s) = 0 \quad (4.33)$$

and ensuring that this solution is a local maximum at $\hat{\theta}(s)$ by checking the following condition is satisfied

$$\left. \frac{\partial S(\theta|s)}{\partial \theta} \right|_{\theta=\hat{\theta}(s)} = \left. \frac{\partial^2 l(\theta|s)}{\partial \theta^2} \right|_{\theta=\hat{\theta}(s)} < 0. \quad (4.34)$$

Now, suppose we have $X, \theta \in \mathbb{R}$ and the following conditions:

$$\frac{\partial^2 \ln[f_\theta(x)]}{\partial \theta^2} \text{ exists for each } x, \quad (4.35)$$

and

$$E_\theta[S(\theta|X)] = \int_{-\infty}^{\infty} \frac{\partial \ln[f_\theta(x)]}{\partial \theta} f_\theta(x) dx = 0 \quad (4.36)$$

along with

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} \left(\frac{\partial \ln[f_\theta(x)]}{\partial \theta} f_\theta(x) \right) dx = 0 \quad (4.37)$$

and finally

$$\int_{-\infty}^{\infty} \left| \frac{\partial^2 \ln[f_\theta(x)]}{\partial \theta^2} \right| f_\theta(x) dx < \infty. \quad (4.38)$$

Noting that in general

$$\frac{\partial f_\theta(x)}{\partial \theta} = \frac{\partial \ln[f_\theta(x)]}{\partial \theta} f_\theta(x) \quad (4.39)$$

we can rewrite (4.36) as

$$\int_{-\infty}^{\infty} \frac{\partial f_\theta(x)}{\partial \theta} dx = 0. \quad (4.40)$$

Furthermore, condition (4.37) can be rewritten

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} \left(\frac{\partial l(\theta|x)}{\partial \theta} f_\theta(x) \right) dx \\ &= \int_{-\infty}^{\infty} \left[\frac{\partial^2 l(\theta|x)}{\partial \theta^2} + \left(\frac{\partial l(\theta|x)}{\partial \theta} \right)^2 \right] f_\theta(x) dx \\ &= \int_{-\infty}^{\infty} \left[\frac{\partial^2 l(\theta|x)}{\partial \theta^2} + S^2(\theta|x) \right] f_\theta(x) dx \\ &= E_\theta \left[\frac{\partial^2 l(\theta|x)}{\partial \theta^2} + S^2(\theta|x) \right]. \end{aligned} \quad (4.41)$$

So with this and conditions (4.36) and (4.38) we can finally rewrite (4.37) as

$$\text{Var}_\theta(S(\theta|X)) = E_\theta[S^2(\theta|X)] = E_\theta \left(-\frac{\partial^2}{\partial \theta^2} l(\theta|x) \right). \quad (4.42)$$

The quantity on the left of (4.42) is defined as the Fisher information $F(\theta) = \text{Var}_\theta(S(\theta|X))$.

Moreover, if $E_\theta[S(\theta|X)] = 0$ the Fisher information is given by

$$\begin{aligned} F(\theta) &= \int_{-\infty}^{\infty} f_\theta(x) \left(\frac{\partial \ln[f_\theta(x)]}{\partial \theta} \right)^2 dx \\ &= \int_{-\infty}^{\infty} \frac{1}{f_\theta(x)} \left(\frac{\partial f_\theta(x)}{\partial \theta} \right)^2 dx. \end{aligned} \quad (4.43)$$

We also highlight the quantity

$$\hat{F}(s) = -\frac{\partial^2 l(\theta|s)}{\partial \theta^2} \Big|_{\theta=\hat{\theta}(s)} \quad (4.44)$$

which is referred to as the *observed* Fisher information, this serves as a measure of how concentrated the log-likelihood function is at its peak. An important result that utilizes the Fisher information is a bound known as the Cramér-Rao bound [60, 61] which states that the variance of an unbiased estimator is at least equal to that of the inverse Fisher information, that is

$$\Delta^2 \hat{\theta} \geq \frac{1}{\nu F(\theta)} \quad (4.45)$$

where ν is the number of repeats of the estimation procedure.

4.3 Quantum Fisher Information

Since the systems we wish to investigate are quantum mechanical, if we wish to use the notion of Fisher information as a quantifier of uncertainty it is necessary to extend its definition to encompass the fundamental aspects of quantum information. Some derivations of the so-called “quantum Fisher information” adopt a geometric interpretation of the quantum measurement process [53, 62] in which the signal information parametrizes a path through the space of quantum states and as such, signal detection becomes a matter of abstract spatial distinguishability. Here we take the approach of Ref. [63] and begin by introducing the symmetric logarithm derivative (SLD) as the self-adjoint operator L_θ that satisfies

$$\frac{\partial \rho_\theta}{\partial \theta} = \frac{1}{2} [L_\theta \rho_\theta + \rho_\theta L_\theta] \quad (4.46)$$

and recall the Born rule $f_\theta(x) = \text{Tr}[\Pi_x \rho_\theta]$ where $\{\Pi_x\}$, $\int dx \Pi_x = \mathbb{I}$ are the elements of a positive operator-valued measure (POVM) and ρ_θ is the density matrix representation of the state of interest parametrised by θ . From this we find $\partial_\theta f_\theta(x) = \text{Tr}[\partial_\theta \rho_\theta \Pi_x] = \text{Re}\{\text{Tr}[\rho_\theta \Pi_x L_\theta]\}$. The Fisher information (4.42) then becomes

$$F(\theta) = \int_{-\infty}^{\infty} dx \frac{\text{Re}\{\text{Tr}[\rho_\theta \Pi_x L_\theta]\}^2}{\text{Tr}[\rho_\theta \Pi_x]} \quad (4.47)$$

then maximising the Fisher information by optimising over the quantum measurements we find

$$\begin{aligned}
F(\theta) &\leq \int_{-\infty}^{\infty} dx \left| \frac{\text{Tr}[\rho_{\theta} \Pi_x L_{\theta}]}{\sqrt{\text{Tr}[\rho_{\theta} \Pi_x]}} \right|^2 \\
&= \int_{-\infty}^{\infty} dx \left| \text{Tr} \left[\frac{\sqrt{\rho_{\theta}} \sqrt{\Pi_x}}{\sqrt{\text{Tr}[\rho_{\theta} \Pi_x]}} \sqrt{\Pi_x} L_{\theta} \sqrt{\rho_{\theta}} \right] \right|^2 \\
&\leq \int_{-\infty}^{\infty} dx \text{Tr}[\Pi_x L_{\theta} \rho_{\theta} L_{\theta}] \\
&= \text{Tr}[L_{\theta} \rho_{\theta} L_{\theta}] \\
&= \text{Tr}[\rho_{\theta} L_{\theta}^2]
\end{aligned} \tag{4.48}$$

where now define $F_Q(\theta) \equiv \text{Tr}[\rho_{\theta} L_{\theta}^2]$ to be the quantum Fisher information. We have thus shown that the Fisher information of any quantum measurement is bounded below by the quantum Fisher information (QFI). Clearly the SLD is a comparatively cumbersome mathematical object to work with hence we aim to derive a more accessible form of the QFI in terms of the eigenvalues and eigenvectors of ρ_{θ} where $\rho_{\theta} |\lambda_i\rangle = \lambda_i |\lambda_i\rangle$ so that

$$\begin{aligned}
\left(\frac{\partial \rho_{\theta}}{\partial \theta} \right)_{ij} &= \langle \lambda_i | \left(\frac{\partial \rho_{\theta}}{\partial \theta} \right) | \lambda_j \rangle \\
&= \frac{1}{2} [\langle \lambda_i | L_{\theta} \rho_{\theta} | \lambda_j \rangle + \langle \lambda_i | \rho_{\theta} L_{\theta} | \lambda_j \rangle] \\
&= \frac{1}{2} [\lambda_j (L_{\theta})_{ij} + \lambda_i (L_{\theta})_{ij}]
\end{aligned} \tag{4.49}$$

which can be solved to reveal the SLD as

$$(L_{\theta})_{ij} = 2 \frac{\langle \lambda_i | \frac{\partial \rho_{\theta}}{\partial \theta} | \lambda_j \rangle}{\lambda_i + \lambda_j} \tag{4.50}$$

then using the fact that

$$F_Q(\theta) = \text{Tr}[\rho_{\theta} L_{\theta}^2] = \frac{1}{2} (\text{Tr}[L_{\theta}^2 \rho_{\theta}] + \text{Tr}[L_{\theta} \rho_{\theta} L_{\theta}]) \tag{4.51}$$

we find the QFI to be

$$F_Q(\theta) = \sum_{i,j} \frac{2}{\lambda_i + \lambda_j} \left| \langle \lambda_i | \frac{\partial \rho_{\theta}}{\partial \theta} | \lambda_j \rangle \right|^2. \tag{4.52}$$

If we are to consider the pure state $\rho_{\theta} = |\psi\rangle \langle \psi| = \rho_{\theta}^2$, then we have

$$\frac{\partial \rho_{\theta}}{\partial \theta} = \frac{\partial \rho_{\theta}^2}{\partial \theta} = \rho_{\theta} \frac{\partial \rho_{\theta}}{\partial \theta} + \frac{\partial \rho_{\theta}}{\partial \theta} \rho_{\theta} \tag{4.53}$$

so from (4.46) we have $L_{\theta} = 2 \frac{\partial \rho_{\theta}}{\partial \theta}$ and

$$\begin{aligned}
F_Q(\theta) &= 4 \text{Tr} \left[\rho \left(\frac{\partial \rho_{\theta}}{\partial \theta} \right)^2 \right] \\
&= 4 [\langle \psi'(\theta) | \psi'(\theta) \rangle - |\langle \psi'(\theta) | \psi(\theta) \rangle|^2]
\end{aligned} \tag{4.54}$$

where $|\psi'(\theta)\rangle = \frac{\partial}{\partial\theta} |\psi(\theta)\rangle$. Note that from (4.48), it immediately follows that

$$\Delta^2\hat{\theta} \geq \frac{1}{\nu F(\theta)} \geq \frac{1}{\nu F_Q(\theta)} \quad (4.55)$$

where the latter most inequality is referred to as the quantum Cramér-Rao bound.

4.4 Mach-Zehnder Interferometer

Interferometry exploits the principle of superposition in order to impart and subsequently infer information of an external physical field of interest. A renowned and early use of this experimental technique is the 1887 Michelson and Morley experiment [64] in which the speed of light was shown to be constant, thus paving the way for special relativity. Our interests lie in a specific kind of interferometer known as the “Mach-Zehnder” interferometer (MZI) [65, 66] as depicted in Fig.4.1. The aim is to measure the phase shift ϕ as precisely as possible which results in the question of what to input into the MZI in order to do so? This is answered in the following.

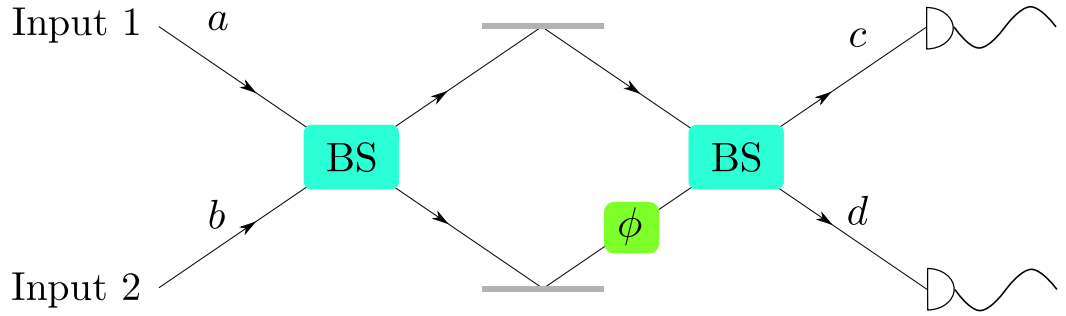


Figure 4.1: A Mach-Zehnder interferometer (MZI) consisting of two inputs states of light, a linear crystal referred to as a beam splitter (BS) which allows the states of light to interact, a phase shift ϕ on one arm (relative to the other arm), a second beam splitter which allows for an interference interaction between the now phase-shifted light and finally two photodetectors output of the MZI.

4.4.1 Independent Photons

To begin answering this question, it is useful to first note that the effect of each beam splitter is to transform the input creation operators according to the following

$$\begin{pmatrix} a_a^\dagger \\ a_b^\dagger \end{pmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} a_a^\dagger \\ a_b^\dagger \end{pmatrix} \quad (\text{up to an arbitrary phase}) \quad (4.56)$$

and the overall effect of the MZI, with a phase shift on path b is thus given by

$$\begin{aligned} \begin{pmatrix} a_a^\dagger \\ a_b^\dagger \end{pmatrix} &\longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} a_c^\dagger \\ a_d^\dagger \end{pmatrix} \\ &= \begin{pmatrix} \sin(\phi/2) & \cos(\phi/2) \\ \cos(\phi/2) & -\sin(\phi/2) \end{pmatrix} \begin{pmatrix} a_c^\dagger \\ a_d^\dagger \end{pmatrix} \end{aligned} \quad (4.57)$$

where, a_c^\dagger, a_d^\dagger are the creation operators at the output. To gain some physical insight into the MZI let us consider sending a single photon into input 1 so that we initially have

$$\begin{aligned} |1, 0\rangle_{a,b} &= a_1^\dagger |0, 0\rangle_{a,b} \longrightarrow \sin(\phi/2) a_c^\dagger |0, 0\rangle_{c,d} + \cos(\phi/2) a_d^\dagger |0, 0\rangle_{c,d} \\ &= \sin(\phi/2) |1, 0\rangle_{c,d} + \cos(\phi/2) |0, 1\rangle_{c,d} \end{aligned} \quad (4.58)$$

thus the probabilities of detecting a photon at outputs c and d are respectively given by

$$P_c = \sin^2(\phi/2), \quad P_d = \cos^2(\phi/2). \quad (4.59)$$

Generalising this to a beam of N independent photons being sent into input 1 is straightforward since the number distribution of the photons at the output port is binomial. The probability mass function thus gives the joint probability that m and $N - m$ particles are detected at ports c and d respectively as

$$P(m, N - m) = \binom{N}{m} P_c^m P_d^{N-m} = \frac{N!}{m!(N-m)!} \sin^{2m}(\phi/2) \cos^{2(N-m)}(\phi/2). \quad (4.60)$$

We also have for a binomially distributed variable x over r experimental repeats with probability of success p , the mean (expected) value and the variance of x are respectively given by

$$\langle x \rangle = r \cdot p, \quad (\Delta x)^2 = r \cdot p(1 - p) \quad (4.61)$$

thus the mean number of photons and respective variance at output c are

$$\langle n_c \rangle = N \sin^2(\phi/2), \quad (\Delta n_c)^2 = N \cos^2(\phi/2) \sin^2(\phi/2) = \frac{\sin^2(\phi)}{4} \quad (4.62)$$

with similar analysis and results for output d . So from (4.62) we are able to infer the phase shift by simply counting the number of photons at each output. Moreover, considering the propagation of errors

$$(\Delta n_i)^2 = \left| \frac{\partial \langle n_i \rangle}{\partial \phi} \right|^2 (\Delta \phi)^2 \quad (4.63)$$

it becomes apparent that we have everything we need to determine the precision to which we know the measured phase. So, using that $\frac{\partial \langle n_c \rangle}{\partial \phi} = \frac{N}{2} \sin(\phi)$, we find

$$\begin{aligned} (\Delta\phi)^2 &= \frac{4}{N^2 \sin^2(\phi)} \cdot N \cos^2(\phi/2) \sin^2(\phi/2) = \frac{1}{N} \\ \Rightarrow \Delta\phi &= \frac{1}{\sqrt{N}} \end{aligned} \quad (4.64)$$

this is the expected classical result otherwise known as the “Standard Quantum Limit” or, as will be referred to hereafter, the “Shot Noise Limit” (SNL). An increase in the number of resources N yields a greater precision.

4.4.2 Coherent State

Here we investigate the effects of sending a coherent state $|\alpha\rangle$ into mode a and not sending anything (which is equivalent to sending the vacuum state $|0\rangle$) into mode b as depicted in Fig.4.2. We also take this opportunity to analyse the unitary operations that model the

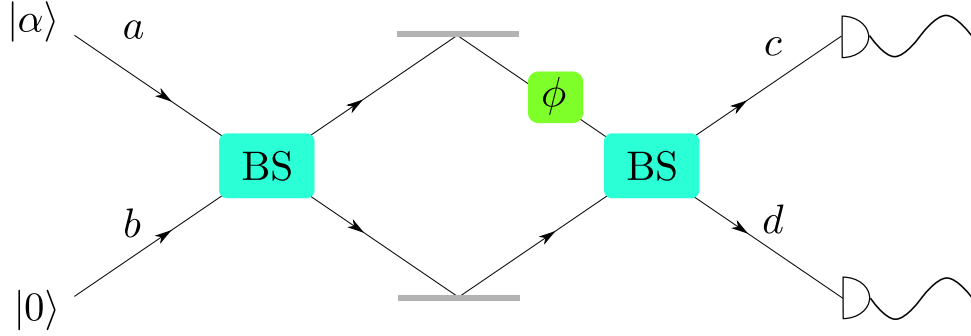


Figure 4.2: An MZI with a coherent state input in mode a and a vacuum state in mode b .

action of the beam splitter and phase shift. We begin with the Hamiltonian of the linear crystal that constitutes the beam splitter which is given by

$$H_0 = \hbar\omega(a^\dagger a + b^\dagger b + 1). \quad (4.65)$$

The beam splitter can be described by the interaction Hamiltonian (in the Schrödinger picture) given by

$$V = \hbar\kappa(e^{i\varphi}a^\dagger b + e^{-i\varphi}ab^\dagger) \quad (4.66)$$

from which we can show that $[H_0, V] = 0$ (which implies total photon number is conserved).

Furthermore, we find

$$V_I(t) = e^{iH_0 t/\hbar} V e^{-iH_0 t/\hbar} = V = \hbar\kappa(e^{i\varphi}a^\dagger b + e^{-i\varphi}ab^\dagger) \quad (4.67)$$

in other words, there is no time dependence. Hence the interaction picture evolution is given by

$$|\psi(t)\rangle = e^{-i\kappa t(e^{i\varphi}a^\dagger b + e^{-i\varphi}ab^\dagger)} |\psi(0)\rangle. \quad (4.68)$$

Noting that the parameter κ is dependent on the properties of the dispersive linear crystal, we can choose the crystal properties to fix this parameter and similarly we can choose the crystal thickness to fix the interaction time t and as such it is useful to define $\theta = \kappa t$. With this we define the unitary evolution operator

$$U_{BS} \equiv e^{-i\theta((e^{i\varphi}a^\dagger b + e^{-i\varphi}ab^\dagger))} \quad (4.69)$$

then setting $\theta = \pi/4$ (which in practise amounts to careful arrangement of κ and t) yields the widely used 50:50 beam splitter. The action of a beam splitter on coherent states $|\alpha\rangle$ and $|\beta\rangle$ (one in each mode) is then given by

$$U_{BS} |\alpha\rangle |\beta\rangle = \exp \left[\frac{\alpha}{\sqrt{2}}(a^\dagger + b^\dagger) - \frac{\alpha^*}{\sqrt{2}}(a + b) \right] \exp \left[\frac{\beta}{\sqrt{2}}(b^\dagger - a^\dagger) - \frac{\beta^*}{\sqrt{2}}(b - a) \right] |0\rangle_a |0\rangle_b \quad (4.70)$$

then using the BCH formula (2.33) we find

$$\begin{aligned} U_{BS} |\alpha\rangle |\beta\rangle &= \exp \left[\frac{\alpha - \beta}{\sqrt{2}}a^\dagger - \frac{\alpha^* - \beta^*}{\sqrt{2}}a \right] \exp \left[\frac{\alpha + \beta}{\sqrt{2}}b^\dagger - \frac{\alpha^* + \beta^*}{\sqrt{2}}b \right] |0\rangle_a |0\rangle_b \\ &= |(\alpha - \beta)/\sqrt{2}\rangle_a |(\alpha + \beta)/\sqrt{2}\rangle_b. \end{aligned} \quad (4.71)$$

Turning our attention to the phase shift operation, we have in the interaction picture

$$H_0 = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \quad (4.72)$$

which, in the Schrödinger picture, gives the interaction Hamiltonian

$$V = -\hbar\kappa a^\dagger a \quad (4.73)$$

and since $[V, H_0] = 0$, photon number is preserved and $V_I = e^{iH_0 t/\hbar} V e^{-iH_0 t/\hbar} = V$ so that the unitary action of V_I is given by

$$U_{PS} = e^{-iV_I t/\hbar} = e^{i\phi a^\dagger a} \quad (4.74)$$

where $\phi = \kappa t$. Returning to the scheme of sending a coherent state $|\alpha\rangle$ into input 1 and the vacuum state $|0\rangle$ into input 2 of the MZI, we now have the initial state $|\psi_i\rangle = |\alpha\rangle_a |0\rangle_b$ and the final state

$$\begin{aligned} |\psi_f\rangle &= U_{BS} U_{PS} U_{BS} |\psi_i\rangle \\ &= U_{BS} U_{PS} |\alpha/\sqrt{2}\rangle_a |\alpha/\sqrt{2}\rangle_b \\ &= U_{BS} |\alpha/\sqrt{2}\rangle_a |e^{i\phi}\alpha/\sqrt{2}\rangle_b \\ &= |\alpha(1 - e^{i\phi})/2\rangle_a |\alpha(1 + e^{i\phi})/2\rangle_b \end{aligned} \quad (4.75)$$

noting that here, $U_{PS} = e^{ib^\dagger b}$ acts only on mode b . The measured intensities at the outputs will then be

$$\begin{aligned} I_c &= \frac{|\alpha|^2}{2}(1 - \cos(\phi)) = |\alpha|^2 \sin^2(\phi/2) \\ I_d &= \frac{|\alpha|^2}{2}(1 + \cos(\phi)) = |\alpha|^2 \cos^2(\phi/2) \end{aligned} \quad (4.76)$$

note that the measured intensity is equivalent to the probability of a detection at the given output. The mean and variance of the detected photon number are found to be

$$\begin{aligned} \langle n \rangle &= |\alpha|^2 \cos(\phi) \\ \Delta^2 n &= |\alpha|^2 \sin^2(\phi) \end{aligned} \quad (4.77)$$

then from (4.63), we find

$$\Delta\phi = \frac{|\alpha| \sin(\phi)}{||\alpha|^2(-\sin(\phi))|} = \frac{1}{|\alpha|} = \frac{1}{\sqrt{\bar{n}}} \quad (4.78)$$

where we have used the fact that the average number of photons in a coherent state is given by $\bar{n} = \langle \hat{n} \rangle = |\alpha|^2$.

4.4.3 Entangled Photons

Since quantum mechanics permits correlations amongst photons we can manipulate and exploit these effects in order to improve upon the precision given by the SNL, indeed this is the aim of quantum enhanced metrology. An insightful such example of these techniques is given by preparing the resources (before the phase shift) in the so-called “NOON state” - a highly correlated state of the original N resources. More concretely, the NOON state is a maximally entangled state comprising of a macroscopic superposition of all N photons on one path of the MZI and all N on the other. The state (before the phase shift) is given by

$$|\psi_{NOON}\rangle = \frac{1}{\sqrt{2}}(|N, 0\rangle_{a,b} + e^{iN\theta} |0, N\rangle_{a,b}) \quad (4.79)$$

where the value of the phase θ is determined by the specific preparation of the state (it is not the phase shift of interest) which is a notoriously difficult process in itself [67] but has been experimentally achieved [68, 69] for up to and including states of size $N = 5$. It then becomes apparent that *all* N photons will be subject to the phase shift on the one path or *nothing* will experience the phase shift (*cf.* independent resource case where half would and half would not)

$$|\psi(\phi)_{NOON}\rangle = \frac{1}{\sqrt{2}}(e^{iN\phi} |N, 0\rangle_{a,b} + e^{iN\theta} |0, N\rangle_{a,b}). \quad (4.80)$$

Choosing $\theta = \pi/2$ and using (4.57), the overall effect of the MZI on the NOON state is that *all* photons will be detected at *either* output *c* or output *d* with probabilities

$$P_c = \sin^2(N\phi/2), \quad P_d = \cos^2(N\phi/2) \quad (4.81)$$

thus the observed interference fringes vary N times faster than the outcome fringes of sending N independent (uncorrelated) resources through the MZI. Indeed the average detected photon number is given by

$$\langle n \rangle = \cos(N\phi) \quad (4.82)$$

and the variance is found to be

$$\Delta^2 n = \sin^2(N\phi) \quad (4.83)$$

then using the propagation of errors (4.63) the precision is found to be

$$\Delta^2 \phi = \frac{\sin^2(N\phi)}{(-N \sin(N\phi))^2} = \frac{1}{N^2} \quad (4.84)$$

hence $\Delta\phi = 1/N$. This is the optimal rate at which the accuracy of a measurement can possibly scale with resources and is known as the Heisenberg limit [70]. This can be mathematically proven by calculating the QFI of the NOON state (as shown below) since the QFI implicitly optimizes over all possible measurements.

4.4.4 QFI Example

So far we have been bounding the precision of the parameter estimates by the propagation of errors given by (4.63). Here we take the NOON state example above and find the associated quantum Cramér-Rao bound via calculation of the QFI. For convenience eqn. (4.54) is given again here

$$F_Q(\phi) = 4 [\langle \psi'(\phi) | \psi'(\phi) \rangle - |\langle \psi'(\phi) | \psi(\phi) \rangle|^2] \quad (4.85)$$

where here we have

$$|\psi(\phi)\rangle = |\psi(\phi)_{NOON}\rangle = \frac{1}{\sqrt{2}}(e^{iN\phi} |N, 0\rangle_{a,b} + e^{iN\theta} |0, N\rangle_{a,b}) \quad (4.86)$$

and

$$|\psi'(\phi)\rangle = \frac{iN}{\sqrt{2}} e^{iN\phi} |N, 0\rangle. \quad (4.87)$$

The relevant state overlaps are then found to be

$$\langle \psi'(\phi) | \psi'(\phi) \rangle = \left(\frac{-iN}{\sqrt{2}} e^{-iN\phi} \right) \langle N, 0 | N, 0 \rangle \left(\frac{iN}{\sqrt{2}} e^{iN\phi} \right) = \frac{N^2}{2} \quad (4.88)$$

and

$$\langle \psi'(\phi) | \psi(\phi) \rangle = \frac{1}{\sqrt{2}} (e^{-iN\phi} \langle N, 0 | - i \langle 0, N |) \frac{iN}{\sqrt{2}} e^{iN\phi} | N, 0 \rangle = \frac{iN}{2} \quad (4.89)$$

so that the QFI is given by

$$F_Q = 4 \left(\frac{N^2}{2} - \left| \frac{iN}{2} \right|^2 \right) = 4 \cdot \frac{N^2}{4} = N^2 \quad (4.90)$$

and the quantum Cramér-Rao bound becomes

$$\Delta\phi \geq \frac{1}{\sqrt{F_Q}} = \frac{1}{N} \quad (4.91)$$

which is the ultimate precision capabilities of the initial NOON state since the QFI implicitly optimises over all possible measurements. At this point we take the opportunity to give a more concrete definition of the Heisenberg limit in conjunction with the Cramér-Rao bound. There is some discrepancy in the literature with regards to what is specifically meant by the Heisenberg limit, the general consensus being that it is the regime in which the precision scales as the inverse of the resources used. We take the following strict definition that for a precision given by

$$\Delta\phi = \frac{k}{N^q} \quad \text{where, } q \leq 1 \quad (4.92)$$

for some constant k and where N is the mean number of resources used, the Heisenberg limit is attained when q saturates its bound. So we cannot achieve a scaling better than $q = 1$ but can aim to minimise k . As evidenced by the NOON state, a $1/\sqrt{N}$ enhancement in precision over the SNL is achievable in principle when using the same number of resources but employing correlations. This seems like a fantastic result but the NOON state has some drastic drawbacks; as previously mentioned it is very difficult to prepare, but moreover, it is extremely susceptible to decoherence once prepared. Indeed if we take a rough model of decoherence to be a measurement made by the environment on the system, the qualitative effect of decoherence becomes apparent - the superposition of the state must collapse onto one of its components

$$|\psi(\phi)_{NOON}\rangle = \frac{1}{\sqrt{2}} (e^{iN\phi} |N, 0\rangle_{a,b} + e^{iN\theta} |0, N\rangle_{a,b}) \longrightarrow \begin{cases} e^{iN\phi} |N, 0\rangle, & (1) \\ e^{iN\theta} |0, N\rangle, & (2) \end{cases} \quad (4.93)$$

so if we obtain case (2) we clearly cannot determine any information about the phase of interest ϕ . However, if case (1) is obtained it may seem like we can infer some information about ϕ but this is now a global phase and thus unmeasurable. Taking the more rigorous

quantum jump approach [21] yields a mathematical description of the state given by the density matrix

$$\begin{aligned} \rho = & \frac{1}{2} e^{-\gamma t} (e^{iN\phi} |N, 0\rangle + e^{iN\theta} |0, N\rangle) (e^{iN\phi} \langle N, 0| + e^{iN\theta} \langle 0, N|) \\ & + \frac{1}{2} (1 - e^{-\gamma t}) \sum_n C_n (|n, 0\rangle \langle n, 0| + |0, n\rangle \langle 0, n|) \end{aligned} \quad (4.94)$$

where, t is the time of evolution and γ is the rate that photons are lost by the field (the emission rate) which here, is dependent on the number of photons N . Our interest lies in the first term of (4.94) containing information on the phase shift of interest. This term is associated with an exponentially decreasing factor dependent on γ hence this term rapidly falls off leaving the much more probable classical result given by the second term which contains no information on the phase shift.

Chapter 5

Quantum Correlations in Quantum Enhanced Metrology

In this chapter, the role of quantum correlations in the setting of optical quantum enhanced metrology protocols are investigated. In particular, two distinct types of quantum correlations used in probe states are identified and contrasted in terms of the precision enhancements and the practicalities of implementation.

This chapter is based on the papers:

- (1) *Practical quantum metrology with large precision gains in the low photon number regime*, P A Knott, T J Proctor, A J Hayes, J P Cooling and J A Dunningham, Physical Review A 93, 033859 (2016)
- (2) *Local versus global strategies in multiparameter estimation*, P A Knott, T J Proctor, A J Hayes, J F Ralph, P Kok and J A Dunningham, Physical Review A 94, 062312 (2016)

5.1 Quantum Correlations

A quintessential application of quantum enhanced metrology involves the utilisation of quantum mechanical correlations (i.e correlations present in nature that cannot be understood through a classical description) among the resources that constitute the probe state (typically particles such as photons or cold atoms), this can result in higher precision measurements with lower particle flux. This is especially useful for biological sensing [20] where radiation incident on the sample could appreciably damage it [71]. Another relevant potential application is that of gravitational wave detection [19] where the mirrors can suffer from distortion if the photon flux is too high [72, 73]. Here we focus on optical MZI

schemes where we investigate how to use quantum correlations in order to improve on the SNL, much like in the NOON state example given in the subsections 4.4.3 and 4.4.4. Moreover, we pay close attention to the type of correlations we employ and in anticipation of this we identify two distinct types of correlations:

- **Intramode correlations** are the interdependence of resources established *within* a spatial mode of the MZI.
- **Intermode correlations** are the interdependence of resources established *between* the spatial modes of the MZI.

In the example of the NOON state, entanglement establishes correlations between spatial modes hence the scheme utilises **intermode** correlations. The reader is also reminded that the practicalities of these particular probe states posit severe limitations in the practical metrological usage of them. A renowned example of the use of **intramode** correlations is

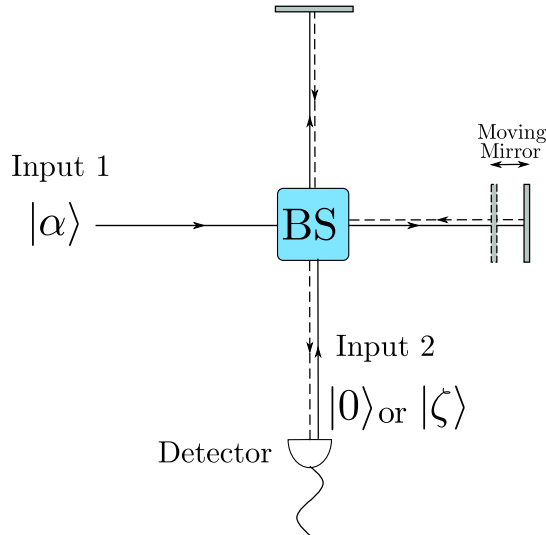


Figure 5.1: The gravitational wave detection scheme as proposed by C. Caves [74] using a Michelson interferometer. A squeezed vacuum state $|\zeta\rangle$ is injected into the previously unused port thus introducing intramode correlations into the scheme.

the scheme put forward by Caves in 1981 [74] which modified the usual usage of a Michelson interferometer (essentially a MZI folded in on itself and is in fact mathematically equivalent in its description). The original scheme involved a coherent state of light being injected into one of the ports of the interferometer and the other port was left unused (equivalent to a vacuum state $|0\rangle$ input) as displayed in Fig.5.1. A gravitational wave imparts a change in the path length of only one arm of the interferometer which consequently shifts the position of the associated mirror and results in a phase difference between the light on different

arms of the interferometer. Caves introduced intramode correlations to this scheme by analysing the effects of injecting a squeezed vacuum state into the previously unused arm of the interferometer. Using this technique with N resource photons and squeezing strength $\zeta = re^{i\theta}$ it can be shown that phase estimates can be made, in principle, to the following precisions

$$\Delta\phi = \begin{cases} \frac{e^{-r}}{\sqrt{N}} & \text{for } N \gg r \\ \frac{1}{N^{3/4}} & \text{optimising over } r \\ \frac{1}{N} & \text{optimising over measurements} \end{cases} \quad (5.1)$$

and furthermore, it has been shown [75] that when photon losses are accounted for in the limit of large photon number N , this scheme is actually optimal which, in stark contrast to the intermode-correlated probe states, demonstrates the robustness of the intramode correlations approach. Given that both the NOON state example and the Caves scheme can in principle achieve the Heisenberg limit and that they rely on inter- and intra-mode correlations respectively, we investigate the usefulness of each type of correlation and put our findings to use by introducing previously unexplored probe states that are constructed based on the following investigation and yield large precision gains.

5.1.1 QFI for Path-Symmetric Pure States

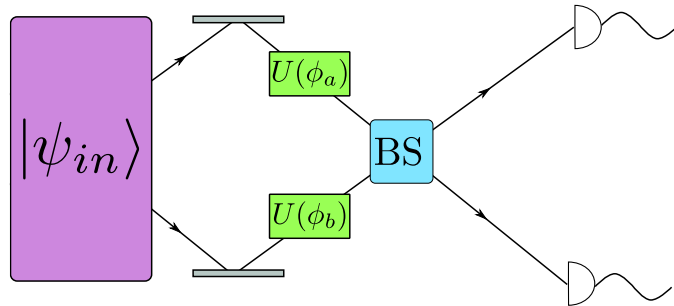


Figure 5.2: The general interferometry scheme used in the following investigation. Path symmetric states are used to exploit the natural symmetries of the MZI.

The following exploits the natural symmetries of the MZI in that we only consider pure, path-symmetric states, as depicted in Fig.5.2. It is first useful to note that if the phase is imprinted onto the probe state through a unitary operator of the form $\hat{U}(\phi) = \exp(i\phi\hat{O})$ we are able to rewrite the QFI (4.85) as

$$F_Q = 4(\langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2) = 4(\Delta^2\hat{O}). \quad (5.2)$$

The probe state in Fig.5.2 undergoes a phase shift that affects each mode and is of the form $U = \exp[i(\phi_a a^\dagger a + \phi_b b^\dagger b)]$, then defining $\hat{O}^\pm = (a^\dagger a \pm b^\dagger b)$ and $\phi^\pm = \phi_a \pm \phi_b$ we have

$$U = \exp[i(\phi^+ \hat{O}^+ + \phi^- \hat{O}^-)] \quad (5.3)$$

noting that the phase difference $\phi^- = \phi_a - \phi_b$ is the quantity of interest. There are some subtleties of phase estimation that must be addressed here and have been detailed in Ref. [76]; if the phases ϕ_a and ϕ_b are both defined relative to some reference phases (independently) then the estimation of ϕ^\pm is a two parameter problem requiring a two-parameter form of the QFI (known as the quantum Fisher information matrix). On the other hand, if no reference phases are available then the sum ϕ^+ is of no relevance and must be averaged over, however for a path-symmetric pure state it has been shown [76] that the QFI is unaffected by such phase averaging and is in fact given by $F_Q = \Delta^2(a^\dagger a - b^\dagger b) \equiv \text{var}[a^\dagger a - b^\dagger b]$ then since the number operators are given by $\hat{n}_a = a^\dagger a$ and $\hat{n}_b = b^\dagger b$, the QFI can be expressed as

$$\begin{aligned} F_Q &= 2 \left(\langle \hat{n}_a^2 \rangle - \langle \hat{n}_a \rangle^2 - \langle \hat{n}_a \otimes \hat{n}_b \rangle + \langle \hat{n}_a \rangle \langle \hat{n}_b \rangle \right) \\ &= 2 (\text{var}[\hat{n}_a] - \text{cov}[\hat{n}_a, \hat{n}_b]). \end{aligned} \quad (5.4)$$

where we can replace \hat{n}_a with \hat{n}_b in the variance term. Furthermore, we note that it has been shown [77] that for all path-symmetric pure states the optimal measurement scheme is to perform mixing of the two modes via a balanced beam splitter and count the photon number at the outputs. From this we can in fact rewrite the QFI once more in a form that plainly reveals the role of intra- and inter-mode correlations. To this end, we introduce the Mandel \mathcal{Q} parameter defined as $\mathcal{Q} = (\text{var}[\hat{n}_a] - \langle \hat{n}_a \rangle) / \langle \hat{n}_a \rangle$ and the mode correlation factor $\mathcal{J} = \text{cov}[\hat{n}_a, \hat{n}_b] / \text{var}[\hat{n}_a]$, it is then just a case of simple rearrangement to arrive at the following form of the QFI

$$F_Q = \bar{n}(1 + \mathcal{Q})(1 - \mathcal{J}) \quad (5.5)$$

where \bar{n} is the average photon number in each mode and as noted in Ref. [78] the mode correlation factor is bounded $-1 < \mathcal{J} < 1$ in contrast to the Mandel \mathcal{Q} factor which has no upper bound. Since the Mandel \mathcal{Q} parameter is a variance based quantity whereas the correlation factor is covariance based, this immediately suggests that the intramode correlations can contribute more to precision gains than the intermode correlations. The following work looks to investigate and exploit this indicative result.

5.2 Squeezed-Entangled State

Since it is clear that both intra- and inter-mode correlations can contribute to precision enhancements, we introduce a state that utilises both types of correlations which we refer to as the Squeezed-Entangled state (SES) and is of the form

$$|\psi_{SES}\rangle = \mathcal{N} (|\zeta, 0\rangle + |0, \zeta\rangle) \quad (5.6)$$

where the normalisation factor is given by

$$\mathcal{N} = \left[2 \left(1 + \frac{1}{\cosh(|\zeta|)} \right) \right]^{-1/2}. \quad (5.7)$$

In order to find the QFI, as given by (5.4), we first evaluate the term

$$\begin{aligned} \text{var}[\hat{n}_a] &= \langle \psi_{SES} | (a^\dagger a)^2 | \psi_{SES} \rangle - (\langle \psi_{SES} | a^\dagger a | \psi_{SES} \rangle)^2 \\ &= \mathcal{N}^2 \langle \zeta | (a^\dagger a)^2 | \zeta \rangle - \mathcal{N}^4 \langle \zeta | a^\dagger a | \zeta \rangle \end{aligned} \quad (5.8)$$

where we have used $a|0\rangle_a = 0$, then noting that the last term is simply the average photon number of the squeezed vacuum state and is well known to be $\langle \zeta | a^\dagger a | \zeta \rangle = \sinh^2(r)$ we are left with the task of evaluating the first term which, using the shorthand $\hat{S}(\zeta) \equiv S$ can be expressed as

$$\langle \zeta | (a^\dagger a)^2 | \zeta \rangle = \langle 0 | S^\dagger a^\dagger S S^\dagger a S S^\dagger a^\dagger S S^\dagger a S | 0 \rangle \quad (5.9)$$

then using the transformations given by (2.49) we are left to deal with a lot of terms but due to the action of the annihilation operator on the vacuum state many of the terms vanish and we find

$$\langle \zeta | (a^\dagger a)^2 | \zeta \rangle = 2 \sinh^2(r) \cosh^2(r) + \sinh^4(r) \quad (5.10)$$

so that

$$\text{var}[\hat{n}_a] = \mathcal{N}^2 \sinh^2(r) [2 \cosh^2(r) + \sinh^2(r)] - \mathcal{N}^4 \sinh^2(r). \quad (5.11)$$

For the covariance term in (5.4), we immediately note that again from the action of the annihilation operator on the vacuum state

$$\langle \psi_{SES} | \hat{n}_a \hat{n}_b | \psi_{SES} \rangle = \mathcal{N}^2 (\langle \zeta, 0 | + \langle 0, \zeta |) a^\dagger a b^\dagger b (|\zeta, 0\rangle + |0, \zeta\rangle) = 0 \quad (5.12)$$

and moreover, since we are dealing with path symmetric states we have $\langle \hat{n}_a \rangle = \langle \hat{n}_b \rangle$. The expression for the covariance is then

$$\begin{aligned} \text{cov}[\hat{n}_a, \hat{n}_b] &= \langle \hat{n}_a \hat{n}_b \rangle - \langle \hat{n}_a \rangle \langle \hat{n}_b \rangle \\ &= -(\langle \psi_{SES} | a^\dagger a | \psi_{SES} \rangle)^2 \\ &= -\mathcal{N}^4 \sinh^2(r) \end{aligned} \quad (5.13)$$

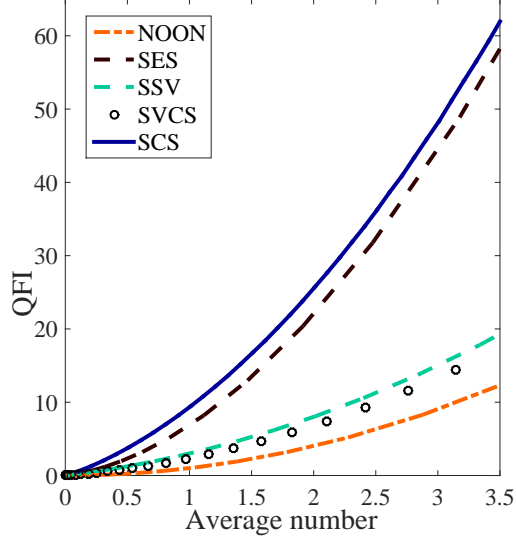


Figure 5.3: [79] Plots of the quantum Fisher information for the NOON state, the Squeezed Entangled State (SES), the Separable Squeezed Vacuum (SSV), the Caves' Squeezed-Vacuum-Coherent-State (SVCS) and the Squeezed Cat State (SCS).

and the QFI is found to be

$$\begin{aligned} F_Q &= 2 (\text{var}[\hat{n}_a] - \text{cov}[\hat{n}_a, \hat{n}_b]) \\ &= 2\mathcal{N}^2 \sinh^2(r) [2 \cosh^2(r) + \sinh^2(r)]. \end{aligned} \quad (5.14)$$

The mean number of particles in the whole state is given by

$$\bar{n} = 2 \langle \psi_{SES} | a^\dagger a | \psi_{SES} \rangle = 2\mathcal{N}^2 \langle \zeta | a^\dagger a | \zeta \rangle = 2\mathcal{N}^2 \sinh^2(r) \quad (5.15)$$

where the factor of 2 is due to the symmetry of the modes. The QFI can then be rewritten as

$$F_Q = \bar{n} \left(2 + \frac{3\bar{n}}{2\mathcal{N}^2} \right) \quad (5.16)$$

then for $r \gg 1$ we have $\mathcal{N}^2 \approx 1/2$ and thus $F_Q \approx \bar{n}(3\bar{n} + 2)$ which, in the asymptotic limit, is a factor 3 improvement over both the NOON state and the probe state used in the Caves scheme. In Fig. 5.3 we compare the QFI of these probe states in addition to another notable state — the separable squeezed vacuum state (SSV) — given by $|\psi_{SSV}\rangle = |\zeta\rangle_a \otimes |\zeta\rangle_b$ which can be made using only Gaussian operations and is the optimal such state with $F_Q = \bar{n}^2 + 2\bar{n}$. From the plots in Fig. 5.3 it is apparent that in the low photon number regime $\bar{n} \approx 1$, the precision gains are even more substantial with $F_Q(\psi_{SES}) \approx 7F(\psi_{NOON})$ and moreover, the potential for precision gains through the use of the SES and in particular, through the use of the combined intra- and inter-mode correlations, is clearly depicted.

5.3 Squeezed-Cat State

Although the SES shows great potential for high precision gains, the practicalities of engineering such a state severely impede it's metrological use. The underlying reasons become apparent when considering the fact that the SES is a coherent superposition of the NOON state and although techniques exist for creating NOON states [67] the nonlinearities required for this are difficult to implement in practise. Furthermore, the result (5.5) suggests that the largest contribution in sensitivity should come from the Mandel \mathcal{Q} parameter i.e intramode correlations. This motivates the investigation into a more practically viable, intramode correlated probe state. In order to find such states, we turn our attention to the non-Gaussian Cat States as given by (2.88) which are superpositions of coherent states and analyse the effect of squeezing on these state. The squeezed cat state (SCS) is then given by

$$|\varphi_{SCS}\rangle = \mathcal{N}\hat{S}(\zeta)(|\alpha\rangle + |-\alpha\rangle) \quad (5.17)$$

where the normalisation factor is given by $\mathcal{N} = (2 + 2e^{-2\alpha^2})^{-1/2}$. For phase estimation as depicted by Fig.5.2, a two-mode state must be considered so we investigate the state

$$|\psi_{SCS}\rangle = |\varphi_{SCS}\rangle_a \otimes |\varphi_{SCS}\rangle_b \quad (5.18)$$

so this probe state does not exhibit any mode entanglement and thus it does not exhibit any intermode correlations. The QFI is now calculated using (5.4) and noting at once that $\text{cov}[\hat{n}_a, \hat{n}_b] = 0$ since the state is mode separable thus $F_Q = 2\text{var}[\hat{n}_a]$. With this, after some algebra we find

$$F_Q = 4(s_1^4 + s_1^2) + 2\alpha^2(\kappa c_4 - s_4) + 2\alpha^4 [c_4 - \kappa s_4 - (\kappa c_2 - s_2)^2] \quad (5.19)$$

using the shorthand where, $s_k \equiv \sinh(k\zeta)$, $c_k \equiv \cosh(k\zeta)$ and $\kappa = (2 - 2e^{-2\alpha^2})(2 + 2e^{-2\alpha^2})^{-1}$. Furthermore, the average total photon number can be found to be

$$\bar{n} = 2s_1^2 + 2\alpha^2(\kappa c_2 - s_2) \quad (5.20)$$

from which it is evident that the QFI cannot be expressed directly in terms of \bar{n} . Note that in the limit of $\alpha \rightarrow 0$ the QFI becomes $F_Q = \bar{n}^2 + 2\bar{n}$ which is exactly that of the SSV. Since the SSV is the optimal Gaussian state, it is clear that this non-Gaussian outperforms all Gaussian states. In Fig.5.3, the QFI is optimised over α and ζ and it is apparent that the SCS provides vast improvements over the NOON state, the Caves state and the SSV. Some small gains over the SES are achieved by the SCS but it's

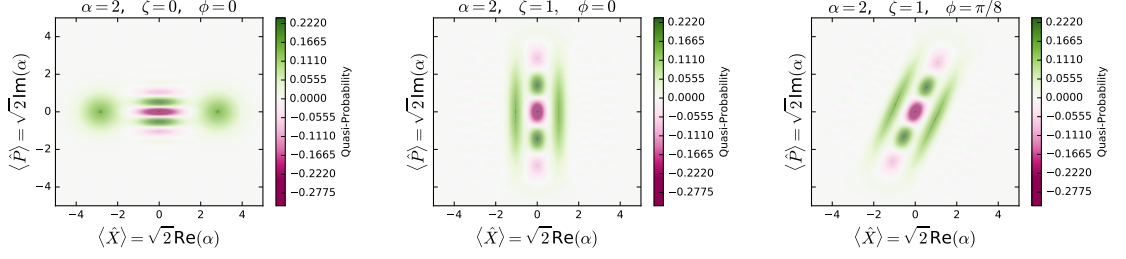


Figure 5.4: Plots of the Wigner functions for the Cat State (left), the Squeezed Cat State (centre) and the phase-shifted Squeezed Cat State (right). It is evident that squeezing the cat state reduces the overlap of the fringes of the states before and after a phase shift.

crucial advantage over the SES is that it has been experimentally realised [80, 81, 82]. For example, the method of Huang *et. al* [82] initially prepares two separate squeezed vacuum states and subjects one to a $\pi/2$ -phase shift before interacting the states in a beam splitter of variable transmissivity, finally a photon number counting measurement is performed on a single mode thus heralding the SCS on the other. The SCSs they are able to produce are of 67% fidelity and of amplitude $|\alpha| = \sqrt{3}$. In principle, one could directly apply squeezing [83, 84, 85] to a Cat State since there are a myriad of techniques to experimentally generate Cat States [86, 87, 88]. In order to gain some intuition on what is generating these high precision gains we turn to the Wigner function, as defined in subsection 2.2.1, to help visualize the measurement protocol. The Wigner function and the QFI are in fact related through a mathematical object known as the fidelity where for the state $|\psi\rangle$ and the infinitesimally phase-shifted state $|\psi(\delta\phi)\rangle$ the fidelity is given by $\mathbb{F} \equiv |\langle\psi(\delta\phi)|\psi\rangle|^2$ [89]. In turn, the fidelity is given in terms of the Wigner function as

$$\mathbb{F} = \pi \int d^2\alpha W_\psi W_{\psi(\delta\phi)} \quad (5.21)$$

thus if the overlap of the Wigner function of the state before the phase shift with the Wigner function of the state after the phase shift is small, the QFI will be large. Hence it is desirable for a phase shift to result in as little overlap of the associated Wigner functions as possible. Indeed from Fig.5.4 it is evident that the precision enhancement comes from the small overlap of the fringes - which is a manifestation of the quantum interference of the superposition state - of the state of the probe before and after the phase shift which is given by a rotation in phase space. Clearly, the greater the squeezing, the smaller the overlap and consequently, the greater the sensitivity. The detection scheme is a combination of mode mixing and photon counting; photon resolving detectors are an area of active research [90] and devices of high resolution in the low photon regime have been demonstrated [91, 92, 93] which is of particular relevance to this the present work. By considering this

classical Fisher information, this measurement scheme can be shown to be optimal [79]. This has demonstrated that intramode correlations alone can provide substantial precision gains and has the advantage that the preparation of these probe states are more practical than their intermode-correlated counterparts. Furthermore, it was shown in Ref. [79] that the SCS demonstrates the potential for robust phase measurements up to 27% loss.

5.4 Multiparameter Estimation

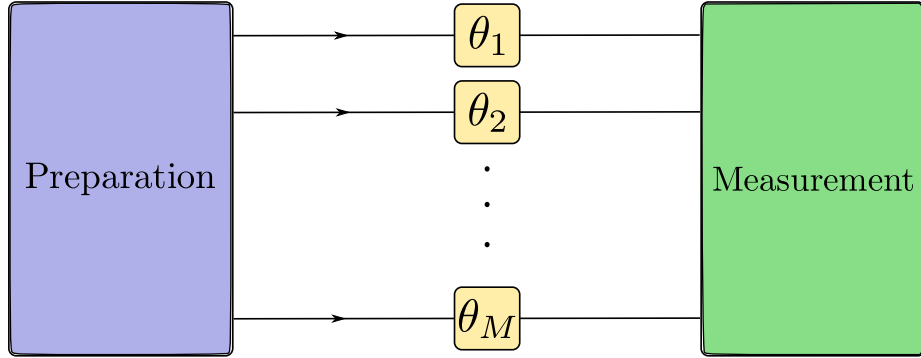


Figure 5.5: The general setup of a quantum optical multiparameter estimation scheme in which the “preparation” stage involves the creation of an M -mode probe state which is subject the $\vec{\theta} = (\theta_1, \dots, \theta_M)$ linear phase shifts and subsequently the multiple parameters ϕ_i for $i \in \{1, \dots, d\}$, which are functions of $\vec{\theta}$, are read out in the “measurement” stage.

It is often desirable to estimate multiple parameters at once, gravitational wave detection is one such instance [94] in which knowledge of quantities such as direction *and* polarisation of the wave are desired. Under this generalisation, we once again ask: which type of quantum correlations can better offer practically attainable, higher precisions? Multimode entanglement schemes which simultaneously measure all parameters have been shown to offer substantial precision gains [95], however this scheme only deals with fixed photon number states (a constraint that is relaxed in the following work) and it has also been shown that multimode entanglement can be detrimental to quantum metrological protocols [96, 97]. The general problem of multiparameter estimation is depicted in Fig. 5.5, the aim is to estimate d parameters given by the vector $\vec{\phi} = (\phi_1, \phi_2, \dots, \phi_d)$. In order to establish a precision bound on the i th parameter, the notion of QFI is extended to encompass multiple parameters. To this end, we introduce the quantum Fisher information matrix (QFIM) which is defined by [63]

$$F_{lm}^Q = \langle \psi(\vec{\phi}) | (L_l L_m + L_m L_l) | \psi(\vec{\phi}) \rangle \quad (5.22)$$

where L_l is the symmetric logarithm derivative, given by

$$L_l = 2 \left(|\partial_l \psi(\vec{\phi})\rangle \langle \psi(\vec{\phi})| + |\psi(\vec{\phi})\rangle \langle \partial_l \psi(\vec{\phi})| \right) \quad (5.23)$$

with $|\partial_l \psi(\vec{\phi})\rangle = \frac{\partial}{\partial \phi_l} |\psi(\vec{\phi})\rangle$. The imprinting of the phase information of multiple parameters onto some ($\vec{\phi}$ independent) multimode probe state $|\psi\rangle$ is mathematically modelled by the unitary operator

$$U(\vec{\phi}) = \exp \left(i \sum_{i=1}^d \phi_i \hat{O}_i \right) \quad (5.24)$$

where the operators denoted by \hat{O}_i are Hermitian and are mutually commutative $[\hat{O}_i, \hat{O}_j] = 0, \forall i, j$. With this, we find the QFIM reduces to

$$F_{lm}^Q = \text{cov}[\hat{O}_i, \hat{O}_j] \quad (5.25)$$

where the covariance is taken with respect to the initial ($\vec{\phi}$ independent) state $|\psi\rangle$. For the case of optical M -optical mode phase estimation, as given in Fig.5.5, we take $\hat{O}_i = \hat{n}_i$, the number operator, so the phase shift generator is given by

$$U(\vec{\theta}) = \exp \left(i \sum_{j=1}^M \theta_j \hat{n}_j \right) \quad (5.26)$$

where, as will become apparent in the specific examples that follow, $d \leq M$ and ϕ_i is a function of θ_j .

5.5 Local vs. Global Strategies

Multiparameter estimation schemes in the context of optical quantum metrology schemes have been shown to offer precision enhancements [95, 98, 99] but the origin of these enhancements is not known. This issue is investigated here through considering how the types of correlations used affect the precision enhancements furthermore, the practicalities of such schemes are highlighted here. As such we define a “local estimation strategy” as a scheme in which the input ($\vec{\phi}$ independent) state is separable and the final measurements made are attained using exclusively local operations. This means that in a local multimode estimation strategy, each parameter can be estimated individually. With this we define a “global estimation strategy” as any strategy that is not local. We now present two examples of global schemes with local analogues and show that for probe state with unfixed photon number, a key resource for high precision gains is a large photon number variance.

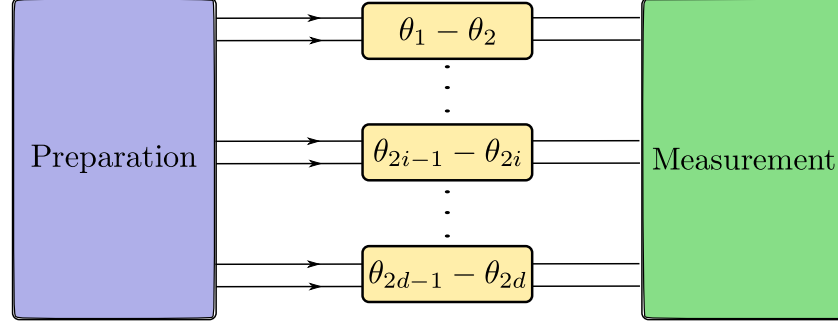


Figure 5.6: A network of quantum sensors modelled by a multimode collection of parallel Mach-Zehnder interferometers, each of which yields a phase difference to be measured.

5.5.1 Parallel Interferometers

The first example considered is essentially a generalisation of the MZI to the multiparameter setting, this can be thought of as a network of quantum sensors and has practical application to gravitational wave detection [94]. Using Fig.5.5 as a starting point, we take $M = 2d$, where the i th interferometer is given by the modes $2i - 1$ and $2i$ with $i \in \{1, \dots, d\}$ and we re-parametrize the phase shifts to the phase sum and difference within the i th interferometer (the latter quantity being the parameter of interest) i.e $\phi_{i\pm} = \theta_{2i-1} \pm \theta_{2i}$ as depicted in Fig.5.6. Consequently, the phase shift operator takes the form of

$$U(\vec{\phi}) = \exp \left(i \sum_{i=1}^d [\phi_{i-} \hat{O}_{i-} + \phi_{i+} \hat{O}_{i+}] \right) \quad (5.27)$$

where, $\vec{\phi} = (\phi_{1-}, \dots, \phi_{d-}, \phi_{1+}, \dots, \phi_{d+})$ and the Hermitian generating operators are given by

$$\hat{O}_{i\pm} = \frac{1}{2}(\hat{n}_{2i-1} \pm \hat{n}_{2i}) \quad (5.28)$$

thus the QFIM takes the form $F_{i\pm j\pm}^Q = 4\text{cov}[\hat{O}_{i\pm}, \hat{O}_{j\pm}]$. Since we are dealing with a network of MZIs, we are again able to exploit the natural symmetries and as such we consider states that are symmetric with respect to swapping interferometer labelling and symmetric with respect to swapping the modes in each interferometer hence the variances of all the modes are equal $\text{var}[\hat{n}_i] = \text{var}[\hat{n}_j] \equiv V$. Furthermore, the covariances of two modes within the same interferometer are equal for any two interferometers $C_{2i-1,2i} = C_{2j-1,2j} \equiv C_{\text{Intra}}$ where $C_{i,j} \equiv \text{cov}[\hat{n}_i, \hat{n}_j]$ and additionally $C_{2i-1,j} = C_{2m-1,n} \equiv C_{\text{Inter}}$ for $j \neq 2i - 1, 2i$ and $n \neq 2m - 1, 2m$. For the case of both parameters taking “ \pm ” the QFIM can then be expressed as

$$F_{i\pm j\pm}^Q = C_{2i-1,2j-1} + C_{2i,2j} \pm C_{2i,2j-1} \pm C_{2i-1,2j} \quad (5.29)$$

and for the case of opposite signs we find

$$F_{i\pm j\mp}^Q = C_{2i-1,2j-1} - C_{2i,2j} \pm C_{2i,2j-1} \mp C_{2i-1,2j} \quad (5.30)$$

from which it is evident that $F_{i\pm j\mp}^Q = F_{i-j-}^Q = 0$, $\forall i \neq j$ and that $F_{i+j+}^Q = 4C_{\text{Intra}}$, $\forall i \neq j$. For the case of $i = j$ we have $F_{i\pm j\mp}^Q = 0$ (off-diagonals of the QFIM) and $F_{i\pm j\mp}^Q = 2(V \pm C_{\text{Intra}})$ (diagonals terms of the QFIM). The QFIM can then be expressed as

$$F^Q = \begin{pmatrix} 2(V - C_{\text{Intra}})\mathbb{I} & 0 \\ 0 & M \end{pmatrix} \quad (5.31)$$

where, \mathbb{I} is the $d \times d$ dimensional identity matrix and $M = \lambda(\mathbb{I} + \omega\mathcal{I})$ with $\lambda = 2(V + C_{\text{Intra}} - 2C_{\text{Inter}})$, $\omega = 2C_{\text{Inter}}/(V + C_{\text{Intra}} - 2C_{\text{Inter}})$ and \mathcal{I} is the $d \times d$ dimensional matrix of all 1's. Calculation of the corresponding quantum Cramér-Rao bound requires finding the inverse of the QFIM and consequently the inverse of M , indeed the inverse of any matrix of the form M is given by

$$M^{-1} = \frac{1}{\lambda} \left(\mathbb{I} - \frac{\omega}{1 + \omega d} \mathcal{I} \right) \quad (5.32)$$

but this term is not of physical consequence as we are interested in the ϕ_{i-} terms only; this quantity is required only for the calculation of the terms of interest. The inverse QFIM is then

$$F^Q = \begin{pmatrix} 2(V - C_{\text{Intra}})^{-1}\mathbb{I} & 0 \\ 0 & M^{-1} \end{pmatrix} \quad (5.33)$$

and the bound on precision for the parameter of interest is found to be

$$\Delta^2 \phi_{i-} \geq \frac{1}{2(V - C_{\text{Intra}})} \quad (5.34)$$

noting that this is for the individual phases. From this we can conclude that the variance of the probe state's photon number and the correlations within an individual interferometer are the only quantities that affect the precision of the phase estimate and hence entanglement between interferometers (sensors) is not necessary. Further understanding can be attained by rewriting (5.34) in terms of the Mandel \mathcal{Q} parameter and the correlation parameter which for multimode schemes, are given by $\mathcal{Q}_i = (V_i - \bar{n}_i)/\bar{n}_i$ and $\mathcal{J}_{ij} = C_{i,j}/\sqrt{V_i V_j}$ respectively. Since in this example all modes have the same \mathcal{Q}_i we denote $\mathcal{Q}_i \equiv \mathcal{Q}$ and we have for the correlation parameter $\mathcal{J}_{ij} = C_{\text{Intra}}/V \equiv \mathcal{J}$ so that (5.34) becomes

$$\Delta^2 \phi_{i-} \geq \frac{1}{2\bar{n}(1 + \mathcal{Q})(1 - \mathcal{J})} \quad (5.35)$$

where, $\bar{n} = \bar{n}_i = \langle n_i \rangle$ is the average photon number in the *individual* i th mode. Noting again that \mathcal{Q} is not bounded above but the correlation factor is bounded by $-1 \leq \mathcal{J} \leq 1$, it is apparent that correlations between modes within each interferometer can provide a maximum of a factor $1/\sqrt{2}$ improvement, so there is far greater potential for precision gains through increased photon number variance which contributes to the \mathcal{Q} parameter.

Generalised Entangled Coherent State

In the established setting of a network of parallel interferometers, we now contrast specific global and local strategies through the analysis of a multimode entangle state and the analagous mode separable state. For this, the generalised entangled coherent state (GECS) is introduced as

$$|\psi_{GC}\rangle = \mathcal{N}_g \sum_{a \in \mathcal{M}} \hat{D}_a(\alpha_g) |\vec{0}\rangle \quad (5.36)$$

where $\hat{D}_a(\alpha_g)$ is the displacement operator (2.32) acting on the a th mode, \mathcal{N}_g is the normalisation factor, \mathcal{M} is the set of $M = 2d$ modes and $|\vec{0}\rangle$ is the vacuum state of the $M = 2d$ states. The precision bound is found to be

$$\Delta^2 \phi_{GC} \geq \frac{d}{\bar{N}_g(|\alpha_g|^2 + 1)} \approx \frac{d}{\bar{N}_g(\bar{N}_g + 1)} \quad (5.37)$$

where the total average photon number is $\bar{N}_g = |\alpha_g|^2 / [1 + (2d - 1)e^{-|\alpha_g|^2}]$ and for $|\alpha_g| \gg 1$ it follows that $\bar{N}_g \approx |\alpha_g|^2$. Noting that if each interferometer is considered separately, the standard quantum enhanced precision is the Heisenberg scaling $\Delta^2 \phi \geq 1/(2\bar{n})^2 = d^2/\bar{N}^2$ where the total average photon number $\bar{N} = 2d\bar{n}$, it is then clear that the GECS yields an order d improvement over the standard quantum enhancement. Now, consider the analagous mode-separable unbalanced cat state (UCS) given by

$$|\psi_{UCS}\rangle = \mathcal{N}_c(|\alpha_c\rangle + \nu|0\rangle)^{\otimes 2d} \quad (5.38)$$

where \mathcal{N}_c is a normalisation factor and $\nu \in \mathbb{R}$ is a weighting parameter. The precision bound is found to be

$$\Delta^2 \phi_{UCS} \geq \frac{d}{\bar{N}_c(|\alpha_c|^2 + 1 - \frac{\bar{N}_c}{2d})} \approx \frac{d}{\bar{N}_c(\frac{\nu^2}{2d}\bar{N}_c + 1)} \quad (5.39)$$

where, $\bar{N}_c = 2d|\alpha_c|^2/(\nu^2 + 1 + 2\nu e^{-|\alpha_c|^2/2})$ and the approximation holds for $|\alpha_c| \gg 1$. For a balanced cat state ($\nu = 1$) the precision bound scales as the standard quantum enhancement, but when taking ν^2 to scale with d the bound scales an order of d better i.e the mode-separable UCS performs just as well as the GECS. Moreover, holding the

total photon numbers equal $\bar{N}_c = \bar{N}_g$ and taking $|\alpha_c|, |\alpha_g| \gg 1$, then for $\nu^2 > 2d$ we have $\Delta^2 \phi_{UCS} < \Delta^2 \phi_{GC}$. Not only can the UCS provide higher precision estimates, all measurement operations are local and can be achieved via two-mode mixing and photon counting which, as previously mentioned, is the optimal measurement for path-symmetric pure states [77]. In contrast to this, global measurement strategies generally require much more practically demanding measurement techniques.

5.5.2 Multimode Quantum Enhanced Imaging

Here we investigate a scheme put forward by Humphreys *et. al* [95] in which d phase shifts are measured relative to one reference mode akin to quantum imaging protocol [100, 101, 102]. Again using Fig.5.5 as a starting point, we take $M = d + 1$ to be the number of modes and $\vec{\phi} = (\phi_1, \dots, \phi_d)$ where, $\phi_i = \theta_i - \theta_{d+1}$. Setting $\theta_{d+1} = 0$ gives the generator of ϕ_i to be \hat{n}_i and as such the QFIM becomes $F_{ij}^Q = 4\text{cov}[\hat{n}_i, \hat{n}_j]$. Once again there is a symmetry in the system that can be exploited, thus we take symmetry between the d modes of the probe state (this does not include the reference mode in general) which gives $V_i = V_j \equiv V, \forall i, j$ and $C_{i,j} = C_{m,n} \equiv C, \forall i \neq j, m \neq n$ so it immediately follows that $F_{ii}^Q = 4V, \forall i$ and $F_{ij}^Q = 4C, \forall i \neq j$. The QFIM then takes the form

$$F^Q = 4(V - C) \left(\mathbb{I} + \frac{C}{V - C} \mathcal{I} \right). \quad (5.40)$$

Again, the inverse of a matrix of this form is given by

$$(F^Q)^{-1} = \frac{1}{4(V - C)} \left(\mathbb{I} - \frac{C}{V + (d - 1)C} \mathcal{I} \right) \quad (5.41)$$

so that the bound on the precision estimate is given by

$$\Delta^2 \phi_i \geq \frac{V + (d - 2)C}{4(V - C)[V + (d - 1)C]}. \quad (5.42)$$

Again, further understanding can be gained through expressing this bound in terms of the Mandel \mathcal{Q} parameter and the mode correlation function $\mathcal{J} = C/V$

$$\Delta^2 \phi_i \geq \frac{f(d, \mathcal{J})}{4\bar{n}(1 + \mathcal{Q})(1 - \mathcal{J})} \quad (5.43)$$

where \bar{n} is the average photon number in the individual mode and

$$f(d, \mathcal{J}) = \frac{1 + (d - 2)\mathcal{J}}{1 + (d - 1)\mathcal{J}} \approx 1 \text{ for } d \gg 1. \quad (5.44)$$

Again, since the correlation factor is bounded and the Mandel \mathcal{Q} parameter is not bounded above, it is clear that multimode correlations can only provide a small constant factor to the precision enhancements on the estimates.

Generalised NOON State

Here we build upon the specific example provided by Humphreys *et. al* which uses the generalised NOON state which is given by

$$|\psi_{GN}\rangle = \frac{1}{\sqrt{d + \gamma^2}} [|N, 0, \dots, 0, 0\rangle + |0, N, \dots, 0, 0\rangle + \dots + |0, 0, \dots, N, 0\rangle + \gamma |0, 0, \dots, 0, N\rangle] \quad (5.45)$$

where $\gamma \in \mathbb{R}$ is a weighting parameter for the reference mode. The precision bound is given by

$$\Delta^2 \phi_{GN} \geq \frac{(d + \gamma^2)(1 + \gamma^2)}{4\gamma^2 N^2} \quad (5.46)$$

which is minimised at $\gamma = d^{1/4}$

$$\Delta^2 \phi_{GN} \geq \frac{(1 + \sqrt{d})^2}{4N^2} \quad (5.47)$$

which is an order d improvement over the standard quantum enhancement, note that $\gamma = 1$ gives equivalent scaling. This demonstrates the benefits of a global estimation strategy for states of fixed *total* photon number, we now consider the larger class of states which have fixed *average* photon number. To this end, the mode-separable unbalanced “NO” (UNO) state is introduced

$$|\psi_{UNO}\rangle = \mathcal{N}_U(|N\rangle + \nu |0\rangle)^{\otimes M} \quad (5.48)$$

where \mathcal{N}_U is the normalisation factor and $\nu \in \mathbb{R}$ is a weighting parameter. Setting $\nu = 1$ gives the same precision bound scaling as M individual NOON states. Setting $\nu = \sqrt{d + \gamma^2} - 1$ (or indeed just demanding $\nu \propto \sqrt{d}$) yields the precision scaling (5.46) given by the GNS. A multimode measurement is required for the GNS [95] whereas the UNO state scheme only requires a collection of single mode measurements to be performed after the phase shifted probes have mixed with the reference mode (the particular mixing mechanism depends on the particular probe state used).

5.5.3 General Procedure

From the previous analysis, it has become apparent that the key characteristic for the large precision gains in multimode optical quantum metrology schemes is strong correlations *within* each mode and as such, multimode entanglement is not an essential property to this end. Considering the GNS, the precision enhancements can be attributed to the scaling of the Mandel \mathcal{Q} parameter where, denoting the order of scaling by $O(\cdot)$, we have

$\mathcal{Q} = O(d\bar{n}) = O(\bar{N})$ as opposed to $O(\bar{n})$, so for fixed \bar{n} the global strategy grows with number of modes d but since the \mathcal{Q} is a local property of the individual mode, the same scaling can be achieved through a judicious choice of a mode-separable state. Indeed, for a general M -mode path-symmetric state $|\Psi\rangle$, it is possible to construct a single-mode state

$$|\psi(\Psi)\rangle = \sum_{n=0}^{\infty} |\langle n|\Psi\rangle| |n\rangle \quad (5.49)$$

so that by construction, the M -mode separable state $|\psi(\Psi)\rangle^{\otimes M}$ and the global state $|\Psi\rangle$ both contain the same average number of photons and $\mathcal{Q}(|\Psi\rangle) = \mathcal{Q}(|\psi\rangle^{\otimes M})$. Therefore, the separable analogue of any multimode scheme can be tuned to outperform the global strategy which can only exhibit a factor of $\sqrt{2}$ improvement over the separable state before any modifications are made to it.

In this chapter, we have shown that local estimation schemes can perform just as well and even surpass global estimation strategies. Moreover, local strategies offer important practical advantages such as flexibility in the distribution of resources in the probe state over the modes, easier probe state preparation as entanglement is not necessary and robustness to local estimation failure; if estimation fails on one mode, the rest of the measurements are still valid whereas this is not the case for a global scheme due to the multimode entanglement.

Chapter 6

Quantum Metrology with Time as a Limited Resource: Concurrent State Preparation and Readout

In this chapter we analyse the traditional quantum metrology protocol under the realistic condition of time being a limited resource and propose a more effective use of this resource. This is achieved through the investigation of an N spin-1/2 system used for magnetic field sensing. We investigate two distinct protocols; one of which explicitly takes into account non-negligible state preparation and sensing times while the other also includes non-negligible readout times. Using the natural transition from the N -spin system to the optical setting provided by the $N \rightarrow \infty$ bosonic limit, we are also able to perform a similar analysis in the optical setting. It is revealed that entanglement does not necessarily improve upon classical sensing schemes and furthermore, that the time limited resource is used more effectively if we are to concurrently sense during state preparation and readout.

This chapter is based on the paper:

Making the most of time in quantum metrology: concurrent state preparation and sensing. A J Hayes, S Dooley, W J Munro, K Nemoto and J A Dunningham, 2018 Quantum Science and Technology, 3(3), 035007.

6.1 Magnetic Field Sensing and Two-Axis Twisting

A quantum metrology protocol is typically ordered into three stages: *i*) Probe state preparation, in which quantum mechanical correlations are introduced to a system that will

be used as a probe. Examples include the generation of spin squeezed states [103] or of Cat States [104, 105]. *ii*) Sensing, in which the probe is subject to, and consequently altered by, a parameter of interest. The quantum mechanical correlations introduced in the preparation stage increase the probe's susceptibility to alterations caused by this parameter beyond classical limits. *iii*) Readout, in which a final measurement is made on the altered probe state enabling estimation of the parameter of interest. The three stages of the protocol take a combined time τ as depicted in Fig.6.1. Usually, the state preparation and readout times are assumed to be negligible, so that the total time τ can be devoted to the sensing stage. If the state preparation and readout times are non-negligible, however, τ should be divided between the three stages [106]. This leads to a trade-off since, for example, too much time given to state preparation subtracts from the available time for sensing, while too little time devoted to state preparation may not provide enough time to generate the most sensitive state. We investigate three schemes *A*, *B* and *C* (as shown

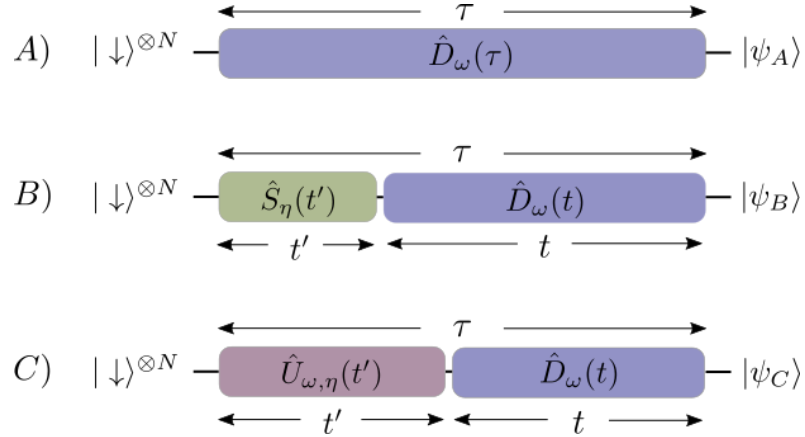


Figure 6.1: In scheme *A*, the magnetic field is applied over the entire time τ , by the operation $\hat{D}_\omega(\tau)$. The spins remain in a separable state throughout. In scheme *B*, the two-axis twisting operation $\hat{S}_\eta(t')$ generates a sensitive entangled state before exposure to the magnetic field through $\hat{D}_\omega(t)$. In scheme *C* the spins are subject to the operation $\hat{U}_{\omega,\eta}(t')$ (as defined in section 6.1.3) which exposes them to the magnetic field during the twisting operation. Each scheme ends with a measurement of the final state $|\psi_i\rangle$ ($i \in \{A, B, C\}$), which we assume can be done in a negligible time. For a fair comparison, between the three schemes, each is constrained by the time τ

in Fig.6.1) which model a classical, a traditional quantum and a “concurrent” quantum metrological protocol respectively. The initial “unprepared” state is taken to be the CSS $|\downarrow\rangle^{\otimes N}$. In order to quantify the effectiveness of each scheme we use the quantum Cramér-

Rao inequality,

$$\delta\omega_j \geq \frac{1}{\sqrt{\nu F_j}} \quad (6.1)$$

where the frequency $\omega = \gamma B$ is proportional to the magnetic field B , so that the problem of estimating ω is the same as the problem of estimating B when the gyromagnetic ratio γ is known, furthermore $j \in \{A, B, C\}$, ν is the number of experimental repeats and F_j is the quantum Fisher information which, for $\omega \approx 0$, is given by

$$F_j = 4 [\langle \partial_\omega \psi_j | \partial_\omega \psi_j \rangle + |\langle \psi_j | \partial_\omega \psi_j \rangle|^2]_{\omega=0} \quad (6.2)$$

with the final state $|\psi_j\rangle$ and $|\partial_\omega \psi_j\rangle = \frac{\partial}{\partial \omega} |\psi_j\rangle$. To make this measure more succinct, we quantify the sensitivity by the dimensionless measure which uses the above to give

$$(\sqrt{\nu}\tau\delta\omega_j)^{-1} \leq \sqrt{F_j}/\tau, \quad (6.3)$$

this bound can be saturated for an optimized measurement of $|\psi_j\rangle$ and $\nu \gg 1$.

6.1.1 Scheme A

The effect of a magnetic field on the initial CSS is described by a rotation operator as discussed in section 3.3 hence, for a magnetic field applied in the y -direction for time τ , the initial state $|\downarrow\rangle^{\otimes N}$ becomes $|\psi_A\rangle = \hat{D}_\omega(\tau) |\downarrow\rangle^{\otimes N}$ where

$$\hat{D}_\omega(\tau) = e^{-i\tau\hat{H}_\omega} \quad \text{and} \quad \hat{H}_\omega = \omega \hat{J}_y / \sqrt{N}. \quad (6.4)$$

Note that for later convenience the Hamiltonian \hat{H}_ω has been scaled by a factor of $1/\sqrt{N}$. The particular effect of this scheme is then to rotate the initial CSS around the y -axis through the angle $\phi = \omega\tau/\sqrt{N}$. Since no entanglement is introduced into the system at any time during this scheme's evolution, this scheme can be regarded as classical. As such, this scheme sets the benchmark for the subsequent quantum metrological schemes to surpass and we find that the dimensionless sensitivity is given by

$$(\sqrt{\nu}\tau\delta\omega_j)^{-1} \leq \sqrt{F_j}/\tau = 1 \quad (6.5)$$

6.1.2 Scheme B

An archetypal aspect of quantum enhanced metrology is the production, and use of, entangled states of a given system in order to improve on schemes that do not exploit this quantum phenomenon such as scheme A. As we have seen in chapters 4 and 5, entanglement can lead to high precision gains but the time taken for preparation of the probe

state and readout is assumed to be negligible. Here, we regard time as a limited resource and as such we take into account the time required for entangled state preparation in scheme B . The particular type of entangled state we investigate is the spin squeezed state generated by TAT (see section 3.4.2). Thus, beginning with the “unprepared” CSS $|\downarrow\rangle^{\otimes N}$, the system evolves under

$$\hat{S}_\eta(t') = \exp[-it' \hat{H}_\eta/\hbar], \text{ where, } \hat{H}_\eta = i\hbar\eta(\hat{J}_-^2 - \hat{J}_+^2)/N \quad (6.6)$$

which, if preparation and readout times are neglected, can lead to sensitivities of ω at the Heisenberg limit $(\sqrt{\nu}\tau\delta\omega)^{-1} = \sqrt{N}$ (which has been scaled here by a factor of $1/\sqrt{N}$ due to the prior scaling introduced in the Hamiltonians). Following the TAT preparation, the state is exposed to the magnetic field for time t giving the final state

$$|\psi_B\rangle = \hat{D}_\omega(t)\hat{S}_\eta(t')|\downarrow\rangle^{\otimes N} \quad (6.7)$$

where we ensure that the total run time of scheme B is limited to τ by constraining $t' = \tau - t$ (so that if $t' \rightarrow 0$ then $t = \tau$ and scheme B converges to scheme A). A Bloch sphere representation of these operations on a coherent state are given in Fig.6.2 to aid understanding of these operations. Note that for ease of visualisation, the z axis of the Bloch sphere has been inverted so the CSS $|\downarrow\rangle^{\otimes N}$ is pictures at the top of the sphere. An analytic expression of the QFI, F_B , is intractable and as such we calculate it numerically. Before doing so we note that scheme B can be completely characterised by the number of spins N , the total run time in units of $1/\eta$ given by $\eta\tau$ and the ratio of sensing time to total run time t/τ ; all three of these parameters are dimensionless. The sensitivities in relation to these parameters are given in Fig.6.3. It is apparent from these plots that there exists values of $\eta\tau$ and N for which scheme B gives no advantage over scheme A . This shows that two-axis twisting is not guaranteed to provide improvements in sensitivity when state preparation time is taken into account. However, values of $\eta\tau$ and N do exist for which it is clear that scheme B does give improvements over scheme A , for judicious choices of sensing time t/τ .

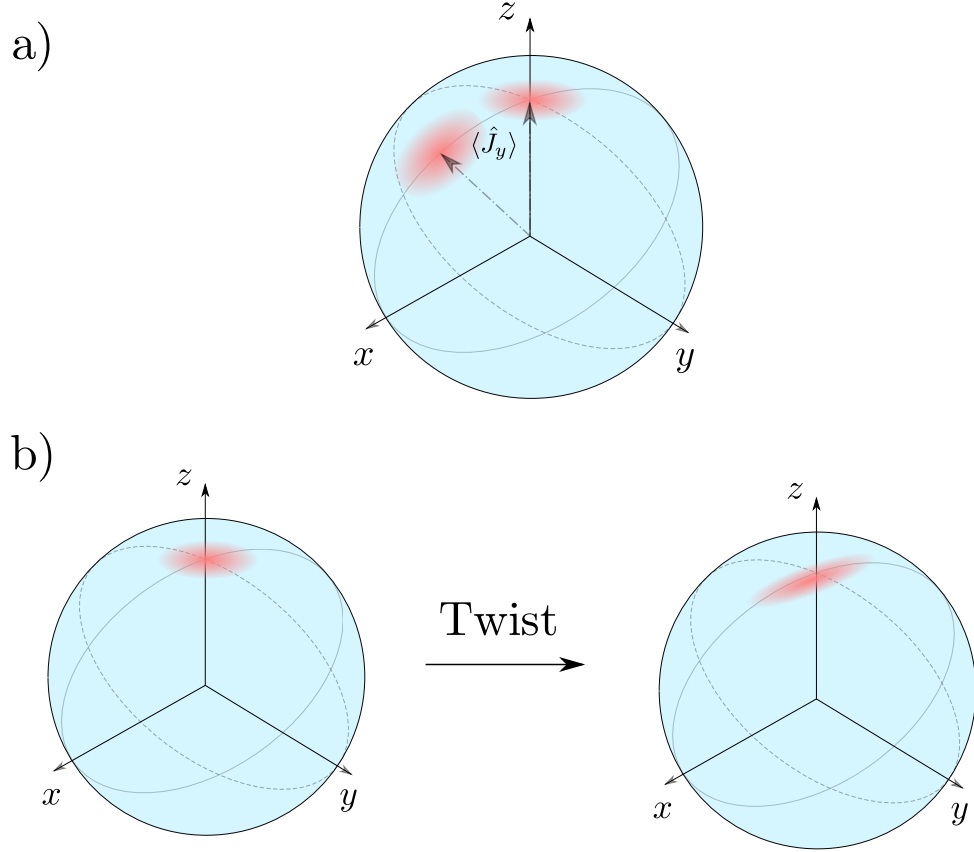


Figure 6.2: Bloch Sphere representation (with inverted z axis for ease of representation) of a) a rotation operation around the y axis, as given by equation (6.4), applied to the CSS $|\downarrow\rangle^{\otimes N}$ and b) a TAT operation as given by equation (6.6).

The size of the parameter space can be reduced by numerically optimizing over t/τ for each $\eta\tau$ and N as displayed in Fig.6.4 . From this we are able to conclude that since the optimal sensing time $(t/\tau)_{\text{opt}} = 1$, the entire time resource τ should be used for sensing i.e there should be no TAT and scheme B reduces to schemes A .

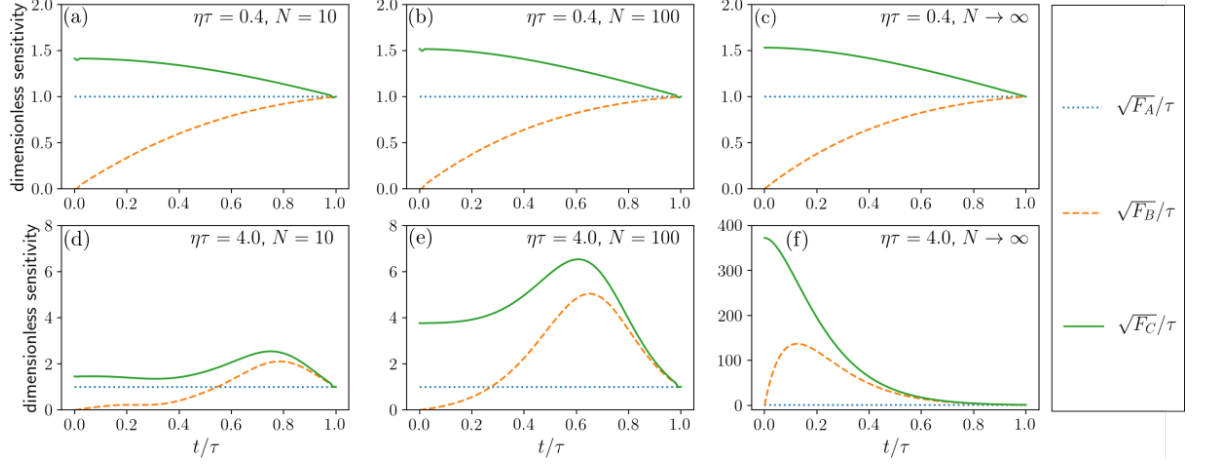


Figure 6.3: These plots show that for a sufficiently small value of $\eta\tau$ (e.g. $\eta\tau = 0.4$ in the upper plots), scheme B gives no improvement over scheme A . For a sufficiently large value of $\eta\tau$ (e.g. $\eta\tau = 4$ in the lower plots), both scheme B and scheme C can give a better sensitivity than scheme A (i.e the two-axis twisting state preparation is worthwhile), if the sensing time t/τ is optimised.

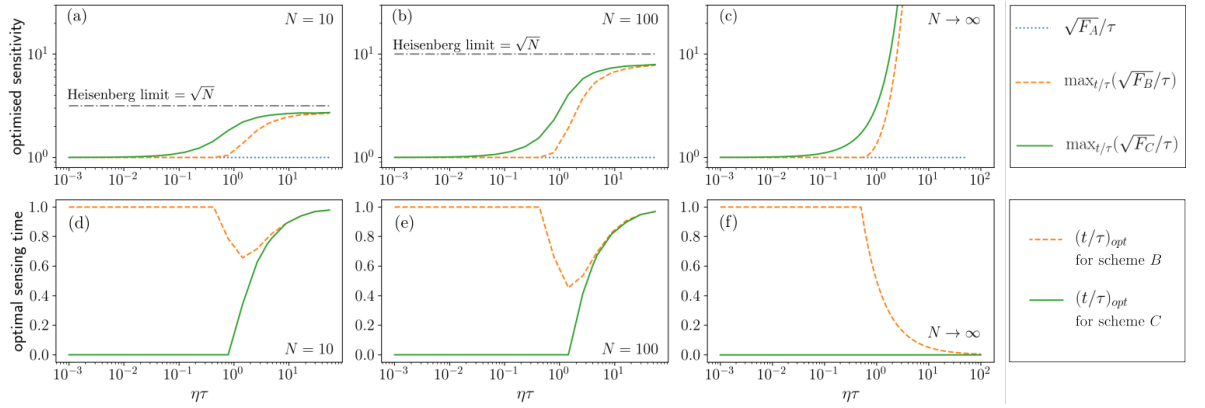


Figure 6.4: The upper plots show the optimised sensitivity bound $\max_{t/\tau}(\sqrt{F_i}/\tau)$ as a function of $\eta\tau$, and the lower plots show the corresponding optimal sensing times, $(t/\tau)_{\text{opt}}$. Comparison of schemes reveals that scheme B gives no advantage over scheme A for $\eta\tau \lesssim 0.5$. Scheme C , however, does better than scheme A for all values of $\eta\tau$, although the advantage vanishes as $\eta\tau \rightarrow 0$.

For extremely large measurement time or extremely strong squeezing ($\eta\tau \rightarrow \infty$) the preparation time t becomes negligible compared to τ and as previously mentioned, this results in sensitivities approaching the Heisenberg limit.

As discussed in section 3.5 there exists a natural transition from the fermionic N spin-1/2

system to the bosonic QHO system. Indeed, in the $N \rightarrow \infty$ limit it is possible to calculate an analytic expression for the sensitivity measure $\sqrt{F_B}/\tau$. Making use of (3.64), (3.65) and (3.68) we have

$$\lim_{N \rightarrow \infty} |\psi_B\rangle = \tilde{D}_\omega(t) \tilde{S}_\eta(t') |0\rangle \quad (6.8)$$

where we have used “tildes” to distinguish these bosonic operators (as given by (2.32) and (2.46)) from their spin operator counterparts. Using (2.48) we find

$$(\sqrt{\nu}\tau\delta\omega_B)^{-1} \leq \sqrt{F_B}/\tau = \frac{t}{\tau} e^{2\eta\tau(1-t/\tau)} \quad (6.9)$$

which we can optimise over t/τ and find, in the $N \rightarrow \infty$ limit, that if $\eta\tau > 0.5$

$$\max_{t/\tau} \left[\sqrt{F_B}/\tau \right] \rightarrow \frac{e^{2\eta\tau-1}}{2\eta\tau} \quad (6.10)$$

$$(t/\tau)_{\text{opt}} \rightarrow \frac{1}{2\eta\tau} \quad (6.11)$$

while for $\eta\tau \leq 0.5$ we find

$$\max_{t/\tau} \left[\sqrt{F_B}/\tau \right] \rightarrow 1 \quad (6.12)$$

$$(t/\tau)_{\text{opt}} \rightarrow 1 \quad (6.13)$$

as depicted in Fig. 6.4. Thus, in the bosonic limit, scheme B improves upon scheme A only if $\eta\tau > 0.5$. If this condition is not met then $(t/\tau)_{\text{opt}} \rightarrow 1$ and scheme B reduces to scheme A . These analytical results for the bosonic system are in broad agreement to the numerical results for the finite N spin-1/2 system.

6.1.3 Scheme C

In scheme B the probe state’s exposure to the magnetic field is compartmentalised into probe state preparation and sensing. Here we consider scheme C in which we aim to utilise the limited time resource τ more effectively by applying the magnetic field and preparation operations concurrently which is in fact representative of a more realistic sensing scenario where it is not possible to switch off the magnetic field during the state preparation stage of the protocol. In order to achieve this, the probe state is first subject to the unitary transformation

$$\hat{U}_{\omega,\eta}(t') \equiv \exp \left[-it'(\hat{H}_\omega + \hat{H}_\eta)/\hbar \right] \quad (6.14)$$

where,

$$\hat{H}_\omega + \hat{H}_\eta = \hbar\omega\hat{J}_y/\sqrt{N} + i\hbar\eta \left(\hat{J}_-^2 - \hat{J}_+^2 \right) / N \quad (6.15)$$

is the sum of the TAT and magnetic field Hamiltonians (6.6) and (6.4). This evolution occurs for a time t' then the TAT is stopped while the probe state continues to evolve under the magnetic field for time t so the final state is given by

$$|\psi_C\rangle = \hat{D}_\omega(t)\hat{U}_{\omega,\eta}(t')|\downarrow\rangle^{\otimes N} \quad (6.16)$$

where again, the total run time is limited to τ by constraining $t' = \tau - t$ so that in the limit $t' \rightarrow 0$, scheme C reduces to scheme A . The numerical plots for $\sqrt{F_C}/\tau$ are given in Fig.6.3 from which it is clear that scheme C consistently outperforms scheme A — even in parameter intervals where scheme B did not. Indeed it is apparent that the concurrent preparation and sensing scheme C outperforms the sequential scheme B and is a more efficient use of the resource τ . Moreover, by optimising over t/τ as shown in Fig.6.4, it is evident that scheme C outperforms both schemes A and $B \forall \eta\tau$ and the sensitivities for all three schemes converge to 1 as $\eta\tau \rightarrow 0$. Additionally, for scheme C it is clear that for small $\eta\tau$ we have $(t/\tau)_{\text{opt}} = 0$ meaning that in such a case the TAT and magnetic field should be applied concurrently throughout the protocol.

It is again feasible to find an analytic expression for the sensitivities in the bosonic limit, we find

$$\lim_{N \rightarrow \infty} |\psi_C\rangle = \tilde{D}_\omega(t)\tilde{U}_{\omega,\eta}(t')|0\rangle \quad (6.17)$$

then using the expansion (2.37) we find the bound on the sensitivity to be

$$(\sqrt{\nu}\tau\delta\omega_C)^{-1} \leq \sqrt{F_C}/\tau = \left(\frac{t}{\tau} + \frac{1}{2\eta\tau}\right) e^{2\eta\tau(1-t/\tau)} - \frac{1}{2\eta\tau} \quad (6.18)$$

which is plotted in Fig.6.3. Optimising over t/τ in the $N \rightarrow \infty$ limit, gives

$$\max_{t/\tau} [\sqrt{F_C}/\tau] \rightarrow \frac{1}{2\eta\tau} (e^{2\eta\tau} - 1) \quad (6.19)$$

$$(t/\tau)_{\text{opt}} \rightarrow 0 \quad (6.20)$$

as demonstrated in Fig.6.4. Furthermore, we can calculate the ratio of these quantities to reveal

$$\frac{\max_{t/\tau} [\sqrt{F_C}/\tau]}{\max_{t/\tau} [\sqrt{F_B}/\tau]} = \begin{cases} e(1 - e^{-2\eta\tau}) & \text{for } \eta\tau > 0.5 \\ \frac{1}{2\eta\tau}(e^{2\eta\tau} - 1) & \text{for } \eta\tau \leq 0.5 \end{cases} \quad (6.21)$$

thus,

$$\frac{\max_{t/\tau} [\sqrt{F_C}/\tau]}{\max_{t/\tau} [\sqrt{F_B}/\tau]} \geq 1, \forall \eta\tau \quad (6.22)$$

from which we infer that in the bosonic limit, scheme C performs just as well, or outperforms scheme B for all values of $\eta\tau$. Moreover, the optimal strategy is again found to have the TAT and the magnetic field operating concurrently throughout the protocol which again, is consistent with our results for finite N . Additionally, the largest possible improvement (as $\eta\tau \rightarrow \infty$) is found to be

$$\frac{\max_{t/\tau} [\sqrt{F_C}/\tau]}{\max_{t/\tau} [\sqrt{F_B}/\tau]} = e \approx 2.7. \quad (6.23)$$

This protocol (in the bosonic limit) has been investigated further to include an alternative method of analytically proving that concurrent state preparation and readout outperforms the analogous sequential scheme (see Appendix.B), and furthermore we analyse the effects of including optical loss for which we find that scheme C persists to yield superior performance than that of schemes A and B for all varying parameters.

6.2 Magnetic Field Sensing and One-Axis Twisting

In the preceding section we demonstrated how concurrent state preparation via TAT and sensing of a magnetic field can outperform traditional sequential sensing schemes. However, as mentioned in section 3.4.2, TAT is challenging to implement in practise and furthermore, optimal readout strategies were assumed which may also be impractical. Additionally, the readouts were assumed to be preformed over a negligible time interval. Here we introduce and investigate modified versions of schemes B and C motivated by practicality and denote these as schemes B' and C' as represented by Fig.6.5. In schemes B'

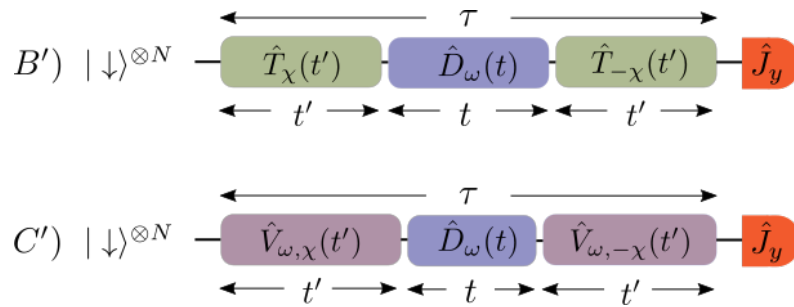


Figure 6.5: In scheme B' , the one-axis twisting operation $\hat{T}_\chi(t')$ generates a spin squeezed state before exposure to the magnetic field through $\hat{D}_\omega(t)$. The “echo” (anti-squeezing) operation $\hat{T}_\chi^\dagger(t') = \hat{T}_{-\chi}(t')$ (as introduced in section 6.2.1) is applied before the final measurement. In scheme C' the spins are exposed to the magnetic field during the OAT and echo operations. For a fair comparison, each protocol is constrained by the time τ .

and C' , OAT is used to generate entanglement in the probe state and as discussed in section

3.4.1 this is a practically viable approach. Moreover, a readout stage is incorporated which is motivated by the work of Davis *et.al* [107] which uses an “echo” readout method. An echo readout is the time reversed (inverse) of the entangling preparation operation which can simplify the final measurement [108], can overcome strict requirements on the resolution of the final measurement [109, 110] and have been demonstrated in recent experiments [107, 111]. We further these echo measurement studies by investigating how sensitivity gains are affected when a limited time resource must be divided between non-negligible state preparation and readout times and the sensing. Restrictions on time as a limited resource could be enforced, for example, by decoherence, by the stability of our equipment or by the fact that the quantity we want to measure is rapidly changing.

6.2.1 Scheme B'

The initial “unprepared” state is again taken to be the CSS $|\downarrow\rangle^{\otimes N}$ which is subject to an OAT operation (as introduced in section 3.4.1) for time t'

$$\hat{T}_\chi(t') = \exp \left[-it' \hat{H}_\chi / \hbar \right], \text{ where } \hat{H}_\chi = \hbar \chi \hat{J}_x^2 / N \quad (6.24)$$

resulting in a spin squeezed state. A Bloch sphere representation of the OAT operation is given in Fig.6.6. The state is then exposed to the magnetic field for time t which amounts to a rotation of the state about the y -axis by $\hat{D}_\omega(t) = \exp \left[-it \hat{H}_\omega / \hbar \right]$. Finally, the inverse of the state preparation OAT operation is applied for time t' , i.e the operator $\hat{T}_\chi^\dagger(t') = \hat{T}_\chi(t')$ is applied for readout. The Final state is then given by

$$|\psi_{B'}\rangle = \hat{T}_{-\chi}(t') \hat{D}_\omega(t) \hat{T}_\chi(t') |\downarrow\rangle^{\otimes N} \quad (6.25)$$

where we impose the constraint $t' = (\tau - t)/2$ in order to limit the run time of this scheme to time τ . After the echo operation the (collective) observable \hat{J}_y is measured then by the propagation of error formula, the error in the estimate of the small scaled magnetic field ω is

$$\delta\omega_{B'} = \frac{1}{\sqrt{\nu}} \left| \frac{\Delta \hat{J}_y}{\partial_\omega \langle \hat{J}_y \rangle} \right|_{\omega=0} \quad (6.26)$$

where, the standard deviation of \hat{J}_y in the state $|\psi_{B'}\rangle$ is $|\Delta \hat{J}_y|_{\omega=0} = \sqrt{N}/2$. In order to calculate the denominator of (6.26), following the derivation given in [107], we first rewrite the expression in the form

$$|\partial_\omega \langle \hat{J}_y \rangle|_{\omega=0} = \left| \frac{it}{\sqrt{N}} \langle \downarrow |^{\otimes N} \left[\hat{T}_{-\chi}(t') \hat{J}_y \hat{T}_\chi(t'), \hat{J}_y \right] | \downarrow \rangle^{\otimes N} \right| \quad (6.27)$$

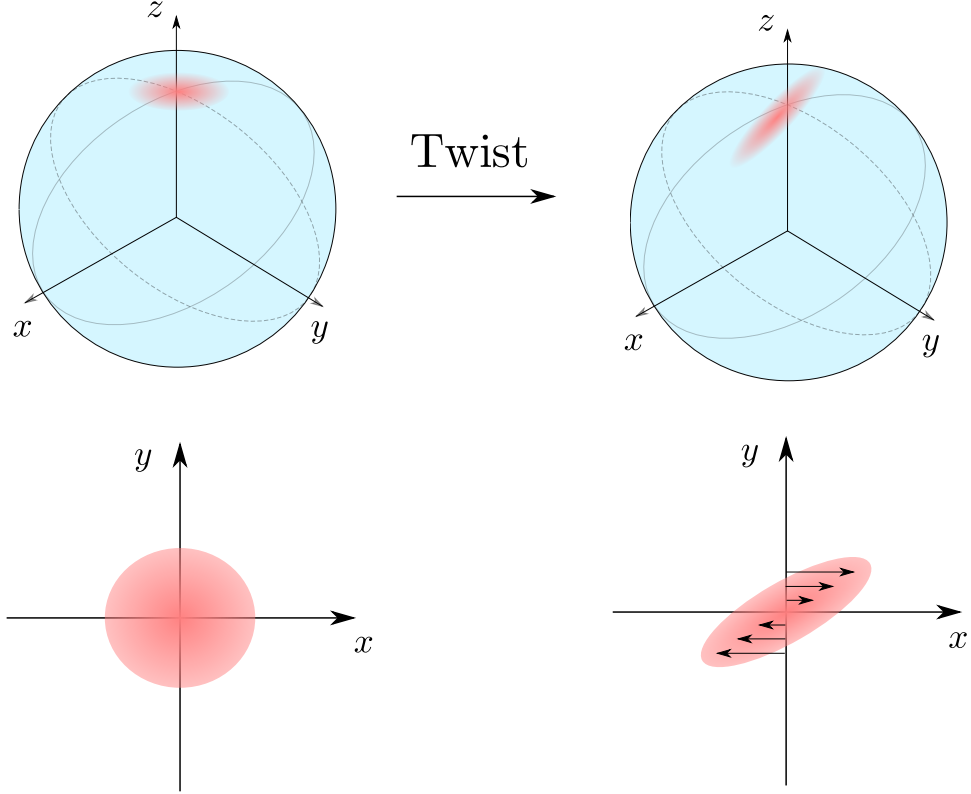


Figure 6.6: Bloch Sphere representation (with inverted z axis for ease of representation) of a OAT operation, as given by equation (6.24), on the CSS $|\downarrow\rangle^{\otimes N}$. In contrast to TAT, the resulting reduction in quantum fluctuations is not in line with any of the primary axes.

and express the operator in the commutator of (6.27) in the form

$$\begin{aligned} \hat{T}_{-\chi}(t')\hat{J}_y\hat{T}_{\chi}(t') &= e^{it'\chi\hat{J}_x^2/N} \left(-\frac{i}{2}\hat{J}_+ + \frac{i}{2}\hat{J}_- \right) e^{-it'\chi\hat{J}_x^2/N} \\ &= e^{-i\pi\hat{J}_y/2} \left(-\frac{i}{2}e^{it'\chi(2\hat{J}_z-1)/N}\hat{J}_+ + \frac{i}{2}\hat{J}_-e^{it'\chi(-2\hat{J}_z-1)/N} \right) e^{i\pi\hat{J}_y/2} \end{aligned} \quad (6.28)$$

then upon substitution of (6.28) into (6.27) we obtain a lengthy expression comprised of eight expectation values such as

$$\langle +|^{\otimes N} \hat{J}_-^2 e^{-2it'\chi\hat{J}_z} |+\rangle^{\otimes N} \quad (6.29)$$

where $|+\rangle^{\otimes N} = e^{i\pi\hat{J}_y/2} |\downarrow\rangle^{\otimes N}$. In order to evaluate this type of expectation value, we follow the method given in the appendix of [32] which takes the approach of differentiating the generating function. Indeed, with the intention of evaluating the example expectation value (6.29), it is shown that

$$\begin{aligned} X_A(\alpha, \beta, \gamma) &\equiv \langle +|^{\otimes N} e^{\gamma\hat{J}_-} e^{\beta\hat{J}_z} e^{\alpha\hat{J}_+} |+\rangle^{\otimes N} \\ &= \left[\frac{1}{2}e^{-\beta/2} + \frac{1}{2}e^{\beta/2}(\alpha+1)(\gamma+1) \right]^N \end{aligned} \quad (6.30)$$

with which we find

$$\begin{aligned} \langle + |^{\otimes N} \hat{J}_-^2 e^{-2it'\chi\hat{J}_z} | + \rangle^{\otimes N} &= \left[\frac{\partial^2}{\partial \gamma^2} X_A \right] \bigg|_{\substack{\alpha=\gamma=0 \\ \beta=-2it'\chi/N}} \\ &= \frac{N(N-1)}{4} \left[\cos \left(\frac{t'\chi}{N} \right) \right]^{N-2} e^{-2it'\chi/N}. \end{aligned} \quad (6.31)$$

Applying this method to all eight terms reveals that

$$|\partial_\omega \langle \hat{J}_y \rangle|_{\omega=0} = \frac{t\sqrt{N}(N-1)}{2} |\sin(\theta(t)) \cos^{N-2}(\theta(t))| \quad (6.32)$$

where $\theta = \chi\tau(1 - t/\tau)/(2N)$. Finally, the sensitivity is found to be

$$(\sqrt{\nu}\tau\delta\omega_{B'})^{-1} = (t/\tau)(N-1) |\sin(\theta(t)) \cos^{N-2}(\theta(t))| \quad (6.33)$$

which, similarly to the TAT schemes, depends only on three dimensionless variables $N, t/\tau$ and $\chi\tau$. Thus, plots of the sensitivity (6.33) against these parameters are given in Fig. 6.7. Furthermore, a numerical optimisation over t/τ is performed and plotted in Fig. 6.8 from which we infer similar behaviour to that of the TAT scheme B; for small $\chi\tau$, OAT does not ensure sensitivity gains over the classical scheme A. In such cases, the limited time resource is best used in the sensing stage. However, it is also evident from Fig. 6.8 that as N increases, the threshold value of $\chi\tau$ for precision gains over the classical scheme decreases; for $N = 10$, scheme B' outperforms scheme A for $\chi\tau \gtrsim 11.5$ whereas for $N = 100$ this value decreases to $\chi\tau \gtrsim 8.2$ from which we can extrapolate that for larger N , scheme B' can outperform scheme A for weaker squeezing strengths or for shorter run times. In the bosonic limit $N \rightarrow \infty$, using (6.33) it is straightforward to find that the sensitivity is

$$\lim_{N \rightarrow \infty} (\sqrt{\nu}\tau\delta\omega_{B'})^{-1} = \frac{\chi\tau(\tau - t)}{2\tau} \quad (6.34)$$

which we can analytically optimise over t/τ and find

$$\max_{t/\tau} [(\sqrt{\nu}\tau\delta\omega_{B'})^{-1}] \rightarrow \frac{\chi\tau}{8} \quad (6.35)$$

$$(t/\tau)_{\text{opt}} \rightarrow \frac{1}{2} \quad (6.36)$$

which reveals that in order for scheme B' to outperform the classical scheme A (which has sensitivity $(\sqrt{\nu}\tau\delta\omega_{B'})^{-1} = 1$) in the bosonic limit, we require $\chi\tau > 8$ otherwise OAT is not worthwhile. Furthermore, from equation (6.36) we can conclude that it is optimal to use half of the run time τ for sensing and a quarter each for preparation and readout i.e it is optimal to use half the run time τ distributed evenly over the echo measurement protocol.

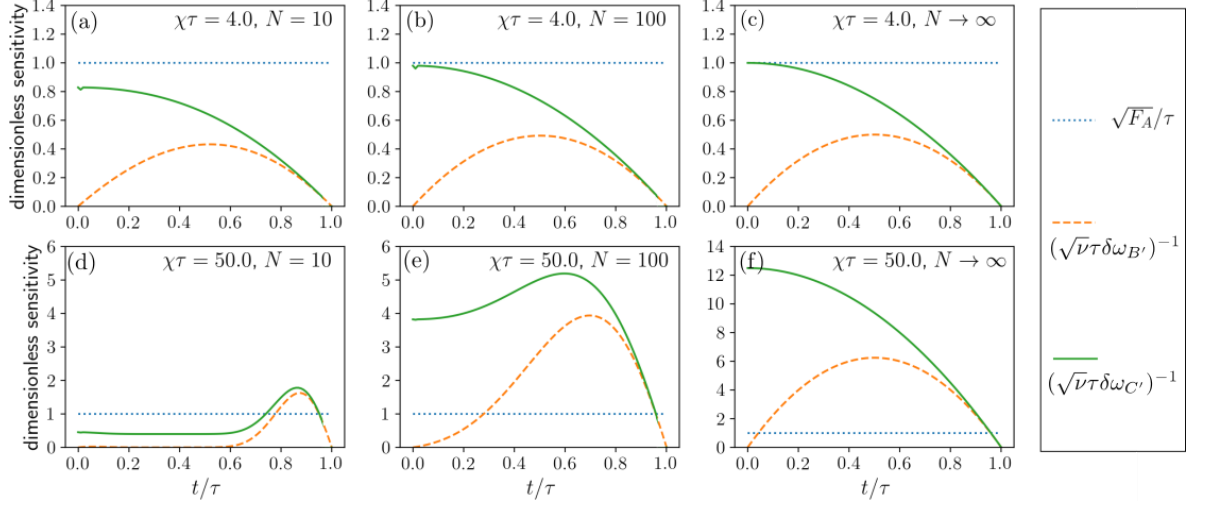


Figure 6.7: These plots show that for a sufficiently small value of $\chi\tau$ (e.g. $\chi\tau = 4$ and $N = 10$), scheme A gives a better sensitivity than scheme B' and scheme C' . For a sufficiently large value of $\chi\tau$ (e.g. $\chi\tau = 50$ and $N = 10$ or $N = 100$), both scheme B' and scheme C' give a better sensitivity than scheme A , i.e the spin squeezing is worthwhile.

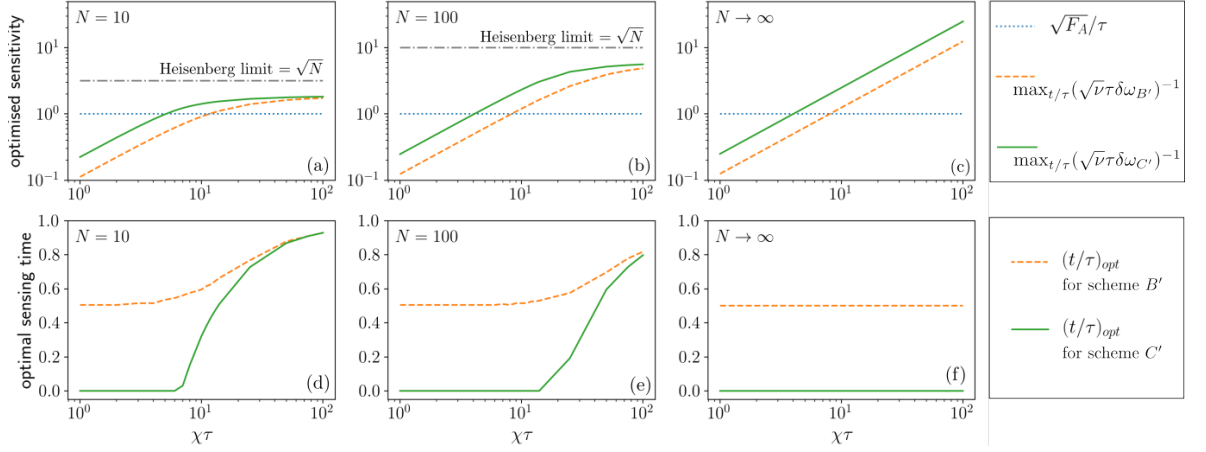


Figure 6.8: The upper plots show the optimised sensitivity bound $\max_{t/\tau} [(\sqrt{\nu}\tau\delta\omega)^{-1}]$ as a function of $\chi\tau$. These plots are optimised over time, but not over measurements, in contrast to Fig.6.3 and Fig.6.4 which are optimised over both. We choose the measurement in line with the scheme of Davis *et.al* [107]. Comparison of schemes reveals that when $N = 10$ scheme A outperforms scheme B' for $\chi\tau \lesssim 11.5$ and scheme C' for $\chi\tau \lesssim 5$. These threshold values decrease for larger N . For very large $\chi\tau$, the sensitivities of schemes B' and C' converge. The lower plots show the optimal sensing time $(t/\tau)_{\text{opt}}$ as a function of $\chi\tau$.

6.2.2 Scheme C'

Analogously to scheme C in which TAT is applied concurrently with sensing, we investigate the effects of extending the sensing time of schemes B' to be concurrent with the echo measurement protocol using OAT. The initial state evolves for time t' by the operation

$$\hat{V}_{\omega,\chi}(t') \equiv \exp \left[-it'(\hat{H}_\omega + \hat{H}_\chi)/\hbar \right] \quad (6.37)$$

where,

$$\hat{H}_\omega + \hat{H}_\chi = \hbar\omega\hat{J}_y/\sqrt{N} + \hbar\chi\hat{J}_x^2/N \quad (6.38)$$

after which the OAT is turned off and the squeezed spin state evolves under the magnetic field for a time t . Lastly, the echo readout is carried out by reversing the OAT component of the evolution but importantly the magnetic field component is not reversed. The final state is then given by

$$|\psi_{C'}\rangle = \hat{V}_{\omega,-\chi}(t')\hat{D}_\omega(t)\hat{V}_{\omega,\chi}(t')|\downarrow\rangle^{\otimes N} \quad (6.39)$$

where, once again the run time τ is constrained by $t' = (\tau - t)/2$. The sensitivity is again quantified by the propagation of error formula, as given by (6.26), however due to the complexities of calculating the numerator, the sensitivities for varying t/τ are calculated numerically and plotted in Fig.6.7. Optimising over t/τ , as given in Fig.6.8, reveals that scheme C' always outperforms scheme B' . Furthermore, it is evident from Fig.6.8 that scheme C' outperforms the classical scheme A for a wider range of $\chi\tau$ than scheme B' e.g for $N = 10$, scheme C' outperforms scheme A for $\chi\tau \gtrsim 5$ compared to $\chi\tau \gtrsim 11.5$ for that of scheme B' .

We now derive an analytic expression for the sensitivity in the bosonic limit. Firstly, we take the $N \rightarrow \infty$ of preparation operator (6.37) using the Holstein-Primakoff transformations

$$\begin{aligned} \tilde{V}_{\omega,\chi}(t') &\equiv \lim_{N \rightarrow \infty} \hat{V}_{\omega,\chi}(t') \\ &= \exp \left[t'\omega(\tilde{a} - \tilde{a}^\dagger) - it'\chi(\tilde{a} + \tilde{a}^\dagger)^2 \right] \end{aligned} \quad (6.40)$$

where we have used “tilde” to distinguish this bosonic operator from the spin operator (6.37). The final state of scheme C' in the bosonic limit is thus

$$\lim_{N \rightarrow \infty} |\psi_{C'}\rangle = \tilde{V}_{\omega,-\chi}(t')\tilde{D}(t)\tilde{V}_{\omega,\chi}(t')|0\rangle. \quad (6.41)$$

Defining $\hat{P} = -ia^\dagger + ia$, it is straightforward to find that the standard deviation of \hat{P} in the state (6.41) is $|\Delta\hat{P}|_{\omega=0} = 1$. Then by iterative use of (2.37) the expectation value is found to be

$$\langle\hat{P}\rangle = \frac{\chi\omega}{4}(\tau^2 - t^2) \quad (6.42)$$

and since we have

$$\lim_{N \rightarrow \infty} \delta\omega_{C'} = \frac{1}{\sqrt{\nu}} \left| \frac{\Delta\hat{P}}{\partial_\omega \langle\hat{P}\rangle} \right|_{\omega=0} \quad (6.43)$$

it becomes apparent that

$$\lim_{N \rightarrow \infty} (\sqrt{\nu}\tau\delta\omega_{C'})^{-1} = \frac{\chi\tau}{4}(1 - t^2/\tau^2) \quad (6.44)$$

then by optimising over t/τ we find

$$\max_{t/\tau} [(\sqrt{\nu}\tau\delta\omega_{C'})^{-1}] \rightarrow \frac{\chi\tau}{4} \quad (6.45)$$

$$(t/\tau)_{\text{opt}} \rightarrow 0. \quad (6.46)$$

Thus in the bosonic limit, scheme C' provides an advantage over scheme A for $\chi\tau > 4$. Furthermore, comparing this to (6.35), we see that in the bosonic limit, scheme C' provides a factor 2 improvement over scheme B' . Additionally, equation (6.46) implies that optimal sensitivity is attained when OAT is applied concurrently with the magnetic field for the entire run time τ .

Chapter 7

Entropy in Information Theory

In ordinary discourse, information is often equated with knowledge or meaning and is thus a property of a single message (or single event or single measurement outcome etc.). In information theory, information is not concerned with the individual messages and their associated content; the focus is on *all* messages a source can *possibly* send. The informative aspect of a single message is not its content or meaning, it is the fact that of *all the possible messages* that could have been sent, *that particular message* was received [112]. Entropy, in the context of information theory, is a measure of the unpredictability of information content. More formally, the entropy of a random variable X measures the amount of uncertainty about X *before* we learn its value. Equivalently, we can take the reverse view; the entropy of a random variable X quantifies how much information we gain about X , on average, *after* we learn its value. In this chapter we shall explore the notion of entropy in the aforementioned context, along with various types of entropy measures and how such measures can be applied in a physical framework.

7.1 Hartley and Shannon Entropy

7.1.1 Hartley Entropy

Entropy in information theory becomes a probabilistic concept. The first to introduce an entropic measure of information was Hartley (1928) [113] who posited that measurement of an element of a set of size N requires $\log_2(N)$ bits of information where one “bit” (the unit of information) is the information needed to distinguish an element of a pair. A simple example of this is the state of a coin, of which there are two possibilities - heads or tails, so to fully describe this system we need $\log_2(2) = 1$ bit of information. A key feature of this measure is the additivity property. Indeed, we demand that the information gained

from independent events (elements) must be the sum of the partial information of each event. If we consider the set E comprising the disjoint union of M lots of the N -tuples E_1, E_2, \dots, E_M we then need $\log_2(M)$ bits of information to determine which of the $E_{i\text{th}}$ sets the element belongs to and a further $\log_2(N)$ bits to determine which element of the E_i set is the considered one. This is encapsulated by Hartley's measure since

$$\log_2(NM) = \log_2(N) + \log_2(M). \quad (7.1)$$

This measure holds if the events we wish to describe are equiprobable, but if we have information on the probabilities of the events i.e if we have knowledge about the likelihood of specific outcomes, then an adjustment can be made to incorporate this into a more accurate model for prediction.

7.1.2 Shannon Entropy

Shannon's aim was to characterize the information gained from an event based on the probability of the event occurring [114, 115]. To see how this was achieved, let E be the disjoint union of the sets E_1, \dots, E_n with N_1, \dots, N_n elements respectively. Suppose that each element of E is equiprobable and we are only interested in finding the $E_{k\text{th}}$ subset in which a given element of E belongs. The information needed to specify a particular element of E is given in two parts; we must first identify the subset E_k which contains the element and second, locate the element within the subset which, by Hartley, requires $\log_2(N_k)$ bits of information. Taking $\sum_{k=1}^n N_k = N$ we also have (again by Hartley) that $\log_2(N)$ bits of information are needed to specify an element of E . Thus by the additivity principle, we have

$$\log_2(N) = I_k + \log_2(N_k) \quad (7.2)$$

where I_k is the amount of information needed to specify the $E_{k\text{th}}$ subset. Hence,

$$I_k = \log_2 \left(\frac{N}{N_k} \right) \quad (7.3)$$

and with this, it is plausible to define the information needed to specify the $E_{k\text{th}}$ set which the element belongs to as the weighted average of I_k

$$I = \sum_{k=1}^n \frac{N_k}{N} I_k \quad (7.4)$$

where $\frac{N_k}{N} \equiv p_k$ is the probability that the element belongs to the subset E_k . Thus we have

$$I = \sum_{k=1}^n p_k \log_2 \left(\frac{1}{p_k} \right) \quad (7.5)$$

which is the *entropy* of the probability distribution $P = (p_1, \dots, p_n)$ or as it's better known, *Shannon's entropy*. So according to Shannon, to characterise an element of the set E where the associated probabilities of its elements are (p_1, \dots, p_n) (under the condition $\sum p_k = 1$) we need $I = \sum_{k=1}^n p_k \log_2 \left(\frac{1}{p_k} \right)$ bits of information. Further understanding of this measure can be gained by applying it to physical quantum mechanical measurements of a random observable where the probability of the k th measurement outcome is given by p_k , this will be our avenue of investigation (predominantly for continuous random variables, for which the notion of Shannon entropy is generalised as given in the following section).

Formally, Shannon's entropy is given for $n \in \mathbb{N}$ and V_m the set of all probability distributions $P = (p_1, \dots, p_n) \equiv (p(m_1), \dots, p(m_n))$ on (m_1, \dots, m_n) as a measure of information given by a function $I(P) : V_m \mapsto \mathbb{R}$ which satisfies the following axioms

1. Continuity: $I(P)$ is continuous in all of its arguments
2. Additivity: The information gained from two independent experiments is the sum of the information from the experiments
3. Monotonicity: For uniform distributions the information increases with n i.e. for $P_U = (1/n, \dots, 1/n)$ and $Q_U = (1/k, \dots, 1/k)$ with $k, n \in \mathbb{N}$ we have $k > n \Rightarrow I(Q_U) > I(P_U)$
4. Branching: The measure of information is independent of how the process is divided into parts. That is, for (p_1, \dots, p_n) , $n > 3$, divide $m = \{m_1, \dots, m_n\}$ into two blocks $A = (m_1, \dots, m_s)$ and $B = (m_{s+1}, \dots, m_n)$, and let $p_A = \sum_{k=1}^s p_k$ and $p_B = \sum_{k=s+1}^n p_k$. Then $I(P) = I(p_A, p_B) + p_A I\left(\frac{p_1}{p_A}, \dots, \frac{p_s}{p_A}\right) + p_B I\left(\frac{p_{s+1}}{p_B}, \dots, \frac{p_n}{p_B}\right)$
5. Bit Normalisation (Convention): The average information gain for two equally likely messages is one "bit" i.e $I(1/2, 1/2) = 1$.

It is apparent that any change towards the equalisation of the probabilities p_1, \dots, p_k leads to an increase in the entropy $I(P)$, reaching a maximum for the uniform distribution $p_1 = \dots = p_n = 1/n$. Shannon went further and asked: what is the minimal amount of physical resources required to store the information being produced by the source on the condition that the information can be reconstructed at a later time? The result, known as Shannon's noiseless coding theorem [114], turns out to be the entropy $I(P)$, that is $I(P)$ bits of information are required per message from the source - a much celebrated result of notable importance. We now introduce an useful entropic measure that will be elaborated

upon in the following section. For now, let p_i and q_i be two probability distributions of the same index i , then the *relative entropy* of the distributions is given by

$$\begin{aligned} I(p_i||q_i) &= \sum_i p_i \log_2 \left(\frac{p_i}{q_i} \right) \\ &= -I(X) - \sum_i \log_2(q_i). \end{aligned} \quad (7.6)$$

This gives an entropy-like measure of the closeness of the two probability distributions. It can be shown that $I(p_i||q_i) \geq 0$ with equality iff $p_i = q_i$ [116]. The relative entropy provides a fundamental building block for other entropic quantities e.g by the non-negativity of the relative entropy it can be shown that for a random variable X with d outcomes, $I(X) \leq \log_2(d)$ which yields the subadditivity of the Shannon entropy;

$$I(X, Y) \leq I(X) + I(Y) \quad (7.7)$$

where

$$I(X, Y) = - \sum_{i,j} p_{i,j} \log_2(p_{i,j}) \quad (7.8)$$

is the *joint entropy* which is a measure of our total uncertainty about the pair of random variable (X, Y) and $p_{i,j} = p_i p_j$ is the joint probability distribution. Considering the two random variables X and Y , we ask how the information content of X is related to the information content of Y ? To answer this we turn to two entropic quantities referred to as the *conditional entropy* and the *mutual entropy*. Suppose we have acquired $I(Y)$ bits of information about the pair (X, Y) , the uncertainty we are left with about this pair is associated with a lack of knowledge of X given that we know Y (which, in general, shares some information with X). The entropy of X , conditional on knowing Y , is given by

$$I(X|Y) = I(X, Y) - I(Y). \quad (7.9)$$

The mutual information of X and Y provides a measure of the amount of information X and Y have in common and is intuitively defined by

$$I(X : Y) = I(X) + I(Y) - I(X, Y) \quad (7.10)$$

there are a myriad of relationships between Shannon's entropic quantities that give insight into overall functionality of entropy and a lot of this understanding can be heuristically encapsulated by the "entropy Venn diagram" (see Fig.7.1). Some final results on these entropic measures:

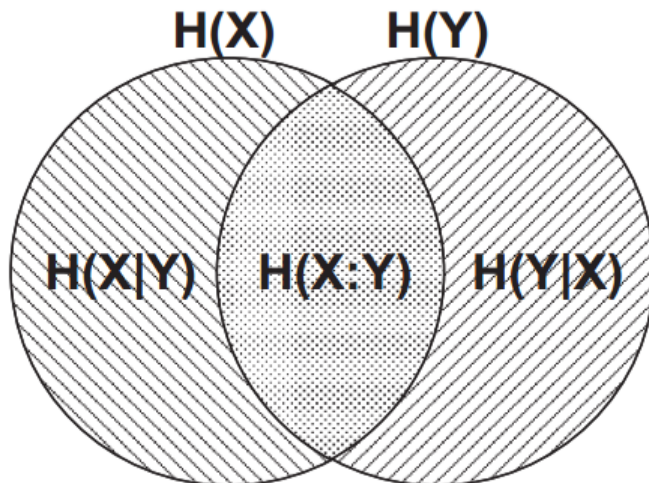


Figure 7.1: [116] The “entropy Venn diagram”, with the discrete Shannon entropy of the random variable X denoted by $H(X)$, allows the heuristic deduction of many of the entropy inequalities but is not a completely reliable guide.

- The *chaining rule for conditional entropies* states for any set of random variables X_1, \dots, X_n and Y , then $I(X_1, \dots, X_n|Y) = \sum_{i=1}^n I(X_i|Y, X_1, \dots, X_{i-1})$
- Mutual information is not always subadditive: $I(X, Y : Z) \not\leq I(X : Z) + I(Y : Z)$
- Mutual information is not always superadditive: $I(X_1 : Y_1) + I(X_2 : Y_2) \not\geq I(X_1, X_2 : Y_1, Y_2)$

7.2 Differential Entropy

The *differential entropy* is a generalisation of the discrete Shannon entropy to what one can consider to be the continuous Shannon entropy; the (linear) entropy of a continuous random variable [117]. We must first clarify what it means for a random variable to be continuous. Let χ be a random variable with a cumulative distribution function $F(x) = \int_{-\infty}^x p(u)du = \Pr(\chi \leq x)$. If $F(x)$ is continuous, the random variable is said to be continuous. If $\int_{-\infty}^{\infty} p(x) = 1$, then $p(x)$ is called the *probability density function* (PDF) for χ . The set where $p(x) > 0$ is called the *support set* of χ . With this, the differential entropy $H(\chi)$ of a continuous random variable χ with a PDF $p(x)$ is defined as

$$H(\chi) = - \int_S p(x) \log_2(p(x)) dx \quad (7.11)$$

where S is the support set of the random variable χ . It is noted that as with every definition involving integrals, we implicitly include the statement “if it exists”. This applies to

both the density function and the entropy integral itself. $H(\chi)$ is often written as $H(p(x))$, a function of probability density since that is where the dependency of the function lies (rather than the random variable). $H(p(x))$ reaches its maximum for a given support set, $\log_2(b-a)$, for the uniform distribution $p(x) = 1/(b-a)$ in $[a, b]$ and zero elsewhere. So, intuitively, every change towards the equalisation of the PDFs $p(x)$ yields an increase in the differential entropy. Note that for $b-a < 1 \Rightarrow \log_2(b-a) < 0$ thus, unlike the discrete Shannon entropy, the continuous Shannon entropy can be negative. The volume of the support set (i.e the continuous analogue of cardinality for the discrete support set) is given by $2^{H(p(x))} = 2^{\log_2(b-a)} = b-a > 0$ as expected. One can interpret the differential entropy as the logarithm of the equivalent side length of the smallest set that contains most of the probability. Hence, a low entropy implies that the random variable is confined to a small effective volume and a high entropy indicates a widely dispersed random variable.

There is an important difference between the discrete and continuous Shannon entropies; the discrete Shannon entropy is uniquely determined by the probability measure over the message, whereas the continuous Shannon entropy's value is relative to the coordinate system. In other words, the continuous Shannon entropy is coordinate dependent and since information measures should not depend on the way we choose to describe a situation [112], extra care must be taken when using the differential entropy. Indeed, if we change from (x_1, \dots, x_n) to (y_1, \dots, y_n) the new entropy is given by

$$H(y) = \int \dots \int p(x_1 \dots x_n) J\left(\frac{x}{y}\right) \log_2 \left(p(x_1 \dots x_n) J\left(\frac{x}{y}\right) \right) dy_1 \dots dy_n \quad (7.12)$$

where $J\left(\frac{x}{y}\right)$ is the Jacobian of the coordinate transformation. Then changing variables to $(x_1 \dots x_n)$ we find

$$H(y) = H(x) - \int \dots \int p(x_1, \dots, x_n) \log_2 \left(J\left(\frac{x}{y}\right) \right) dx_1 \dots dx_n. \quad (7.13)$$

In the continuous case the entropy can be considered a measure of randomness relative to an assumed standard, namely the coordinate system chosen with each small element dx_1, \dots, dx_n given equal weight. However, *entropy differences* are coordinate independent in the continuous case and thus provide a measure for differences in information. For random variables χ, \mathcal{Y} with joint PDF $p(x, y)$, we can define the *conditional differential entropy* $H(\chi|\mathcal{Y})$ to be

$$\begin{aligned} H(\chi|\mathcal{Y}) &= - \int p(x, y) \log_2(p(x|y)) dx dy \\ &= H(\chi, \mathcal{Y}) - H(\mathcal{Y}) \end{aligned} \quad (7.14)$$

as $p(x|y) = p(x, y)/p(y)$. The *relative entropy* $D(p||q)$ between the two PDFs p and q is defined by

$$D(p||q) = \int p \log_2 \left(\frac{p}{q} \right) \quad (7.15)$$

noting that $D(p||q)$ is finite iff the support set of p is contained within q . The *mutual information* $I(\chi; \mathcal{Y})$ between two random variables with joint PDF $p(x, y)$ is defined by

$$\begin{aligned} I(\chi; \mathcal{Y}) &= \int p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) dx dy \\ &= D(p(x, y) || p(x)p(y)) \\ &= H(\chi) - H(\chi|\mathcal{Y}) \\ &= H(\mathcal{Y}) - H(\mathcal{Y}|\chi). \end{aligned} \quad (7.16)$$

It can be shown that the properties of the relative and mutual information for the continuous case are the same as for the discrete case. The relationship between continuous and discrete entropy can be established via the following; consider a random variable χ with a range divided into increments of length Δ and with a PDF $p(x)$ (see Fig. 7.2). Assuming the PDF is continuous within each increment, the mean value theorem yields a value x_i within each increment such that

$$p(x_i)\Delta = \int_{i\Delta}^{(i+1)\Delta} p(x) dx. \quad (7.17)$$

Introducing the quantised random variable

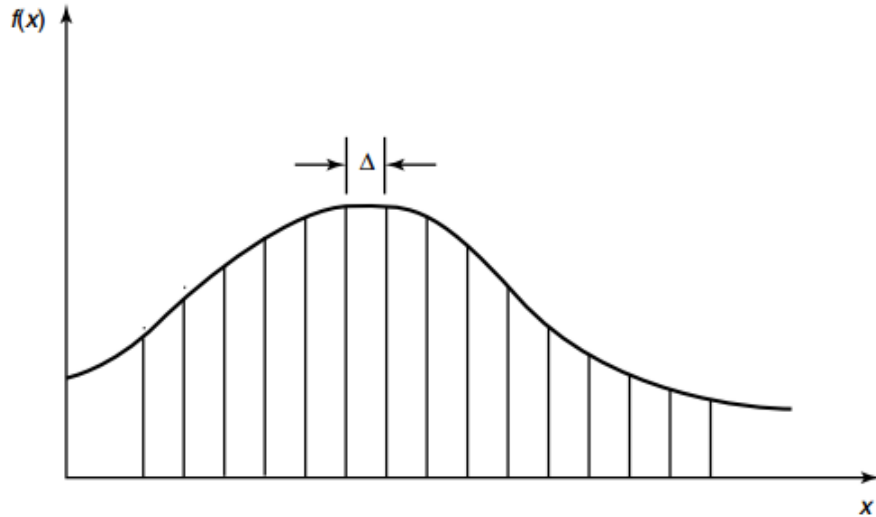


Figure 7.2: [117] Quantisation of a continuous random variable.

$$\chi^\Delta \equiv x_i, \quad \text{if } i\Delta \leq \chi < (i+1)\Delta \quad (7.18)$$

then the probability that $\chi^\Delta = x_i$ is

$$P_i = \int_{i\Delta}^{(i+1)\Delta} p(x)dx = p(x_i)\Delta. \quad (7.19)$$

The entropy of the quantised version is given by

$$\begin{aligned} I(\chi^\Delta) &= - \sum_{-\infty}^{\infty} P_i \log_2(P_i) \\ &= - \sum \Delta p(x_i) \log_2(p(x_i)) - \sum p(x_i)\Delta \log_2(\Delta) \\ &= - \sum \Delta p(x_i) \log_2(p(x_i)) - \log_2(\Delta) \end{aligned} \quad (7.20)$$

noting that $\sum p(x_i)\Delta = \int p(x) = 1$. The first term in 7.20 then approaches $-\int p(x) \log_2(p(x))$ as $\Delta \rightarrow 0$ by the definition of Riemann integrability (iff the integral is well defined). It then follows that

$$I(\chi^\Delta) + \log_2(\Delta) \rightarrow H(P), \quad \text{as } \Delta \rightarrow 0 \quad (7.21)$$

i.e the entropy of an n -bit quantisation of a continuous random variable is approximately $H(\chi) + n$.

7.3 Rényi Entropy

7.3.1 Discrete Rényi Entropy

Rényi's view of the formulation was slightly distilled in the sense that he considered (7.5) attainable through two (main) postulates; additivity and linear averaging of the information measure [118]. Additivity demands the total information received from the outcomes of two independent events is equal to the sum of the two partial events thus an information measure must satisfy

$$\mathcal{I}(pq) = \mathcal{I}(p) + \mathcal{I}(q) \quad (7.22)$$

assuming further that $\mathcal{I}(1/2) = 1$ and $\mathcal{I}(p)$ is monotonic then it follows that

$$\mathcal{I}(p) = \log_2 \left(\frac{1}{p} \right). \quad (7.23)$$

The postulate of linear averaging implies that the total amount of information received will be the average of the individual pieces of information received, weighted by the probabilities of their occurrences. More formally, for the possible measurement outcomes A_1, \dots, A_n

with the respective probabilities $P = (p_1, \dots, p_n)$ where outcome A_k yields \mathcal{I}_k bits of information then, on average, we receive

$$\mathcal{I}(P, \mathbb{I}) = \sum_{k=1}^n p_k \mathcal{I}_k \quad (7.24)$$

bits of information, where $\mathbb{I} = (\mathcal{I}_1, \dots, \mathcal{I}_n)$. Note that (7.24) is independent of any outcome A_k . It is then obvious that (7.5) is a consequence of these two postulates. It is well known that the linear mean is a widely used way of averaging but not the only way. Rényi's insight was to employ the general theory of means to the information measure. The general theory of means [119] states that a mean of the real numbers with weights p_1, \dots, p_n (such that $p_k > 0$, $\sum_{k=1}^n p_k = 1$) is an expression of the form

$$\varphi^{-1} \left(\sum_{k=1}^n \varphi(x_k) p_k \right) \quad (7.25)$$

where, $\varphi(x)$ is an arbitrary, strictly monotone function on the reals. Following the formulation of the Shannon entropy but using (7.25) as opposed to the linear mean then the amount of information associated with the probability distribution $P = (p_1, \dots, p_n)$ is then

$$\mathcal{I}(p_k) = \varphi^{-1} \left(\sum_{k=1}^n \varphi \left(\log_2 \left(\frac{1}{p_k} \right) \right) p_k \right) \quad (7.26)$$

and similarly

$$\mathcal{I}(P, \mathbb{I}) = \varphi^{-1} \left(\sum_{k=1}^n \varphi(\mathcal{I}_k) p_k \right) \quad (7.27)$$

This gives rise to the question; does an arbitrary choice of $\varphi(x)$ actually produce a reasonable measure of information that adheres to the postulate of additivity? Indeed, the postulate places a critical restriction on the choice of φ . If $\varphi(x)$ is linear then (7.26) and (7.27) reduce to (7.5) and (7.24) respectively. The only other functions satisfying this are the exponential functions. To gain some insight on this we consider \mathcal{E} to be the union of the two independent events $\mathcal{E}_1, \mathcal{E}_2$ from which we obtain \mathcal{I}_h and \mathcal{J}_k bits of information with probabilities p_h and q_k respectively. Thus the total information received is $\mathcal{I}_h + \mathcal{J}_k$ with probability $p_h q_k$ ($h = 1, \dots, m ; k = 1, \dots, n$). Taking the average amount of information obtained from the union of the two events \mathcal{E} to be the sum of the average amounts of information received from the two independent events, then (7.27) yields

$$\varphi^{-1} \left(\sum_{h=1}^m \sum_{k=1}^n p_h q_k \varphi(\mathcal{I}_h + \mathcal{J}_k) \right) = \varphi^{-1} \left(\sum_{h=1}^m p_h \varphi(\mathcal{I}_h) \right) + \varphi^{-1} \left(\sum_{k=1}^n q_k \varphi(\mathcal{J}_k) \right). \quad (7.28)$$

This must hold for arbitrary, finite, discrete probability distributions $P = \{p_h\}$ and $Q = \{q_k\}$ with arbitrary associated numbers \mathcal{I}_h and \mathcal{J}_k . If we choose $\mathcal{J}_k = \mathcal{J}$ independently of

k then from (7.25) we have

$$\varphi^{-1} \left(\sum_{h=1}^m p_h \varphi(\mathcal{I}_h + \mathcal{J}) \right) = \varphi^{-1} \left(\sum_{h=1}^m p_h \varphi(\mathcal{I}_h) \right) + \mathcal{J}. \quad (7.29)$$

Then from the theory of means [119] we have that (7.29) can hold only for linear or exponential functions, where the former results in Shannon's entropy. Let us then take $\varphi(x)$ to be an exponential function of the form $\varphi(x) = 2^{1-\alpha}x$ where $\alpha \neq 1$. From this we obtain

$$\mathcal{I}_\alpha(P) = \frac{1}{1-\alpha} \log_2 \left(\sum_{k=1}^n p_k^\alpha \right). \quad (7.30)$$

This can be considered a measure of information for $\alpha \neq 1$. However, for $\alpha < 0$ this measure will become infinite as $p_1 \rightarrow 0$ i.e. this measure becomes overly sensitive to small probabilities. Furthermore, the $\alpha = 0$ case must be excluded as it yields an expression that is independent of the probability distribution ($\mathcal{I}_0(P) = \log_2(n)$, dependent only on the number of events). Thus for $\alpha \in \mathbb{R}^+ \setminus \{0, 1\}$, $\mathcal{I}_\alpha(P)$ is a measure of information of order α with the probability distribution $P = (p_1, \dots, p_n)$. This is how Rényi referred to the measure (7.30), now it is more commonly referred to as the *Rényi entropy*. Note that

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \mathcal{I}_\alpha(P) &= \lim_{\alpha \rightarrow 1} \frac{\frac{d}{d\alpha} \log_2 \left(\sum_{k=1}^n p_k^\alpha \right)}{\frac{d}{d\alpha} (1-\alpha)} \\ &= \left(\sum_{k=1}^n \log_2(p_k) p_k^\alpha \right) \left(\sum_{k=1}^n p_k^\alpha \right)^{-1} \Big|_{\alpha=1} \cdot (-1)^{-1} \\ &= \sum_{k=1}^n p_k \log_2 \left(\frac{1}{p_k} \right) \\ &= I \end{aligned} \quad (7.31)$$

hence, Rényi's entropy is a generalisation of Shannon's entropy (7.5).

7.3.2 Geometric Interpretation of the Discrete Rényi Entropy

Probability mass functions (PMFs) can be visualized geometrically as points in a vector space known as the simplex with the axis given by the normalised probabilities [120]. For an n -dimensional random variable, the simplex Δ_n consists of all possible probability distributions;

$$\Delta_n = \left\{ p = (p_1, \dots, p_n)^T \in \mathbb{R}^n, p_i \geq 0, \sum_i p_i = 1, \forall i \right\}. \quad (7.32)$$

To illustrate this, consider the 3-dimensional simplex for the variables (x, y, z) . The space of all such distributions is a tetrahedron with vertices $(1, 0, 0), (0, 1, 0), (0, 0, 1)$. This is

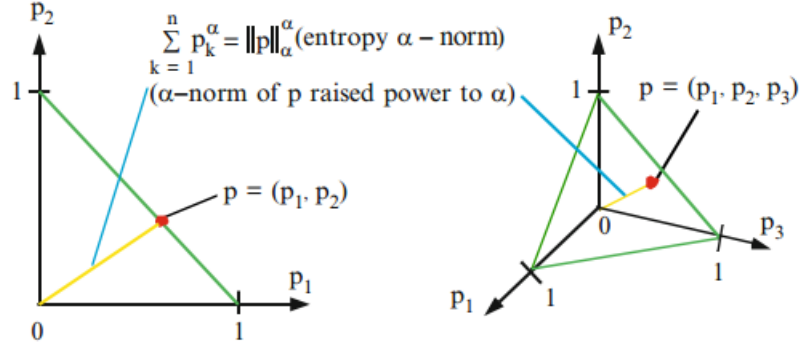


Figure 7.3: [120] The 2- and 3-dimensional representations of the probability simplex with the entropy α -norm of arbitrary point (or PMF) p .

demonstrated in Fig.7.3 along side the $n = 2$ -dimensional case. Any point in the simplex represents a different PMF and thus each PMF can be characterised by it's distance to the origin. Then defining

$$\begin{aligned} \|p(x)\|_\alpha &= \left(\sum_{k=1}^n p_k^\alpha \right)^{\frac{1}{\alpha}} \\ &= \sqrt[\alpha]{V_\alpha(X)} \end{aligned} \quad (7.33)$$

which implies

$$V_\alpha(X) = \|p(x)\|_\alpha^\alpha. \quad (7.34)$$

Recalling that the p -norm of an m -dimensional vector \vec{x} is given by

$$\|\vec{x}\|_p = \left(\sum_{i=1}^m |x_i|^p \right)^{\frac{1}{p}}. \quad (7.35)$$

we have that the α -information potential $V_\alpha(x)$ can be interpreted as the α power of the PMF α -norm. More formally, the discrete Rényi entropy takes the $\alpha - 1$ root of $V_\alpha(x)$ and rescales it by the logarithm. From this it becomes apparent that the role of α is to specify the norm to measure the distance of $p(x)$ to the origin within the simplex. Then from the theory of norms [121] it is evident that α is in fact altering the importance of the small values versus the large values in the set. With this we can interpret the Shannon entropy ($\alpha \rightarrow 1$) as the functional value of the 1-norm of the probability density. In fact, the 1-norm of any probability density is, by definition, always 1. As Rényi's entropy is a scalar that characterises densities, it is useful, in order to gain further insight, to display the contours of equal Rényi entropy in the simplex for several α . In Fig.7.4 the isoentropy

contours are plotted as a function of α . Notice that as α get closer to zero, the values inside the simplex exhibit little change. For higher values of α we see that the contours have rotated by 180 degrees and as α increases the changes with respect to the origin are emphasised. For $\alpha > 1$, the Rényi entropies are monotonically decreasing functions of the information potential V_α and for $\alpha \leq 1$, Rényi's entropies are monotonically increasing functions of V_α . In other words, the entropy maximisation is equivalent to maximisation of the information potential and entropy minimisation is equivalent to information potential minimisation.

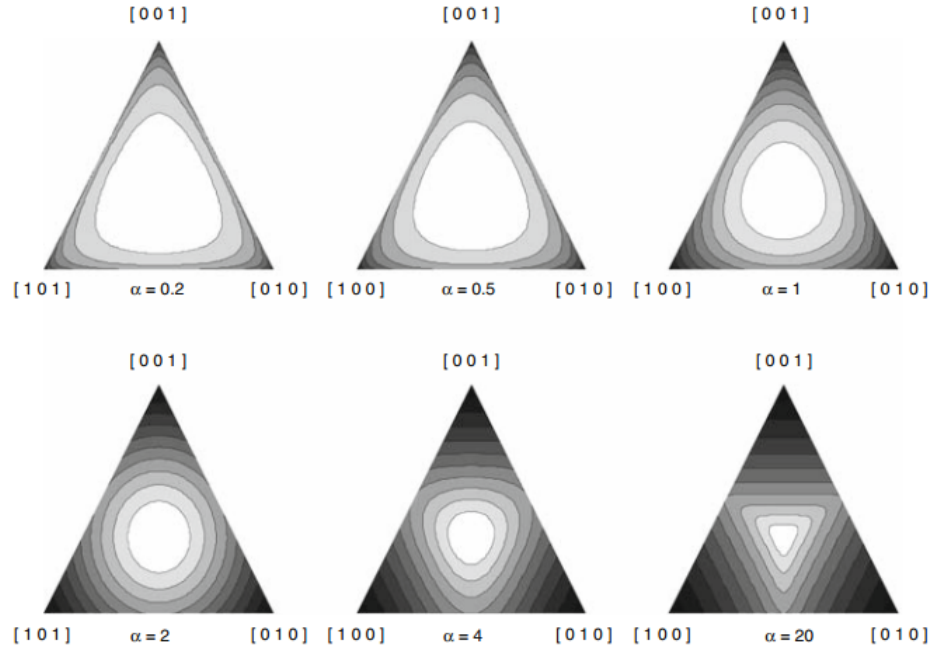


Figure 7.4: [120] The 3-dimensional probability simplex isoentropy contours (contours of equal Rényi entropy) for various values of α .

7.3.3 Continuous Rényi Entropy

Similarly to the Shannon entropy, it is possible to define the information measure of order α for the case of continuous distributions confined to a finite interval on \mathbb{R} . Let $f(x)$ be an arbitrary positive density function on the interval $[a, b]$, let $F_{nk} = \int_{k/n}^{(k+1)/n} f(x)dx$ and furthermore, let us denote the information quantity of order p for the corresponding discrete distribution $\mathcal{F}_n = \{F_{nk}\}$ by $\mathcal{I}(\mathcal{F}_n)$ then it follows that

$$\mathcal{I}_p(\mathcal{F}_n) = \frac{1}{1-p} \log_2 \left(\sum_{k=1}^n F_{nk}^p \right). \quad (7.36)$$

Just as for $p = 1$ [118], it can be easily shown that if the integral $\int_a^b f(x)dx$ exists (as it does for $0 < p < 1$) then for $p > 0$

$$\mathcal{I}_p(f(x)) \equiv \lim_{n \rightarrow \infty} (\mathcal{I}_p(\mathcal{F}_n) - \log_2(n)) = \frac{1}{1-p} \log_2 \left(\int_a^b f(x)^p dx \right) \quad (7.37)$$

the proof of which can be found in Ref.[118]. Since $\log_2(n)$ represents the entropy of the uniform distribution, we can interpret the continuous Rényi entropy as the gain in information provided by the experimental results \mathcal{F}_n relative to the uniform distribution. The right hand side of (7.37) is formally referred to as the information quantity of order p assigned to the PDF $f(x)$. We refer to this quantity as the continuous Rényi entropy. The continuous Rényi entropy has properties similar to the corresponding differential entropy ($p \rightarrow 1$) e.g for $p = 1$, $\mathcal{I}_p(f(x))$ can be negative as well.

Chapter 8

A One-Parameter Class of Uncertainty Relations

In this chapter we explore the notion of uncertainty relations within quantum mechanics. In particular we take an information-theoretic approach to the investigation of applications of entropic measures within this setting. This culminates in a novel derivation and application of an infinite family of uncertainty relations which are in fact a generalisation of the familiar uncertainty relations derived and used in quantum theory.

This chapter is based on the papers:

- (1) *New class of entropy-power-based uncertainty relations*, P Jizba, A Hayes and J A Dunningham, 2017 J. Phys.: Conf. Ser. 880 012054
- (2) *One-parameter class of uncertainty relations based on entropy power*, Petr Jizba, Yue Ma, Anthony Hayes and Jacob A Dunningham, Physical Review E 93, 060104(R) (2016)

8.1 Uncertainty Relations

8.1.1 Variance-Based Uncertainty Relations

In section 2.1 the uncertainty relation (2.17) (in units of \hbar)

$$\Delta^2 \hat{X}_\lambda \Delta^2 \hat{X}_{\lambda+\pi/2} \geq \frac{1}{4} \quad (8.1)$$

was introduced as a mathematical consequence of the non-commutativity of the operators involved. Historically, this concept was first introduced in a quantum mechanical context by Heisenberg (1927) [122] and is famously known as the Heisenberg uncertainty principle however this often conflates two subtly different ideas; the concept of quantum mechan-

ical uncertainty relations and the idea that a measurement of a physical system must necessarily disturb the system itself. While the latter is true, and indeed the notion that Heisenberg posited, this is not the underlying concept of uncertainty relations that we investigate in the following. Uncertainty relations emerge as a consequence of fundamental properties of quantum mechanics and they impose a trade-off on the precision between position and momentum when a *simultaneous* measurement is performed. The underlying reasoning can be interpreted via wave mechanics in that a function and its Fourier transform cannot both be sharply localised or it can be interpreted through matrix mechanics in that non-commutative observables cannot have simultaneous eigenstates. These ideas were developed independently by Kennard (1927) [123] within the mathematical formalism of quantum mechanics, this initial result states that for the pair of conjugate variables position x and momentum p , the amount of information we have on one of the variables places a fundamental limit on the information attainable on the other and reads

$$\sigma_x \sigma_p \geq \frac{\hbar}{2} \quad (8.2)$$

where, σ_x and σ_p are the standard deviations of the position and momentum respectively. We now refer to Heisenberg's original result as the “error-disturbance” relation [124] in order to distinguish it from Kennard's result. A pivotal development was the generalisation of Kennard's relation by Robertson (1929) [125] to encompass arbitrary conjugate variables

$$\sigma(\hat{A})\sigma(\hat{B}) \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle \right| \quad (8.3)$$

which was in turn generalised by Schrödinger (1930) [126] to include statistical correlations captured by non-zero covariance terms between the observables

$$\Delta^2 \hat{A} \Delta^2 \hat{B} \geq \left(\frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle \right)^2 - \left| \frac{1}{2} \langle \{\hat{A}, \hat{B}\} \rangle \right|^2 \quad (8.4)$$

where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ is the anti-commutator. The next major development concerning this type of uncertainty relations came from Arthurs and Kelly (1965) [127] who incorporated a quantum mechanical apparatus as discussed in section 4.1.4 into the measurement scheme and Arthurs and Goodman who generalised the measurement scheme for arbitrary complementary observables. This analysis included the condition of unbiasedness, that is to say that the expectation values of the system of interest can be directly estimated from the outcomes of measurements performed on the apparatus. With this, it was found that the measurement outcome is composed of two distinct types of uncertainty; the inherent quantum mechanical fluctuations and the error in the measurement, this is

encapsulated in the following

$$\sigma'(\hat{A}')\sigma'(\hat{B}') \geq \sigma(\hat{A})\sigma(\hat{B}) + \epsilon(\hat{A})\epsilon(\hat{B}) \geq |\langle [\hat{A}, \hat{B}] \rangle| \quad (8.5)$$

where $\sigma'(\hat{A}')$ and $\sigma'(\hat{B}')$ denote the standard deviation of the apparatus for the arbitrary complementary observables and $\epsilon(\cdot)$ is the error in the measurement. The unbiasedness condition was later removed by Ozawa (2003) [128]. This allowed Ozawa to claim that the original Heisenberg error-disturbance relation can be violated but it can be shown [129] that this is a result of the definition of the error and disturbance having no correspondence to the accuracy of estimation.

8.1.2 Entropic Uncertainty Relations

Prompted by the shared properties of the Fisher information and the Shannon entropy, Stam (1965) conjectured that the Shannon entropy could be used to form an uncertainty relation. This was confirmed independently by Hirschman (1957) [130] by proving that the sum of the Shannon entropies of the absolute value squared of any function $f(x) \in L^2$ and its Fourier transform is non-negative, i.e for

$$f(x) = \int_{-\infty}^{\infty} \exp(i2\pi xy)g(y)dy \quad (8.6)$$

with a convergent integral in L^2 , it follows that

$$H(|f|^2) + H(|g|^2) \geq 0. \quad (8.7)$$

Additionally, Hirschman conjectured a tighter bound (as did Everret [131])

$$H(|f|^2) + H(|g|^2) \geq \ln\left(\frac{e}{2}\right) \quad (8.8)$$

which was later proven by Beckner (1975) [132]. This was adapted to a quantum mechanical uncertainty relation by Bialynicki-Birula [133] which showed that the Shrödinger (8.4) and Robertson VURs (8.3) can be derived from this entropic uncertainty relation.

8.1.3 Entropy Power Uncertainty Relations

Here, with the aim of constructing a generalised uncertainty relation, we introduce another mathematical object related to entropy that was first introduced by Shannon [114] in the context of communication theory, it is known as *entropy power* (EP). For the differential entropy, as given by (7.11), the entropy power $N(\chi)$ of the random variable $\chi \in \mathbb{R}^D$ is defined by the quantity that satisfies

$$H(\chi) = H\left(\sqrt{N(\chi)} \cdot Z_G\right) \quad (8.9)$$

where \mathcal{Z}_G is the Gaussian vector with zero mean and unit covariance matrix. That is to say, the EP is the variance of what would be a Gaussian random variable that has the equivalent differential entropy as the random variable of interest. Hence if the variable of interest is in fact Gaussian, the EP is simply the variance of that Gaussian. However we are not limited to the analysis random variables which submit Gaussian PDFs as we shall see in the following and to facilitate this we first generalise the notion of EP using the Rényi entropy. In units of bits, the p th Rényi entropy power (REP) $N_p(\chi)$ of the random variable $\chi \in \mathbb{R}^D$ is found to be [114, 134]

$$N_p(\chi) = \frac{1}{2\pi} p^{-p'/p} 2^{\frac{2}{D} \mathcal{I}_p(\chi)} \quad (8.10)$$

where $1/p + 1/p' = 1$ with $p \in \mathbb{R}^+$. This is the solution of

$$\mathcal{I}_p(\chi) = \mathcal{I}_p \left(\sqrt{N_p(\chi)} \cdot \mathcal{Z}_G \right) \quad (8.11)$$

where \mathcal{I}_p is the continuous Rényi entropy as defined by (7.37). This also has the property that $N_p(\chi) \rightarrow N(\chi)$ as $p \rightarrow 1$. With the aim of deriving an uncertainty relation we now introduce the Beckner-Babenko theorem [132, 135]: let

$$f^{(2)}(\vec{x}) = \int_{\mathbb{R}^D} e^{2\pi i \vec{x} \vec{y}} f^{(1)}(\vec{y}) d\vec{y} \quad (8.12)$$

then for $p \in [1, 2]$ the following inequality holds

$$|(p')^{D/2}|^{1/p'} \|f^{(2)}\|_{p'} \leq |p^{D/2}|^{1/p} \|f^{(1)}\|_p \quad (8.13)$$

where, p and p' are Hölder conjugates and

$$\|F\|_p \equiv \left(\int_{\mathbb{R}^D} |F(\vec{y})|^p d\vec{y} \right)^{1/p} \quad (8.14)$$

for any $F \in L^p(\mathbb{R}^D)$ where L^p is the function space defined using the p -norm for finite dimensional vector spaces. Choosing $\sqrt{\mathcal{F}(\vec{y})} \equiv |f(\vec{y})|$ we can rewrite inequality (8.13) as

$$\left(\int_{\mathbb{R}^D} |\mathcal{F}^{(2)}(\vec{y})|^{1+t} d\vec{y} \right)^{1/t} \left(\int_{\mathbb{R}^D} |\mathcal{F}^{(1)}(\vec{y})|^{1+r} d\vec{y} \right)^{1/r} \leq [2(1+t)]^D |t/r|^{D/2r} \quad (8.15)$$

where $r = p/2 - 1$, $t = p'/2 - 1$ and since $1/p + 1/p' = 1$ we have that $t = -r/(2r + 1)$. Furthermore, given that $p \in [1, 2]$ we have $r \in [-1/2, 0]$ and $t \in [0, \infty)$. Taking the logarithm (in base 2) of each side of the inequality and multiplying through by -1 gives

$$\mathcal{I}_{1+t}(\mathcal{F}^{(2)}) + \mathcal{I}_{1+r}(\mathcal{F}^{(1)}) \geq \frac{1}{r} \log_2 [2(1+r)]^{D/2} + \frac{1}{t} \log_2 [2(1+t)]^{D/2} \quad (8.16)$$

which, in the limit of $t \rightarrow 0_+$ and $r \rightarrow 0_-$, reduces to the Hirschman conjecture for the differential Shannon entropy [130]

$$H(\mathcal{F}^{(2)}) + H(\mathcal{F}^{(1)}) \geq \log_2 \left(\frac{e}{2} \right)^D. \quad (8.17)$$

The semidefiniteness of the RE ($\mathcal{I}_\alpha(\chi) \geq 0, \forall \alpha, \chi$) renders (8.16) impractical. Inequality (8.16) can be rewritten in terms of REPs to give

$$N_{1+t}(\mathcal{F}^{(2)})N_{1+r}(\mathcal{F}^{(1)}) \geq \frac{1}{16\pi^2} \quad (8.18)$$

and since $t = -r/(2r+1)$, this constitutes a family of inequalities that are characterised by a single parameter. The right hand side of (8.18) constitutes a universal lower bound since it is not dependent on any parameter, this is not the case for (8.16). We now consider this bound in a quantum mechanical setting by taking the position and momentum wavefunctions $\psi(\vec{x})$ and $\hat{\psi}(\vec{p})$, these are state vectors that are related to each other through the fact that they are Fourier transform duals

$$\psi(\vec{x}) = \int_{\mathbb{R}^D} e^{i\vec{x}\cdot\vec{p}/\hbar} \hat{\psi}(\vec{p}) \frac{d\vec{p}}{(2\pi\hbar)^{D/2}} \quad (8.19)$$

for which [119] $\|\psi\|_2 = \|\hat{\psi}\|_2 = 1$. Defining

$$f^{(2)}(\vec{x}) = (2\pi\hbar)^{D/4} \psi(\sqrt{2\pi\hbar}\vec{x}), \quad f^{(1)}(\vec{p}) = (2\pi\hbar)^{D/4} \hat{\psi}(\sqrt{2\pi\hbar}\vec{p}) \quad (8.20)$$

where the factors of $(2\pi\hbar)^{D/4}$ are included so the wavefunctions adhere to the form of the Fourier transform as given in the Beckner-Babenko theorem. Then using the fact that

$$\mathcal{I}_p(|f^{(1)}|^2) = \mathcal{I}_p(|\hat{\psi}|^2) - \frac{D}{2} \log_2(2\pi\hbar), \quad (8.21)$$

inequality (8.16) then becomes

$$\mathcal{I}_{1+t}(|\psi|^2) + \mathcal{I}_{1+r}(|\hat{\psi}|^2) \geq \frac{1}{r} \log_2 \left(\frac{1+r}{\pi\hbar} \right)^{D/2} + \frac{1}{t} \log_2 \left(\frac{1+t}{\pi\hbar} \right)^{D/2} \quad (8.22)$$

which in terms of REPs gives

$$N_{1+t}(|\psi|^2)N_{1+r}(|\hat{\psi}|^2) \geq \frac{\hbar^2}{4} \quad (8.23)$$

which resembles the familiar Robertson-Schrödinger VUR with the crucial difference of being a family of uncertainty relations characterised by a single parameter (since t and r are dependent on each other).

8.2 REPUR for Gaussian States

In the following, we intend to investigate the uncertainty relation (8.23). The inspection of the bound (8.23) for the various Gaussian states requires the PDFs of the general quadrature variable x_λ i.e the eigenvalue of the relation $\hat{X}_\lambda |x_\lambda\rangle = x_\lambda |x_\lambda\rangle$ where the quadrature operator is given by (2.15). Consider a general Gaussian function

$$f(x) = a \exp \left(\frac{-(x-b)^2}{2c^2} \right) \quad (8.24)$$

where a is the amplitude of the distribution, b is the mean value at which the distribution is centred and c is the standard deviation. The entropy power of (8.24) is then given by

$$N_{1+t}(f(x)) = \frac{1}{2\pi} (1+t)^{-\frac{1}{t}} 2^{-\frac{2}{t} \log_2 \left[\int_{-\infty}^{\infty} dx \left(a \exp \left(\frac{-(x-b)^2}{2c^2} \right) \right)^{(1+t)} \right]} \quad (8.25)$$

then let $y = x - b \Rightarrow dy = dx$, (and integration limits remain unchanged), thus

$$N_{1+t}(f(x)) = \frac{(1+t)^{-\frac{1}{t}}}{2\pi} 2^{-\frac{2}{t} \log_2 \left[a^{(1+t)} \int_{-\infty}^{\infty} dy \exp \left(\frac{-y^2(1+t)}{2c^2} \right) \right]}. \quad (8.26)$$

We have

$$\int_{-\infty}^{\infty} e^{-kx^2} dx = \sqrt{\frac{\pi}{k}} \quad \text{for } \text{Re}(k) > 0 \quad (8.27)$$

so with $k = \frac{1+t}{2c^2}$ (and hence $\text{Re}(k) > 0$ as $t, r \in [-\frac{1}{2}, \infty)$ and $c > 0$ by definition) we have

$$\begin{aligned} N_{1+t}(f(x)) &= \frac{(1+t)^{-\frac{1}{t}}}{2\pi} 2^{-\frac{2}{t} \log_2 \left[a^{(1+t)} \pi^{\frac{1}{2}} (1+t)^{-\frac{1}{2}} 2^{\frac{1}{2}} c \right]} \\ &= \frac{(1+t)^{-\frac{1}{t}}}{2\pi} a^{-\frac{2}{t}} a^{-2} (2\pi)^{-\frac{1}{t}} (1+t)^{\frac{1}{t}} c^{-\frac{2}{t}} \\ &= \frac{1}{2\pi} (c^2 2\pi a^2)^{-\frac{1}{t}} a^{-2} \end{aligned} \quad (8.28)$$

thus $N_{1+t}(f(x))$ is independent of the parameter t iff $c^2 2\pi a^2 = 1$ as $c, a \in \mathbb{R}$. To inspect the conjugate variable, we must take the entropy power of the Fourier transform of the probability *amplitude* of $f(x)$ to give $\hat{f}(p)$ however, we have that the Fourier transform of a Gaussian distribution is also a Gaussian distribution. The above argument then holds exactly for the entropy power of the conjugate variable's PDF (trivially replacing “ t ” with “ r ”). Then we find for a general Gaussian PDF

$$N_{1+t}(f(x)) N_{1+r}(\hat{f}(p)) = \frac{1}{4\pi^2} (c^2 2\pi a^2)^{-\frac{1}{t}} (\hat{c}^2 2\pi \hat{a}^2)^{-\frac{1}{r}} a^{-2} \hat{a}^{-2} \geq \frac{1}{4} \quad (8.29)$$

where, $\hat{f}(p)$ is the Fourier transform with amplitude \hat{a} and standard deviation \hat{c} . So for general Gaussian PDFs the bound is saturated under the condition $(c^2 2\pi a^2)^{-\frac{2}{t}} (\pi a^2)^{-2} = 1$. Furthermore, for Gaussian *states* the PDF must be normalised i.e.

$$\int_{-\infty}^{\infty} f(x) dx = a \int_{-\infty}^{\infty} dx \exp \left(\frac{-(x-b)^2}{2c^2} \right) = 1 \quad (8.30)$$

so, making the usual substitution $y = x - b$ and using (8.27), we see

$$a\sqrt{2\pi}c = 1 \quad (8.31)$$

thus for Gaussian states the entropy power is given by

$$N_{1+t}(f(x)) = \frac{1}{2\pi} a^{-2} \quad (8.32)$$

independent of the parameter t . The bound (8.23) becomes

$$\begin{aligned}
N_{1+t}(f(x))N_{1+r}(\hat{f}(p)) &= \frac{1}{4\pi^2}a^{-2}\hat{a}^{-2} \\
&= \frac{1}{4\pi^2}(2\pi c^2)(2\pi\hat{c}^2) \\
&= (c\hat{c})^2 \\
&\geq \frac{1}{4}.
\end{aligned} \tag{8.33}$$

Considering a probability amplitude of the PDF described by (8.24) we have for general quadrature variable

$$\psi(x_\lambda) = \sqrt{a} \exp\left(\frac{(x_\lambda - b)^2}{4c^2}\right) \tag{8.34}$$

with the conjugate probability amplitude given by the Fourier transform which we define as

$$\hat{\psi}(x_{\lambda+\pi/2}) = \sqrt{\hat{a}} \exp\left(\frac{(x_{\lambda+\pi/2} - b)^2}{4\hat{c}^2}\right) \tag{8.35}$$

which, due to the Gaussian nature of the probability amplitudes of the conjugate quadrature variables satisfies the relation $\sigma_\lambda \sigma_{\lambda+\pi/2} = 1$ where $\sigma_\lambda = \sqrt{2}c$ and $\sigma_{\lambda+\pi/2} = \sqrt{2}\hat{c}$. This is equivalent to $c\hat{c} = 1/2$, so substituting this into (8.33) then enables us to conclude

$$N_{1+t}(\mathcal{F}_N^G(x_\lambda))N_{1+r}(\mathcal{F}_N^G(x_{\lambda+\pi/2})) = \frac{1}{4} \tag{8.36}$$

i.e all Gaussian states saturate the bound (8.23) independent of the family of parameters t and r . In the following, we shall demonstrate this result by applying it to Gaussian states prominent in quantum metrology.

Coherent State

From equation (2.76) which is restated here for convenience, the overlap of the general quadrature state with a coherent state is given by

$$\langle x_\lambda | e^{i\phi} \beta \rangle = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(x_\lambda^2 + \beta^2)} e^{-\frac{1}{2}e^{2i(\lambda+\phi)}\beta^2 + \sqrt{2}e^{i(\lambda+\phi)}\beta x_\lambda}. \tag{8.37}$$

Setting $\phi = 0$, the PDFs of a coherent state for the conjugate quadrature variables x_0 and $x_{\pi/2}$ are given by

$$\mathcal{F}(x_0) = |\langle x_0 | \beta \rangle|^2 = \pi^{-\frac{1}{2}} e^{-(x_0 - \sqrt{2}\beta)^2} \tag{8.38}$$

$$\mathcal{F}(x_{\pi/2}) = |\langle x_{\pi/2} | \beta \rangle|^2 = \pi^{-\frac{1}{2}} e^{-x_{\pi/2}^2} \tag{8.39}$$

as illustrated in Fig.8.1. The Shannon entropy, which corresponds to $r = t = 0$, is found

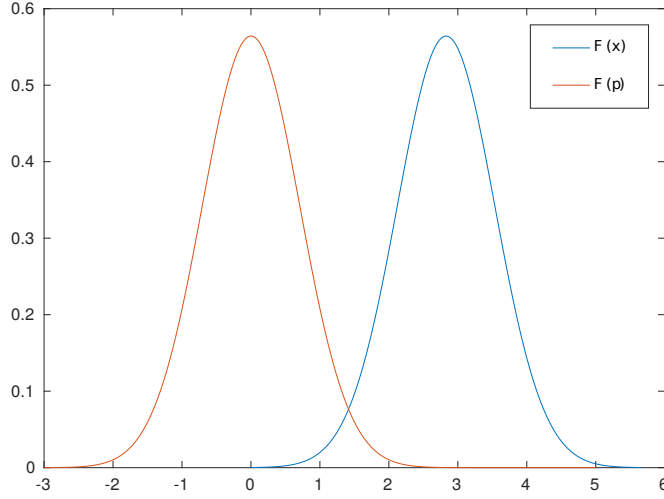


Figure 8.1: Plots of position and momentum probability distributions for a coherent state with $\beta = 2$. Both are Gaussian PDFs.

to be

$$\begin{aligned}
 \mathcal{I}_1(x_0) &= - \int_{-\infty}^{+\infty} |\langle x_0 | \beta \rangle|^2 \log_2 (|\langle x_0 | \beta \rangle|^2) dx_0 \\
 &= - \int_{-\infty}^{+\infty} \left| \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(x_0^2 + \beta^2)} e^{-\frac{1}{2}\beta^2 + \sqrt{2}\beta x_0} \right|^2 \log_2 \left(\left| \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(x_0^2 + \beta^2)} e^{-\frac{1}{2}\beta^2 + \sqrt{2}\beta x_0} \right|^2 \right) dx_0 \\
 &= - \int_{-\infty}^{+\infty} \pi^{-\frac{1}{2}} e^{-(x_0 - \sqrt{2}\beta)^2} \log_2 \left(\pi^{-\frac{1}{2}} e^{-(x_0 - \sqrt{2}\beta)^2} \right) dx_0. \tag{8.40}
 \end{aligned}$$

Then making the change of variable $x_0 = \tilde{x}_0 - \sqrt{2}\beta \Rightarrow dx_0 = d\tilde{x}_0$, we have

$$\begin{aligned}
 \mathcal{I}_1(\tilde{x}_0) &= - \int_{-\infty}^{+\infty} \pi^{-\frac{1}{2}} e^{-\tilde{x}_0^2} \log_2 \left(\pi^{-\frac{1}{2}} e^{-\tilde{x}_0^2} \right) d\tilde{x}_0 \\
 &= \frac{1}{2} \pi^{-\frac{1}{2}} \log_2(\pi) \int_{-\infty}^{+\infty} e^{-\tilde{x}_0^2} d\tilde{x}_0 + (\log_2 e) \pi^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-\tilde{x}_0^2} \tilde{x}_0^2 d\tilde{x}_0 \\
 &= \frac{1}{2\sqrt{\pi}} (\log_2(\pi)) \sqrt{\pi} + (\log_2(e)) \frac{1}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} \\
 &= \frac{1}{2} \log_2(e\pi) \tag{8.41}
 \end{aligned}$$

where we have made use of the Gaussian integral results $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ and $\int_{-\infty}^{+\infty} e^{-x^2} x^2 dx = \frac{\sqrt{\pi}}{2}$. Similarly, we find for the conjugate variable

$$\begin{aligned}
\mathcal{I}_1(x_{\pi/2}) &= - \int_{-\infty}^{+\infty} |\langle x_{\pi/2} | \beta \rangle|^2 \log_2 |\langle x_{\pi/2} | \beta \rangle|^2 dx_{\pi/2} \\
&= - \int_{-\infty}^{+\infty} \left| \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(x_{\pi/2}^2 + \beta^2)} e^{-\frac{1}{2}e^{i\pi}\beta^2 + \sqrt{2}e^{i\frac{\pi}{2}}\beta x_{\pi/2}} \right|^2 \\
&\quad \cdot \log_2 \left(\left| \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(x_{\pi/2}^2 + \beta^2)} e^{-\frac{1}{2}e^{i\pi}\beta^2 + \sqrt{2}e^{i\frac{\pi}{2}}\beta x_{\pi/2}} \right|^2 \right) dx_{\pi/2} \\
&= - \int_{-\infty}^{+\infty} \pi^{-\frac{1}{2}} e^{-x_{\pi/2}^2} \log_2(\pi^{-\frac{1}{2}} e^{-x_{\pi/2}^2}) dx_{\pi/2} \\
&= \frac{1}{2} \log_2(e\pi)
\end{aligned} \tag{8.42}$$

thus the entropy powers for both conjugate variables are given by

$$N_1(x_0) = N_1(x_{\pi/2}) = \frac{1}{2\pi e} 2^{2(\frac{1}{2} \log_2(e\pi))} = \frac{1}{2} \tag{8.43}$$

so that $N_1(x_0)N_1(x_{\pi/2}) = \frac{1}{4}$ saturating the bound as expected.

We now calculate the so called “min-entropy” with respect to the momentum-like quadrature variable ($r \rightarrow \infty$, which forces $t = -1/2$):

$$\begin{aligned}
\mathcal{I}_\infty(x_{\pi/2}) &= \lim_{\alpha \rightarrow \infty} \frac{1}{1-\alpha} \log_2 \left(\int_{-\infty}^{+\infty} dx_{\pi/2} (\mathcal{F}(x_{\pi/2}))^\alpha \right) \\
&= \lim_{\alpha \rightarrow \infty} \frac{1}{1-\alpha} \log_2 (dx_{\pi/2} \cdot \max[\mathcal{F}(x_{\pi/2})^\alpha]) \\
&= \lim_{\alpha \rightarrow \infty} \frac{1}{1-\alpha} (\log_2(dx_{\pi/2}) + \alpha \log_2(\max[\mathcal{F}(x_{\pi/2})])) \\
&= -\log_2(\max[\mathcal{F}(x_{\pi/2})])
\end{aligned} \tag{8.44}$$

where, from (8.37), $\mathcal{F}(x_{\pi/2}) = |\langle x_{\pi/2} | \beta \rangle|^2 = \pi^{-\frac{1}{2}} e^{-x_{\pi/2}^2}$ and hence, $\max[\mathcal{F}(x_{\pi/2})] = \pi^{-\frac{1}{2}}$. With this, equation (8.44) yields

$$\mathcal{I}_\infty(x_{\pi/2}) = \frac{1}{2} \log_2(\pi). \tag{8.45}$$

For the conjugate variable we find

$$\begin{aligned}
\mathcal{I}_{1/2}(x_0) &= \frac{1}{1-\frac{1}{2}} \log_2 \left(\int_{-\infty}^{+\infty} dx_0 (|\langle x_0 | \beta \rangle|^2)^{\frac{1}{2}} \right) \\
&= 2 \log_2 \left(\int_{-\infty}^{+\infty} dx_0 (\pi^{-\frac{1}{2}} e^{-(x_0 - \sqrt{2}\beta)^2})^{\frac{1}{2}} \right) \\
&= \log_2(2\pi^{\frac{1}{2}})
\end{aligned} \tag{8.46}$$

and the according entropy powers are found to be

$$N_{\frac{1}{2}}(x_0) = \frac{1}{2}, \quad N_\infty(x_{\pi/2}) = \frac{1}{2} \tag{8.47}$$

thus

$$N_{\frac{1}{2}}(x_0)N_{\infty}(x_{\pi/2}) = \frac{1}{4}. \quad (8.48)$$

again saturating the bound (8.23) as expected. For the min-entropy with respect to the position-like quadrature variable ($t \rightarrow \infty$, which forces $r = -1/2$), we find

$$\begin{aligned} \mathcal{I}_{\infty}(x_0) &= -\log_2(\max[\mathcal{F}(x_0)]) \\ &= \frac{1}{2} \log_2(\pi) \end{aligned} \quad (8.49)$$

$$\begin{aligned} \mathcal{I}_{1/2}(x_{\pi/2}) &= 2 \log_2 \left(\int_{-\infty}^{+\infty} dx_{\pi/2} (|\langle x_{\pi/2} | \beta \rangle|^2)^{\frac{1}{2}} \right) \\ &= \log_2(2\pi^{\frac{1}{2}}) \end{aligned} \quad (8.50)$$

hence the entropy powers are given by

$$N_{\frac{1}{2}}(x_{\pi/2}) = \frac{1}{2} \quad N_{\infty}(x_0) = \frac{1}{2} \quad (8.51)$$

thus

$$N_{\frac{1}{2}}(x_{\pi/2})N_{\infty}(x_0) = \frac{1}{4} \quad (8.52)$$

once again, saturating the bound (8.23).

Squeezed State

From equation (2.56), the squeezed vacuum state $|\zeta\rangle$ with $\zeta = ze^{i\phi}$ has the following Fock basis representation

$$|\zeta\rangle = \sum_{m=0}^{\infty} (-1)^m \frac{\sqrt{(2m)!}}{2^m m!} \frac{(\tanh(z))^m}{\sqrt{\cosh(z)}} e^{i2m\phi} |2m\rangle \quad (8.53)$$

and from [22] we have

$$\langle x_{\lambda} | \zeta \rangle = (2\pi \Delta^2 \hat{X}_{\lambda})^{-\frac{1}{4}} \exp \left(-\frac{x_{\lambda}^2}{4\Delta^2 \hat{X}_{\lambda}} [1 - i \sin(2\lambda - \phi) \sinh(2z)] \right) \quad (8.54)$$

where the variance is given by

$$\Delta^2 \hat{X}_{\lambda} = \frac{1}{2} \left[e^{2z} \sin^2 \left(\lambda - \frac{\phi}{2} \right) + e^{-2z} \cos^2 \left(\lambda - \frac{\phi}{2} \right) \right] \quad (8.55)$$

hence the PDFs for the squeezed state for the conjugate quadrature variables x_0 and $x_{\pi/2}$ are given by

$$\mathcal{F}(x_0) = \pi^{-1/2} \exp(-x_0^2 e^{2z} + z) \quad (8.56)$$

$$\mathcal{F}(x_{\pi/2}) = \pi^{-1/2} e^{-x_{\pi/2}^2} \quad (8.57)$$

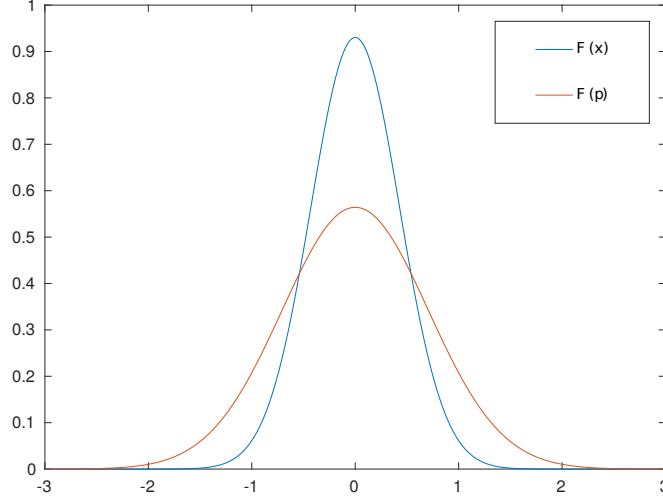


Figure 8.2: Plots of position and momentum probability distributions for a squeezed vacuum state with $z = 1/2$ and $\phi = 0$. Both are Gaussian PDFs.

as displayed in Fig.8.2. The Shannon entropy is evaluated via the following

$$\begin{aligned}
 \mathcal{I}_1(x_0) &= - \int_{-\infty}^{+\infty} |\langle x_0 | \zeta \rangle|^2 \log_2 |\langle x_0 | \zeta \rangle|^2 dx_0 \\
 &= - \int_{-\infty}^{+\infty} \left| \pi^{-\frac{1}{4}} e^{\frac{z}{2}} e^{-\frac{x_0^2 e^{2z}}{2}} \right|^2 \log_2 \left(\left| \pi^{-\frac{1}{4}} e^{\frac{z}{2}} e^{-\frac{x_0^2 e^{2z}}{2}} \right|^2 \right) dx_0 \\
 &= - \int_{-\infty}^{+\infty} \pi^{-\frac{1}{2}} e^z e^{-x_0^2 e^{2z}} \log_2 (\pi^{-\frac{1}{2}} e^z e^{-x_0^2 e^{2z}}) dx_0 \\
 &= -\pi^{-\frac{1}{2}} e^z \int_{-\infty}^{+\infty} e^{-x_0^2 e^{2z}} \left(\log_2 (\pi^{-\frac{1}{2}} e^z) + \log_2 (e^{-x_0^2 e^{2z}}) \right) dx_0 \\
 &= -\pi^{-\frac{1}{2}} \left(\log_2 (\pi^{-\frac{1}{2}} e^z) \int_{-\infty}^{+\infty} e^{-(x_0 e^z)^2} d(x_0 e^z) - \log_2 e \int_{-\infty}^{+\infty} e^{-(x_0 e^z)^2} (x_0 e^z)^2 d(x_0 e^z) \right) \\
 &= -\pi^{-\frac{1}{2}} \log_2 (\pi^{-\frac{1}{2}} e^z) \int_{-\infty}^{+\infty} e^{-x_0^2} dx_0 + \pi^{-\frac{1}{2}} (\log_2 e) \int_{-\infty}^{+\infty} e^{-x_0^2} x_0^2 dx_0 \\
 &= \log_2 \left((e\pi)^{\frac{1}{2}} e^{-z} \right)
 \end{aligned} \tag{8.58}$$

and similarly for the conjugate variable

$$\begin{aligned}
 \mathcal{I}_1(x_{\pi/2}) &= - \int_{-\infty}^{+\infty} |\langle x_{\pi/2} | \zeta \rangle|^2 \log_2 |\langle x_{\pi/2} | \zeta \rangle|^2 dx_{\pi/2} \\
 &= - \int_{-\infty}^{+\infty} \left| \pi^{-\frac{1}{4}} e^{-\frac{z}{2}} e^{-\frac{x_{\pi/2}^2 e^{-2z}}{2}} \right|^2 \log_2 \left(\left| \pi^{-\frac{1}{4}} e^{-\frac{z}{2}} e^{-\frac{x_{\pi/2}^2 e^{-2z}}{2}} \right|^2 \right) dx_{\pi/2} \\
 &= \log_2 \left((e\pi)^{\frac{1}{2}} e^z \right).
 \end{aligned} \tag{8.59}$$

Hence, the entropy powers are found to be

$$N_1(x_0) = \frac{e^{-2z}}{2} \quad N_1(x_{\pi/2}) = \frac{e^{2z}}{2} \tag{8.60}$$

so that

$$N_1(x_0)N(x_{\pi/2}) = \frac{1}{4} \quad (8.61)$$

saturating the bound (8.23) as predicted.

The min-entropy with respect to the momentum-like quadrature variable ($r \rightarrow \infty$, which forces $t = -1/2$) is found to be

$$\begin{aligned} \mathcal{I}_\infty(x_{\pi/2}) &= -\log_2 \left(\max [|\langle x_{\pi/2} | \zeta \rangle|^2] \right) \\ &= -\log_2 \left(\max \left[\pi^{-\frac{1}{2}} e^{-z} e^{-\frac{x_{\pi/2}^2}{2}} e^{-2z} \right] \right) \\ &= -\log_2 \left(\pi^{-\frac{1}{2}} e^{-z} \right) \\ &= \log_2 \left(\pi^{\frac{1}{2}} e^z \right) \end{aligned} \quad (8.62)$$

and for the conjugate variable

$$\begin{aligned} \mathcal{I}_{1/2}(x_0) &= 2 \log_2 \left(\int_{-\infty}^{+\infty} dx_0 (|\langle x_0 | \zeta \rangle|^2)^{\frac{1}{2}} \right) \\ &= 2 \log_2 \left(\int_{-\infty}^{+\infty} dx_0 \pi^{-\frac{1}{4}} e^{\frac{z}{2}} e^{-\frac{x_0^2}{2}} e^{2z} \right) \\ &= \log_2 \left(2\pi^{\frac{1}{2}} e^{-z} \right) \end{aligned} \quad (8.63)$$

where we have used the result $\int_{-\infty}^{\infty} e^{-\frac{kx^2}{2}} dx = \sqrt{\frac{2\pi}{k}}$.

$$N_{1/2}(x_0) = \frac{e^{-2z}}{2}, \quad N_\infty(x_{\pi/2}) = \frac{e^{2z}}{2} \quad (8.64)$$

thus

$$N_{1/2}(x_0)N_\infty(x_{\pi/2}) = \frac{1}{4} \quad (8.65)$$

again saturating the bound (8.23). The min-entropy with respect to the position-like quadrature variable ($t \rightarrow \infty$, which forces $r = -1/2$) is given by

$$\begin{aligned} \mathcal{I}_{1/2}(x_{\pi/2}) &= 2 \log_2 \left(\int_{-\infty}^{+\infty} (|\langle x_{\pi/2} | \zeta \rangle|^2)^{\frac{1}{2}} dx_{\pi/2} \right) \\ &= 2 \log_2 \left(\int_{-\infty}^{+\infty} \pi^{-\frac{1}{4}} e^{-\frac{z}{2}} e^{-\frac{x_{\pi/2}^2}{2}} e^{-2z} dx_{\pi/2} \right) \\ &= \log_2 \left(2\pi^{\frac{1}{2}} e^z \right) \end{aligned} \quad (8.66)$$

$$\begin{aligned} \mathcal{I}_\infty(x_0) &= -\log_2 \left(\max [|\langle x_0 | \zeta \rangle|^2] \right) \\ &= -\log_2 \left(\max \left[\pi^{-\frac{1}{2}} e^z e^{-x_0^2} e^{2z} \right] \right) \\ &= -\log_2 \left(\pi^{-\frac{1}{2}} e^z \right) \\ &= \log_2 \left(\pi^{\frac{1}{2}} e^{-z} \right) \end{aligned} \quad (8.67)$$

yielding the entropy powers

$$N_{1/2}(x_{\pi/2}) = \frac{e^{2z}}{2} \quad N_{\infty}(x_0) = \frac{e^{-2z}}{2} \quad (8.68)$$

hence

$$N_{1/2}(x_{\pi/2})N_{\infty}(x_0) = \frac{1}{4} \quad (8.69)$$

saturating the bound (8.23) once again.

8.3 REPUR for Non-Gaussian States

8.3.1 Cat States

The first non-Gaussian state we inspect is the Cat State; a macroscopic superposition of coherent states with opposite phases as given by (2.95) and restated here for convenience

$$|\psi_C\rangle = \mathcal{N}_C(|i\beta\rangle + |-i\beta\rangle). \quad (8.70)$$

From (2.97) and (2.98), the PDFs in the conjugate quadratures x_0 and $x_{\pi/2}$ are given by

$$\mathcal{F}^+(x_0, \pi/2, -\pi/2) = 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 e^{-x_0^2} \cos^2(\sqrt{2}\beta x_0) \quad (8.71)$$

$$\mathcal{F}^+(x_{\pi/2}, \pi/2, -\pi/2) = 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 e^{-x_{\pi/2}^2 - 2\beta^2} \cosh^2(\sqrt{2}\beta x_{\pi/2}). \quad (8.72)$$

With this we can now calculate the min-entropies with respect to the position variable. This requires integrating the quantity $\lim_{p \rightarrow \infty} (\mathcal{F}(x))^p$, the maximum value of the probability density function will dominate in this limit so we need only concern ourselves with the value

$$\lim_{p \rightarrow \infty} \int_M dx (\mathcal{F}(x))^p = \lim_{p \rightarrow \infty} dx \max [\mathcal{F}(x)]^p. \quad (8.73)$$

This in turn requires evaluating the maximum values of (2.97) and (2.98) for which we have $\mathcal{F}^+(x_0, \pi/2, -\pi/2) \equiv \mathcal{F}(x_0) = C e^{-x_0^2} \cos^2(\sqrt{2}\beta x_0)$ where, $C = 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2$ is constant (for fixed β). We have $\max[e^{-x_0}] = 1$ at $x_0 = 0$ and $\max[\cos^2(x_0)] = 1$ at $x_0 = 0$. Therefore, $\max[\mathcal{F}(x_0)] = C = 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2$ at $x_0 = 0$. The maximum of the probability density function for the conjugate variable $\mathcal{F}^+(x_{\pi/2}, \pi/2, -\pi/2) \equiv \mathcal{F}(x_{\pi/2})$ can be determined by

arranging (2.98) in the following suggestive form

$$\begin{aligned}
\mathcal{F}(x_{\pi/2}) &= 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 e^{-x_{\pi/2}^2 - 2\beta^2} \cosh^2(\sqrt{2}\beta x_{\pi/2}) \\
&= 4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 \left(e^{-\frac{x_{\pi/2}^2}{2} - \beta^2} \cosh(\sqrt{2}\beta x_{\pi/2}) \right)^2 \\
&= \pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 \left[e^{-\frac{x_{\pi/2}^2}{2} - \beta^2} \left(e^{\sqrt{2}\beta x_{\pi/2}} + e^{-\sqrt{2}\beta x_{\pi/2}} \right)^2 \right] \\
&= \pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 \left(e^{-\left(\frac{x_{\pi/2}}{\sqrt{2}} - \beta\right)^2} + e^{-\left(\frac{x_{\pi/2}}{\sqrt{2}} + \beta\right)^2} \right)^2
\end{aligned} \tag{8.74}$$

from which it is apparent that $\mathcal{F}(x_{\pi/2})$ is the sum of two Gaussian distributions. Maximising $\mathcal{F}(x_{\pi/2})$ over $x_{\pi/2}$ will yield a function dependent on β : the maximum at small β (when the two Gaussians overlap and “interfere”), due to the symmetry of Gaussian functions, will clearly be located at $x_{\pi/2} = 0$. For large β (when the two Gaussians are completely separated, roughly 3 standard deviations from the $\beta = 0$ case) we see the common maximum of the distribution will be located at the means of the constituent Gaussians i.e at $x_{\pi/2} = \pm\sqrt{2}\beta$. Using this we find

$$\max[\mathcal{F}(x_{\pi/2})] = \begin{cases} 2\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 e^{-2\beta^2} & \text{at } x = 0, \text{ for small } \beta \\ \pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 (1 + e^{-4\beta^2})^2 & \text{at } |\bar{x}_{\pi/2}| = \sqrt{2}\beta, \text{ for large } \beta. \end{cases} \tag{8.75}$$

The entropies for the $p \rightarrow \infty$ cases are in general given by

$$\begin{aligned}
\mathcal{I}_{\infty}(\mathcal{F}(x)) &= \lim_{p \rightarrow \infty} \frac{1}{1-p} \log_2 \left(\int_{-\infty}^{\infty} dx (\mathcal{F}(x))^p \right) \\
&= \lim_{p \rightarrow \infty} \frac{1}{1-p} \log_2 (dx [\max[\mathcal{F}(x)]]^p) \\
&= \lim_{p \rightarrow \infty} \frac{1}{1-p} [\log_2(dx) + \log_2(\max[\mathcal{F}(x)]^p)] \\
&= \lim_{p \rightarrow \infty} \left[\frac{\log_2(dx)}{1-p} + \frac{p}{1-p} \log_2(\mathcal{F}(x)) \right] \\
&= -\log_2(\max[\mathcal{F}(x)])
\end{aligned} \tag{8.76}$$

thus

$$\begin{aligned}
\mathcal{I}_{\infty}(\mathcal{F}(x_0)) &= \lim_{p \rightarrow \infty} \frac{1}{1-p} \log_2 \left(\int_{-\infty}^{\infty} dx_0 (\mathcal{F}(x_0))^p \right) \\
&= \lim_{p \rightarrow \infty} \frac{1}{1-p} \log_2 \left(dx_0 \left[4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 \right]^p \right) \\
&= -\log_2 \left(4\pi^{-\frac{1}{2}}(\mathcal{N}_{\beta}^+)^2 \right)
\end{aligned} \tag{8.77}$$

and similarly

$$\mathcal{I}_\infty(\mathcal{F}(x_{\pi/2})) = \begin{cases} -\log_2 \left(2\pi^{\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-2\beta^2} \right) & \text{at } x = 0, \text{ for small } \beta \\ -\log_2 \left[\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 \left(1 + e^{-4\beta^2} \right)^2 \right] & \text{at } |\bar{x}_{\pi/2}| = \sqrt{2}\beta, \text{ for large } \beta. \end{cases} \quad (8.78)$$

We now calculate the Rényi entropy for the probability density functions (2.97) and (2.98) for the case $p = 1/2$

$$\begin{aligned} \mathcal{I}_{1/2}(\mathcal{F}(x_{\pi/2})) &= \frac{1}{1 - \frac{1}{2}} \log_2 \left(\int_{-\infty}^{\infty} dx_{\pi/2} [4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_0^2 - 2\beta^2} \cosh^2(\sqrt{2}\beta x_{\pi/2})]^{\frac{1}{2}} \right) \\ &= 2 \log_2 \left(2\pi^{-\frac{1}{4}} \mathcal{N}_\beta^+ e^{-\beta^2} \int_{-\infty}^{\infty} e^{-\frac{x_{\pi/2}^2}{2}} \cosh(\sqrt{2}\beta x_{\pi/2}) dx_{\pi/2} \right) \end{aligned} \quad (8.79)$$

we have

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cosh(kx) dx = e^{\frac{k^2}{2}} \sqrt{2\pi} \quad (8.80)$$

thus,

$$\begin{aligned} \mathcal{I}_{1/2}(\mathcal{F}(x_{\pi/2})) &= 2 \log_2 \left(2\pi^{-\frac{1}{4}} \mathcal{N}_\beta^+ e^{-\beta^2} e^{\beta^2} \sqrt{2\pi} \right) \\ &= \log_2 \left(2^3 \pi^{\frac{1}{2}} (\mathcal{N}_\beta^+)^2 \right). \end{aligned} \quad (8.81)$$

Now for the remaining Rényi entropy we have

$$\begin{aligned} \mathcal{I}_{1/2}(\mathcal{F}(x_0)) &= \frac{1}{1 - \frac{1}{2}} \log_2 \left(\int_{-\infty}^{\infty} dx_0 [4\pi^{-\frac{1}{2}} (\mathcal{N}_\beta^+)^2 e^{-x_0^2} \cos^2(\sqrt{2}\beta x_0)]^{\frac{1}{2}} \right) \\ &= 2 \log_2 \left(2\pi^{-\frac{1}{4}} \mathcal{N}_\beta^+ \int_{-\infty}^{\infty} e^{-\frac{x_0^2}{2}} |\cos(\sqrt{2}\beta x_0)| dx_0 \right). \end{aligned} \quad (8.82)$$

Due to the complex nature of this integral, we use numerical methods to evaluate this.

The general expressions for the REPs can be written as

$$N_{\frac{1}{2}}(x_{\pi/2}) = \frac{1}{8\pi} 2^{2\mathcal{I}_{1/2}(\mathcal{F}(x_{\pi/2}))}, \quad N_\infty(x_0) = \frac{1}{2\pi} 2^{2\mathcal{I}_\infty(\mathcal{F}(x_0))} \quad (8.83)$$

which gives the REPUR

$$\begin{aligned} N_{\frac{1}{2}}(x_{\pi/2}) N_\infty(x_0) &= \frac{1}{16\pi^2} \left(2^{\mathcal{I}_{1/2}(\mathcal{F}(x_{\pi/2})) + \mathcal{I}_\infty(\mathcal{F}(x_0))} \right)^2 \\ &= \frac{1}{4} \end{aligned} \quad (8.84)$$

thus saturating the bound. On the other side, we find

$$N_{\frac{1}{2}}(x_0) = \frac{1}{8\pi} 2^{2\mathcal{I}_{1/2}(\mathcal{F}(x_0))}, \quad N_\infty(x_{\pi/2}) = \frac{1}{2\pi} 2^{2\mathcal{I}_\infty(\mathcal{F}(x_{\pi/2}))} \quad (8.85)$$

hence the REPUR becomes

$$N_{\frac{1}{2}}(x_0)N_{\infty}(x_{\pi/2}) = \frac{1}{16\pi^2} \left(2^{\mathcal{I}_{1/2}(\mathcal{F}(x_0)) + \mathcal{I}_{\infty}(\mathcal{F}(x_{\pi/2}))} \right)^2 \quad (8.86)$$

for which, due to the complicated nature of (8.82), we numerically plot against the Rényi parameter in Fig.8.3 which confirms equation (8.84) and suggests that (8.86) plateaus for

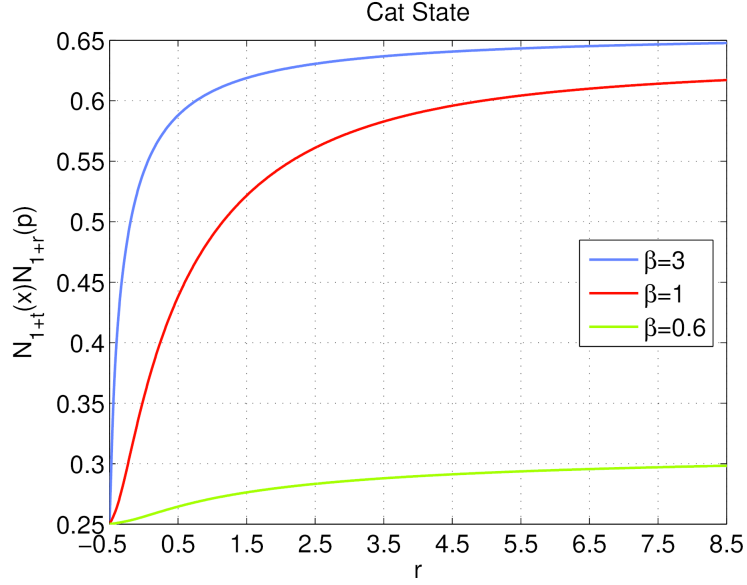


Figure 8.3: Plot of REPUR against the Rényi parameter r for the Cat State. It is evident that the bound is saturated for $r = -1/2$, begins to diverge as r increases then plateaus for large r .

increasing r with constant β .

8.3.2 Superstition State of Vacuum and Squeezed Vacuum

Here we investigate the REPUR for the superposition of a vacuum state with a squeezed vacuum (VSV) state

$$|\psi_V\rangle = \mathcal{N}_V(|0\rangle + |\zeta\rangle) \quad (8.87)$$

where, $|\zeta\rangle$ is given by (8.53) and

$$\mathcal{N}_V = \left[2 + 2 \cosh(r)^{-\frac{1}{2}} \right]^{-\frac{1}{2}}. \quad (8.88)$$

As before, we must first find the probability density functions

$$\mathcal{F}(x_0) = |\langle x_0|\psi_V\rangle|^2, \quad \mathcal{F}(x_{\pi/2}) = |\langle x_{\pi/2}|\psi_V\rangle|^2. \quad (8.89)$$

The general quadrature eigenstate overlap with (8.87) is given by

$$\langle x_\lambda|\psi_V\rangle = \mathcal{N}_V(\langle x_\lambda|0\rangle + \langle x_\lambda|\zeta\rangle) \quad (8.90)$$

where, from equation (2.76), it is straight forward to find

$$\langle x_\lambda | 0 \rangle = \pi^{-\frac{1}{4}} e^{-\frac{x_\lambda^2}{2}} \quad (8.91)$$

and from equations (8.54) and (8.55) we find

$$\langle x_0 | 0 \rangle = \pi^{-\frac{1}{4}} e^{-\frac{x_0^2}{2}} \quad (8.92)$$

and

$$\begin{aligned} \langle x_0 | \zeta \rangle &= \left(2\pi \cdot \frac{1}{2} e^{-2z} \right)^{-\frac{1}{4}} \exp \left(-\frac{x_0^2}{4 \cdot \frac{1}{2} e^{-2z}} \right) \\ &= \pi^{-\frac{1}{4}} e^{\frac{z}{2}} \exp \left(-\frac{x_0^2 e^{2z}}{2} \right). \end{aligned} \quad (8.93)$$

Similarly, we find

$$\langle x_{\pi/2} | 0 \rangle = \pi^{-\frac{1}{4}} e^{-\frac{x_{\pi/2}^2}{2}} \quad (8.94)$$

and

$$\begin{aligned} \langle x_{\pi/2} | \zeta \rangle &= \left(2\pi \cdot \frac{1}{2} e^{2z} \right)^{-\frac{1}{4}} \exp \left(-\frac{x_{\pi/2}^2}{4 \cdot \frac{1}{2} e^{2z}} \right) \\ &= \pi^{-\frac{1}{4}} e^{-\frac{z}{2}} \exp \left(-\frac{x_{\pi/2}^2 e^{-2z}}{2} \right) \end{aligned} \quad (8.95)$$

thus we find the probability density functions to be

$$\mathcal{F}(x_0) = \left| \mathcal{N}_V \pi^{-\frac{1}{4}} \left[e^{-\frac{x_0^2}{2}} + e^{\frac{z}{2}} \exp \left(-\frac{x_0^2 e^{2z}}{2} \right) \right] \right|^2 \quad (8.96)$$

and

$$\mathcal{F}(x_{\pi/2}) = \left| \mathcal{N}_V \pi^{-\frac{1}{4}} \left[e^{-\frac{x_{\pi/2}^2}{2}} + e^{-\frac{z}{2}} \exp \left(-\frac{x_{\pi/2}^2 e^{-2z}}{2} \right) \right] \right|^2. \quad (8.97)$$

These PDFs are plotted in Fig.8.4 which depict the overall non-Gaussian structure. With (8.96) and (8.97) the REPs can be numerically calculated and plotted as shown in Fig. 8.5, from which it becomes clear that the bound is saturated for both $N_{1/2}(x_0)N_\infty(x_{\pi/2})$ and $N_\infty(x_0)N_{1/2}(x_{\pi/2})$ independent of the value of the squeezing parameter. Furthermore we see that the Shannon entropy power uncertainty relation is the furthest from saturating the bound. These results can be understood by the fact that the non-linear nature of the RE emphasizes different parts of the PDF depending on the RE-parameter i.e for $p > 1$ the more probable parts of the PDF are emphasized while for $p < 1$, the less probable parts are emphasized - this corresponds to the peak and tails respectively. This result is a manifestation of the extra information supplied by the higher order statical moments provided through the Rényi entropy, this point will be expanded upon in the discussion of this section.

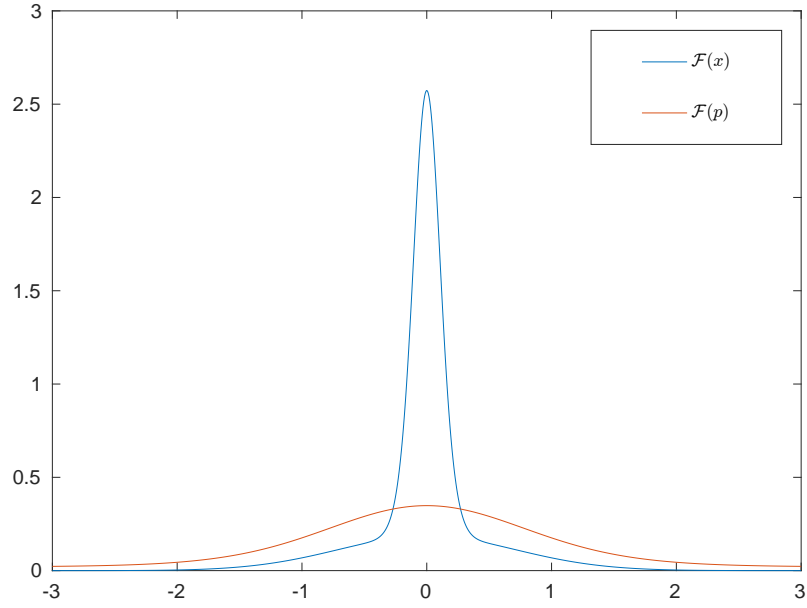


Figure 8.4: Plots of the position and momentum PDFs for the superposition state consisting of the vacuum and squeezed vacuum with $z = 2$.

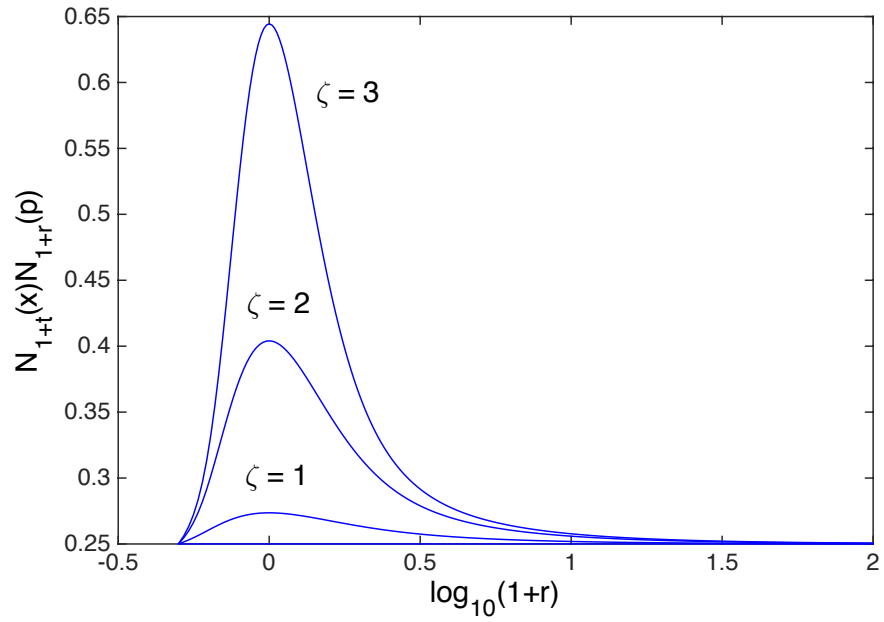


Figure 8.5: Plots of the REPUR (in units of \hbar) for the superposition state consisting of the vacuum and squeezed vacuum against $\log_{10}(1+r)$ for squeezing parameters $\zeta \in \{1, 2, 3\}$. The REPUR bound is saturated in both $N_{\infty}(x)N_{1/2}(p)$ and $N_{1/2}(x)N_{\infty}(p)$ cases. Deviation from this bound saturation is observed for all other values of r with maximal deviation at $r = 0$ corresponding to the Shannon EP.

8.3.3 Uninformative VUR

The superior nature of the entropic uncertainty relations over variance based counterparts is evidenced here. To do so, we evaluate the variance

$$\Delta^2 \hat{X}_\lambda = \langle \hat{X}_\lambda^2 \rangle - \langle \hat{X}_\lambda \rangle^2 \quad (8.98)$$

for $\lambda = 0$ and $\lambda = \pi/2$ where

$$\hat{X}_0^2 = \frac{1}{2}(a^\dagger a^\dagger + 2a^\dagger a + 1 + aa), \quad \hat{X}_{\pi/2}^2 = \frac{1}{2}(2a^\dagger a + 1 - a^\dagger a^\dagger - aa) \quad (8.99)$$

thus

$$\begin{aligned} \langle \hat{X}_0^2 \rangle_{\psi_V} &= \mathcal{N}_V^2 (\langle 0| + \langle \zeta|) \hat{X}_0^2 (|0\rangle + |\zeta\rangle) \\ &= \mathcal{N}_V^2 \left(\langle 0| \hat{X}_0^2 |0\rangle + \langle 0| \hat{X}_0^2 |\zeta\rangle + \langle \zeta| \hat{X}_0^2 |0\rangle + \langle \zeta| \hat{X}_0^2 |\zeta\rangle \right). \end{aligned} \quad (8.100)$$

Considering each term in (8.100) separately and recalling that $a|n\rangle = \sqrt{n}|n-1\rangle$ and $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, we find

$$\begin{aligned} \langle 0| \hat{X}_0^2 |0\rangle &= \frac{1}{2} \langle 0| (a^\dagger a^\dagger + 2a^\dagger a + 1 + aa) |0\rangle \\ &= \frac{1}{2} \end{aligned} \quad (8.101)$$

along with

$$\begin{aligned} \langle 0| \hat{X}_0^2 |\zeta\rangle &= \frac{1}{2} \langle 0| (a^\dagger a^\dagger + 2a^\dagger a + 1 + aa) |\zeta\rangle \\ &= \frac{1}{2} (\langle 0|\zeta\rangle + \langle 0|aa|\zeta\rangle) \\ &= \frac{1}{2} (\langle 0|\zeta\rangle + \sqrt{2}\langle 2|\zeta\rangle) \\ &= \frac{1}{2} \sqrt{\text{sech}(z)} (1 - \tanh(z)) \end{aligned} \quad (8.102)$$

where we have used (2.56) to evaluate the individual overlaps. Next we see that

$$\begin{aligned} \langle \zeta| \hat{X}_0^2 |0\rangle &= \frac{1}{2} \langle \zeta| (a^\dagger a^\dagger + 2a^\dagger a + 1 + aa) |0\rangle \\ &= \frac{1}{2} (\langle \zeta| a^\dagger a^\dagger |0\rangle + \langle \zeta|0\rangle) \\ &= \frac{1}{2} (\langle 0|\zeta\rangle + \sqrt{2}\langle 2|\zeta\rangle)^\dagger \\ &= \frac{1}{2} \sqrt{\text{sech}(z)} (1 - \tanh(z)) \end{aligned} \quad (8.103)$$

and finally, we have from [22] that in general

$$\langle \zeta| \hat{X}_\lambda^2 |\zeta\rangle = \frac{1}{2} \left(e^{2z} \sin^2 \left(\lambda - \frac{\phi}{2} \right) + e^{-2z} \cos^2 \left(\lambda - \frac{\phi}{2} \right) \right) \quad (8.104)$$

so setting $\phi = \lambda = 0$ we have

$$\langle \zeta | \hat{X}_0^2 | \zeta \rangle = \frac{1}{2} e^{-2z}. \quad (8.105)$$

We now have all the parts needed to evaluate (8.100), yielding

$$\langle \hat{X}_0^2 \rangle_{\psi_V} = \mathcal{N}_V^2 \left(\frac{1}{2} (1 + e^{-2z}) + \sqrt{\text{sech}(z)} (1 - \tanh(z)) \right). \quad (8.106)$$

Through similar working we find

$$\langle \hat{X}_{\pi/2}^2 \rangle_{\psi_V} = \mathcal{N}_V^2 \left(\frac{1}{2} (1 + e^{2z}) + \sqrt{\text{sech}(z)} (1 + \tanh(z)) \right). \quad (8.107)$$

Now we consider the second term in (8.98), we have

$$\begin{aligned} \langle \hat{X}_0 \rangle_{\psi_V} &= \mathcal{N}_V^2 (\langle 0 | + \langle \zeta |) \hat{X}_0 (|0\rangle + |\zeta\rangle) \\ &= \mathcal{N}_V^2 (\langle 0 | \hat{X}_0 |0\rangle + \langle 0 | \hat{X}_0 | \zeta \rangle + \langle \zeta | \hat{X}_0 |0\rangle + \langle \zeta | \hat{X}_0 | \zeta \rangle). \end{aligned} \quad (8.108)$$

Again, we shall consider each term separately while noting that we have, from [22],

$\langle \zeta | \hat{X}_\lambda | \zeta \rangle = 0 \Rightarrow \langle \zeta | \hat{X}_0 | \zeta \rangle = \langle 0 | \hat{X}_0 |0\rangle = 0$ along with

$$\begin{aligned} \langle 0 | \hat{X}_0 | \zeta \rangle &= \frac{1}{\sqrt{2}} \langle 0 | a | \zeta \rangle \\ &= \frac{1}{\sqrt{2}} \langle 1 | \zeta \rangle \\ &= 0, \quad \text{as } m \in \mathbb{N} \end{aligned} \quad (8.109)$$

and similarly

$$\langle \zeta | \hat{X}_0 |0\rangle = 0 \quad (8.110)$$

from which we conclude

$$\langle \hat{X}_0 \rangle_{\psi_V} = 0. \quad (8.111)$$

Now considering the case where $\lambda = \pi/2$, we again have $\langle \zeta | \hat{X}_{\pi/2} | \zeta \rangle = \langle 0 | \hat{X}_{\pi/2} |0\rangle = 0$

and

$$\langle 0 | \hat{X}_{\pi/2} | \zeta \rangle = \langle \zeta | \hat{X}_{\pi/2} |0\rangle = 0 \quad (8.112)$$

thus

$$\langle \hat{X}_{\pi/2} \rangle_{\psi_V} = 0. \quad (8.113)$$

So from (8.106), (8.107), (8.111) and (8.113) we are able to determine the variances of interest

$$\begin{aligned}\Delta^2 \hat{X}_0 &= \langle \hat{X}_0^2 \rangle_{\psi_V} - \langle \hat{X}_0 \rangle_{\psi_V}^2 \\ &= \mathcal{N}_V^2 \left(\frac{1}{2}(1 + e^{-2z}) + \sqrt{\text{sech}(z)}(1 - \tanh(z)) \right)\end{aligned}\quad (8.114)$$

and

$$\begin{aligned}\Delta^2 \hat{X}_{\pi/2} &= \langle \hat{X}_{\pi/2}^2 \rangle_{\psi_V} - \langle \hat{X}_{\pi/2} \rangle_{\psi_V}^2 \\ &= \mathcal{N}_V^2 \left(\frac{1}{2}(1 + e^{2z}) + \sqrt{\text{sech}(z)}(1 + \tanh(z)) \right).\end{aligned}\quad (8.115)$$

With these expressions the product of the variances of the conjugate quadrature variables are plotted against the squeezing parameter in Fig.8.6 and it is clear that the bound is saturated for $z = 0$ as expected (since the vacuum state is Gaussian) but blows up rapidly as the squeezing parameter increases rendering the VUR uninformative.

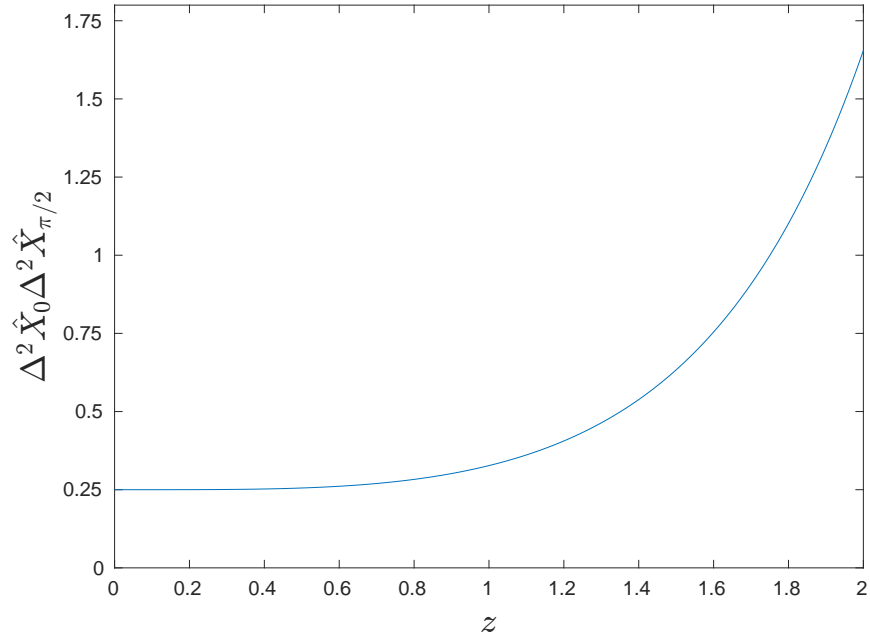


Figure 8.6: Plot of the variance based uncertainty relation for the superposition state of a vacuum and squeezed vacuum. The bound is saturated for $z = 0$ as this corresponds to the Gaussian vacuum state but blows up rapidly as z increases.

8.3.4 Discussion

From the preceding analysis it is evident that the REPURs supply far more information on the underlying PDFs than VURs and Shannon-based entropy power uncertainty relations

alone. Indeed the REPURs are a generalisation of these relations, the non-linear nature of the RE emphasises the more probable parts of a distribution for $p > 1$ and emphasises the less probable parts of a distribution for $p < 1$. Furthermore, it has been shown that the only class of distribution that saturates the REPUR for all values of p are the Gaussian distributions. With these two aspects of the REPUR established, the results displayed in Fig.8.5 can be better understood; for this superposition state of the vacuum and squeezed vacuum, the peak and the tails in both the position and momentum quadrature are close to Gaussian thus the REPUR approaches its lower bound. Moreover, the overall structure is non-Gaussian and it is apparent that for the Shannon entropy the bound is maximally divergent, this is due to the Shannon entropy providing no particular emphasis on any specific part of the PDF. Thus from the REPURs displayed in Fig.8.5 we can glean that the underlying distribution and its Fourier transform is Gaussian in both the peaks and tails but has an overall non-Gaussian Structure which is readily confirmed by inspection of (8.96) and (8.97) and displayed in Fig.8.4.

This interpretation is bolstered by the REPURs of the Cat State depicted in Fig.8.3 we see bound saturation for $r = -1/2$ implying that the peak in the momentum quadrature and the tails in the position quadrature behave as a Gaussian but the divergence of the bound for increasing r implies that the peak in the position quadrature and the tails in the momentum are non-Gaussian. This is easily confirmed by inspection of equations (8.71) and (8.72) along with Fig.2.8 in which the latter case displays a clear non-Gaussian behaviour via the multi-peaked distribution while the former case does indeed exhibit Gaussian behaviour. The overall non-Gaussian structure is also captured here by the SEP being far from the lower bound. This notion of inferring information on quantum mechanical state distributions through the REPs is expanded upon in the following chapter.

Chapter 9

Information Scan and Reconstruction of Quantum States

In the previous chapter, we saw that variance-based uncertainty measures can be rendered uninformative for some quantum states. Furthermore, it became evident that Rényi-based entropic measures can provide far more information on the underlying state of the system under investigation. In this chapter we elaborate on these points and introduce a method of reconstructing the underlying probability distribution of the quantum mechanical state through knowledge of the Rényi entropy powers. The underlying features of this “information scan” are detailed and the method of reconstruction is demonstrated on the associated probability distributions of specific quantum mechanical states.

This chapter is based on the paper:

Information scan of quantum states based on entropy-power uncertainty relations. P Jizba, A J Hayes and J A Dunningham (2018). Submitted to Physical Rev. A.

9.1 Variance is a Deceptive Measure of Uncertainty

In this section we investigate two emblematic situations which exemplify the failings of variance as a measure of uncertainty. We then further this discussion by applying similar reasoning to Cat States which displays features analogous to the heuristic examples.

9.1.1 Example 1

Following the work of [136] we consider a particle whose position is constrained to the x dimension and four boxes of equal size situated along the x axis but not necessarily

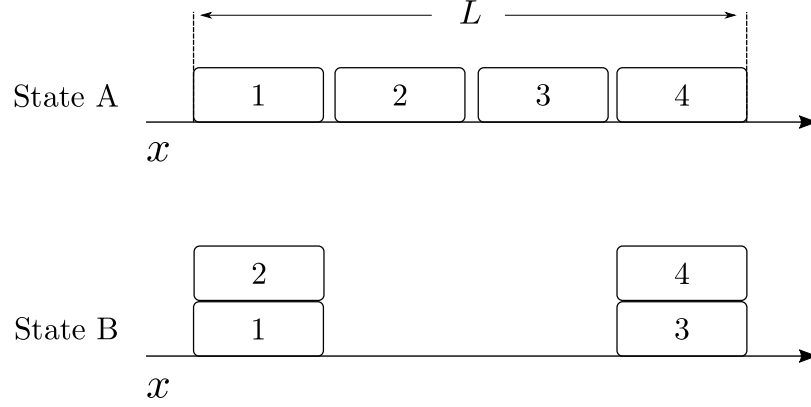


Figure 9.1: A heuristic example depicting the failings of variance as a measure of uncertainty in which a particle in state A can be found over length L and in state B it can be found over a length $L/2$. Clearly we know more about position in state B but the variance suggests otherwise.

constrained to it (i.e the boxes can be stacked on top of each other). Imposing the condition that the particle must be localised to be within one of the boxes, we consider the two arrangements given in Fig.9.1. The wave functions of each of these state configurations are given by

$$\psi_A = \begin{cases} 1/\sqrt{L} & x \in [0, L] \\ 0 & \text{otherwise} \end{cases} \quad \psi_B = \begin{cases} \sqrt{2/L} & x \in [0, L/4] \\ \sqrt{2/L} & x \in [3L/4, L] \\ 0 & \text{otherwise} \end{cases}$$

then, by using

$$\Delta^2 x = \int \psi^*(x, t) x^2 \psi(x, t) dx - \left(\int \psi^*(x, t) x \psi(x, t) dx \right)^2$$

we can find that the variance for each state is given by

$$\Delta^2 x_A = \frac{L^2}{12}, \quad \Delta^2 x_B = \frac{7}{4} \cdot \frac{L^2}{12}.$$

Clearly the uncertainty of the particle's location should be greater in state A than it is in state B yet we have $\Delta^2 x_A < \Delta^2 x_B$. In connection with this, we introduce the Balanced Cat State (BCS)

$$|\psi_B\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\alpha\rangle), \quad (9.1)$$

the PDF of this state can be found by (2.94) and in the position-like quadrature the PDF mimics that of the example Fig.9.1 as demonstrated in Fig.9.2.

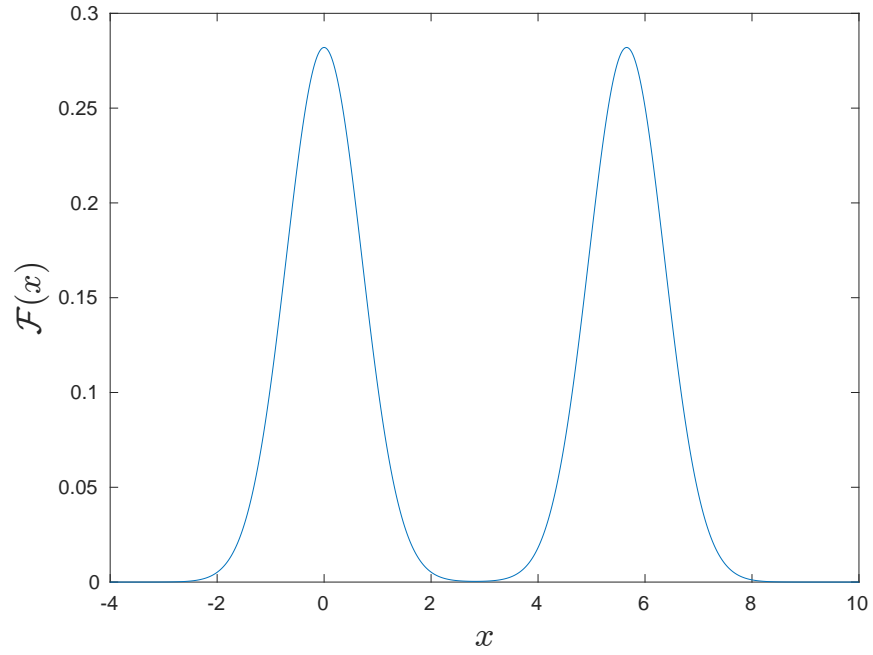


Figure 9.2: Plot of the PDF of the Balanced Cat state in the x -quadrature with $\alpha = 4$. This is a physical analogue of State B.

9.1.2 Example 2

An extremised example that highlights the defects of variance as a measure of uncertainty is depicted in Fig.9.3 in which the position of a particle in the x dimension is again constrained but this time to be within a wide box and a narrow strip. The *probability*

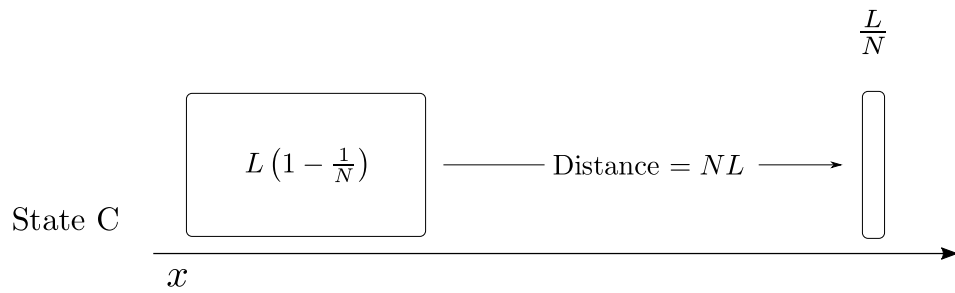


Figure 9.3: An alternative example depicting the failings of variance as a measure of uncertainty.

distribution of this configuration is given by

$$|\psi_C(x)|^2 = \begin{cases} 1/L & x \in \text{the wide box} \\ 1/L & x \in \text{the narrow strip} \\ 0 & \text{elsewhere} \end{cases}$$

then using (9.1) we find

$$\Delta^2 x_C = L^2 \left(N - \frac{1}{N} + \frac{1}{12} \right) \quad (9.2)$$

thus the variance tends to infinity with increasing N while the probability of finding the particle in the wide box should tend toward certainty or in other words, the variance increases as the uncertainty should decrease! In connection with this we introduce the unbalanced cat state (UCS)

$$|\psi_{UCS}\rangle = \mathcal{N}_{UCS} (|0\rangle + \nu |\alpha/\nu\rangle). \quad (9.3)$$

The PDF of this state can also be found using (2.94) and as displayed in Fig.9.4, the position quadrature PDF is analogous to that of the example illustrated in Fig.9.3. The

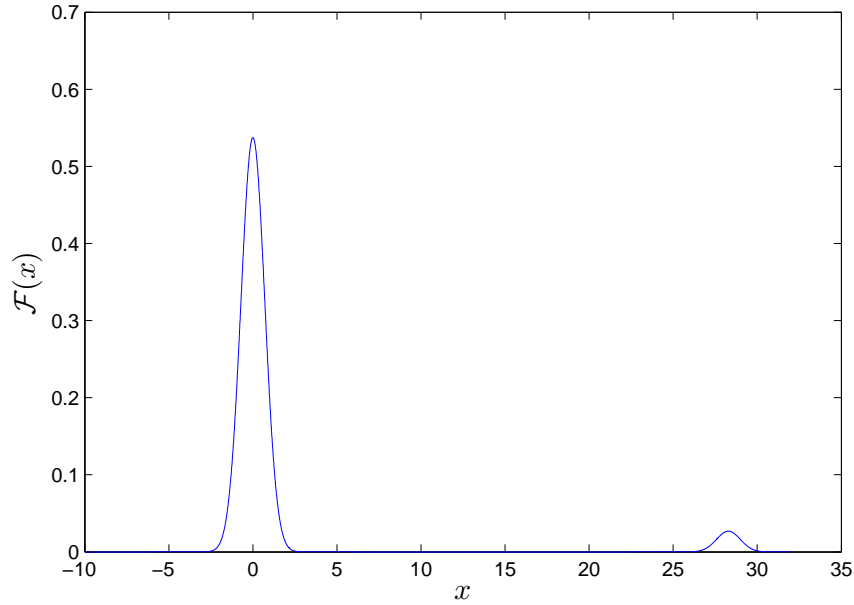


Figure 9.4: Plot of the PDF for the unbalanced cat state in the x -quadrature with $\alpha = 1$ and $\nu = 0.2$. This is a physical analogue of State C.

UCS is a superposition of the vacuum state and a small coherent state (displaced vacuum). The coherent state is weighted by a factor of ν so choosing this such that $\nu^2 \ll 1$ we are far more likely to observe the UCS positioned around its constituent vacuum state than its coherent state. Here the vacuum state is analogous to the wide box in Fig.9.3 and the coherent state, with almost negligible weighting, is analogous to the narrow strip and the parameter N is the inverse analogue of the weighting ν in that *decreasing* ν results in $\Delta^2 x \rightarrow \infty$ as the peaks separate (both vertically and horizontally).

9.1.3 Entropies of Heuristic Examples

Since the information entropy of a random variable tells us the number of binary question (in base 2) needed to fully determine the random variable, the lack of information, i.e. uncertainty, can be depicted by simply reversing the sign of the information entropy. Taking the “coarse-grained” approach of Bialynicki-Birula [136] in which the region of interest is divided into bins of size δx (this bin size can be regarded as the experimental error), the probability of finding the particle in the i -th bin is given by

$$q_i = \int_{(i-1/2)\delta x}^{(i+1/2)\delta x} |\psi(x)|^2 dx \quad (9.4)$$

and the corresponding Shannon entropy is given by

$$H^{(x)} = - \sum_i q_i \ln(q_i). \quad (9.5)$$

Clearly in Example 1 the bin size is $\delta x = L/4$ and the probability distribution of State A is given by $|\psi_A(x)|^2 = 1/L$ while that of State B is given by $|\psi_B(x)|^2 = 2/L$. Applying (9.4) to State A yields

$$\begin{aligned} q_i^A &= \int_{(i-1/2)L/4}^{(i+1/2)L/4} \frac{1}{L} dx \\ &= \frac{1}{L} \cdot \frac{L}{4} \left(i + \frac{1}{2} - i + \frac{1}{2} \right) \\ &= \frac{1}{4} \end{aligned} \quad (9.6)$$

where $i \in \{1, 2, 3, 4\}$. Then the Shannon Entropy for State A is given by

$$\begin{aligned} H_A^{(x)} &= - \sum_{i=1}^4 \frac{1}{4} \ln \left(\frac{1}{4} \right) \\ &= -4 \cdot \frac{1}{4} \ln \left(\frac{1}{4} \right) = -\ln \left(\frac{1}{4} \right) = \ln(2^2) = 2 \ln(2) \end{aligned} \quad (9.7)$$

while the probability of finding the particle in the i -th bin for State B is given by

$$\begin{aligned} q_i^B &= \int_{(i-1/2)L/4}^{(i+1/2)L/4} \frac{2}{L} dx \\ &= \frac{2}{L} \cdot \frac{L}{4} \left(i + \frac{1}{2} - i + \frac{1}{2} \right) \\ &= \frac{1}{2} \end{aligned} \quad (9.8)$$

where $i \in \{1, 2\}$. The Shannon entropy is then found to be

$$\begin{aligned} H_B^{(x)} &= - \sum_{i=1}^2 \frac{1}{2} \ln \left(\frac{1}{2} \right) \\ &= -2 \cdot \frac{1}{2} \ln \frac{1}{2} = \ln(2). \end{aligned} \quad (9.9)$$

It is evident that $H_A^{(x)} = 2 \cdot H_B^{(x)}$ i.e the uncertainty of the particles position in state A is twice that of State B - a much more intuitive result than that of the variance-based measure! The foundation of this result is the invariance of entropies under piecewise rearrangements of the probability distributions, a useful property which we investigate further and utilize in the following.

9.2 The Information Scan

Here we show that knowledge of the family of Rényi entropies allows for the reconstruction of the underlying PDFs. For this, it is helpful to use the following form of the continuous Rényi entropy

$$\mathcal{I}_p(\chi) = \frac{1}{(1-p)} \log_2 \mathbb{E} \left[2^{(1-p)i_\chi} \right] \quad (9.10)$$

where, $\mathbb{E}[\cdot]$ is the expectation value and $i_\chi(\vec{y}) = -\log_2(\mathcal{F}(\vec{y}))$ is the information in \vec{y} with respect to the PDF $\mathcal{F}(\vec{y})$. From this it is clear that the continuous Rényi entropy can be viewed as a reparametrized version of the cumulant generating function of the information random variable $i_\chi(\vec{y})$. We can draw the connection between the familiar PDF $F(\vec{y})$ and the information random variable's PDF, which we shall denote as $g(x)$, by first considering the cumulative distribution function (CDF) of $i_\chi(\vec{y})$

$$f(x) = \int_{-\infty}^x df(i_\chi) = \int_{\mathbb{R}^D} \mathcal{F}(\vec{y}) \theta[\log_2(\mathcal{F}(\vec{y})) + x] d\vec{y} \quad (9.11)$$

where $\theta[\cdot]$ is the Heaviside step function. It then becomes apparent that the CDF (9.11) is effectively a discontinuous function given by

$$f(x) = \begin{cases} \int_{\mathbb{R}^D} \mathcal{F}(\vec{y}) & \text{if } \mathcal{F}(\vec{y}) \geq 2^{-x} \\ 0 & \text{otherwise.} \end{cases}$$

Upon inspection of this we have that (9.2) in fact requires us to integrate the PDF $\mathcal{F}(\vec{y})$ between the solutions of

$$\begin{aligned} \log_2(\mathcal{F}(\vec{y}_{\text{limits}})) + x &= 0 \\ \implies \vec{y}_{\text{limits}} &= \mathcal{F}^{-1}(2^{-x}). \end{aligned} \quad (9.12)$$

This is shown pictorially (with $D = 1$) in Fig.9.5. The interpretation of (9.11) can be alternatively rephrased as the probability that the random variable $i_\chi(\vec{y})$ is less than or equal to x with $df(i_\chi)$ denoting the probability measure. The Laplace transform of a function $F(t)$ is given by

$$\mathcal{L}\{F(t)\}(s) \equiv \int_0^\infty F(t) e^{-st} dt \quad (9.13)$$

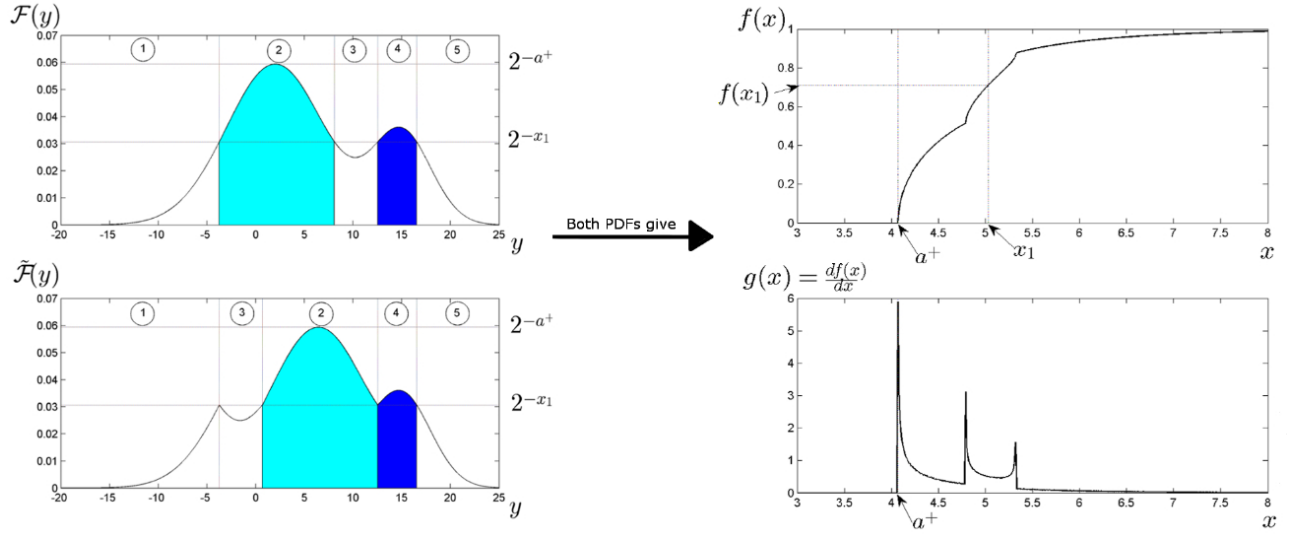


Figure 9.5: Information scan of $\mathcal{F}(y)$ and the equimeasurably rearranged PDF $\tilde{\mathcal{F}}(y)$. Cumulative distribution function $f(x)$ measures the area of $\mathcal{F}(y)$ within the limits dictated by the intercept line $\mathcal{F}(y) = 2^{-x}$ with $\mathcal{F}(y)$ (shaded area). For the entropy measured in nats $2^{-x} \rightarrow e^{-x}$. The information PDF $g(x)$ represents the rate of change of the area of the cumulative distribution $f(x)$. Note that $f(x)$ and $g(x)$ are identical for the equimeasurably rearranged PDFs $\mathcal{F}(y)$ and $\tilde{\mathcal{F}}(y)$. The 3-peak structure of $g(x)$ is one of the invariant characteristics of the equimeasurable family of PDFs.

thus taking the Laplace transform of (9.11), we have

$$\begin{aligned} \mathcal{L}\{f\}(s) &= \int_{\mathbb{R}^D} \mathcal{F}(\vec{y}) \frac{1}{s} \exp(s \log_2(\mathcal{F}(\vec{y}))) d\vec{y} \\ &= \frac{1}{s} \mathbb{E} [\exp(s \log_2(\mathcal{F}))]. \end{aligned} \quad (9.14)$$

The PDF of the random variable is then given by

$$g(x) = \frac{df(x)}{dx} = \mathcal{L}^{-1} \{ \mathbb{E} [\exp(s \log_2(\mathcal{F}))] \}(x) \quad (9.15)$$

and thus

$$\mathcal{L}\{g\}(s = (p-1) \ln(2)) = \mathbb{E} [2^{(1-p)i_x}] \quad (9.16)$$

then comparing equation (9.16) with (9.10) and the usual expression for the Rényi entropy

$$\mathcal{I}_p(\chi) = \frac{1}{1-p} \log_2 \left(\int_M d\vec{y} \mathcal{F}^p(\vec{y}) \right) \quad (9.17)$$

it is clear that (9.16) can be equivalently expressed as

$$\int_{\mathbb{R}} g(x) 2^{(1-p)x} dx = \int_{\mathbb{R}^D} \mathcal{F}^p(\vec{y}) d\vec{y}. \quad (9.18)$$

Now, since $\mathcal{L}\{g\}(s)$ is the moment-generating function of the random variable $i_\chi(\vec{y})$ with the associated PDF $g(x)$ one can find all moments (assuming they exist) by taking the derivatives of $\mathcal{L}\{g\}$ w.r.t s and furthermore, we can find all cumulants for the random variable $i_\chi(\vec{y})$ and the associated PDF $g(x)$ by taking the binary logarithm of both sides of (9.16) and the subsequent derivatives. From (9.16) it is then clear that these cumulants can be rephrased in terms of $\mathbb{E}[(\log_2 \mathcal{F})^m]$ which is in close relation to the continuous Rényi entropy, as displayed by (9.10). Lastly, using the identity

$$\begin{aligned} \mathcal{I}_p(\chi) &= \mathcal{I}_p \left(\sqrt{N_p(\chi)} \cdot \mathcal{Z}_G \right) \\ &= \frac{D}{2} \log_2 \left[2\pi p^{-1/(p+1)} N_p(\chi) \right] \end{aligned} \quad (9.19)$$

it is inferred that the continuous Rényi entropy can be phrased in terms of the entropy powers N_p of \mathcal{F} . Hence, the right hand side of (9.16) provides the means to obtain the PDF $\mathcal{F}(\vec{y})$ through rephrasing in terms of the entropy powers moreover, it is clearly related to the PDF $g(x)$ - all the statistical information of which is available through the moments, or equivalently, the cumulants. In Fig.9.6 we plot $f(x)$ and $g(x)$ for the state $|\psi_V\rangle = \mathcal{N}_V(|0\rangle + |\zeta\rangle)$ as introduced in section 8.3.2. From the multi-peak structure of $g(x)$ one can determine the number and height of the stationary points. These are invariant characteristics of a given family of *equimeasurable* PDFs where the notion of equimeasurably rearranged PDFs $\mathcal{F}(y)$ and $\tilde{\mathcal{F}}(y)$ is depicted in Fig.9.5; $f(x)$ and $g(x)$ are identical for such distributions. With this relationship between $g(x)$ and $\mathcal{F}(\vec{y})$ established, it is evident that complete knowledge of $g(x)$ gives an information scan of the PDF $\mathcal{F}(\vec{y})$. Our aim is now to construct the information PDF $g(x)$ from knowledge of the Rényi entropy powers.

9.3 Reconstruction of Quantum States

Here we show that given measured Rényi Entropy powers (REPs) we are able to find the cumulants of the PDF of the information random variable $i_\chi(\vec{y})$. Using these cumulants, the information PDF can be reconstructed through the Gram-Charlier A series expansion [137]. Furthermore, a direct link has been established between the information PDF and the class of equimeasurable PDFs, in which the PDF of interest $\mathcal{F}(y)$ exists. We use the following equation to reconstruct the information PDF

$$g(x) = \exp \left[\sum_{k=2}^{\infty} (\kappa_k - \gamma_k) (-1)^k \frac{d^k/dx^k}{k!} \right] G(x) \quad (9.20)$$

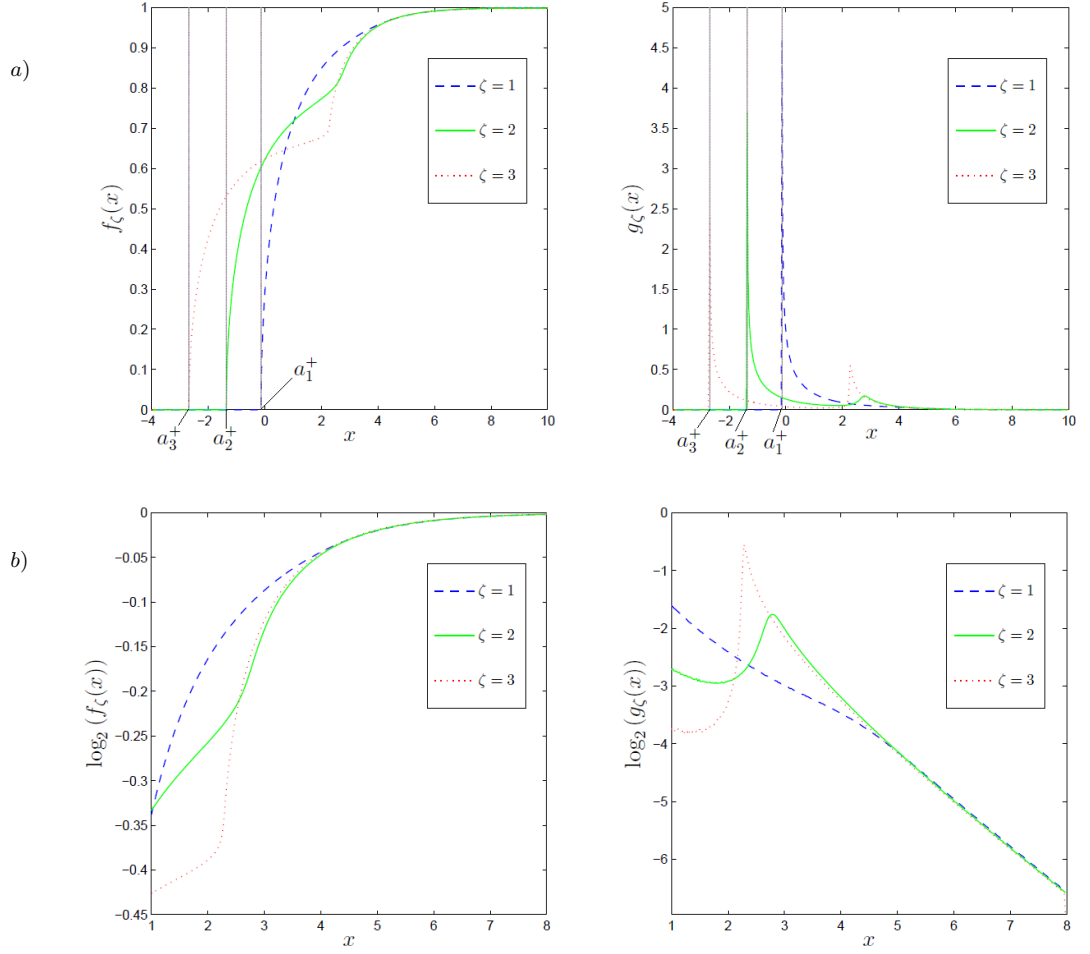


Figure 9.6: (a) Cumulative distribution functions $f_\zeta(x)$ and information PDFs $g_\zeta(x)$ for the superposition state of a vacuum with a squeezed vacuum. (b) The logarithmic scaling depicting the tail behaviour (corresponding to $x > 1$) of $\mathcal{F}(y) = |\psi_V(y)|^2$. From (a) it is apparent that the larger the value of ζ , the higher the peak of $\mathcal{F}(y)$, and the peaks have heights $2^{-a_\zeta^+}$. For $\zeta = 2$ and $\zeta = 3$ we see a second peak near $x = 3$ which is due to a sharp change in the overall shape of the PDF at height $\mathcal{F}(y) = 2^{-x}$. From (b) the tail behaviour is displayed and the best-fit analysis reveals the tails to be Gaussian.

where, $G(x)$ is a known “reference” PDF whose cumulants are given by γ_k , the quantities κ_n are the cumulants of the $g(x)$. Expanding the exponential of (9.20) gives

$$\begin{aligned}
 g(x) = & \left[1 + \left(\sum_{k=2}^{\infty} (\kappa_k - \gamma_k) (-1)^k \frac{d^k/dx^k}{k!} \right) + \frac{1}{2!} \left(\sum_{k=2}^{\infty} (\kappa_k - \gamma_k) (-1)^k \frac{d^k/dx^k}{k!} \right)^2 \right. \\
 & \left. + \frac{1}{3!} \left(\sum_{k=2}^{\infty} (\kappa_k - \gamma_k) (-1)^k \frac{d^k/dx^k}{k!} \right)^3 + \dots \right] G(x)
 \end{aligned} \tag{9.21}$$

and going a step further by expanding the sum in each of the terms of order greater than or equal to 2 gives

$$g(x) = \left[1 + \left((\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2} \frac{d^2}{dx^2} + (\kappa_3 - \gamma_3)(-1)^3 \frac{1}{3!} \frac{d^3}{dx^3} + \dots \right) \right. \\ \left. + \frac{1}{2} \left((\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2} \frac{d^2}{dx^2} + (\kappa_3 - \gamma_3)(-1)^3 \frac{1}{3!} \frac{d^3}{dx^3} + \dots \right)^2 \right. \\ \left. + \frac{1}{3!} \left((\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2} \frac{d^2}{dx^2} + (\kappa_3 - \gamma_3)(-1)^3 \frac{1}{3!} \frac{d^3}{dx^3} + \dots \right)^3 + \dots \right] G(x) \quad (9.22)$$

then by labelling

$$t_2 = (\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2} \frac{d^2}{dx^2} \\ t_3 = (\kappa_3 - \gamma_3)(-1)^3 \frac{1}{3!} \frac{d^3}{dx^3} \\ \vdots \\ t_m = (\kappa_m - \gamma_m)(-1)^m \frac{1}{m!} \frac{d^m}{dx^m} \quad (9.23)$$

it is apparent that $g(x)$ approximated to m th order in the sum and N th order in the exponential is given by

$$g(x) \approx \left[1 + (t_2 + t_3 + \dots t_m) + \frac{1}{2!} (t_2 + t_3 + \dots t_m)^2 + \frac{1}{3!} (t_2 + t_3 + \dots t_m)^3 + \dots \right. \\ \left. \dots + \frac{1}{N!} (t_2 + t_3 + \dots t_m)^N \right] G(x). \quad (9.24)$$

The multinomial theorem states

$$(t_2 + t_3 + \dots + t_m)^n = \sum_{l_2, l_3, \dots, l_m} \frac{n!}{l_2! l_3! \dots l_m!} t_2^{l_2} t_3^{l_3} \dots t_m^{l_m} \quad (9.25)$$

where, $l_2 + l_3 + \dots + l_m = n$. Then $g(x)$ can be approximated by

$$g(x) \approx \left[1 + \sum_{n_1=0}^1 \sum_{n_3=0}^{1-n_2} \dots \sum_{n_m=0}^{1-n_2-\dots-n_{m-1}} \frac{1! t_2^{n_2} t_3^{n_3} \dots t_{m-1}^{n_{m-1}} t_m^{1-n_2-n_3-\dots-n_{m-1}}}{n_2! n_3! \dots n_{m-1}! (1-n_2-n_3-\dots-n_{m-1})!} + \right. \\ \left. + \frac{1}{2} \sum_{n_2=0}^2 \sum_{n_3=0}^{2-n_2} \dots \sum_{n_m=0}^{2-n_2-\dots-n_{m-1}} \frac{2! t_2^{n_2} t_3^{n_3} \dots t_{m-1}^{n_{m-1}} t_m^{2-n_2-n_3-\dots-n_{m-1}}}{n_2! n_3! \dots n_{m-1}! (2-n_2-n_3-\dots-n_{m-1})!} + \right. \\ \vdots \\ \left. + \frac{1}{N!} \sum_{n_2=0}^N \sum_{n_3=0}^{N-n_2} \dots \sum_{n_m=0}^{N-n_2-\dots-n_{m-1}} \frac{N! t_2^{n_2} t_3^{n_3} \dots t_{m-1}^{n_{m-1}} t_m^{N-n_2-n_3-\dots-n_{m-1}}}{n_2! n_3! \dots n_{m-1}! (N-n_2-n_3-\dots-n_{m-1})!} \right] G(x) \quad (9.26)$$

and by using (9.23) to rewrite in terms of cumulants, $g(x)$ is finally approximated by

$$\begin{aligned}
g(x) \approx & \left[1 + \sum_{n_2=0}^1 \sum_{n_3=0}^{1-n_2} \cdots \sum_{n_m=0}^{1-n_2-\dots-n_{m-1}} \frac{1! [(\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2!}]^{n_2} \cdots [(\kappa_m - \gamma_m)(-1)^m \frac{1}{m!}]^{1-n_2-\dots-n_{m-1}}}{n_2! \cdots n_{m-1}! (1 - n_2 - \dots - n_{m-1})!} \right. \\
& \times \frac{d^{(2n_2+3n_3+\dots+m(1-n_2-\dots-n_{m-1}))}}{dx^{(2n_2+\dots+m(1-n_2-\dots-n_{m-1}))}} \\
& + \frac{1}{2} \sum_{n_2=0}^2 \sum_{n_3=0}^{2-n_2} \cdots \sum_{n_m=0}^{2-n_2-\dots-n_{m-1}} \frac{2! [(\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2!}]^{n_2} \cdots [(\kappa_m - \gamma_m)(-1)^m \frac{1}{m!}]^{2-n_2-\dots-n_{m-1}}}{n_2! \cdots n_{m-1}! (2 - n_2 - \dots - n_{m-1})!} \\
& \times \frac{d^{(2n_2+3n_3+\dots+m(2-n_2-\dots-n_{m-1}))}}{dx^{(2n_2+\dots+m(2-n_2-\dots-n_{m-1}))}} \\
& \vdots \\
& + \frac{1}{N!} \sum_{n_2=0}^N \sum_{n_3=0}^{N-n_2} \cdots \sum_{n_m=0}^{N-\dots-n_{m-1}} \frac{N! [(\kappa_2 - \gamma_2)(-1)^2 \frac{1}{2!}]^{n_2} \cdots [(\kappa_m - \gamma_m)(-1)^m \frac{1}{m!}]^{N-\dots-n_{m-1}}}{n_2! \cdots n_{m-1}! (N - n_2 - \dots - n_{m-1})!} \\
& \times \frac{d^{(2n_2+3n_3+\dots+m(N-n_2-\dots-n_{m-1}))}}{dx^{(2n_2+\dots+m(N-n_2-\dots-n_{m-1}))}} \Big] G(x) \quad (9.27)
\end{aligned}$$

in other words, this is the expanded form of $g(x)$ approximated up to the m th cumulant and the N th order term of the exponential.

9.3.1 Relating REPs to Cumulants

To relate κ_n to the REPs we begin with the cumulant expansion of the RE

$$p\mathcal{I}_{1-p} = \log_2(e) \sum_{n=1}^{\infty} \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2(e)} \right)^n \quad (9.28)$$

and use the identity that relates the RE to the REP

$$\begin{aligned}
\mathcal{I}_{1-p}(\chi) &= \frac{D}{2} \log_2 \left[2\pi(1-p)^{-1/p} N_{1-p}(\chi) \right] \\
&\Rightarrow \frac{D}{2} \log_2(N_{1-p}(\chi)) = \mathcal{I}_{1-p}(\chi) - \frac{D}{2} \log_2 \left[\frac{2\pi}{(1-p)^{1/p}} \right] \\
&\Rightarrow \log_2(N_{1-p}(\chi)) = \frac{2}{pD} \log_2(e) \sum_{n=1}^{\infty} \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2(e)} \right)^n + \log_2 \left[\frac{(1-p)^{1/p}}{2\pi} \right] \quad (9.29)
\end{aligned}$$

then taking the $(n-1)$ th derivative with respect to p , at $p=0$ of both sides

$$\begin{aligned}
& \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \{ \log_2(N_{1-p}(\chi)) \} = \\
&= \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \log_2 \left[\frac{(1-p)^{1/p}}{2\pi} \right] \right\} + \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \frac{2}{D} \sum_{n=1}^{\infty} \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2(e)} \right)^{n-1} \right\} \\
&= \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \log_2 \left[\frac{(1-p)^{1/p}}{2\pi} \right] \right\} + \frac{2}{D} \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \kappa_1(\chi) + \frac{\kappa_2(\chi)}{2} \left(\frac{p}{\log_2(e)} \right) \right. \\
&\quad \left. + \frac{\kappa_3(\chi)}{3!} \left(\frac{p}{\log_2(e)} \right)^2 + \dots + \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2(e)} \right)^{n-1} + \frac{\kappa_{n+1}(\chi)}{(n+1)!} \left(\frac{p}{\log_2(e)} \right)^n + \dots \right\} \quad (9.30)
\end{aligned}$$

where we have for the first term in (9.30)

$$\begin{aligned}
\left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \log_2 \left[\frac{(1-p)^{1/p}}{2\pi} \right] \right\} &= \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \log_2 \left[(1-p)^{1/p} \right] - \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \log_2(2\pi) \\
&= -\frac{(n-1)!}{n} \log_2(e) - \delta_{1,n} \log_2(2\pi) \\
&= -\log_2(e) \left(\frac{(n-1)!}{n} + \delta_{1,n} \ln(2\pi) \right) \tag{9.31}
\end{aligned}$$

and for the second term

$$\begin{aligned}
\left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} (p^{n-1}) &= (n-1)! \\
\Rightarrow \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} (p^n) &= (n-1)!p = 0 \\
\Rightarrow \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} (p^{n+1}) &= (n-1)!p^2 = 0 \dots \tag{9.32}
\end{aligned}$$

and of course

$$\left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} (p^k) = 0 \quad \forall k < n-1 \tag{9.33}$$

which gives

$$\begin{aligned}
\left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \left\{ \frac{2}{D} \sum_{n=1}^{\infty} \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2(e)} \right)^{n-1} \right\} &= \frac{2}{D} \frac{\kappa_n(\chi)}{n!} \frac{1}{(\log_2(e))^{n-1}} (n-1)! \\
&= \frac{2}{nD} \frac{\kappa_n(\chi)}{(\log_2(e))^{n-1}} \tag{9.34}
\end{aligned}$$

thus

$$\left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \{ \log_2(N_{1-p}(\chi)) \} = -\log_2(e) \left(\frac{(n-1)!}{n} + \delta_{1,n} \ln(2\pi) \right) + \frac{2}{nD} \frac{\kappa_n(\chi)}{(\log_2(e))^{n-1}} \tag{9.35}$$

and rearranging for $\kappa_n(\chi)$ (and noting that $\delta_{1,n}(\log_2(e))^n = \delta_{1,n} \log_2(e)$) gives

$$\kappa_n(\chi) = \frac{nD}{2} (\log_2(e))^{n-1} \left. \frac{d^{n-1}}{dp^{n-1}} \right|_{p=0} \{ \log_2(N_{1-p}(\chi)) \} + \frac{nD}{2} (\log_2(e))^n \left(\frac{(n-1)!}{n} + \delta_{1,n} \ln(2\pi) \right). \tag{9.36}$$

Due to the complexity of taking a direct analytic derivative in first term of (9.36) we turn to Newton's difference quotient to obtain a more manageable expression for the cumulants

$$\frac{f(a+h) - f(a)}{h} \tag{9.37}$$

which in the limit of $h \rightarrow 0$ represents the derivative of f at a . From this we extrapolate that the second derivative is given by

$$\begin{aligned} f''(x) &= \lim_{h_2 \rightarrow 0} \frac{\lim_{h_1 \rightarrow 0} \frac{f(x+h_1+h_2) - f(x+h_1)}{h_1} - \lim_{h_1 \rightarrow 0} \frac{f(x+h_1)}{h_1}}{h_2} \\ &= \lim_{h \rightarrow 0} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2} \end{aligned} \quad (9.38)$$

where, we have assumed h_1 and h_2 converge synchronously. Extrapolating further, we have for the n th derivative

$$f^{(n)}(x) = \lim_{h \rightarrow 0} \frac{1}{h^n} \sum_{m=0}^n (-1)^m \binom{n}{m} f(x + (n-m)h). \quad (9.39)$$

Equation (9.39) is known as the Grünwald-Letnikov derivative formula which, applied to (9.36) and noting $n\delta_{1,n} = \delta_{1,n}$ yields the form of the cumulants used in the following to relate them to REPs

$$\kappa_n(\chi) = \lim_{\Delta \rightarrow 0} \frac{nD}{2} \frac{(\log_2 e)^n}{h^{n-1}} \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \ln [N_{1-(n-1-k)h}(\chi)] + \frac{D}{2} (\log_2 e)^n [(n-1)! + \delta_{1,n} \ln(2\pi)]. \quad (9.40)$$

An alternative method for obtaining accurate numerical approximations of higher order derivatives is given in Appendix.A which can be applied directly to equation 9.36.

So far we have shown how the PDF of interest $\mathcal{F}(y)$ relates to the information PDF $g(x)$ and that it's associated cumulants $\kappa_n(\chi)$ can be expressed in terms of the REPs $N_p(\chi)$. Thus in principle, given a set of REPs, one can approximate the information PDF associated with the equimeasurable set in which the underlying PDF of interest exists. This is demonstrated in the following section by applying this method to the states introduced in section 9.1.

9.4 BCS (Equimeasurable to Gaussian) Example

To gain further insight into the working of this information scan technique we consider an example where $\mathcal{F}(y)$ is taken to be that of the BCS. This is in fact a piecewise rearrangement of a Gaussian PDF yet has an overall, double peaked non-Gaussian structure, as depicted in Fig.9.2, thus $N_{1-p}(\chi) = \sigma^2$, $\forall p$, where σ^2 is the variance of the ‘would be Gaussian’. It is sufficient to analyse a Gaussian PDF of zero mean and variance σ^2 .

Firstly, the analytic form of the information PDF is found by using

$$g(x) = \frac{df(x)}{dx} = \frac{2\sigma^2}{\log_2(e)} \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \frac{[\delta(y - \sigma\sqrt{z(x)}) + \delta(y + \sigma\sqrt{z(x)})]}{2\sigma\sqrt{z(x)}} dy = \frac{2e^{-z(x)/2}}{\log_2(e)\sqrt{2\pi z(x)}} \quad (9.41)$$

where $z(x) = 2x/\log_2(e) - \log(2\pi\sigma^2)$. This is known as the shifted gamma distribution.

The cumulants of the shifted gamma distribution are [138]

$$\kappa_n = \begin{cases} \frac{1}{2} \log_2(e) + \frac{1}{2} \log_2(2\pi\sigma^2) & \text{for } n = 1 \\ \frac{1}{2} (\log_2(e))^n \Gamma(n) & \text{for } n \geq 2. \end{cases}$$

We now use the Gram-Charlier A series (9.20) to reconstruct the information PDF and demonstrate the tail convergence of the reconstructed PDF to the analytic solution. Taking the reference PDF to be a shifted gamma distribution

$$G(x) = \mathcal{G}(x|a, \alpha, \beta) = \frac{e^{-(x-a)/\beta} (x-a)^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \quad (9.42)$$

with $a = \log_2(2\pi\tilde{\sigma}^2)$, $\alpha = 1/2$ and $\beta = \log_2(e)$ where $\tilde{\sigma} \neq \sigma$, we can use the fact that

$$\frac{d^k}{dx^k} \mathcal{G}(x|a, 1/2, \beta) = \frac{k!}{\sqrt{\beta}(x-a)^k} L_k^{(-1/2-k)} \left(\frac{x-a}{\beta} \right) \mathcal{G}(x|a, 1/2, \beta) \quad (9.43)$$

where L_p^q is an associated Laguerre polynomial of order p with parameter q . To first order in the exponential function in the expression for the Gram Charlier A series (9.20), the expression for the information PDF becomes

$$g(x) = \mathcal{G}(x|a, 1/2, \beta) \left[1 + \frac{(\kappa_2 - \gamma_2)}{\sqrt{\beta}(x-a)^2} L_2^{-5/2} \left(\frac{x-a}{\beta} \right) - \frac{(\kappa_3 - \gamma_3)}{\sqrt{\beta}(x-a)^3} L_3^{-7/2} \left(\frac{x-a}{\beta} \right) + \dots \right]. \quad (9.44)$$

Adding higher order terms in the expansion of the exponential function gives a tighter convergence of the reconstructed PDF and since only the first order cumulants contains non-trivial information on the shape of the Gaussian, the reconstructed PDF can be rewritten as

$$g(x) = \left[\sum_{m=0}^{\infty} \frac{1}{m!} \left(-(\kappa_1 - \gamma_1) \frac{d}{dx} \right)^m \right] G(x). \quad (9.45)$$

The affect of adding higher order terms in the expansion of the exponential function is shown in Fig.9.7. Note that in practise we approximate the information PDF, to first order, by a Gaussian so that the first cumulant of the reference PDF $G(x)$ is exactly that of the information PDF.

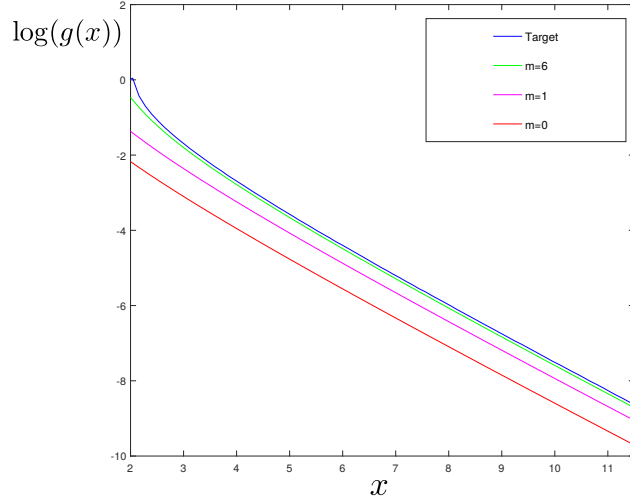


Figure 9.7: Tail convergence of the reconstructed information distribution of a Gaussian for higher order terms of the exponential expansion denoted by m . Only the first cumulant contributes to the reconstruction of a Gaussian PDF.

9.5 Analytic Expressions for the Cumulants of the UCS

The PDF of the general Cat State is given by (2.94), by taking $\theta = 0$, relabelling $x_0 = y$ and $\alpha \in \mathbb{R}$, we find the PDF of the UCS to be

$$\mathcal{F}(y) = \frac{\mathcal{N}^2}{\sqrt{\pi}} \left(e^{-y^2} + \nu^2 e^{-(y-\sqrt{y}\alpha/\nu)^2} + 2\nu e^{-y^2/2} e^{-\frac{1}{2}(y-\sqrt{2}\alpha/\nu)^2} \right) \quad (9.46)$$

then by completing the square of the argument of the last term in (9.46)

$$\begin{aligned} \frac{-y^2}{2} - \frac{1}{2} \left(y - \frac{\sqrt{2}\alpha}{\nu} \right)^2 &= -y^2 + \sqrt{2}y\frac{\alpha}{\nu} - \frac{\alpha^2}{\nu^2} \\ &= - \left(y - \frac{\alpha}{\sqrt{2}\nu} \right)^2 - \frac{\alpha^2}{2\nu^2} \end{aligned} \quad (9.47)$$

thus (9.46) can be expressed as

$$\mathcal{F}(y) = \frac{\mathcal{N}^2}{\sqrt{\pi}} \left(e^{-y^2} + \nu^2 e^{-(y-\sqrt{y}\alpha/\nu)^2} + 2\nu e^{-\frac{\alpha^2}{2\nu^2}} e^{\left(y - \frac{\alpha}{\sqrt{2}\nu}\right)^2} \right) \quad (9.48)$$

which is evidently a weighted sum of three Gaussian's. As we have seen in section 9.4, for Gaussian PDFs the analytic form of it's reconstructed information PDF $g(x)$ is the shifted gamma distribution. Similarly, the reconstructed information PDF of (9.48) is a weighted sum of three shifted gamma distributions

$$g(x) = \frac{\mathcal{N}^2}{\sqrt{\pi}} \left(\mathcal{G}(x|a_1, \tilde{\alpha}, \beta) + \nu^2 \mathcal{G}(x|a_2, \tilde{\alpha}, \beta) + 2\nu e^{-\alpha^2/2\nu^2} \mathcal{G}(x|a_3, \tilde{\alpha}, \beta) \right). \quad (9.49)$$

The k th moment of the *unshifted* gamma distribution $\mathcal{G}(x|0, \tilde{\alpha}, \beta)$ is known to be

$$\mathbb{E}[x^k] = \frac{\Gamma(\tilde{\alpha} + k)}{\Gamma(\tilde{\alpha})} = (\log_2)^k \frac{\Gamma(k + 1/2)}{\sqrt{\pi}} \quad (9.50)$$

which can be transformed to give the moments of the *shifted* gamma distribution $\mathcal{G}(x|a, \tilde{\alpha}, \beta)$ via

$$\begin{aligned} \mathbb{E}_s[x^k] &= \int_a^\infty x^k \mathcal{G}(x|0, \tilde{\alpha}, \beta) dx \\ &= \int_0^\infty (\tilde{x} + a)^k \mathcal{G}(\tilde{x}|0, \tilde{\alpha}, \beta) d\tilde{x} \\ &= \sum_{l=0}^k \binom{k}{l} a^{k-l} \mathbb{E}[x^l] \end{aligned} \quad (9.51)$$

where we have used $\tilde{x} = x - a$, the binomial theorem and $\int_0^\infty \tilde{x}^k \mathcal{G}(\tilde{x}|0, \tilde{\alpha}, \beta) d\tilde{x} = \mathbb{E}[x^k]$. The moments of (9.49) are then found to be

$$\mathbb{E}_g[x^k] = \sum_{l=0}^k \binom{k}{l} \mathbb{E}[x^l] \left(a_1^{k-l} + \nu^2 a_2^{k-l} + 2\nu^2 e^{-\frac{\alpha^2}{2\nu^2}} a_3^{k-l} \right). \quad (9.52)$$

Finally, we are able to obtain expressions for the cumulants via the standard formulae

$$\begin{aligned} \kappa_1 &= \mu_1 \\ \kappa_2 &= \mu_2 - \mu_1^2 \\ \kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \\ &\vdots \end{aligned} \quad (9.53)$$

which can be neatly expressed by the recursive formula

$$\kappa_k = \mu_k - \sum_{m=1}^{k-1} \binom{k-1}{m-1} \kappa_m \mu_{k-m} \quad (9.54)$$

where μ_k are the moments found using (9.52). Using this method bypasses the unstable numerical differentiation and in fact reveals the desired cumulants directly while bypassing the evaluation of all integrals and derivatives. Practically speaking, the cumulants can be obtained through measurements of REP's as described in (9.3.1).

9.5.1 Reconstruction

Given the cumulants of the information PDF the Gram Charlier A series (9.20) can be used to reconstruct the desired result. Here we demonstrate how another expansion known as the Edgeworth series [139] can be used to accomplish this. The Edgeworth series provides an additional advantage by being error controlled so the expansion is guaranteed to asymptote. This expansion takes the form

$$g(x) = \exp \left(n \sum_{j=2}^n (\kappa_j - \gamma_j) \frac{(-1)^j}{j!} \frac{d^j}{dx^j} n^{-j/2} \right) G(x). \quad (9.55)$$

With the Edgeworth series the expansion is grouped by orders of the power of n i.e for n^k we group terms of equal k . It is useful to note that for $j = 2$ there are an infinite number of such terms which come about from the expansion of the exponential function however, for $j \geq 3$ there is always a finite number of terms for each k . As such, it is useful to rewrite (9.55) as

$$g(x) = \exp \left(n \sum_{j=3}^n (\kappa_j - \gamma_j) \frac{(-1)^j}{j!} \frac{d^j}{dx^j} n^{-j/2} \right) \exp \left((\kappa_2 - \gamma_2) \frac{1}{2} \frac{d^2}{dx^2} \right) G(x) \quad (9.56)$$

and truncate the $j = 2$ by matching the order of magnitude of the smallest term to the smallest terms of the $j \geq 3$ expansion truncated to a desired order k . In Fig.9.8, the reference PDF is again taken to be a shifted Gamma distribution with $\tilde{\alpha} = \frac{1}{2}$, $\beta = \log_2(e)$ and $a = \kappa_1 - \frac{1}{2} \log_2(e)$, the expansion is taken to $n^{-3/2}$ and as demonstrated, the reconstructed information PDF displays convergence to the analytic information PDF.

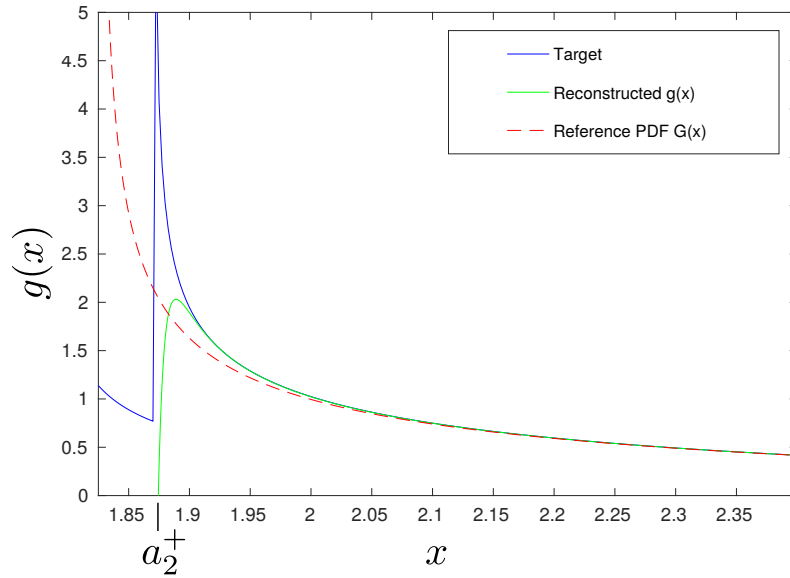


Figure 9.8: Convergence of the reconstructed information distribution of an Unbalanced Cat State (UCS) with $\nu = 0.97$ and $\alpha = 10$. The value a_2^+ corresponds to the value of x at the point of intersection with the second (lower) peak of the $\mathcal{F}(y)$ for the UCS. The Edgeworth expansion has been used here to order $n^{-3/2}$ requiring control of the first 5 cumulants.

Chapter 10

Conclusion

In this thesis we have shown that the statistical correlations employed in quantum metrology protocols can be identified as intra- and inter-mode correlations which can contribute to the sensitivity in different amounts. In particular we have explored the effects of combining the two types of correlation in the Squeezed-Entangled state $|\psi_{SES}\rangle = \mathcal{N}(|\zeta, 0\rangle + |0, \zeta\rangle)$, this was shown to provide high precision gains over previously considered probe states however, this state is not easily produced in practise. Moreover, by analysing the quantum Fisher information as a metric for measurement precision we found that intramode correlations can contribute far more to the sensing capabilities of a scheme than intermode correlations. These findings motivated the investigation into the practically viable, intramode correlated Squeezed Cat State $|\varphi_{SCS}\rangle = \mathcal{N}\hat{S}(\zeta)(|\alpha\rangle + |-\alpha\rangle)$ which was also shown to outperform previously considered states such as the NOON state and separable two-mode squeezed vacuum state. Further work includes applications of sensitivity metrics other than the Fisher information, such as entropic measures, to these probe states. Moreover, the Fisher information, and consequently Cram -Rao bound, is not guaranteed to be a reliable measure of sensitivity in the low particle number regime hence the use of more robust bounds, such as the Ziv-Zakai bound [140] which takes into account prior information, would be a natural extension of this work. We also explored multiparameter estimation for which we identified local and global estimation strategies. Through the investigation of two distinct measurement schemes, one consisting of parallel interferometers and the other consisting of multi-mode quantum enhanced imaging, we demonstrated that the global strategies can be matched and even outperformed by local estimation strategies which have the additional benefit of being more practical than the previously proposed global strategy analogues. Nonetheless, in practise the precision scaling with photon number may not be the most relevant measure whereas the absolute

precision that can be obtained for a given total photon number is [19, 71], in this thesis we have exhibited probe states that improve on the absolute precision of NOON states and thus present candidates for further investigation in the multiparameter local estimation strategies considered.

We also addressed the usual approach of considering quantum metrological protocol to be comprised of three stages: probe preparation, sensing and readout, in which the time taken to perform the first and last stages is often neglected. By regarding time as a non-negligible, limited resource in the protocol of magnetic field sensing with an N spin-1/2 system, we introduced two measurement schemes which employ concurrent sensing during state preparation and readout, one of which used Two-Axis Twisting as state preparation and another which used One-Axis Twisting as an echo measurement scheme. Our investigation of these schemes revealed that the traditional quantum metrological prepare-sense-readout protocol can yield measurement precisions that are worse than even classical schemes and moreover, the concurrent sensing schemes always outperform their sequential counterparts. Future work would include the effects of decoherence in the state preparation, sensing and readout stages and some preliminary results are given in Appendix.B. Another related avenue of investigation would build upon the work of Refs [110, 141] which considers echo measurements in the presence of detection noise and finds that in the absence of detection noise the optimal sensitivities come about by devoting no time to the echo unitary i.e the echo does not give any theoretical advantage over a projection measurement but in the presence of detection noise the echo measurement scheme makes the sensitivity much more robust to noise. Further work could also include the application of concurrent sensing during preparation and readout to the experiment demonstrated by M. Penasa *et.al* [142] in which an echo measurement protocol is employed to estimate the amplitude of a small displacement acting on a cavity field. The notable difference in the scheme of Penasa and the schemes that were analysed here is that execution of preparation and readout takes the form of atom-cavity interactions in order to create, and undo the creation of, optical cat states.

In the latter part of this thesis we took an information theoretic approach to quantum mechanical states and uncertainty relations, in particular we introduced the Rényi entropy from which we derived a generalised uncertainty relation, this was given in terms of the associated entropy power and constitutes an infinite class of uncertainty relations.

We applied this Rényi entropy power uncertainty relation (REPUR) to two non-Gaussian states; a Cat State $|\psi_C\rangle = \mathcal{N}_C(|i\beta\rangle + |-i\beta\rangle)$ and the superposition state of the vacuum and squeezed vacuum $|\psi_V\rangle = \mathcal{N}_V(|0\rangle + |\zeta\rangle)$. From this we were able to infer that the extra information provided by the REPUR gives a measure of “Gaussianity” of the PDFs of quantum states. Furthermore, we were able show that where variance based uncertainty measures breakdown the REPURs are still informative. A scheme for the direct measurement of Rényi entropies using existing AMO and solid state platforms has recently been proposed by the group of P. Zoller [143], motivating future collaborative work with experimentalists. Building upon the REPUR result, we then gave explicit examples of known entropy measures yielding informative results where variance based measure of uncertainty breakdown. The result is based upon the invariance of entropies under piecewise rearrangements of the underlying probability distributions. We further this notion to include equimeasurable PDFs through which we introduce key properties of a technique akin to quantum tomography which we refer to as an “information scan”. Moreover, we gave a numerical simulation of this information scan being performed on the superposition state of a vacuum and a squeezed vacuum $|\psi_V\rangle$. We then provided a detailed derivation of how this information scan depends on the Rényi entropy powers and in turn how the Rényi entropy powers relate to the cumulants of the information PDF of the quantum states. We followed this by presenting an analytical application of this reconstruction theorem to two distinct types of cat state and showed that the analytical and numerical results converge. This work can be extended to designing quantum states through their associated PDFs via this reconstruction theorem, this would provide crucial insight into the structure of quantum states enabling them to be tailored for specific metrology tasks.

Bibliography

- [1] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?,” *Physical Review*, vol. 47, no. 10, p. 777, 1935. [2](#)
- [2] E. Schrödinger, “Die gegenwärtige situation in der quantenmechanik,” *Naturwissenschaften*, vol. 23, no. 48, pp. 807–812, 1935. [2](#)
- [3] S. J. Freedman and J. F. Clauser, “Experimental test of local hidden-variable theories,” *Physical Review Letters*, vol. 28, no. 14, p. 938, 1972. [3](#)
- [4] A. Aspect, P. Grangier, and G. Roger, “Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: a new violation of Bell’s inequalities,” *Physical Review Letters*, vol. 49, no. 2, p. 91, 1982. [3](#)
- [5] J. S. Bell, “On the einstein podolsky rosen paradox,” 1964. [3](#)
- [6] H. Bennett Ch and G. Brassard, “Quantum cryptography: public key distribution and coin tossing int,” in *Conf. on Computers, Systems and Signal Processing (Bangalore, India, Dec. 1984)*, pp. 175–9, 1984. [3](#)
- [7] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, “Experimental quantum cryptography,” *Journal of Cryptology*, vol. 5, no. 1, pp. 3–28, 1992. [3](#)
- [8] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and einstein-podolsky-rozen channels,” *Physical Review Letters*, vol. 70, no. 13, p. 1895, 1993. [3](#)
- [9] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, “Experimental quantum teleportation,” *Nature*, vol. 390, no. 6660, pp. 575–579, 1997. [3](#)

- [10] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, *et al.*, “Experimental demonstration of free-space decoy-state quantum key distribution over 144 km,” *Physical Review Letters*, vol. 98, no. 1, p. 010504, 2007. [3](#)
- [11] R. Raussendorf and H. J. Briegel, “A one-way quantum computer,” *Physical Review Letters*, vol. 86, no. 22, p. 5188, 2001. [3](#)
- [12] V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology,” *Nature Photonics*, vol. 5, no. 4, p. 222, 2011. [3](#)
- [13] J. Dunningham, A. Rau, and K. Burnett, “From pedigree cats to fluffy-bunnies,” *Science*, vol. 307, no. 5711, pp. 872–875, 2005. [3](#)
- [14] C. J. Isham, *Lectures on quantum theory Mathematical and structural foundations*. Allied Publishers, 2001. [5](#), [7](#), [47](#)
- [15] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information*. Cambridge University Press, 2002. [9](#)
- [16] G. E. Idang, “Thales, Anaximander and Anaximenes as pathfinders of modern science,” 2013. [10](#)
- [17] A. Kaykov, T. Taillefumier, A. Bensimon, and P. Nurse, “Molecular combing of single dna molecules on the 10 megabase scale,” *Scientific Reports*, vol. 6, p. 19636, 2016. [11](#)
- [18] S. Lloyd, “Enhanced sensitivity of photodetection via quantum illumination,” *Science*, vol. 321, no. 5895, pp. 1463–1465, 2008. [11](#)
- [19] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, R. Adhikari, *et al.*, “Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light,” *Nature Photonics*, vol. 7, no. 8, p. 613, 2013. [12](#), [66](#), [150](#)
- [20] M. A. Taylor and W. P. Bowen, “Quantum metrology and its application in biology,” *Physics Reports*, vol. 615, pp. 1–59, 2016. [12](#), [66](#)
- [21] C. Gerry and P. Knight, *Introductory Quantum Optics*. Cambridge University Press, 2005. [13](#), [17](#), [19](#), [24](#), [65](#)

- [22] S. M. Barnett and P. M. Radmore, *Methods in theoretical quantum optics*. Oxford University Press, 2002. [13](#), [20](#), [27](#), [119](#), [128](#), [129](#)
- [23] M. Born, W. Heisenberg, and P. Jordan, “Zur quantenmechanik. ii.,” *Zeitschrift für Physik*, vol. 35, no. 8-9, pp. 557–615, 1926. [14](#)
- [24] C. W. Gardiner and Others, *Handbook of stochastic methods*, vol. 3. Springer Berlin, 1985. [14](#)
- [25] R. J. Glauber, “Coherent and incoherent states of the radiation field,” *Physical Review*, vol. 131, no. 6, p. 2766, 1963. [16](#)
- [26] D. Stoler, “Equivalence classes of minimum uncertainty packets,” *Physical Review D*, vol. 1, no. 12, p. 3217, 1970. [17](#)
- [27] E. Wigner, “On the quantum correction for thermodynamic equilibrium,” *Physical Review*, vol. 40, no. 5, p. 749, 1932. [17](#)
- [28] V. Fock, “Konfigurationsraum und zweite quantelung,” *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 75, no. 9, pp. 622–647, 1932. [27](#)
- [29] R. Eisberg, R. Resnick, and J. Brown, “Quantum physics of atoms, molecules, solids, nuclei, and particles,” *Physics Today*, vol. 39, p. 110, 1986. [32](#)
- [30] W. Gerlach and O. Stern, “Der experimentelle nachweis der richtungsquantelung im magnetfeld,” *Zeitschrift für Physik*, vol. 9, no. 1, pp. 349–352, 1922. [32](#)
- [31] J. J. Sakurai and E. D. Commins, “Modern quantum mechanics, revised edition,” 1995. [35](#)
- [32] F. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, “Atomic coherent states in quantum optics,” *Physical Review A*, vol. 6, no. 6, p. 2211, 1972. [37](#), [42](#), [45](#), [92](#)
- [33] J. Radcliffe, “Some properties of coherent spin states,” *Journal of Physics A: General Physics*, vol. 4, no. 3, p. 313, 1971. [37](#), [45](#)
- [34] J. Ma, X. Wang, C.-P. Sun, and F. Nori, “Quantum spin squeezing,” *Physics Reports*, vol. 509, no. 2, pp. 89–165, 2011. [40](#), [42](#)
- [35] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. Heinzen, “Squeezed atomic states and projection noise in spectroscopy,” *Physical Review A*, vol. 50, no. 1, p. 67, 1994. [41](#)

- [36] M. Kitagawa and M. Ueda, “Squeezed spin states,” *Physical Review A*, vol. 47, no. 6, p. 5138, 1993. [41](#), [42](#), [43](#)
- [37] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. Moore, and D. Heinzen, “Spin squeezing and reduced quantum noise in spectroscopy,” *Physical Review A*, vol. 46, no. 11, p. R6797, 1992. [41](#)
- [38] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, “14-qubit entanglement: Creation and coherence,” *Physical Review Letters*, vol. 106, no. 13, p. 130506, 2011. [42](#)
- [39] A. Sørensen, L.-M. Duan, J. Cirac, and P. Zoller, “Many-particle entanglement with bose-einstein condensates,” *Nature*, vol. 409, no. 6816, pp. 63–66, 2001. [42](#)
- [40] C. Orzel, A. Tuchman, M. Fenselau, M. Yasuda, and M. Kasevich, “Squeezed states in a bose-einstein condensate,” *Science*, vol. 291, no. 5512, pp. 2386–2389, 2001. [42](#)
- [41] C. Gross, T. Zibold, E. Nicklas, J. Esteve, and M. K. Oberthaler, “Nonlinear atom interferometer surpasses classical precision limit,” *Nature*, vol. 464, no. 7292, pp. 1165–1169, 2010. [42](#)
- [42] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, “Atom-chip-based generation of entanglement for quantum metrology,” *Nature*, vol. 464, no. 7292, pp. 1170–1173, 2010. [42](#)
- [43] T. Fernholz, H. Krauter, K. Jensen, J. F. Sherson, A. S. Sørensen, and E. S. Polzik, “Spin squeezing of atomic ensembles via nuclear-electronic spin entanglement,” *Physical Review Letters*, vol. 101, no. 7, p. 073601, 2008. [42](#)
- [44] S. Chaudhury, S. Merkel, T. Herr, A. Silberfarb, I. H. Deutsch, and P. S. Jessen, “Quantum control of the hyperfine spin of a cs atom ensemble,” *Physical Review Letters*, vol. 99, no. 16, p. 163002, 2007. [42](#)
- [45] I. D. Leroux, M. H. Schleier-Smith, and V. Vuletić, “Implementation of cavity squeezing of a collective atomic spin,” *Physical Review Letters*, vol. 104, no. 7, p. 073602, 2010. [42](#)
- [46] M. H. Schleier-Smith, I. D. Leroux, and V. Vuletić, “Squeezing the collective spin of a dilute atomic ensemble by cavity feedback,” *Physical Review A*, vol. 81, no. 2, p. 021804, 2010. [42](#)

- [47] G. S. Agarwal, R. Puri, and R. Singh, “Atomic schrödinger cat states,” *Physical Review A*, vol. 56, no. 3, p. 2249, 1997. [42](#)
- [48] J. P. Dowling, G. S. Agarwal, and W. P. Schleich, “Wigner distribution of a general angular-momentum state: Applications to a collection of two-level atoms,” *Physical Review A*, vol. 49, no. 5, p. 4101, 1994. [42](#)
- [49] K. Helmerson and L. You, “Creating massive entanglement of bose-einstein condensed atoms,” *Physical Review Letters*, vol. 87, no. 17, p. 170402, 2001. [43](#)
- [50] M. Jafarpour and A. Akhound, “Entanglement and squeezing of multi-qubit systems using a two-axis countertwisting hamiltonian with an external field,” *Physics Letters A*, vol. 372, no. 14, pp. 2374–2379, 2008. [43](#)
- [51] A. Andre and M. Lukin, “Atom correlations and spin squeezing near the heisenberg limit: Finite-size effect and decoherence,” *Physical Review A*, vol. 65, no. 5, p. 053819, 2002. [43](#)
- [52] T. Holstein and H. Primakoff, “Field dependence of the intrinsic domain magnetization of a ferromagnet,” *Physical Review*, vol. 58, no. 12, p. 1098, 1940. [44](#)
- [53] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control*. Cambridge University Press, 2009. [46](#), [47](#), [48](#), [52](#), [56](#), [174](#)
- [54] D. Bohm and H. P. Stapp, “The undivided universe: An ontological interpretation of quantum theory,” 1994. [47](#)
- [55] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, “Chapter four-quantum limits in optical interferometry,” *Progress in Optics*, vol. 60, pp. 345–435, 2015. [49](#)
- [56] B. Tamir and E. Cohen, “Introduction to weak measurements and weak values,” *Quanta*, vol. 2, no. 1, pp. 7–17, 2013. [50](#)
- [57] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, vol. 1. Springer Science & Business Media, 2011. [52](#)
- [58] C. W. Helstrom, *Quantum detection and estimation theory*. Academic Press, 1976. [52](#)
- [59] M. J. Evans and J. S. Rosenthal, *Probability and statistics: The science of uncertainty*. Macmillan, 2004. [54](#)

- [60] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*, vol. 9. Princeton University Press, 2016. [56](#)
- [61] C. R. Rao, “Information and the accuracy attainable in the estimation of statistical parameters,” in *Breakthroughs in statistics*, pp. 235–247, Springer, 1992. [56](#)
- [62] S. L. Braunstein and C. M. Caves, “Statistical distance and the geometry of quantum states,” *Physical Review Letters*, vol. 72, no. 22, p. 3439, 1994. [56](#)
- [63] M. G. Paris, “Quantum estimation for quantum technology,” *International Journal of Quantum Information*, vol. 7, no. supp01, pp. 125–137, 2009. [56](#), [74](#)
- [64] A. A. Michelson and E. W. Morley, “On the relative motion of the earth and the luminiferous ether,” *American Journal of Science*, vol. 34, pp. 333–345, 1887. [58](#)
- [65] L. Zehnder, *Ein neuer interferenzrefraktor*. Springer, 1891. [58](#)
- [66] L. Mach, “Ueber einen interferenzrefraktor,” *Zeitschrift für Instrumentenkunde*, vol. 12, pp. 89–93, 1892. [58](#)
- [67] J. Dunningham and T. Kim, “Using quantum interferometers to make measurements at the heisenberg limit,” *Journal of Modern Optics*, vol. 53, no. 4, pp. 557–571, 2006. [62](#), [72](#)
- [68] P. Walther, J.-W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, “De broglie wavelength of a non-local four-photon state,” *Nature*, vol. 429, no. 6988, pp. 158–161, 2004. [62](#)
- [69] I. Afek, O. Ambar, and Y. Silberberg, “High-noon states by mixing quantum and classical light,” *Science*, vol. 328, no. 5980, pp. 879–881, 2010. [62](#)
- [70] V. Giovannetti, S. Lloyd, and L. Maccone, “Quantum-enhanced measurements: beating the standard quantum limit,” *Science*, vol. 306, no. 5700, pp. 1330–1336, 2004. [63](#)
- [71] M. Taylor, *Quantum Microscopy of Biological Systems*. Springer, 2015. [66](#), [150](#)
- [72] M. Punturo, “M. punturo et al., classical quantum gravity 27, 194002 (2010).,” *Classical Quantum Gravity*, vol. 27, p. 194002, 2010. [66](#)
- [73] T. P. Purdy, R. W. Peterson, and C. Regal, “Observation of radiation pressure shot noise on a macroscopic object,” *Science*, vol. 339, no. 6121, pp. 801–804, 2013. [66](#)

- [74] C. M. Caves, “Quantum-mechanical noise in an interferometer,” *Physical Review D*, vol. 23, no. 8, p. 1693, 1981. [x](#), [67](#)
- [75] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, “Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector geo 600,” *Physical Review A*, vol. 88, no. 4, p. 041802, 2013. [68](#)
- [76] M. Jarzyna and R. Demkowicz-Dobrzański, “Quantum interferometry with and without an external phase reference,” *Physical Review A*, vol. 85, no. 1, p. 011801, 2012. [69](#)
- [77] H. F. Hofmann, “All path-symmetric pure states achieve their maximal phase sensitivity in conventional two-path interferometry,” *Physical Review A*, vol. 79, no. 3, p. 033822, 2009. [69](#), [79](#)
- [78] J. Sahota and N. Quesada, “Quantum correlations in optical metrology: Heisenberg-limited phase estimation without mode entanglement,” *Physical Review A*, vol. 91, no. 1, p. 013808, 2015. [69](#)
- [79] P. Knott, T. Proctor, A. Hayes, J. Cooling, and J. Dunningham, “Practical quantum metrology with large precision gains in the low-photon-number regime,” *Physical Review A*, vol. 93, no. 3, p. 033859, 2016. [x](#), [71](#), [74](#)
- [80] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier, “Generation of optical schrödinger cats from photon number states,” *Nature*, vol. 448, no. 7155, pp. 784–786, 2007. [73](#)
- [81] J. Etesse, M. Bouillard, B. Kanseri, and R. Tualle-Brouri, “Experimental generation of squeezed cat states with an operation allowing iterative growth,” *Physical Review Letters*, vol. 114, no. 19, p. 193602, 2015. [73](#)
- [82] K. Huang, H. Le Jeannic, J. Ruaudel, V. Verma, M. Shaw, F. Marsili, S. Nam, E. Wu, H. Zeng, Y.-C. Jeong, *et al.*, “Optical synthesis of large-amplitude squeezed coherent-state superpositions with minimal resources,” *Physical Review Letters*, vol. 115, no. 2, p. 023602, 2015. [73](#)
- [83] R. Guzmán, J. Retamal, E. Solano, and N. Zagury, “Field squeeze operators in optical cavities with atomic ensembles,” *Physical Review Letters*, vol. 96, no. 1, p. 010502, 2006. [73](#)

- [84] D. W. Brooks, T. Botter, S. Schreppler, T. P. Purdy, N. Brahms, and D. M. Stamper-Kurn, “Non-classical light generated by quantum-noise-driven cavity optomechanics,” *Nature*, vol. 488, no. 7412, pp. 476–480, 2012. [73](#)
- [85] A. H. Safavi-Naeini, S. Gröblacher, J. T. Hill, J. Chan, M. Aspelmeyer, and O. Painter, “Squeezed light from a silicon micromechanical resonator,” *Nature*, vol. 500, no. 7461, pp. 185–189, 2013. [73](#)
- [86] M. Brune, E. Hagley, J. Dreyer, X. Maitre, A. Maali, C. Wunderlich, J. Raimond, and S. Haroche, “Observing the progressive decoherence of the meter in a quantum measurement,” *Physical Review Letters*, vol. 77, no. 24, p. 4887, 1996. [73](#)
- [87] A. Lund, H. Jeong, T. Ralph, and M. Kim, “Conditional production of superpositions of coherent states with inefficient photon detection,” *Physical Review A*, vol. 70, no. 2, p. 020101, 2004. [73](#)
- [88] T. J. Bartley, G. Donati, J. B. Spring, X.-M. Jin, M. Barbieri, A. Datta, B. J. Smith, and I. A. Walmsley, “Multiphoton state engineering by heralded interference between single photons and coherent states,” *Physical Review A*, vol. 86, no. 4, p. 043820, 2012. [73](#)
- [89] C. M. Caves and K. Wódkiewicz, “Fidelity of gaussian channels,” *Open Systems & Information Dynamics*, vol. 11, no. 04, pp. 309–323, 2004. [73](#)
- [90] M. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, “Invited review article: Single-photon sources and detectors,” *Review of Scientific Instruments*, vol. 82, no. 7, p. 071101, 2011. [73](#)
- [91] B. Calkins, P. L. Mennea, A. E. Lita, B. J. Metcalf, W. S. Kolthammer, A. Lamas-Linares, J. B. Spring, P. C. Humphreys, R. P. Mirin, J. C. Gates, *et al.*, “High quantum-efficiency photon-number-resolving detector for photonic on-chip information processing,” *Optics Express*, vol. 21, no. 19, pp. 22657–22670, 2013. [73](#)
- [92] D. Fukuda, G. Fujii, T. Numata, K. Amemiya, A. Yoshizawa, H. Tsuchida, H. Fujino, H. Ishii, T. Itatani, S. Inoue, *et al.*, “Titanium-based transition-edge photon number resolving detector with 98% detection efficiency with index-matched small-gap fiber coupling,” *Optics express*, vol. 19, no. 2, pp. 870–875, 2011. [73](#)
- [93] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva, *et al.*, “Superconducting nanowire photon-

- number-resolving detector at telecommunication wavelengths,” *Nature Photonics*, vol. 2, no. 5, pp. 302–306, 2008. [73](#)
- [94] A. Freise, S. Chelkowski, S. Hild, W. Del Pozzo, A. Perreca, and A. Vecchio, “Triple michelson interferometer for a third-generation gravitational wave detector,” *Classical and Quantum Gravity*, vol. 26, no. 8, p. 085012, 2009. [74](#), [76](#)
- [95] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, “Quantum enhanced multiple phase estimation,” *Physical Review Letters*, vol. 111, no. 7, p. 070403, 2013. [74](#), [75](#), [79](#), [80](#)
- [96] P. Kok, J. Dunningham, and J. F. Ralph, “Quantum estimation of coupled parameters and the role of entanglement,” *arXiv preprint arXiv:1505.06321*, 2015. [74](#)
- [97] P. Knott, T. Proctor, K. Nemoto, J. Dunningham, and W. Munro, “Effect of multimode entanglement on lossy optical quantum metrology,” *Physical Review A*, vol. 90, no. 3, p. 033846, 2014. [74](#)
- [98] J. Liu, X.-M. Lu, Z. Sun, and X. Wang, “Quantum multiparameter metrology with generalized entangled coherent state,” *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 11, p. 115302, 2016. [75](#)
- [99] J.-D. Yue, Y.-R. Zhang, and H. Fan, “Quantum-enhanced metrology for multiple phase estimation with noise,” *Scientific Reports*, vol. 4, 2014. [75](#)
- [100] V. Giovannetti, S. Lloyd, L. Maccone, and J. H. Shapiro, “Sub-rayleigh-diffraction-bound quantum imaging,” *Physical Review A*, vol. 79, no. 1, p. 013827, 2009. [79](#)
- [101] M. Tsang, “Quantum imaging beyond the diffraction limit by optical centroid measurements,” *Physical Review Letters*, vol. 102, no. 25, p. 253601, 2009. [79](#)
- [102] H. Shin, K. W. C. Chan, H. J. Chang, and R. W. Boyd, “Quantum spatial superresolution by optical centroid measurements,” *Physical Review Letters*, vol. 107, no. 8, p. 083603, 2011. [79](#)
- [103] S. Dooley, E. Yukawa, Y. Matsuzaki, G. C. Knee, W. J. Munro, and K. Nemoto, “A hybrid-systems approach to spin squeezing using a highly dissipative ancillary system,” *New Journal of Physics*, vol. 18, no. 5, p. 053011, 2016. [83](#)
- [104] S. Dooley, F. McCrossan, D. Harland, M. J. Everitt, and T. P. Spiller, “Collapse and revival and cat states with an n-spin system,” *Physical Review A*, vol. 87, no. 5, p. 052323, 2013. [83](#)

- [105] W. J. Munro, K. Nemoto, G. J. Milburn, and S. L. Braunstein, “Weak-force detection with superposed coherent states,” *Physical Review A*, vol. 66, no. 2, p. 023819, 2002. [83](#), [168](#)
- [106] S. Dooley, W. J. Munro, and K. Nemoto, “Quantum metrology including state preparation and readout times,” *Physical Review A*, vol. 94, no. 5, p. 052320, 2016. [83](#)
- [107] E. Davis, G. Bentsen, and M. Schleier-Smith, “Approaching the Heisenberg limit without single-particle detection,” *Physical Review Letters*, vol. 116, no. 5, p. 053601, 2016. [xii](#), [91](#), [94](#)
- [108] T. Macrì, A. Smerzi, and L. Pezzè, “Loschmidt echo for quantum metrology,” *Physical Review A*, vol. 94, no. 1, p. 010102, 2016. [91](#)
- [109] O. Hosten, R. Krishnakumar, N. Engelsen, and M. Kasevich, “Quantum phase magnification,” *Science*, vol. 352, no. 6293, pp. 1552–1555, 2016. [91](#)
- [110] S. P. Nolan, S. S. Szegedi, and S. A. Haine, “Optimal and robust quantum metrology using interaction-based readouts,” *Physical Review Letters*, vol. 119, no. 19, p. 193601, 2017. [91](#), [150](#)
- [111] D. Linnemann, H. Strobel, W. Muessel, J. Schulz, R. J. Lewis-Swan, K. V. Khurutsyan, and M. K. Oberthaler, “Quantum-enhanced sensing based on time reversal of nonlinear dynamics,” *Physical Review Letters*, vol. 117, no. 1, p. 013001, 2016. [91](#)
- [112] R. Frigg and C. Werndl, “Entropy - a guide for the perplexed,” *Probabilities in Physics*, 2011. [97](#), [102](#)
- [113] R. V. Hartley, “Transmission of information1,” *Bell System Technical Journal*, vol. 7, no. 3, pp. 535–563, 1928. [97](#)
- [114] C. E. Shannon, “A mathematical theory of communication,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001. [98](#), [99](#), [112](#), [113](#)
- [115] C. E. Shannon and W. Weaver, *The mathematical theory of communication*. University of Illinois press, 2015. [98](#)
- [116] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press, 2010. [xii](#), [100](#), [101](#)

- [117] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012. [xii](#), [101](#), [103](#)
- [118] A. Rényi, “On the dimension and entropy of probability distributions,” *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 10, no. 1-2, pp. 193–215, 1959. [104](#), [109](#)
- [119] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1952. [105](#), [106](#), [114](#)
- [120] D. Xu and D. Erdogmus, “Renyis entropy, divergence and their nonparametric estimators,” in *Information Theoretic Learning*, pp. 47–102, Springer, 2010. [xiii](#), [106](#), [107](#), [108](#)
- [121] N. Bourbaki, “Topological vector spaces, elements of mathematics,” 1987. [107](#)
- [122] W. Heisenberg, “Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik,” in *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pp. 478–504, Springer, 1927. [110](#)
- [123] E. H. Kennard, “Zur quantenmechanik einfacher bewegungstypen,” *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 44, no. 4, pp. 326–352, 1927. [111](#)
- [124] W. Heisenberg, *The physical principles of the quantum theory*. Courier Corporation, 1949. [111](#)
- [125] H. P. Robertson, “The uncertainty principle,” *Physical Review*, vol. 34, no. 1, p. 163, 1929. [111](#)
- [126] E. Schrodinger, “Über die kraftfreie bewegung in der relativistischen quantenmechanik,” *Sitzung Phys.-Math.*, vol. 31, p. 418, 1930. [111](#)
- [127] E. Arthurs and J. Kelly, “Bstj briefs: On the simultaneous measurement of a pair of conjugate observables,” *The Bell System Technical Journal*, vol. 44, no. 4, pp. 725–729, 1965. [111](#)
- [128] M. Ozawa, “Universally valid reformulation of the heisenberg uncertainty principle on noise and disturbance in measurement,” *Physical Review A*, vol. 67, no. 4, p. 042105, 2003. [112](#)

- [129] Y. Watanabe, *Formulation of Uncertainty Relation Between Error and Disturbance in Quantum Measurement by Using Quantum Estimation Theory*. Springer Science & Business Media, 2013. [112](#)
- [130] I. I. Hirschman, “A note on entropy,” *American Journal of Mathematics*, vol. 79, no. 1, pp. 152–156, 1957. [112](#), [113](#)
- [131] H. Everett III, “The theory of the universal wave function (1956),” *Princeton University Dissertation*, 1973. [112](#)
- [132] W. Beckner, “Inequalities in fourier analysis,” *Annals of Mathematics*, pp. 159–182, 1975. [112](#), [113](#)
- [133] I. Białynicki-Birula and J. Mycielski, “Uncertainty relations for information entropy in wave mechanics,” *Communications in Mathematical Physics*, vol. 44, no. 2, pp. 129–132, 1975. [112](#)
- [134] M. Costa, “A new entropy power inequality,” *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 751–760, 1985. [113](#)
- [135] K. I. Babenko, “An inequality in the theory of fourier integrals,” *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 25, no. 4, pp. 531–542, 1961. [113](#)
- [136] I. Białynicki-Birula and Ł. Rudnicki, “Entropic uncertainty relations in quantum physics,” in *Statistical Complexity*, pp. 1–34, Springer, 2011. [132](#), [136](#)
- [137] W. BUCKLAND, “Advanced theory of statistics-kendall, mg and stuart, a,” 1970. [139](#)
- [138] N. Pal, C. Jin, and W. K. Lim, *Handbook of exponential and related distributions for engineers and scientists*. CRC Press, 2005. [145](#)
- [139] P. Hall, *The bootstrap and Edgeworth expansion*. Springer Science & Business Media, 2013. [147](#)
- [140] M. Tsang, “Ziv-zakai error bounds for quantum parameter estimation,” *Physical Review Letters*, vol. 108, no. 23, p. 230401, 2012. [149](#)
- [141] S. S. Mirkhalaf, S. P. Nolan, and S. A. Haine, “Robustifying twist-and-turn entanglement with interaction-based readout,” *arXiv preprint arXiv:1803.08789*, 2018. [150](#)

- [142] M. Penasa, S. Gerlich, T. Rybarczyk, V. Métillon, M. Brune, J. Raimond, S. Haroche, L. Davidovich, and I. Dotsenko, “Measurement of a microwave field amplitude beyond the standard quantum limit,” *Physical Review A*, vol. 94, no. 2, p. 022313, 2016. [150](#)
- [143] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, “Rényi entropies from random quenches in atomic hubbard and spin models,” *Physical Review Letters*, vol. 120, no. 5, p. 050406, 2018. [151](#)
- [144] M. G. Paris, “Displacement operator by beam splitter,” *Physics Letters A*, vol. 217, no. 2-3, pp. 78–80, 1996. [168](#)

Appendix A

Accurate Numerical Higher Order Derivatives

Clearly the cumulants can also be obtained through taking the $(n - 1)$ th derivative of the Rényi Entropy power with respect to p directly as depicted by (9.36). However, as previously mentioned, the analytical form of such derivatives are extremely cumbersome to work with. As such, a method for accurate numerical derivatives is presented here - the “method of undetermined coefficients”. Starting with the difference approximation

$$D^n u(\vec{x}) = \sum_{k=0}^n c_k u(\vec{x} - kh) \quad (\text{A.1})$$

so that

$$\begin{aligned} D^0 u(\vec{x}) &= c_0 u(\vec{x}) \\ D^1 u(\vec{x}) &= c_0 u(\vec{x}) + c_1 u(\vec{x} - h) \\ D^2 u(\vec{x}) &= c_0 u(\vec{x}) + c_1 u(\vec{x} - h) + c_2 u(\vec{x} - 2h) \\ D^3 u(\vec{x}) &= c_0 u(\vec{x}) + c_1 u(\vec{x} - h) + c_2 u(\vec{x} - 2h) + c_3 u(\vec{x} - 3h) \\ &\vdots \\ D^n u(\vec{x}) &= c_0 u(\vec{x}) + c_1 u(\vec{x} - h) \dots c_n u(\vec{x} - nh) \end{aligned} \quad (\text{A.2})$$

and the n th derivative is given by

$$u^{(n)}(\vec{x}) = \lim_{h \rightarrow 0} \frac{D^n u(\vec{x})}{h^n}. \quad (\text{A.3})$$

The aim is then to find the $n + 1$ coefficients c_0, c_1, \dots, c_n which requires at least $n + 1$ equations. To acquire these equations, the Taylor expansion of $u(\vec{x} - kh)$ can be taken to

n th order at $h = 0$

$$\begin{aligned} u(\vec{x} - kh) &= \sum_{m=0}^n (-1)^m \frac{(kh)^m}{m!} u^{(m)}(\vec{x}) \\ &= u(\vec{x}) - kh u'(\vec{x}) + \frac{1}{2} (kh)^2 u''(\vec{x}) - \frac{1}{3!} (kh)^3 u'''(\vec{x}) + \dots \end{aligned} \quad (\text{A.4})$$

which gives

$$\begin{aligned} D^0 u(\vec{x}) &= c_0 u(\vec{x}) \\ D^1 u(\vec{x}) &= c_0 u(\vec{x}) + c_1 \left[u(\vec{x}) - hu'(\vec{x}) + \frac{1}{2} (h)^2 u''(\vec{x}) - \dots + (-1)^n h \frac{h^n}{n!} u^{(n)}(\vec{x}) \right] \\ D^2 u(\vec{x}) &= c_0 u(\vec{x}) + c_1 \left[u(\vec{x}) - hu'(\vec{x}) + \frac{1}{2} (h)^2 u''(\vec{x}) - \dots + (-1)^n h \frac{h^n}{n!} u^{(n)}(\vec{x}) \right] \\ &\quad + c_2 \left[u(\vec{x}) - 2hu'(\vec{x}) + \frac{1}{2} (2h)^2 u''(\vec{x}) - \dots + (-1)^n h \frac{(2h)^n}{n!} u^{(n)}(\vec{x}) \right] \\ &\quad \vdots \\ D^n u(\vec{x}) &= D^{n-1} u(\vec{x}) + c_n \left[u(\vec{x}) - nh u'(\vec{x}) + \frac{1}{2} (nh)^2 u''(\vec{x}) - \dots + (-1)^n \frac{(nh)^n}{n!} u^{(n)}(\vec{x}) \right] \end{aligned} \quad (\text{A.5})$$

then grouping coefficients of similar order in h (or equivalently of the same order derivative of $u(\vec{x})$) it becomes apparent that the n th order difference approximation is given by

$$\begin{aligned} D^n u(\vec{x}) &= (c_0 + c_1 + c_2 + \dots + c_n) - (c_1 + c_2 + c_3 + \dots + c_n) h u'(\vec{x}) \\ &\quad + \frac{1}{2} (c_1 + 2^2 c_2 + 3^2 c_3 + \dots + n^2 c_n) h^2 u''(\vec{x}) \\ &\quad - \frac{1}{3!} (c_1 + 2^3 c_2 + 3^3 c_3 + \dots + n^3 c_n) h^3 u'''(\vec{x}) \\ &\quad \vdots \\ &\quad \frac{(-1)^n}{n!} (c_1 + 2^n c_2 + 3^n c_3 + \dots + n^n c_n) h^n u^{(n)}(\vec{x}) \end{aligned} \quad (\text{A.6})$$

now only the n th term should contribute to the n th derivative of $u(\vec{x})$ so for $k < n$ all of the coefficients of $h^k u^{(k)}(\vec{x}) = 0$ in (A.6) and the coefficient of $h^n u^{(n)}(\vec{x}) = \frac{1}{h^n}$ thus

$$\begin{aligned} 0 &= c_0 + c_1 + c_2 + \dots + c_n \\ 0 &= -(c_1 + c_2 + c_3 + \dots + c_n) \\ 0 &= \frac{1}{2} (c_1 + 2^2 c_2 + 3^2 c_3 + \dots + n^2 c_n) \\ 0 &= -\frac{1}{3!} (c_1 + 2^3 c_2 + 3^3 c_3 + \dots + n^3 c_n) \\ &\quad \vdots \\ \frac{1}{h^n} &= \frac{(-1)^n}{n!} (c_1 + 2^n c_2 + 3^n c_3 + \dots + n^n c_n) \end{aligned} \quad (\text{A.7})$$

Then solving this system of equations for c_k in terms of h and substituting these values back into the original expression for $D^n u(\vec{x})$ we have an accurate approximation for the n th derivative. Note that in practice, setting $h \approx 10^{-2}$ is sufficient for accurate results.

Appendix B

Optical System

Here we present an alternative method which demonstrates of the sensing capabilities in the bosonic limit of the TAT schemes introduced in section 3.4.2. This entails considering the more realistic case of the run time of the quantum metrology protocol being a limited resource, with this we show that concurrent state preparation and sensing outperforms the usual approach of sequential state preparation and readout. Furthermore, we shall include and analyse the effects of loss and show that concurrent state preparation and readout outperforms the analogous sequential scheme. We demonstrate this alternative approach in the absence of optical loss.

B.1 Lossless

In the optical setting, non-classical states of light can be used to achieve high precision gains for the detection of a weak classical force $F(t)$ [105] which acts for time t . The effect of this force on a harmonic oscillator is a displacement of the complex amplitude of the oscillator in phase space, indeed the result of this force acting on the harmonic oscillator can be modelled in the interaction picture as the displacement operator (2.32) [144]. In the following all displacements are taken to be along the position axis of phase space x_0 which allows us to take advantage of the relation (2.40). The preparation of the non-classical state is performed by the bosonic squeezing operation (2.46) where the squeezing parameter is taken to be real $\xi \in \mathbb{R}^+ \Rightarrow \xi = r$. The aim here is to produce a unitary time evolution operator that simultaneously mimics the action of both of the aforementioned operators i.e concurrent state preparation and sensing. With this the following Hamiltonian is introduced

$$H = i\hbar\eta(a^2 - (a^\dagger)^2) \pm i\hbar\omega(a^\dagger - a) \quad (\text{B.1})$$

where the " \pm " dictates the direction of the displacement along the position axis and ω is the parameter we wish to estimate. This gives the following unitary

$$\begin{aligned}
U(t) &= \exp\left(\frac{r}{2}[a^2 - (a^\dagger)^2] \mp \omega t(a^\dagger - a)\right) \\
&= \exp\left[\frac{r}{2}a^2 \mp a\omega t + \frac{1}{2r}\omega^2 t^2 - \left(\frac{r}{2}(a^\dagger)^2 \mp a^\dagger\omega t + \frac{1}{2r}\omega^2 t^2\right)\right] \\
&= \exp\left[\left(\sqrt{\frac{r}{2}}a \mp \frac{1}{2}\sqrt{\frac{2}{r}}\omega t\right)^2 - \left(\sqrt{\frac{r}{2}}a^\dagger \mp \frac{1}{2}\sqrt{\frac{2}{r}}\omega t\right)^2\right] \\
&= \exp\left(\frac{r}{2}\left[\left(a \mp \frac{\omega t}{r}\right)^2 - \left(a^\dagger \mp \frac{\omega t}{r}\right)^2\right]\right) \tag{B.2}
\end{aligned}$$

expanding the exponential then gives

$$\begin{aligned}
U(t) &= 1 + \frac{r}{2}\left[\left(a \mp \frac{\omega t}{r}\right)^2 - \left(a^\dagger \mp \frac{\omega t}{r}\right)^2\right] + \frac{1}{2}\frac{r^2}{4}\left[\left(a \mp \frac{\omega t}{r}\right)^2 - \left(a^\dagger \mp \frac{\omega t}{r}\right)^2\right]^2 + \dots \\
&= 1 + \frac{r}{2}\left[D^\dagger(\mp\omega t/r)a^2D(\mp\omega t/r) - D^\dagger(\mp\omega t/r)(a^\dagger)^2D(\mp\omega t/r)\right] + \\
&\quad + \frac{1}{2}\frac{r^2}{4}\left[D^\dagger(\mp\omega t/r)a^2D(\mp\omega t/r) - D^\dagger(\mp\omega t/r)(a^\dagger)^2D(\mp\omega t/r)\right]^2 + \dots \\
&= D^\dagger(\mp\omega t/r)\left[1 + \frac{r}{2}[a^2 - (a^\dagger)^2] + \frac{1}{2}\frac{r^2}{4}[a^2 - (a^\dagger)^2]^2 + \dots\right]D(\mp\omega t/r) \\
&= D^\dagger(\mp\omega t/r)\exp\left[\frac{r}{2}(a^2 - (a^\dagger)^2)\right]D(\mp\omega t/r) \\
&= D^\dagger(\mp\omega t/r)S(r)D(\mp\omega t/r). \tag{B.3}
\end{aligned}$$

Allowing this evolution for time t_1 followed by the usual sensing via a displacement operation acting for time t_2 gives the final state

$$\begin{aligned}
|\psi_f\rangle &= D(\pm\omega t_2)U(t_1)|0\rangle \\
&= D(\pm\omega t_2)D^\dagger(\mp\omega t_1/r)S(r)D(\mp\omega t_1/r)|0\rangle \tag{B.4}
\end{aligned}$$

noting that $r = 2\eta t_1$. Recalling that the squeezing and displacement operators are non-commutative, we use result (2.61) to find

$$\begin{aligned}
|\psi_f\rangle &= D(\pm\omega t_2)D^\dagger(\mp\omega t_1/r)D[(\mp\omega t_1/r)\cosh(r) - (\mp\omega t_1/r)\sinh(r)]S(r)|0\rangle \\
&= D(\pm\omega t_2)D^\dagger(\mp\omega t_1/r)D[(\mp\omega t_1/r)(\cosh(r) - \sinh(r))]S(r)|0\rangle \\
&= D(\pm\omega t_2)D(\pm\omega t_1/r)D[(\mp\omega t_1/r)e^{-r}]S(r)|0\rangle \\
&= D\left[\pm\omega(T - t_1) \pm \frac{\omega}{2\eta}(1 - e^{-2\eta t_1})\right]S(r)|0\rangle \tag{B.5}
\end{aligned}$$

where $T = t_1 + t_2$ is a fixed total time. We note that taking magnitude of the squeezing to zero should be equivalent to taking $t_1 \rightarrow 0$ and thus $t_2 = T$. We find

$$\lim_{t_1 \rightarrow 0} |\psi_f\rangle = D(\pm\omega T)|0\rangle \tag{B.6}$$

so if there is no squeezing (no preparation) we are indeed displacing (sensing) for the total time. Furthermore, we note that

$$(1 - e^{-2\eta t_1}) \in [0, 1] \quad (\text{B.7})$$

thus an extra displacement, in the desired direction for enhanced sensitivity, is guaranteed by this scheme. As a measure of sensitivity of the displacement of the oscillator along the position quadrature x_0 we use the propagation of error formula which gives

$$\delta^2 \omega = \frac{\text{Var}[x_0]}{\nu \left| \frac{\partial \langle x_0 \rangle}{\partial \omega} \right|^2} \quad (\text{B.8})$$

where ν is the number of experimental repeats so if the total time for all repeats of the experiment is given by T_{Tot} then $\nu = T_{Tot}/T$. We also have, for a displaced-squeezed state defined by $|\alpha, \zeta\rangle = D(\alpha)S(\zeta)|0\rangle$ with $\alpha = |\alpha|e^{i\theta}$ and $\zeta = re^{i\varphi}$, the expectation value of the general quadrature operator x_λ is given by

$$\langle \alpha, \zeta | x_\lambda | \alpha, \zeta \rangle = \frac{1}{2}(\alpha e^{-i\lambda} + \alpha^* e^{i\lambda}) \quad (\text{B.9})$$

and the variance is given by

$$\Delta^2 x_\lambda = \frac{1}{4} (e^{2r} \sin^2(\lambda - \varphi/2) + e^{-2r} \cos^2(\lambda - \varphi/2)) \quad (\text{B.10})$$

thus for $\lambda = \varphi = \theta$, the expectation value of the displaced-squeezed state in the position quadrature for the simultaneous scheme is given by

$$\langle x_0 \rangle = \alpha, \quad \Delta^2 x_0 = \frac{1}{4} e^{-2r} \quad (\text{B.11})$$

hence, for the sequential squeeze-then-sense scheme, where clearly $|\psi_f\rangle = D[\pm\omega(T - t_1)]S(r)|0\rangle$, the expectation value and variance in the position quadrature are given by

$$\langle x_0 \rangle_s = \pm\omega(T - t_1) \quad (\text{B.12})$$

and

$$\Delta_s^2 x_0 = \frac{1}{4} e^{-4\eta t_1}. \quad (\text{B.13})$$

For the final state of the concurrent preparation and sensing scheme (B.5), the expectation value and variance in the position quadrature are given by

$$\langle x_0 \rangle_c = \pm\omega \left((T - t_1) + \frac{1}{2\eta} (1 - e^{-2\eta t_1}) \right) \quad (\text{B.14})$$

and

$$\Delta_c^2 x_0 = \frac{1}{4} e^{-4\eta t_1}. \quad (\text{B.15})$$

Note that $\Delta_c^2 x_0 = \Delta_s^2 x_0$ so that

$$\begin{aligned} \delta^2 \omega_c &= \frac{\Delta_c^2 x_0}{\nu \left| \frac{\partial \langle x_0 \rangle_c}{\partial \omega} \right|^2} \\ &= \frac{\Delta_s^2 x_0}{\nu \left| \left((T - t_1) + \frac{1}{2\eta} (1 - e^{-2\eta t_1}) \right) \right|^2} \\ &= \frac{\Delta_s^2 x_0}{\nu |(T - t_1)|^2 \left| \left(1 + \frac{(1 - e^{-2\eta t_1})}{2\eta(T - t_1)} \right) \right|^2} \\ &= \frac{\Delta_s^2 x_0}{\nu \left| \frac{\partial \langle x_0 \rangle_s}{\partial \omega} \right|^2 \left| \left(1 + \frac{(1 - e^{-2\eta t_1})}{2\eta(T - t_1)} \right) \right|^2} \\ &= \delta^2 \omega_s \left| \left(1 + \frac{(1 - e^{-2\eta t_1})}{2\eta(T - t_1)} \right) \right|^{-2} \end{aligned} \quad (\text{B.16})$$

thus

$$\frac{\delta \omega_s}{\delta \omega_c} = 1 + \frac{1 - e^{-2\eta t_1}}{2\eta(T - t_1)} \quad (\text{B.17})$$

then since $T \geq t_1 \geq 0$ and $\eta \geq 0$, this implies

$$\frac{1 - e^{-2\eta t_1}}{2\eta(T - t_1)} \geq 0 \quad (\text{B.18})$$

hence, for all parameters

$$\frac{\delta \omega_s}{\delta \omega_c} \geq 1 \quad (\text{B.19})$$

thus we find that the concurrent state preparation and readout scheme will perform as well or outperform the analogous sequential scheme in agreement with the results of section 3.4.2.

B.1.1 Physical Setup

A system that submits the contrived Hamiltonian depicted by (B.1) is introduced here. We consider a nonlinear medium subject to a classical pump field and a coherent signal field. The initially fully quantized Hamiltonian is given by

$$\begin{aligned} H &= \hbar \omega_s a^\dagger a + \hbar \omega_p b^\dagger b + i \hbar \chi^{(2)} [a^2 b^\dagger - (a^\dagger)^2 b] \\ &\quad + \hbar \epsilon_p (b^\dagger + b) \sin(\omega_p t + \varphi) + \hbar \epsilon_s (a^\dagger + a) \sin(\omega_s t + \phi) \end{aligned} \quad (\text{B.20})$$

where, a is the driving field mode, b is the pump mode and the terms (from left to right) correspond to the free energy of the driving and pump field, the nonlinear coupling of the fields, the pump driving and the signal driving. Furthermore, the frequencies ω_s and ω_p correspond to the signal and pump fields respectively, $\chi^{(2)}$ is the 2nd order susceptibility, φ is a controlled phase and ϵ_p and ϵ_s are the amplitudes of the pump and signal respectively. Taking the pump field to be a coherent classical field that loses no photons over time t , i.e taking the parametric approximation, the operators b and b^\dagger become $\beta e^{-i\omega_p t}$ and $\beta^* e^{i\omega_p t}$ respectively, where $\beta = |\beta|e^{i\theta}$ with $\theta \in [0, 2\pi]$. Note that this essentially amounts to stating that for an operator $\hat{A} \approx A_0 + \delta\hat{A}$ where A_0 is a constant, we have $\langle A_0 \rangle \gg \langle \delta\hat{A} \rangle$ the operator \hat{A} is effectively a constant. Thus, ignoring the irrelevant constant terms, the Hamiltonian (B.20) becomes

$$H = \hbar\omega_s a^\dagger a + i\hbar[\eta^* e^{i\omega_p t} a^2 - \eta(a^\dagger)^2 e^{-i\omega_p t}] + \hbar\epsilon_s(a^\dagger + a) \sin(\omega_s t + \phi). \quad (\text{B.21})$$

where $\eta = \chi^{(2)}\beta$. Defining $H_0 = \hbar\omega_s a^\dagger a$, $H_1 = i\hbar[\eta^* e^{i\omega_p t} a^2 - \eta(a^\dagger)^2 e^{-i\omega_p t}]$ and $H_2 = \hbar\epsilon_s(a^\dagger + a) \sin(\omega_s t + \phi)$, the Hamiltonian in the interaction picture is given by

$$\begin{aligned} H_I &= e^{iH_0 t/\hbar} (H - H_0) e^{-iH_0 t/\hbar} \\ &= e^{iH_0 t/\hbar} (H_1 + H_2) e^{-iH_0 t/\hbar}. \end{aligned} \quad (\text{B.22})$$

Using the operator identity (2.37) we find

$$e^{i\omega_s t a^\dagger a} a e^{-i\omega_s t a^\dagger a} = a e^{-i\omega_s t} \quad \text{and} \quad e^{i\omega_s t a^\dagger a} a^\dagger e^{-i\omega_s t a^\dagger a} = a^\dagger e^{i\omega_s t} \quad (\text{B.23})$$

and using this result with the exponential form of the sine function, the Hamiltonian becomes

$$\begin{aligned} H_I &= i\hbar \left[\eta^* e^{i\omega_p t - 2i\omega_s t} a^2 - \eta e^{-i\omega_p t + 2i\omega_s t} (a^\dagger)^2 \right] + \hbar\epsilon \left(a^\dagger e^{i\omega_s t} + a e^{-i\omega_s t} \right) \left[\frac{e^{i(\omega_s t + \phi)} - e^{-i(\omega_s t + \phi)}}{2i} \right] \\ &= i\hbar \left[\eta^* e^{i\omega_p t - 2i\omega_s t} a^2 - \eta e^{-i\omega_p t + 2i\omega_s t} (a^\dagger)^2 \right] - \frac{i\hbar\epsilon_s}{2} \left(a^\dagger e^{i(2\omega_s t + \phi)} - a^\dagger e^{i\phi} + a e^{i\phi} - a e^{-i(2\omega_s t + \phi)} \right). \end{aligned} \quad (\text{B.24})$$

Finally, setting $\omega_p = 2\omega_s$, $\phi = 0$, taking $\eta, \epsilon_s \in \mathbb{R}^+$ and applying the rotating wave approximation, we arrive at the Hamiltonian

$$H_I = i\hbar\eta \left[a^2 - (a^\dagger)^2 \right] \pm \frac{i\hbar\epsilon_s}{2} (a^\dagger - a) \quad (\text{B.25})$$

where, the " \pm " allows for the displacement to be applied in the positive or negative direction along the position axis in phase space. Simply substituting $\epsilon_s/2 = \omega$ retrieves the desired Hamiltonian (B.1)

B.2 Lossy

Here we introduce loss into the optical system that was analysed in the preceding section and find that even in the presence of loss, concurrent state preparation and readout is a more effective use of the limited time resource.

B.2.1 Sequential Scheme

To include optical loss the system is coupled to a thermal bath which can be taken to be a multitude of harmonic oscillator modes of the electromagnetic field. The system-bath coupling is given by

$$V = \sum_k \lambda_k (ab_k^\dagger + a^\dagger b_k) \quad (\text{B.26})$$

which can be represented as k fictional beam-splitters coupling the system mode to the k field modes. The full Hamiltonian is then given by

$$H = H_0 + H_s(t) + V \quad (\text{B.27})$$

where the system and bath free energy is

$$H_0 = \omega_s a^\dagger a + \sum_k \omega_k b_k^\dagger b_k + \omega_p c^\dagger c \quad (\text{B.28})$$

and the squeeze-then-sense system Hamiltonian is given by

$$H_s(t) = \begin{cases} \hbar\omega_p c^\dagger c + i\hbar\chi[a^2 c^\dagger - (a^\dagger)^2 c] + \hbar\epsilon_p(c^\dagger + c) \sin(\omega_p t + \varphi) & \text{for } t \leq t_1 \\ \hbar\epsilon_s(a^\dagger + a) \sin(\omega_s t + \phi) & \text{for } t_1 \leq t \leq T \end{cases}$$

in which a is the system mode, b_k are the bath modes and c is a pump field taken to be a coherent classical field that loses no photons over the total time T . With this $c \approx \beta e^{-i\omega_p t}$ and $c^\dagger \approx \beta^* e^{i\omega_p t}$ where $\beta = |\beta|e^{i\theta}$, $\theta \in [0, 2\pi]$ then transforming to the interaction picture, in units of \hbar , yields

$$\begin{aligned} H_I(t) &= e^{iH_0 t} (H - H_0) e^{-iH_0 t} \\ &= e^{iH_0 t} H_s(t) e^{-iH_0 t} + e^{iH_0 t} V e^{-iH_0 t} \\ &= H'_s(t) + e^{iH_0 t} \sum_k \lambda_k (ab_k^\dagger + a^\dagger b_k) e^{-iH_0 t}. \end{aligned} \quad (\text{B.29})$$

Through the same reasoning given in the previous section, taking the RWA and setting $\omega_p = 2\omega_s$ gives

$$H'_s(t) = \begin{cases} i\eta[a^2 - (a^\dagger)^2] & \text{for } t \leq t_1 \\ \pm \frac{i\epsilon_s}{2}(a^\dagger - a) & \text{for } t_1 \leq t \leq T. \end{cases}$$

For the system-bath coupling term, transforming to the interaction frame gives

$$\begin{aligned}
V' &= e^{i\omega_s t a^\dagger a} a e^{-i\omega_s t a^\dagger a} \sum_k \lambda_k \exp\left(i \sum_{k'} b_k^\dagger b_{k'}\right) b_k^\dagger \exp\left(-i \sum_{k'} b_k^\dagger b_{k'} a\right) \\
&\quad + e^{i\omega_s t a^\dagger a} a^\dagger e^{-i\omega_s t a^\dagger a} \sum_k \lambda_k \exp\left(i \sum_{k'} b_k^\dagger b_{k'}\right) b_k \exp\left(-i \sum_{k'} b_k^\dagger b_{k'} a\right) \\
&= a \sum_k \lambda_k b_k^\dagger e^{i(\omega_s - \omega_k)t} + a^\dagger \sum_k \lambda_k b_k e^{-i(\omega_s - \omega_k)t} \\
&= a B^\dagger(t) + a^\dagger B(t)
\end{aligned} \tag{B.30}$$

where we have defined $B(t) \equiv \sum_k \lambda_k b_k e^{-i(\omega_s - \omega_k)t}$ noting that

$$[B(t), B^\dagger(t')] = \gamma \delta(t - t') \tag{B.31}$$

This is a consequence of the Markov approximation [53] which is a useful tool in analysing open quantum systems. An open quantum system consists of the system of interest and it's environment (in the above case, this environment is taken to be a thermal bath). Since the system will typically entangle with it's environment, even if initially described by a pure state it will evolve to a mixed state ρ . The Markov approximation comes about from the largeness of the bath (more strictly, the closeness of the energy levels), which ensures that from one moment to the next the system effectively interacts with a different part of the environment. In other words, the bath has no history. The Markov approximation is often used in combination with the Born approximation when analysing the dynamics of open quantum systems. The Born approximation is that the system-environment coupling is taken to be weak. This ensures that the affect of the system on the environment is negligible. Then rather than attempting to find a quantum state of the system - and by proxy, the environment - the approximate evolution of the system state alone is found. In other words, the Born-Markov approximations are used in order to find a differential equation for ρ . Following [53] the Heisenberg-picture dynamics of an the position system operator are found using

$$\dot{s}(t) = -i[H, s(t)] \tag{B.32}$$

where $s(t)$ is an arbitrary system operator (here, choose $s(t) = x(t)$) which can be set to the system's density matrix operator ρ thus revealing the master equation. However, the commutation relation (B.31) complicates this and as such, the infinitesimal evolution of $s(t)$ (in the Heisenberg picture) is considered. Furthermore, the infinitesimal operator is defined as

$$dB(t) = B(t)dt \tag{B.33}$$

so that the unitary operator for an infinitesimal evolution generated by V' (in the interaction frame) is given by

$$U_I(t + dt, t) = \exp \left[-i \left(adB^\dagger(t) + a^\dagger dB(t) \right) \right]. \quad (\text{B.34})$$

and commutation relation is found to be

$$[dB(t), dB^\dagger(t)] = dt \quad (\text{B.35})$$

using the heuristic notion that dt is the smallest increment of time and the area under a delta function is 1 thus $\delta(0)dt = 1 \Rightarrow \delta(0) = 1/dt$. From this it is apparent that $dB(t)$ is not of order dt but in fact of order \sqrt{dt} thus it is necessary to expand (B.34) to second order in it's argument when inspecting the infinitesimal evolution of $s(t)$ in the Heisenberg picture which is given by

$$\begin{aligned} \tilde{s}(t + dt) &= U_I^\dagger(t + dt, t') s(t') U_I(t + dt, t') \\ &= U_I^\dagger(t, t') U_I^\dagger(t + dt, t) U_I(t, t') U_I^\dagger(t, t') s(t') U_I(t, t') U_I^\dagger(t, t') U_I(t + dt, t') \\ &= \tilde{U}^\dagger(t + dt, t) \tilde{s}(t) \tilde{U}(t + dt, t) \end{aligned} \quad (\text{B.36})$$

where, “tilde” denotes Heisenberg picture operators. Then it is apparent that

$$\begin{aligned} \tilde{U}(t + dt, t) &= U_I^\dagger(t, t') e^{-iH_I(t)dt} U_I(t, t') \\ &\approx U_I^\dagger(t, t') \left(1 - iH_I(t)dt - \frac{1}{2}H_I^2(t)dt^2 \right) U_I(t, t') \\ &= U_I^\dagger(t, t') U_I(t, t') - iU_I^\dagger(t, t') H_I(t) U_I(t, t') dt \\ &\quad - \frac{1}{2}U_I^\dagger(t, t') H_I(t) U_I(t, t') U_I^\dagger(t, t') H_I(t) U_I(t, t') dt^2 \\ &= 1 - i\tilde{H}_I(t)dt - \frac{1}{2}\tilde{H}_I^2(t)dt^2 \\ &= 1 - i \left(\tilde{H}'_s(t) + \tilde{a}\tilde{B}^\dagger(t) + \tilde{a}^\dagger\tilde{B}(t) \right) dt - \frac{1}{2} \left(\tilde{H}'_s(t) + \tilde{a}\tilde{B}^\dagger(t) + \tilde{a}^\dagger\tilde{B}(t) \right)^2 dt^2 \\ &= 1 - i \left(\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t) \right) - \frac{1}{2} \left(\tilde{H}'_s(t)^2 dt^2 + \tilde{H}'_s(t)\tilde{a}d\tilde{B}^\dagger(t)dt \right. \\ &\quad \left. + \tilde{H}'_s(t)\tilde{a}^\dagger d\tilde{B}(t)dt + \tilde{a}d\tilde{B}^\dagger(t)\tilde{H}'_s(t)dt + \tilde{a}^\dagger d\tilde{B}(t)\tilde{H}'_s(t)dt \right. \\ &\quad \left. + \tilde{a}\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) + \tilde{a}^\dagger\tilde{a}d\tilde{B}(t)d\tilde{B}^\dagger(t) + (\tilde{a}d\tilde{B}^\dagger(t))^2 + (\tilde{a}^\dagger d\tilde{B}(t))^2 \right) \\ &= 1 - i \left(\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t) \right) - \frac{1}{2} \left(\tilde{a}\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) + \tilde{a}^\dagger\tilde{a}d\tilde{B}(t)d\tilde{B}^\dagger(t) \right. \\ &\quad \left. + (\tilde{a}d\tilde{B}^\dagger(t))^2 + (\tilde{a}^\dagger d\tilde{B}(t))^2 \right) + \mathcal{O}(dt^{3/2}) + \mathcal{O}(dt^2) \end{aligned} \quad (\text{B.37})$$

thus,

$$\begin{aligned}
\tilde{s}(t+dt) &= \tilde{s}(t) - \tilde{s}(t)i \left(\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t) \right) - \tilde{s}(t)\frac{1}{2} \left(\tilde{a}\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) + \tilde{a}^\dagger \tilde{a} d\tilde{B}(t)d\tilde{B}^\dagger(t) \right. \\
&\quad \left. + (\tilde{a}d\tilde{B}^\dagger(t))^2 + (\tilde{a}^\dagger d\tilde{B}(t))^2 \right) + i \left(\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t) \right) \tilde{s}(t) \\
&\quad + \left(\tilde{a}^\dagger d\tilde{B}(t) + \tilde{a}d\tilde{B}^\dagger(t) \right) \tilde{s}(t) \left(\tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t) \right) - \frac{1}{2} \left(\tilde{a}\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) \right. \\
&\quad \left. + \tilde{a}^\dagger \tilde{a} d\tilde{B}(t)d\tilde{B}^\dagger(t) + (\tilde{a}d\tilde{B}^\dagger(t))^2 + (\tilde{a}^\dagger d\tilde{B}(t))^2 \right) + \mathcal{O}(dt^{3/2}) + \mathcal{O}(dt^2) \\
&= \tilde{s}(t) + i[\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t), \tilde{s}(t)] + \left(\tilde{a}^\dagger \tilde{s}(t)\tilde{a}d\tilde{B}(t)d\tilde{B}^\dagger(t) + \tilde{a}^\dagger \tilde{s}(t)\tilde{a}^\dagger (d\tilde{B}(t))^2 \right. \\
&\quad \left. + \tilde{a}\tilde{s}(t)\tilde{a}(d\tilde{B}^\dagger(t))^2 + \tilde{a}\tilde{s}(t)\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) \right) \\
&\quad - \frac{1}{2} \left\{ \tilde{a}\tilde{a}^\dagger d\tilde{B}^\dagger(t)d\tilde{B}(t) + \tilde{a}^\dagger \tilde{a} d\tilde{B}(t)d\tilde{B}^\dagger(t) + (\tilde{a}d\tilde{B}^\dagger(t))^2 + (\tilde{a}^\dagger d\tilde{B}(t))^2, \tilde{s}(t) \right\}
\end{aligned} \tag{B.38}$$

ignoring terms of order $\mathcal{O}(dt^{3/2})$ where we have used $[\tilde{s}(t), d\tilde{B}(t)] = 0$ and $\{\cdot, \cdot\}$ denotes the anti-commutator. We have for a general bath

$$d\tilde{B}^\dagger(t)d\tilde{B}(t) = \gamma N dt \tag{B.39}$$

$$d\tilde{B}(t)d\tilde{B}^\dagger(t) = \gamma(N+1)dt \tag{B.40}$$

$$d\tilde{B}(t)d\tilde{B}(t) = \gamma M dt \tag{B.41}$$

$$\langle d\tilde{B}(t) \rangle = \beta dt \tag{B.42}$$

where a $M = 0$ and $N = (\exp[\hbar\omega_k/k_B T] - 1)^{-1}$ for a white noise bath - which is a good approximation of a thermal bath. This implies

$$\begin{aligned}
\tilde{s}(t+dt) &= \tilde{s}(t) + i[\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t), \tilde{s}(t)] + \tilde{a}^\dagger \tilde{s}(t)\tilde{a}\gamma(N+1)dt + \tilde{a}\tilde{s}(t)\tilde{a}^\dagger \gamma N dt \\
&\quad - \frac{1}{2} \left\{ \tilde{a}^\dagger \tilde{a}\gamma(N+1)dt + \tilde{a}\tilde{a}^\dagger \gamma N dt, \tilde{s}(t) \right\} \\
&= \tilde{s}(t) + i[\tilde{H}'_s(t)dt + \tilde{a}d\tilde{B}^\dagger(t) + \tilde{a}^\dagger d\tilde{B}(t), \tilde{s}(t)] + \tilde{a}^\dagger \tilde{s}(t)\tilde{a}\gamma(N+1)dt + \tilde{a}\tilde{s}(t)\tilde{a}^\dagger \gamma N dt \\
&\quad - \frac{dt}{2} \gamma(N+1) \left(\tilde{a}^\dagger \tilde{a}\tilde{s}(t) + \tilde{s}(t)\tilde{s}(t)\tilde{a}^\dagger \tilde{a} \right) - \frac{dt}{2} \gamma N \left(\tilde{a}\tilde{a}^\dagger \tilde{s}(t) + \tilde{s}(t)\tilde{a}\tilde{a}^\dagger \right)
\end{aligned} \tag{B.43}$$

so finally, the rate of change of a general system observable is given by

$$\begin{aligned}
\frac{d\tilde{s}}{dt} &= \frac{\tilde{s}(t+dt) - \tilde{s}(t)}{dt} = i[\tilde{H}'_s(t) + \tilde{a}\tilde{B}^\dagger(t) + \tilde{a}^\dagger \tilde{B}(t), \tilde{s}(t)] + \gamma N \left(\tilde{a}\tilde{s}(t)\tilde{a}^\dagger - \frac{1}{2}\tilde{a}\tilde{a}^\dagger \tilde{s}(t) - \frac{1}{2}\tilde{s}(t)\tilde{a}\tilde{a}^\dagger \right) \\
&\quad + \gamma(N+1) \left(\tilde{a}^\dagger \tilde{s}(t)\tilde{a} - \frac{1}{2}\tilde{a}^\dagger \tilde{a}\tilde{s}(t) - \frac{1}{2}\tilde{s}(t)\tilde{a}^\dagger \tilde{a} \right).
\end{aligned} \tag{B.44}$$

The position quadrature operator is given by

$$x(t) = \frac{1}{2}(a(t) + a^\dagger(t)) \tag{B.45}$$

thus the rate of change (in the Heisenberg picture) is given by

$$\frac{d\tilde{x}}{dt} = \frac{1}{2} \left(\frac{d\tilde{a}}{dt} + \frac{d\tilde{a}^\dagger}{dt} \right) \quad (\text{B.46})$$

then from (B.44)

$$\begin{aligned} \frac{d\tilde{a}}{dt} &= i[\tilde{H}'_s(t) + \tilde{a}\tilde{B}^\dagger(t) + \tilde{a}^\dagger\tilde{B}(t), \tilde{a}] + \gamma N \left(\tilde{a}\tilde{a}\tilde{a}^\dagger - \frac{1}{2}\tilde{a}\tilde{a}^\dagger\tilde{a} - \frac{1}{2}\tilde{a}\tilde{a}\tilde{a}^\dagger \right) \\ &\quad + \gamma(N+1) \left(\tilde{a}^\dagger\tilde{a}\tilde{a} - \frac{1}{2}\tilde{a}^\dagger\tilde{a}\tilde{a} - \frac{1}{2}\tilde{a}\tilde{a}^\dagger\tilde{a} \right) \\ &= i \left(\tilde{H}'_s(t)\tilde{a} + \tilde{a}\tilde{B}^\dagger(t)\tilde{a} + \tilde{a}^\dagger\tilde{B}(t)\tilde{a} \right) - i \left(\tilde{a}\tilde{H}'_s(t) + \tilde{a}\tilde{B}^\dagger(t) + \tilde{a}\tilde{a}^\dagger\tilde{B}(t) \right) \\ &\quad + \gamma N \left(\tilde{a}\frac{1}{2}(\tilde{a}\tilde{a}^\dagger - \tilde{a}^\dagger\tilde{a}) \right) + \gamma(N+1) \left((\tilde{a}^\dagger\tilde{a} - \tilde{a}\tilde{a}^\dagger)\tilde{a}\frac{1}{2} \right) \\ &= i\tilde{H}'_s(t)\tilde{a} - i\tilde{a}\tilde{H}'_s(t) + i\tilde{a}^\dagger\tilde{a}\tilde{B}(t) - i\tilde{a}\tilde{a}^\dagger\tilde{B}(t) \\ &\quad + i\tilde{a}\tilde{a}\tilde{B}^\dagger(t) - i\tilde{a}\tilde{a}\tilde{B}^\dagger(t) + \gamma N\frac{\tilde{a}}{2} - \gamma(N+1)\frac{\tilde{a}}{2} \\ &= i[\tilde{H}_s(t), \tilde{a}] + i\tilde{B}(t)[\tilde{a}^\dagger, \tilde{a}] + \gamma\frac{\tilde{a}}{2}(N - N - 1) \\ &= i[\tilde{H}_s(t), \tilde{a}] - i\tilde{B}(t) - \frac{\gamma\tilde{a}}{2} \end{aligned} \quad (\text{B.47})$$

and

$$\begin{aligned} \frac{d\tilde{a}^\dagger}{dt} &= -i[\tilde{H}_s(t), \tilde{a}]^\dagger + i\tilde{B}^\dagger(t) - \frac{\gamma\tilde{a}}{2} \\ &= -i \left(\left(\tilde{H}'_s(t)\tilde{a} \right)^\dagger - \left(\tilde{a}\tilde{H}'_s(t) \right)^\dagger \right) + i\tilde{B}^\dagger(t) - \frac{\gamma\tilde{a}}{2} \\ &= -i \left(\tilde{a}^\dagger\tilde{H}'_s(t) - \tilde{H}'_s(t)\tilde{a} \right) + i\tilde{B}^\dagger(t) - \frac{\gamma\tilde{a}}{2} \\ &= i[\tilde{H}_s(t), \tilde{a}^\dagger] + i\tilde{B}^\dagger(t) - \frac{\gamma\tilde{a}}{2}. \end{aligned} \quad (\text{B.48})$$

The commutation relations give

$$[\tilde{H}'_s(t), \tilde{a}] = \begin{cases} [i\eta(\tilde{a}^2 - (\tilde{a}^\dagger)^2), \tilde{a}] & \text{for } t \leq t_1 \\ [\pm \frac{i\epsilon_s}{2}(\tilde{a}^\dagger - \tilde{a}), \tilde{a}] & \text{for } t_1 \leq t \leq T \end{cases}$$

then using that

$$[\tilde{a}, \tilde{a}] = 0 \quad (\text{B.49})$$

$$[\tilde{a}^\dagger, \tilde{a}] = -1 \quad (\text{B.50})$$

$$[(\tilde{a}^\dagger)^2, \tilde{a}] = -2\tilde{a}^\dagger \quad (\text{B.51})$$

it is evident that

$$\begin{aligned} [\tilde{H}'_s(t), \tilde{a}] &= \begin{cases} [-i\eta((\tilde{a}^\dagger)^2), \tilde{a}] & \text{for } t \leq t_1 \\ [\pm \frac{i\epsilon_s}{2}(\tilde{a}^\dagger), \tilde{a}] & \text{for } t_1 \leq t \leq T \end{cases} \\ &= \begin{cases} 2i\eta\tilde{a}^\dagger & \text{for } t \leq t_1 \\ \mp \frac{i\epsilon_s}{2} & \text{for } t_1 \leq t \leq T \end{cases} \end{aligned} \quad (\text{B.52})$$

and similarly

$$[\tilde{H}'_s(t), \tilde{a}^\dagger] = \begin{cases} 2i\eta\tilde{a} & \text{for } t \leq t_1 \\ \mp \frac{i\epsilon_s}{2} & \text{for } t_1 \leq t \leq T. \end{cases}$$

so that

$$\begin{aligned} \frac{d\tilde{x}}{dt} &= \begin{cases} \frac{1}{2} \left(-2\tilde{a}^\dagger\eta - i\tilde{B}(t) - \frac{\gamma}{2}\tilde{a} - 2\tilde{a}\eta + i\tilde{B}^\dagger(t) - \frac{\gamma}{2}\tilde{a}^\dagger \right) & \text{for } t \leq t_1 \\ \frac{1}{2} \left(\pm \frac{\epsilon_s}{2} - i\tilde{B}(t) - \frac{\gamma}{2}\tilde{a} \pm \frac{\epsilon_s}{2} + i\tilde{B}^\dagger(t) - \frac{\gamma}{2}\tilde{a}^\dagger \right) & \text{for } t_1 \leq t \leq T. \end{cases} \\ &= \begin{cases} -2\eta\tilde{x}(t) - \frac{\gamma}{2}\tilde{x}(t) + \frac{i}{2} \left(\tilde{B}^\dagger(t) - \tilde{B}(t) \right) & \text{for } t \leq t_1 \\ \pm \frac{\epsilon_s}{2} - \frac{\gamma}{2}\tilde{x}(t) + \frac{i}{2} \left(\tilde{B}^\dagger(t) - \tilde{B}(t) \right) & \text{for } t_1 \leq t \leq T. \end{cases} \\ &= \kappa(t)\tilde{x}(t) + v(t) \end{aligned} \quad (\text{B.53})$$

where the following have been defined

$$\kappa(t) \equiv \begin{cases} -2\eta - \frac{\gamma}{2} & \text{for } t \leq t_1 \\ -\frac{\gamma}{2} & \text{for } t_1 \leq t \leq T. \end{cases} \quad (\text{B.54})$$

$$v(t) \equiv \begin{cases} \frac{i}{2} \left(\tilde{B}^\dagger(t) - \tilde{B}(t) \right) & \text{for } t \leq t_1 \\ \pm \frac{\epsilon_s}{2} + \frac{i}{2} \left(\tilde{B}^\dagger(t) - \tilde{B}(t) \right) & \text{for } t_1 \leq t \leq T. \end{cases} \quad (\text{B.55})$$

Relabelling the time increment to $t_0 \leq t \leq t_1$ for the squeezing (preparation) time and $t_1 \leq t \leq t_2$ for the sensing time (the previous notation corresponds to $t_0 = 0$ and $t_2 = T$), the solution of this differential equation for the system position operator at the time t is given by

$$\tilde{x}(t) = e^{\int_{t_0}^{t_2} \kappa(t'') dt''} \left[+ \int_{t_0}^{t_2} e^{\int_{t'}^{t_2} \kappa(t'') dt''} v(t') dt' + C \right] \quad (\text{B.56})$$

setting $t = t_0$ implies $\tilde{x}_0(t) = C$ and thus the solution at final time t_2 is given by

$$\tilde{x}(t_2) = \tilde{x}(t_0) e^{\int_{t_0}^{t_2} \kappa(t'') dt''} + \int_{t_0}^{t_2} e^{\int_{t'}^{t_2} \kappa(t'') dt''} v(t') dt' \quad (\text{B.57})$$

where,

$$\begin{aligned}
\int_{t_0}^{t_2} \kappa(t'') dt'' &= \int_{t_0}^{t_1} \kappa(t'') dt'' + \int_{t_1}^{t_2} \kappa(t'') dt'' \\
&= \left(-2\eta - \frac{\gamma}{2}\right) (t_1 - t_0) + \left(-\frac{\gamma}{2}\right) (t_2 - t_1) \\
&= -2\eta t_1 = \frac{\gamma}{2} t_1 - 2\eta t_0 + \frac{\gamma}{2} t_2 + \frac{\gamma}{2} t_1 \\
&= -2\eta(t_1 - t_0) - \frac{\gamma}{2}(t_2 - t_0)
\end{aligned} \tag{B.58}$$

noting that

$$\begin{aligned}
e^{\int_{t_0}^{t_2} \kappa(t'') dt''} e^{-\int_{t_0}^{t'} \kappa(t'') dt''} &= \exp \left[\int_{t'}^{t_0} \kappa(t'') dt'' + \int_{t_0}^{t_2} \kappa(t'') dt'' \right] \\
&= \exp \left[\int_{t'}^{t_2} \kappa(t'') dt'' \right].
\end{aligned} \tag{B.59}$$

Furthermore,

$$\begin{aligned}
\int_{t_0}^{t_2} e^{\int_{t'}^{t_2} \kappa(t'') dt''} v(t') dt' &= \int_{t_0}^{t_1} e^{\int_{t'}^{t_1} \kappa(t'') dt'' + \int_{t_1}^{t_2} \kappa(t'') dt''} v(t') dt' + \int_{t_1}^{t_2} e^{\int_{t'}^{t_2} \kappa(t'') dt''} v(t') dt' \\
&= \int_{t_0}^{t_1} e^{-(2\eta + \frac{\gamma}{2})(t_1 - t') - \frac{\gamma}{2}(t_2 - t_1)} v(t') dt' + \int_{t_1}^{t_2} e^{-\frac{\gamma}{2}(t_2 - t')} v(t') dt'
\end{aligned} \tag{B.60}$$

noting that $t' \in [t_0, t_2]$ and $t'' \in [t', t_2]$ which can be split into the two cases where $t' \in [t_0, t_1] \Rightarrow t'' \in [t_0, t_2]$ and $t' \in [t_1, t_2] \Rightarrow t'' \in [t_1, t_2]$, i.e if t' does not exist between t_0 and t_1 , then neither can t'' . The expression for the system position operator in the Heisenberg picture is then

$$\begin{aligned}
\tilde{X}(t_2) &= \tilde{X}(t_0) e^{-2\eta(t_1 - t_0) - \gamma(t_2 - t_0)/2} + \int_{t_0}^{t_1} e^{(-2\eta - \gamma/2)(t_1 - t') - \gamma(t_2 - t_1)/2} \frac{i}{2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) dt' \\
&\quad + \int_{t_1}^{t_2} e^{-\gamma(t_2 - t')/2} \left(\frac{\epsilon_s}{2} + \frac{i}{2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) \right) dt'.
\end{aligned} \tag{B.61}$$

Recall that a thermal state in the Fock basis is given by

$$\rho = \sum_n \frac{\langle n \rangle^n}{(1 + \langle n \rangle)^{n+1}} |n\rangle \langle n| = \sum_n c_n |n\rangle \langle n| \tag{B.62}$$

thus for the thermal state of our bath

$$\langle \tilde{B}(t) \rangle = \text{Tr}[\rho \tilde{B}(t)] = \text{Tr} \left[\sum_n c_n |n\rangle \langle n| \tilde{B}(t) \right] = \sum_n c_n \langle n| \tilde{B}(t) |n\rangle = 0 \tag{B.63}$$

and similarly $\langle \tilde{B}^\dagger(t) \rangle = 0$. Descriptively; because the thermal state is diagonal in the Fock basis, the expectation values of it's associated ladder operator will vanish. Consequently,

$$\begin{aligned}
\langle \tilde{X}(t_2) \rangle &= \langle \tilde{X}(t_0) \rangle e^{-2\eta(t_1 - t_0) - \gamma(t_2 - t_0)/2} + \frac{\epsilon_s}{2} \int_{t_1}^{t_2} e^{-\gamma(t_2 - t')/2} dt' \\
&= \langle \tilde{X}(t_0) \rangle e^{-2\eta(t_1 - t_0) - \gamma(t_2 - t_0)/2} + \frac{\epsilon_s}{\gamma} \left(1 - e^{-\gamma(t_2 - t_1)/2} \right)
\end{aligned} \tag{B.64}$$

and further

$$\begin{aligned}
\langle \tilde{X}^2(t_2) \rangle &= \left\langle \left(\tilde{X}(t_0) e^{-2\eta(t_1-t_0)-\gamma(t_2-t_0)/2} + \int_{t_0}^{t_1} e^{(-2\eta-\gamma/2)(t_1-t')-\gamma(t_2-t_1)/2} \frac{i}{2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) dt' \right. \right. \\
&\quad \left. \left. + \int_{t_1}^{t_2} e^{-\gamma(t_2-t')/2} \left(\frac{\epsilon_s}{2} + \frac{i}{2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) \right) dt' \right)^2 \right\rangle \\
&= \langle \tilde{X}^2(t) \rangle e^{-4\eta(t_1-t_0)-\gamma(t_2-t_0)} + 2 \langle \tilde{X}(t_0) \rangle e^{-2\eta(t_1-t_0)-\gamma(t_2-t_1)/2} \frac{\epsilon_s}{\gamma} (1 - e^{-\gamma(t_2-t_1)/2}) \\
&\quad + \frac{1}{4} \gamma (2N+1) \left[\frac{1}{\gamma+4\eta} (e^{-\gamma(t_2-t_1)} - e^{-\gamma(t_2-t_0)-4\eta(t_1-t_0)}) + \frac{1}{4} (1 - e^{-\gamma(t_2-t_1)}) \right] \\
&\quad + \frac{\epsilon_s^2}{\gamma^2} (1 - e^{-\gamma(t_2-t_1)})^2 \tag{B.65}
\end{aligned}$$

where we have use the fact that terms like the following

$$\begin{aligned}
& -\frac{1}{4} \int_{t_0}^{t_1} e^{(-2\eta-\gamma/2)(t_1-t')-\gamma(t_2-t_1)/2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) dt' \int_{t_1}^{t_2} e^{-\gamma(t_2-t')/2} (\tilde{B}^\dagger(t) - \tilde{B}(t)) dt' = \\
& \int_{t_0}^{t_1} dt' \int_{t_1}^{t_2} dt'' e^{-2\eta(t_1-t')-\gamma(t_2-t')/2-\gamma(t_2-t'')/2} (\tilde{B}^\dagger(t') \tilde{B}^\dagger(t'') + \tilde{B}^\dagger(t') \tilde{B}(t'') + \tilde{B}(t') \tilde{B}^\dagger(t'') + \tilde{B}(t') \tilde{B}(t'')) \\
& = \int_{t_0}^{t_1} dt' \int_{t_1}^{t_2} dt'' e^{-2\eta(t_1-t')-\gamma(t_2-t')/2-\gamma(t_2-t'')/2} (\gamma N \delta(t' - t'') + \gamma(N+1) \delta(t' - t'')) \\
& = \int_{t_1}^{t_1} dt' e^{-2\eta(t_1-t')-\gamma(t_2-t')/2} \gamma(2N+1) = 0 \tag{B.66}
\end{aligned}$$

and $\tilde{B}(t') \tilde{B}^\dagger(t'') = \gamma N \delta(t' - t'')$ etc. The variance is then found to be

$$\begin{aligned}
\Delta_s^2 \tilde{X}(t_2) &= \Delta_s^2 \tilde{X}(t_0) e^{-4\eta(t_1-t_0)-\gamma(t_2-t_0)} + \frac{2N+1}{4} (1 - e^{-\gamma(t_2-t_1)}) \\
&\quad + \frac{\gamma(2N+1)}{4(\gamma+4\eta)} (e^{-\gamma(t_2-t_1)} - e^{-\gamma(t_2-t_0)-4\eta(t_1-t_0)}) \tag{B.67}
\end{aligned}$$

where for $|\psi(t_0)\rangle = |0\rangle \Rightarrow \Delta_s^2 \tilde{X}(t_0) = \frac{1}{4} \langle 0 | \tilde{a}^2 + \tilde{a} \tilde{a}^\dagger + \tilde{a}^\dagger \tilde{a} + (\tilde{a}^\dagger)^2 | 0 \rangle = \frac{1}{4}$

B.2.2 Simultaneous

The (Heisenberg) system Hamiltonian in the interaction picture for the simultaneous scheme is given by

$$\tilde{H}'_s(t) = \begin{cases} i\eta(\tilde{a}^2 + (\tilde{a}^\dagger)^2) + \frac{i\epsilon_s}{2}(\tilde{a}^\dagger - \tilde{a}) & \text{for } t_0 \leq t \leq t_1 \\ \frac{i\epsilon_s}{2}(\tilde{a}^\dagger - \tilde{a}) & \text{for } t_1 \leq t \leq t_2 \end{cases} \tag{B.68}$$

noting that the resulting displacements in the $t \in [t_0, t_1]$ and $t \in [t_1, t_2]$ will be in opposite directions as desired (*c.f.* lossless regime). As such

$$\begin{aligned}
[\tilde{H}'_s(t), \tilde{a}] &= \begin{cases} [i\eta(\tilde{a}^2 + (\tilde{a}^\dagger)^2) + \frac{i\epsilon_s}{2}(\tilde{a}^\dagger - \tilde{a}), \tilde{a}] & \text{for } t_0 \leq t \leq t_1 \\ [\frac{i\epsilon_s}{2}(\tilde{a}^\dagger - \tilde{a}), \tilde{a}] & \text{for } t_1 \leq t \leq t_2 \end{cases} \\
&= \begin{cases} -i\eta[(\tilde{a}^\dagger)^2, \tilde{a}] + \frac{i\epsilon_s}{2}[\tilde{a}^\dagger, \tilde{a}] & \text{for } t_0 \leq t \leq t_1 \\ \frac{i\epsilon_s}{2}[\tilde{a}^\dagger, \tilde{a}] & \text{for } t_1 \leq t \leq t_2 \end{cases} \\
&= \begin{cases} 2i\eta\tilde{a}^\dagger - \frac{i\epsilon_s}{2} & \text{for } t_0 \leq t \leq t_1 \\ -\frac{i\epsilon_s}{2} & \text{for } t_1 \leq t \leq t_2 \end{cases} \tag{B.69}
\end{aligned}$$

and further

$$\begin{aligned}
[\tilde{H}'_s(t), \tilde{a}]^\dagger &= -[\tilde{H}'_s(t), \tilde{a}^\dagger] = \begin{cases} -2i\eta\tilde{a}^\dagger + \frac{i\epsilon_s}{2} & \text{for } t_0 \leq t \leq t_1 \\ +\frac{i\epsilon_s}{2} & \text{for } t_1 \leq t \leq t_2 \end{cases} \\
\Rightarrow [\tilde{H}'_s(t), \tilde{a}^\dagger] &= \begin{cases} 2i\eta\tilde{a} - \frac{i\epsilon_s}{2} & \text{for } t_0 \leq t \leq t_1 \\ -\frac{i\epsilon_s}{2} & \text{for } t_1 \leq t \leq t_2. \end{cases} \tag{B.70}
\end{aligned}$$

From this and equations (B.46-B.48) the rate of change of the position operator is given by

$$\begin{aligned}
\frac{d\tilde{x}}{dt} &= \begin{cases} -2\eta\tilde{x}(t) - \frac{\gamma}{2}\tilde{x}(t) + \frac{\epsilon_s}{2} + \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(t)) & \text{for } t_0 \leq t \leq t_1 \\ \frac{\epsilon_s}{2} - \frac{\gamma}{2}\tilde{x}(t) + \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(t)) & \text{for } t_1 \leq t \leq t_2. \end{cases} \\
&= \kappa(t)\tilde{x}(t) + v(t) \tag{B.71}
\end{aligned}$$

where $\kappa(t)$ is given by (B.55) and

$$v(t) \equiv \begin{cases} \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(t)) & \text{for } t_0 \leq t \leq t_1 \\ \pm\frac{\epsilon_s}{2} + \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(t)) & \text{for } t_1 \leq t \leq t_2. \end{cases} \tag{B.72}$$

so equations (B.56-B.60) hold as before but with (B.72) so that

$$\begin{aligned}
\tilde{X}(t_2) &= \tilde{X}(t_0)e^{-2\eta(t_1-t_0)-\gamma(t_2-t_0)/2} + \int_{t_0}^{t_1} e^{(-2\eta-\gamma/2)(t_1-t')-\gamma(t_2-t_1)/2} \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(t))dt' \\
&\quad + \int_{t_1}^{t_2} e^{-\gamma(t_2-t')/2} \left(\frac{\epsilon_s}{2} + \frac{i}{2}(\tilde{B}^\dagger(t) - \tilde{B}(1t)) \right) dt' + \frac{\epsilon_s}{2} \int_{t_0}^{t_1} e^{(-2\eta-\gamma/2)(t_1-t')-\gamma(t_2-t_1)} dt'. \tag{B.73}
\end{aligned}$$

where the final term is the contribution due to the simultaneous preparation and sensing; an additional displacement. The expectation value is then found to be

$$\langle \tilde{X}(t_2) \rangle = \langle \tilde{X}(t_0) \rangle + \frac{\epsilon_s}{\gamma} (1 - e^{-\gamma(t_2-t_1)/2}) \left[1 + \left(\frac{1}{1 + 4\eta/\gamma} \right) (1 - e^{-(2\eta+\gamma/2)(t_1-t_0)}) \right] \quad (\text{B.74})$$

where we have used (B.63). Furthermore

$$\begin{aligned} \langle \tilde{X}^2(t_2) \rangle = & \langle \tilde{X}^2(t_0) \rangle e^{-4\eta(t_1-t_0)-\gamma(t_2-t_0)} \\ & + \epsilon_s^2 \left[\frac{1}{\gamma + 4\eta} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)/2-2\eta(t_1-t_0)} + \frac{1}{\gamma} (1 - e^{-\gamma(t_2-t_0)}) \right) \right]^2 \\ & + 2 \langle \tilde{X}^2(t_0) \rangle e^{-2\eta(t_1-t_0)-\gamma(t_2-t_0)/2} \frac{\epsilon_s}{\gamma + 4\eta} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)/2-2\eta(t_1-t_0)} + \frac{1}{\gamma} (1 - e^{-\gamma(t_2-t_0)}) \right) \\ & + \frac{1}{4} \int_{t_0}^{t_1} dt' \int_{t_0}^{t_1} dt'' e^{(-2\eta-\gamma/2)(t_1-t')-\gamma(t_2-t_1)+(-2\eta-\gamma/2)(t_1-t'')-\gamma(t_2-t_1)/2} [\gamma(N+1)\delta(t'-t'') \\ & \quad \quad \quad + \gamma N \delta(t'-t'')] \\ & - \frac{1}{4} \int_{t_1}^{t_2} dt' \int_{t_1}^{t_2} dt'' e^{-\gamma(t_2-t')/2-\gamma(t_2-t'')} (-\gamma(N+1)\delta(t'-t'') - \gamma N \delta(t'-t'')) \end{aligned} \quad (\text{B.75})$$

which reveals

$$\begin{aligned} \langle \tilde{X}^2(t_2) \rangle = & \langle \tilde{X}^2(t_0) \rangle e^{-4\eta(t_1-t_0)-\gamma(t_2-t_0)} + \epsilon_s^2 \left[\frac{1}{\gamma + 4\eta} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)/2-2\eta(t_1-t_0)} + \frac{1}{\gamma} (1 - e^{-\gamma(t_2-t_0)}) \right) \right]^2 \\ & + 2 \langle \tilde{X}^2(t_0) \rangle e^{-2\eta(t_1-t_0)-\gamma(t_2-t_0)/2} \frac{\epsilon_s}{\gamma + 4\eta} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)/2-2\eta(t_1-t_0)} + \frac{1}{\gamma} (1 - e^{-\gamma(t_2-t_0)}) \right) \\ = & \frac{\gamma(2N+1)}{4(\gamma+4\eta)} \left(e^{-\gamma(t_2-t_1)} - e^{-\gamma(t_2-t_0)-4\eta(t_1-t_0)} \right) + \frac{\gamma(2N+1)}{4\gamma} \left(1 - e^{-\gamma(t_2-t_1)} \right) \end{aligned} \quad (\text{B.76})$$

so that the variance is given by

$$\begin{aligned} \Delta_c^2 \tilde{X}(t_2) = & \Delta_c^2 \tilde{X}(t_0) e^{-4\eta(t_1-t_0)-\gamma(t_2-t_0)} + \frac{2N+1}{4} \left(1 - e^{-\gamma(t_2-t_1)} \right) \\ & + \frac{\gamma(2N+1)}{4(\gamma+4\eta)} \left(e^{-\gamma(t_2-t_1)} - e^{-\gamma(t_2-t_0)-4\eta(t_1-t_0)} \right). \end{aligned} \quad (\text{B.77})$$

Comparison and Checks

Now we have from the error of propagation formula

$$\delta^2 \epsilon_s^c = \frac{(t_2 - t_0)}{T} \frac{\Delta_s^2 \tilde{X}(t_2)}{\left| \frac{\partial \langle \tilde{X}(t_2) \rangle_c}{\partial \epsilon_s} \right|^2} \quad (\text{B.78})$$

and we note

$$\frac{\partial \langle \tilde{X}(t_2) \rangle_s}{\partial \epsilon_s} = \frac{1}{\gamma} \left(1 - e^{-\gamma(t_2-t_1)/2} \right) \quad (\text{B.79})$$

and

$$\begin{aligned}
& \left| \frac{\partial \langle \tilde{X}(t_2) \rangle_c}{\partial \epsilon_s} \right|^2 = \\
& \left| \left(1 + \frac{1}{1 + 4\eta/\gamma} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)-2\eta(t_1-t_0)} \right) \left[\left(1 - e^{-\gamma(t_2-t_1)} \right) \right]^{-1} \right) \right|^2 \left| \frac{1}{\gamma} \left(1 - e^{-\gamma(t_2-t_1)} \right) \right|^2 \\
& \left| \left(1 + \frac{1}{1 + 4\eta/\gamma} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)-2\eta(t_1-t_0)} \right) \left[\left(1 - e^{-\gamma(t_2-t_1)} \right) \right]^{-1} \right) \right|^2 \cdot \left| \frac{\partial \langle \tilde{X}(t_2) \rangle_s}{\partial \epsilon_s} \right|^2
\end{aligned} \tag{B.80}$$

then since $\Delta_s^2 \tilde{X}(t_2) = \Delta_c^2 \tilde{X}(t_2)$

$$\begin{aligned}
& \delta^2 \epsilon_s^c = \\
& \frac{(t_2 - t_0)}{T} \frac{\Delta_s \tilde{X}(t_2)}{\left| \frac{\partial \langle \tilde{X}(t_2) \rangle_s}{\partial \epsilon_s} \right|^2} \left| \left(1 + \frac{1}{1 + 4\eta/\gamma} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)-2\eta(t_1-t_0)} \right) \left[\left(1 - e^{-\gamma(t_2-t_1)} \right) \right]^{-1} \right) \right|^2 \\
& = \delta^2 \epsilon_s^s \left| \left(1 + \frac{1}{1 + 4\eta/\gamma} \left(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)-2\eta(t_1-t_0)} \right) \left[\left(1 - e^{-\gamma(t_2-t_1)} \right) \right]^{-1} \right) \right|^2
\end{aligned} \tag{B.81}$$

hence,

$$\begin{aligned}
\frac{\delta \epsilon_s^s}{\delta \epsilon_s^c} &= 1 + \frac{(e^{-\gamma(t_2-t_1)/2} - e^{-\gamma(t_2-t_0)-2\eta(t_1-t_0)})}{(1 + 4\eta/\gamma) (1 - e^{-\gamma(t_2-t_1)/2})} \\
&= 1 + \frac{(e^{-\gamma(t_2-t_1)/2} - e^{-(2\eta+\gamma/2)(t_1-t_0)-\gamma(t_2-t_1)/2})}{(1 + 4\eta/\gamma) (1 - e^{-\gamma(t_2-t_1)/2})} \\
&= 1 + \frac{(1 - e^{-(2\eta+\gamma/2)(t_1-t_0)}) e^{-\gamma(t_2-t_1)/2}}{(1 + 4\eta/\gamma) (1 - e^{-\gamma(t_2-t_1)/2})} \\
&\geq 1 \quad \forall t_2, t_1, t_0, \eta \text{ and } \gamma
\end{aligned} \tag{B.82}$$

and we conclude $\delta \epsilon_s^c \leq \delta \epsilon_s^s$ for all parameters i.e the precision of the measurement of parameter ϵ_s provided by the concurrent preparation and sensing scheme is superior to that provided by the analagous sequential scheme even under the effects of optical loss.

CHECKS: In the limit of no squeezing, i.e when $\eta \rightarrow 0$ we find

$$\frac{\delta \epsilon_s^s}{\delta \epsilon_s^c} = 1 + \frac{(1 - e^{-\gamma(t_1-t_0)/2}) e^{-\gamma(t_2-t_1)/2}}{(1 - e^{-\gamma(t_2-t_1)/2})} \tag{B.83}$$

noting that zero squeezing implies $t_1 = t_0$ it is the clear that $\frac{\delta \epsilon_s^s}{\delta \epsilon_s^c} = 1$, as expected.

Furthermore, in the limit of vanishing loss rate we find

$$\lim_{\gamma \rightarrow 0} \frac{\delta \epsilon_s^s}{\delta \epsilon_s^c} = 1 + \frac{1 - e^{-2\eta(t_1-t_0)}}{2\eta(t_2 - t_1)} \tag{B.84}$$

where we have used the fact that

$$\lim_{\gamma \rightarrow 0} (1 + 4\eta/\gamma)(1 - e^{-\gamma(t_2 - t_1)/2}) = 2\eta(t_2 - t_1). \quad (\text{B.85})$$

Comparing equations (B.84) and (B.17) it is clear the expressions are equivalent (up to arbitrary choices of parameters e.g taking $t_2 \equiv T$ and $t_0 = 0$ as is the case in (B.17)).