Quantum phase estimation in noisy systems

By

Hossein Tavakoli Dinani

A thesis submitted to Macquarie University
for the degree of Doctor of Philosophy
Department of Physics and Astronomy
August 2016

MACQUARIE
University
SYDNEY · AUSTRALIA

© Hossein Tavakoli Dinani, 2016.

Typeset in $\LaTeX\,2_\varepsilon$.

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

Hossein Tavakoli Dinani

To Dad and Mom

Acknowledgements

I would like to thank my supervisor, Dr. Dominic Berry, for giving me the opportunity to be here, and being always available to answer my questions. I'm also grateful to him for being persistent and patient with my writings and his great inputs. I would like to thank Prof. Jason Twamley for his support for my research travels. Thanks to Prof. Jonathan Dowling for the idea of the quantum spectroscopy work, and his hospitality during my visit at Louisiana State University. Also thanks to Hubert for keeping in touch, and being encouraging.

Thanks to all the admin people at Macqurie, especially Lisa and Carol, for being helpful all the time.

I've been fortunate to meet some great people in Sydney. Thanks to all my friends in Sydney who made my stay here such an amazing experience. I would especially like to mention my dear friends, Tommaso and Mauro, who ignited a great journey for me, Sukhi who's been a great friend and also let me use his place for some parts of writing this thesis, my dear flatmates and friends, Andrea and Veronica, and all my friends in "The comrades, le commarie" group. I would also like to thank Alexander, Nora, and Babatunde.

Finally, I would like to thank my parents Ali and Fatemeh, my brothers Saeid and Masoud and my sister Maryam for their support and love all the time.

List of Publications

- H. T. Dinani, M. K. Gupta, J. P. Dowling, and D. W. Berry, *Quantum-enhanced spectroscopy with entangled multiphoton states*. Phys. Rev. A **93**, 063804 (2016).
- C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, L. Markham, D. J. Twichen, and R. Hanson, *Optimized quantum sensing with a single electron spin using real-time adaptive measurements*. Nature Nanotechnology **11**, 247 (2016).
- H. T. Dinani, D. W. Berry, *Loss-resistant unambiguous phase measurement*. Phys. Rev. A **90**, 023856 (2014).

Abstract

The precision in estimation of unknown parameters can go beyond the classical limits by using nonclassical properties of quantum mechanics. However, nonclassical properties are very sensitive to the interaction with environment. This interaction could result in the loss of photons and low visibility. The purpose of this thesis is to find states and measurement schemes which give enhancement in estimation accuracy over the classical methods in the presence of such effects.

We begin with estimation of an unknown phase in a Mach-Zehnder interferometer in the presence of photon loss. We propose a scheme to produce states which are loss-resistant and perform very close to optimal states. We propose sequences of such states combined with single photon states to obtain an unambiguous phase estimate with better accuracy than the standard quantum limit.

We then consider the case that the loss is due to the interaction of the beam with an ensemble of atoms. Traditionally, the transition frequencies of atoms are measured via absorption, but there is also information available in the phase shift. We numerically find states that give maximum information about the transition frequency, obtained from both the absorption and the phase and quantified by the Fisher information.

We also consider phase estimation in a Ramsey interferometer using an NV centre to measure an unknown time-independent magnetic field. The low visibility in Ramsey measurements requires different measurement schemes than optical interferometers. We find an optimised adaptive scheme which reaches the limit analogous to the Heisenberg-like scaling in the estimation of the magnetic field, and outperforms the optimal non-adaptive scheme.

Finally, we consider continuous measurement on a single spatial mode field with a time-varying phase. We consider a phase which varies in time with power law spectral density. We show that by using squeezed coherent states in an adaptive homodyne measurement scheme we can estimate the phase with accuracy scaling at the Heisenberg limit.

Contributions

The details of the contributions of the other authors of the papers upon which this thesis is based are given below. This is listed according to the structure of the thesis as presented in the layout.

Loss resistant unambiguous phase measurement
The scheme to prepare loss-resistant states is the work of Dr. Dominic Berry. I verified Dr. Berry's calculations and via numerical simulations found the optimal sequence of loss-resistant states that gives variances smaller than the standard quantum limit for the estimation of an unknown phase. These results have been published in Phys. Rev. A **90**, 023856 (2014).

Quantum enhanced spectroscopy
The idea of this work is from Prof. Jonathan Dowling. I did the calculations and numerical simulations to find the optimal states. M. K. Gupta verified my calculations. To analyse the data I consulted with Dr. Berry and Prof. Dowling. This work is published in Phys. Rev. A **93**, 063804 (2016).

Magnetometry with an NV centre
This is a collaborative work with the group of Prof. Hanson. In this work, Dr. Cristian Bonato, Dr. Machiel Blok and Prof. Ronald Hanson conceived the experiment. Dr. Bonato and Dr. Blok performed the measurements and processed the data. They also performed numerical simulations for the non-adaptive, limited-adaptive and full-adaptive protocols. Dr. Berry and I advised on the theoretical part of the experiment. I independently performed the numerical simulations for the non-adaptive, limited-adaptive and full-adaptive protocols. I also performed numerical simulations for the optimised adaptive protocol and provided the experimental group with the feedback phases to use for the optimised adaptive protocol. This work is published in Nature Nanotechnology **11**, 247 (2016).

Contents

Acknowledgements **v**

List of Publications **vii**

Abstract **ix**

Contributions **xi**

Contents **xiii**

List of Figures **xv**

1 Introduction **1**
- 1.1 Estimation theory 1
  - 1.1.1 Classical parameter estimation 2
  - 1.1.2 Quantum parameter estimation 5
- 1.2 Mach-Zehnder interferometer 5
  - 1.2.1 Single photons 5
  - 1.2.2 NOON states 9
- 1.3 Quantum phase estimation algorithm 12
- 1.4 Ramsey interferometry 13
- 1.5 Dyne measurements 14
- 1.6 Thesis layout 16

2 Loss resistant unambiguous phase measurement **19**
- 2.1 Introduction 19
- 2.2 The state preparation scheme 20
- 2.3 Photon loss 24
- 2.4 The measurement scheme 26
- 2.5 Conclusion 32

3 Quantum enhanced spectroscopy **33**
- 3.1 Introduction 33
- 3.2 Interferometric scheme 34
- 3.3 Optimised states 37
  - 3.3.1 Large $\mathcal{N}_a$ 38
  - 3.3.2 Small $\mathcal{N}_a$ 40
- 3.4 Conclusion 41

4 Magnetometry with an NV centre **43**
4.1 Introduction 43
4.2 Ramsey interferometry 43
4.3 Phase estimation 45
4.3.1 Non-adaptive protocol 47
4.3.2 Adaptive protocols 47
4.4 Conclusion 54

5 Estimation of a time varying phase **57**
5.1 Introduction 57
5.2 Adaptive homodyne measurement with squeezed states 58
5.3 System phase time variation 61
5.4 Feedback phase 63
5.5 Numerical results 64
5.6 Conclusion 69

6 Conclusions **71**

A Appendices **75**
A.1 Optimal sequence for loss resistant phase estimation 75
A.2 Particle swarm optimisation 76

References **79**

List of Figures

1.1 Mach-Zehnder interferometer. It consists of two beam splitters (BS), two mirrors (M) and two phase shifters, φ and θ. We have labeled the modes after each BS and phase shifter. 6

1.2 The probability distribution for the system phase after 100 measurements using single-photon states. This is an example for randomly generated measurement results, and the actual system phase is $\pi/4 \approx 0.8$. 8

1.3 The output conditional probability. Green, solid line: three-photon NOON state. Black, dashed line: single photon state. 11

1.4 Quantum phase estimation algorithm diagram to estimate a phase with four binary digits. Double lines represent classical rails. 13

1.5 Ramsey interferometry. A $\pi/2$ pulse prepares an equal superposition of the ground and excited state of a two level atom, initially in the ground state. The atom then goes through a free evolution. This is followed by another $\pi/2$ pulse which projects the atom into one of the states depending on the phase φ imprinted on the superposition state. 14

1.6 Adaptive measurement of the phase φ imposed on a single-mode beam. This scheme approximates the homodyne measurement. The single mode is combined with a strong local oscillator (LO) on a 50/50 beam splitter (BS). Based on the result of detectors D_1 and D_2 the processor adjusts the phase of the local oscillator θ. 15

2.1 A linear optical device (LOD) with two SPDC states and vacuum modes as input. $|\psi\rangle_{\text{out}}$ is the desired output state, and there is postselection on vacuum in the other modes. The two SPDC input states can be simplified to one. . . 20

2.2 (a) A triangular array of beam splitters and phase shifts fed with a two-mode input state $|\psi\rangle_{\text{in}}$ and three vacuum modes, $|0\rangle$. The phase shifts are included with the beam splitters (shown as thick black lines) for simplicity. Output photons are detected in two of the modes, $|\psi\rangle_{\text{out}}$, and vacuum is postselected in three of the output modes. All the beams below the dashed line contain vacuum. (b) The five-port LOD can be simplified to a four-port one by keeping only the beam splitters above the dotted line. 22

2.3 The state preparation scheme for loss-resistant states. It is a three-port interferometer consisting of three beam splitters with reflectivities R_1, R_2, R_3 and two phase shifters φ_1, φ_2. The initial dual Fock state $|n\rangle\,|n\rangle$ gives the $2n$-photon loss-resistant state in the output, as indicated in Eq. (2.16). . . . 22

2.4 Adaptive phase measurement scheme. φ is the unknown phase, θ is the controlled phase and D_1 and D_2 are the photon detectors in the two outputs. The grey diagonal lines in the two arms are the fictitious beam splitters with transmissivity η modeling photon loss. 25

2.5 Maximum of classical Fisher information $F(\varphi,\theta)$ versus efficiency η. ○ : four-photon exact optimal state, Eq. (2.20), maximized over χ_1', χ_2', φ and θ. ▲: four-photon loss-resistant states, Eq. (2.19), maximized over χ, φ and θ. +: four single-photon states. 28

2.6 Fisher information, $F(\varphi,\theta)$ and Holevo phase variance, V_{H}, versus χ for $\eta = 0.6$. Solid green line: Fisher information for $\varphi = \pi/4$, $\theta = 0$ versus χ calculated using Eq. (2.30) for the two-photon state given in Eq. (2.18). Dashed black line: phase variance versus χ for the sequence of seven single photons followed by one two-photon loss-resistant state in the adaptive measurement protocol. The Fisher information gives a lower bound to the phase variance, rather than an exact phase variance, so the value of χ that maximises the Fisher information need not be the value that minimises the phase variance. 29

2.7 The scaled phase variance $\mathcal{N} V_{\mathrm{H}}$ versus total number of photons $\mathcal{N}$ for $\eta = 0.6$ for: only single photon states (SQL) (solid line) and optimal sequence of single photon states combined with two- and four-photon loss-resistant states (shown by ×). 31

2.8 The phase variance V_{H} versus total number of photons $\mathcal{N}$ for $\eta = 0.6$ for three different input states. ×: optimal sequence of single photon states combined with two- and four-photon loss-resistant states. Solid line: only single photon states (SQL). •: the scheme proposed in [1]. 31

3.1 A Mach-Zehnder interferometer with an ensemble of atoms placed in the upper arm. D_1 and D_2 are photon number detectors in the output modes. . . 34

3.2 The real (solid line) and the imaginary (dashed line) parts of susceptibility, χ' and χ'' respectively, calculated using Eq. (3.1) for the D1 transition line of sodium. 35

3.3 Beam splitter model to model the interaction of photons with the ensemble of atoms. 35

3.4 The transmissivity T (dashed line) and the phase shift φ (solid line) versus detuning Δ for the D1 transition line of sodium, $L = 1$ cm and $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3}. 36

3.5 Fisher information $F(\Delta)$ versus detuning Δ for $N = 2$ photons [(a) and (b)] and $N = 10$ [(c) and (d)]. In (a) and (c), the solid black lines are for N independent single photons $|1,0\rangle^{\otimes N}$, and the dashed red lines are for N-photon NOON states $(|N,0\rangle + |0,N\rangle)/\sqrt{2}$. In (b) and (d) the dashed green lines are for N-photon optimal states, and the solid orange lines are for N copies of single-photon NOON states $(|1,0\rangle + |0,1\rangle)^{\otimes N}/\sqrt{2^N}$. 39

3.6 Coefficients ψ_k of the optimal states for number density of atoms $\mathcal{N} = 2.5 \times 10^{17}$ m^{-3} for four values of the total photon number N. 39

3.7 Transmissivity T and phase shift φ versus Δ, for a range of values of number density of atoms $\mathcal{N}$. Solid-red line: $\mathcal{N} = 2.5 \times 10^{15}$ m^{-3}. Dashed-green line: $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3}. Dotted-black line: $\mathcal{N} = 2.5 \times 10^{17}$ m^{-3}. 40

3.8 Fisher information versus Δ for total number of photons $N = 2$, for the optimal states of a range of number densities of atoms $\mathcal{N}$. Solid-red line: N independent single photons with $\mathcal{N} = 2.5 \times 10^{15}$ m^{-3}. Dashed-green line: N independent single photons with $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3}. Dotted-black line: numerically optimized state with $\mathcal{N} = 2.5 \times 10^{17}$ m^{-3}. 41

4.1 (a) An NV centre is formed by a substitutional nitrogen atom with a vacancy in its adjacent lattice position in a diamond lattice. This figure is adapted from Ref. [2] (b) The energy level structure of a negatively charged NV centre. (c) The pulse sequence used to estimate the magnetic field. The black line represents the microwave (MW) field used for coherent control of the spin. The phase θ of the second $\pi/2$ pulse is the controlled phase which determines the measurement basis. The orange line represents the optical field used for initialisation and measurement. 44

4.2 The scaled variance for a non-adaptive protocol where the number of repetitions for each interaction time varies as in Eq. (4.4), with $F = 5$ for a range of values of G. 47

4.3 The scaled variance, $V_{\mathrm{H}}T/\tau$, versus T/τ for the non-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F. Here we have used the experimental parameters $f_0 = 0.88$, $f_1 = 0.993$ and $T_2^* = 96\mu$s. . . 48

4.4 The scaled variance, $V_{\mathrm{H}}T/\tau$, versus T/τ for the limited-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F. This figure is plotted for $f_0 = 0.88$, $f_1 = 0.993$ and $T_2^* = 96\mu$s. 49

4.5 Minimum of the scaled variance versus F for two protocols; limited-adaptive and non-adaptive. 50

4.6 The scaled variance, $V_{\mathrm{H}}T/\tau$, versus T/τ for the full-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F. 52

4.7 Scaled variance of the optimised non-adaptive protocol ($G = 1$, $F = 5$) compared to the best limited- and full-adaptive protocols. 53

4.8 Minimum of the scaled variance versus F for full-, limited- and non-adaptive protocols. 53

4.9 The scaled variance $V_{\mathrm{H}}T/\tau$ versus T/τ_0 comparing the optimised adaptive protocol with a range of values of G and F with the optimised non-adaptive protocol $G = 1$, $F = 5$. 53

4.10 The scaled variance including the overhead time; $V_{\mathrm{H}}T_{\mathrm{total}}/\tau$ versus T_{total}/τ where $T_{\mathrm{total}} = T + T_{\mathrm{overhead}}$ with T_{overhead} given in Eq. 4.14. 54

4.11 The minimum of the scaled variance versus F for optimised, full-, limited- and non-adaptive protocols. 54

5.1 The scheme for adaptive homodyne measurement of the phase φ imposed on a squeezed coherent state generated by a cavity with decay constant γ. D_1 and D_2 are the photodetectors. $I(t)$ is the difference photocurrent between the two outputs of the 50/50 beam splitter (BS). The processor adjusts the phase of the local oscillator (LO) labeled by θ based on $I(t)$. 58

5.2 A Gaussian random process with power law spectral density $1/(\omega^p + \Gamma^p)$, with $p = 2$, $\Gamma = 10^{-3}$, and $\delta t = 10^{-3}$. 63

5.3 The optimal values of the parameters χ and e^r versus $\mathcal{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$. 66

5.4 The optimal values of the parameters γ and δ versus $\bar{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$. 67

5.5 The scaled MSE versus $\bar{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$. 68
5.6 The scaling constant of the MSE for the Heisenberg limit (green line), c_z given in Eq. (5.49), the pulsed measurement (black line), c_A given in Eq. (5.50), and the homodyne scheme with a continuous squeezed state (red line). . . 68
5.7 System phase (black line) and the estimated phase (green line) for $\bar{N}/\kappa = 10^8$. Top: $p = 1.5$, bottom: $p = 3$. 69

1

Introduction

The science of high precision measurement using quantum mechanics to describe the physical system is known as quantum metrology. Although quantum mechanics imposes limits on the possible precision, far better precision can be achieved by taking advantage of nonclassical properties of quantum mechanical states such as entanglement and squeezing [3–6]. One notable example is the use of squeezed states of light in gravitational wave detectors. This was first proposed by Caves [6] and has recently been successfully demonstrated in different prototypes [7–10]. The direct detection of gravitational waves by the LIGO observatories, announced in February 2016, marks a significant advance in high precision measurements [11]. That was achieved without squeezing, but injecting the squeezed light in the next generation of gravitational wave detectors could result in more and frequent discoveries of gravitational wave sources [12].

One serious impediment to obtaining improved precision with quantum states is that they are extremely fragile. As a result of the interaction of the system with the surrounding environment the system becomes entangled with the environment. To provide a description of the system on its own, the environment needs to be traced out. As a result the system state becomes a mixed state. This effect of the environment on the system is equivalent to noise acting on the system. If the state is an entangled state, the noise tends to destroy the entanglement, and if the state is a squeezed state, the squeezing will be reduced.

In this thesis, we address the effects of interaction with the environment in metrological schemes. We propose quantum states and measurement schemes that are less affected by this interaction. In this chapter, we set the foundations by reviewing the essential background. We start with a quick review of both classical and quantum estimation theory. Then, we review phase estimation in Mach-Zehnder and Ramsey interferometers. This is followed by a review of dyne measurements and a summary of the thesis layout.

1.1 Estimation theory

For many quantities in quantum mechanics there is no suitable Hermitian operator to form an observable [13]. These quantities need to be estimated based on the data obtained from measurement of observables [14–16]. As a result of the statistical nature of quantum mechanics the precision in parameter estimation is limited by statistical bounds.

1.1.1 Classical parameter estimation

In random processes we need to estimate what we wish to know from a finite set of measurements. Suppose we have measurement results $\{x_1, x_2, ..., x_n\}$ which depend on an unknown parameter φ. Each measurement result is obtained with the conditional probability $P(x|\varphi)$, i.e. the probability of measuring x given the parameter φ. For now, take φ to be an arbitrary scalar real-valued parameter with unbounded range (in contrast to phase, where the range is bounded). We label our estimate of φ based on the measurement results by $\breve{\varphi}(x_1, x_2, ..., x_n)$. The estimate is a function of the random measurement results. Therefore, it is a random variable itself and is different for each set of the measurement results. For brevity we do not write the dependence of the estimate on the measurement results explicitly.

An estimator is called unbiased if

$$\langle\breve{\varphi}\rangle = \sum_x \breve{\varphi}\, P(x|\varphi) = \varphi, \tag{1.1}$$

otherwise it is called biased. The summation is over all the possible measurement results. The bias is defined as

$$B(\varphi) \equiv \langle\breve{\varphi}\rangle - \varphi. \tag{1.2}$$

Suppose we have obtained N samples $\{\breve{\varphi}_1, \breve{\varphi}_2, ..., \breve{\varphi}_N\}$ of the unknown parameter φ. Normally, we don't know the exact probability distribution of the samples. Therefore, the distribution mean is estimated by the mean of the samples $\bar{\breve{\varphi}}$ as

$$\bar{\breve{\varphi}} = \frac{1}{N}\sum_{i=1}^{N} \breve{\varphi}_i. \tag{1.3}$$

The variance of the distribution, $V(\varphi) = \left\langle(\breve{\varphi} - \langle\breve{\varphi}\rangle)^2\right\rangle = \sum_x (\breve{\varphi} - \langle\breve{\varphi}\rangle)^2 P(x|\varphi)$, may be simply estimated from the samples using

$$\frac{1}{N}\sum_{i=1}^{N}\left(\breve{\varphi}_i - \bar{\breve{\varphi}}\right)^2. \tag{1.4}$$

This is a slightly biased estimate of the variance, because the mean is being estimated from the samples. (This is a different issue from bias in the estimate of φ.) An unbiased estimate may be obtained simply by replacing N with $N-1$, so the estimate is

$$\frac{1}{N-1}\sum_{i=1}^{N}\left(\breve{\varphi}_i - \bar{\breve{\varphi}}\right)^2. \tag{1.5}$$

The goal is to provide an estimate that is best, in the sense of being closest to the actual value of φ. For such an estimate the distribution is centred at the actual value of the parameter, and its peak is as narrow as possible. The quality of an estimator may be quantified by the mean square error (MSE) defined as

$$\mathrm{MSE}(\varphi) = \left\langle(\breve{\varphi} - \varphi)^2\right\rangle = V(\varphi) + (B(\varphi))^2. \tag{1.6}$$

As is clear, for unbiased estimates, i.e. for estimates where $B(\varphi) = 0$, the MSE and the variance are the same. If one were to just consider the variance of the estimators, biased estimators would appear to give results that are as good as unbiased estimators. One could place an additional restriction that the estimator is unbiased. Instead the MSE penalises biased estimators in a natural way.

Note that, for pathological estimators which always give the same value, the MSE is zero for a particular value of the parameter φ. However, if we consider the performance of this estimator for other values of φ the MSE won't be zero. Therefore, to realistically evaluate the performance of the estimate we should take the average of the MSE over the full range of the parameter φ. On the other hand, for covariant measurements [17] the probability distribution of the MSE in the estimate is independent of the parameter. Therefore, we do not need to take the average for covariant measurements.

Holevo variance

We will consider the specific case of phase, which is bounded to the range $[0, 2\pi)$ or more generally $[\theta_0, \theta_0 + 2\pi)$, in which case there are problems with the usual definition of the variance. One of the problems is that in the limit of a flat distribution we expect the variance to go to infinity. However, because of the limited range for the phase distribution, we obtain a finite variance with the usual definition of the variance [18, 19]. Furthermore, if the mean of the distribution is at one bound of the phase range, the usual definition of the variance results in very large values.

To avoid these problems, alternative figures of merit are considered which are naturally modulo 2π. Many of these measures are based on the average of the exponential of the phase, i.e. $\langle \exp(i\varphi) \rangle$ [17, 20–22]. One such measure is the Holevo phase variance defined as [17]

$$V_{\mathrm{H}} = \left| \left\langle e^{i(\check{\varphi}-\varphi)} \right\rangle \right|^{-2} - 1. \tag{1.7}$$

To penalise biased estimates the Holevo variance can be modified to

$$V_{\mathrm{H}} = \left[\mathrm{Re} \left\langle e^{i(\check{\varphi}-\varphi)} \right\rangle \right]^{-2} - 1, \tag{1.8}$$

which is the equivalent of the MSE for the Holevo variance. These two definitions of the Holevo variance are equivalent if the measurement is unbiased in the sense that [23]

$$\varphi = \arg \left\langle e^{i\check{\varphi}} \right\rangle. \tag{1.9}$$

The Holevo phase variance is approximately the same as the usual variance for distributions sharply peaked well away from the bounds of the phase range. For unbiased estimates, Holevo variance can be estimated from the samples as

$$\left| \frac{1}{N} \sum_{j=1}^{N} e^{i(\check{\varphi}_j - \varphi)} \right|^{-2} - 1. \tag{1.10}$$

Cramér-Rao bound

Our desired estimate is the one that minimises the MSE. The Cramér-Rao bound gives a lower bound on the variance of an unbiased estimate (so the variance is equal to the MSE).

To derive the lower bound, consider the measurement results $\{x_1, x_2, ..., x_n\}$ obtained with the conditional probabilities $P(x|\varphi)$. We have

$$\sum_x P(x|\varphi) = 1. \tag{1.11}$$

For an unbiased estimator we can write

$$\langle\check{\varphi}\rangle = \varphi = \sum_x \check{\varphi} P(x|\varphi). \tag{1.12}$$

Taking the derivative of Eqs. (1.11) and (1.12) with respect to the parameter φ we obtain

$$\sum_x \frac{\partial P(x|\varphi)}{\partial\varphi} = \sum_x P(x|\varphi)\frac{\partial \ln P(x|\varphi)}{\partial\varphi} = \left\langle \frac{\partial \ln P(x|\varphi)}{\partial\varphi} \right\rangle = 0 \tag{1.13}$$

$$\frac{\partial\langle\check{\varphi}\rangle}{\partial\varphi} = \left\langle \check{\varphi}\frac{\partial \ln P(x|\varphi)}{\partial\varphi} \right\rangle = 1. \tag{1.14}$$

Using the Cauchy-Schwarz inequality for $\partial \ln P(x|\varphi)/\partial\varphi$ and $\check{\varphi}$ we obtain

$$V(\varphi) \geq \frac{1}{F(\varphi)}, \qquad F(\varphi) = \left\langle \left(\frac{\partial \ln p(x|\varphi)}{\partial\varphi} \right)^2 \right\rangle \tag{1.15}$$

where $F(\varphi)$ is called Fisher information. It is a measure of the information that the experimental setup reveals about φ. Equation (1.15) is called the Cramér-Rao bound. According to this equation, the lower bound on the variance of an unbiased estimator is set by the inverse of the Fisher information. Repeating the experiment M times increases the Fisher information by a factor of M and therefore

$$V(\varphi) \geq \frac{1}{MF(\varphi)}. \tag{1.16}$$

Note that, the Cramér-Rao bound only gives a lower bound on the variance of an estimate. This means the variance obtained from the experiment does not necessarily reach this lower bound. It is possible to saturate the Cramér-Rao bound in the limit of many measurements, but only for probability distributions with a unique global maximum and finite Fisher information [24]. When the Fisher information is infinite the Cramér-Rao bound is zero, which is not achievable. The Fisher information can be infinite for a finite mean photon number [25]. This happens when the variance of the photon number is infinite (despite the mean photon number being bounded). The reason why the mean photon number should be bounded is that, if it is not bounded, then the average energy is infinite, which is unphysical. It is trivial to see that the Fisher information can be infinite in the limit of infinite mean photon number.

Caution is needed in using the Cramér-Rao bound for phase measurements, because it is not possible for phase estimates to be completely unbiased in the sense that both $B(\varphi)$ and $\partial B(\varphi)/\partial\varphi$ are zero for all φ [26]. For biased measurements there is an alternative form for the Cramér-Rao bound [23, 27].

When using the Cramer-Rao bound, the MSE is calculated for a single value of the parameter φ, which is distinct from the approach mentioned above where the MSE is averaged over the parameter φ. When the MSE is averaged over φ, the Cramer-Rao bound does not hold. In the average MSE approach, one allows any estimator, but estimators that have unreasonable behaviour, like the pathological example above, will be penalised because they give large MSE for some values of the parameter. In contrast, when using the Cramer-Rao bound the requirement that the estimator is unbiased ensures that the estimator is reasonable and excludes pathological cases. This means that the MSE accurately quantifies the performance of the estimator, even when it is only evaluated at a single value of φ.

1.1.2 Quantum parameter estimation

The Fisher information is a function of the conditional probability $P(x|\varphi)$. In quantum mechanics, the probability of measurement result x is given via a positive-operator valued measure (POVM), which consists of a set of positive semidefinite Hermitian operators $\{E_x\}$ which sum to the identity. The probability for measurement result x is then given by $\mathrm{Tr}(E_x\rho)$. In our case, the state is a function of the parameter φ, so the probability for x as a function of φ is given by

$$P(x|\varphi) = \mathrm{Tr}(E_x\rho(\varphi)). \tag{1.17}$$

The Fisher information can be maximised by optimising the state of the system and the POVM. The Fisher information optimised over all the possible quantum measurements is the quantum Fisher information F_Q. The inverse of the quantum Fisher information gives a further lower bound for the variance, known as the quantum Cramér-Rao bound [28]. For M repetitions of the experiment we have

$$V(\varphi) \geq \frac{1}{MF(\varphi)} \geq \frac{1}{MF_Q(\varphi)}. \tag{1.18}$$

In many cases, including phase measurement, the unknown parameter φ is imprinted on the probe state via a unitary operation $U_\varphi = \exp(-i\varphi H)$, where H is a known Hermitian operator. For phase, H would be the photon number operator. For a pure probe state the quantum Fisher information F_Q is equal to $4V(H)$, where $V(H)$ is the variance of H for the probe state [28, 29]. In this case, the quantum Cramér-Rao bound for M repetitions of the process is of the form [28, 29]

$$V(\varphi) \geq \frac{1}{4MV(H)}. \tag{1.19}$$

In the estimation of a general parameter X, if X can be represented by an observable, the above equation corresponds to the Heisenberg uncertainty relation for the conjugated variables [30–32]. There is no phase operator, so the above form of the Heisenberg uncertainty relation does not hold for phase. In this case the uncertainty relation can be identified with the Mandelstam-Tamm bound [33].

The unknown parameter, which could be length difference, magnetic field, temperature or other physical quantities, is imprinted as a phase shift on the probe state. Therefore, the unknown parameter estimation is a phase estimation problem. To be able to obtain a quantum enhancement in parameter estimation we need to use nonclassical features such as squeezing and entanglement. This can be implemented with interferometric or dyne measurements which we review in the following sections.

1.2 Mach-Zehnder interferometer

A Mach-Zehnder interferometer is shown in Fig. 1.1. It consists of two 50/50 beam splitters, an unknown phase shift φ which we would like to measure, another phase shift θ which is known and can be controlled, and two mirrors. In this section we review the Bayesian estimation techniques introduced in Refs. [34, 35].

1.2.1 Single photons

First consider having a single photon in one of the input modes and vacuum in the other mode, represented by the state $|1,0\rangle$. The 50/50 beam splitter (BS) transforms the input creation operators to

$$\hat{a}_0^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(\hat{b}_1^\dagger + i\hat{b}_0^\dagger\right), \qquad \hat{a}_1^\dagger \rightarrow \frac{1}{\sqrt{2}}\left(\hat{b}_0^\dagger + i\hat{b}_1^\dagger\right), \tag{1.20}$$

and the phase shifter φ transforms the creation operator of the upper mode to

$$\hat{b}_0^\dagger \rightarrow e^{i\varphi}\hat{c}_0^\dagger. \tag{1.21}$$

Therefore

$$\begin{aligned} |1,0\rangle &= \hat{a}_0^\dagger\,|0,0\rangle \\ &\mapsto \frac{1}{\sqrt{2}}\left(\hat{b}_1^\dagger + i\hat{b}_0^\dagger\right)|0,0\rangle \\ &\mapsto \frac{1}{\sqrt{2}}\left(e^{i\theta}\hat{c}_1^\dagger + ie^{i\varphi}\hat{c}_0^\dagger\right)|0,0\rangle \\ &\mapsto \frac{1}{2}\left[e^{i\theta}\left(\hat{d}_0^\dagger + i\hat{d}_1^\dagger\right) + ie^{i\varphi}\left(\hat{d}_1^\dagger + i\hat{d}_0^\dagger\right)\right]|0,0\rangle \\ &= \frac{1}{2}e^{i\theta}\left[\left(1 - e^{i(\varphi-\theta)}\right)|1,0\rangle + i\left(1 + e^{i(\varphi-\theta)}\right)|0,1\rangle\right]. \end{aligned} \tag{1.22}$$

The conditional probability of detecting a photon in the output modes d_0 and d_1, corresponding to $u = 0$ and $u = 1$, respectively, given the phases φ and θ is

$$P(u|\varphi,\theta) = \frac{1 + (-1)^{u+1}\cos(\varphi - \theta)}{2}. \tag{1.23}$$

In order to estimate the system phase we calculate the probability distribution of the system phase φ. We repeat the experiment N times and use Bayes' theorem to calculate this probability distribution. After each run we update our knowledge of the phase by using Bayes' theorem which can be stated as [36]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{1.24}$$

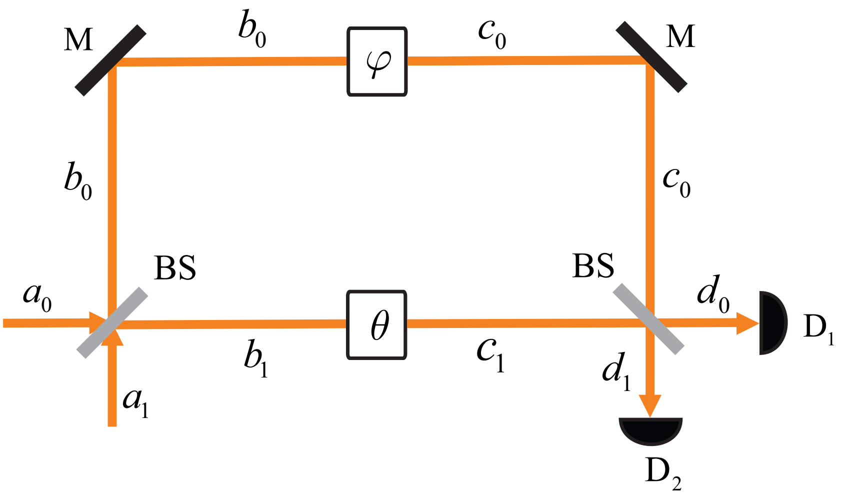


Figure 1.1: Mach-Zehnder interferometer. It consists of two beam splitters (BS), two mirrors (M) and two phase shifters, φ and θ. We have labeled the modes after each BS and phase shifter.

That is, Bayes' theorem gives the conditional probability of A given B, $P(A|B)$, from the conditional probability of B given A, $P(B|A)$, and the probabilities of A and B, $P(A)$ and $P(B)$. For phase we need a probability density (the integral over a finite range of phase gives a probability). The form of Bayes' theorem is unchanged for probability density. Assuming we have no initial knowledge of the phase, the initial probability density for the phase is flat, i.e. $P(\varphi) = 1/2\pi$. If we label the measurement result for the first run by u_1, according to Bayes' theorem we can write

$$P(\varphi|u_1, \theta_1) \propto P(u_1|\varphi, \theta_1) \tag{1.25}$$

where the proportional to symbol is due to omitting a normalization factor. After the second run we update our knowledge of the phase based on the measurement result u_2 and the already obtained knowledge $P(\varphi|u_1, \theta_1)$, using Bayes' rule

$$P(\varphi|u_2, \theta_2) \propto P(u_2|\varphi, \theta_2)P(\varphi|u_1, \theta_1). \tag{1.26}$$

After repeating this process N times we obtain the probability of φ given all the measurement results. In the limit of many measurements this probability distribution is a Gaussian centred at the actual system phase with the variance $1/(NF(\varphi))$ [37]. However, we stress that here we use Bayes' theorem in order to find the phase estimate. We do not quantify the performance of the estimate in terms of the spread of this probability distribution.

In general the controlled phase θ may be changed after each measurement, and we have denoted the initial controlled phase by θ_1, the controlled phase for the second measurement by θ_2, and so forth. The controlled phases can be independent of the measurement results (non-adaptive updating) or they can be changed based on the previous measurement results (adaptive updating). Figure 1.2 shows the probability of the system phase $\varphi = \pi/4$ given the measurement results after 100 measurements using single-photon states. In this figure, we have incremented the controlled phase by steps of $\pi/100$ after each run. If we were fixing the controlled phase through the whole process, we would obtain two peaks, one centred at $\varphi = \pi/4$ and the another peak, if $\theta = 0$, centred at $7\pi/4$. Incrementing the controlled phase by $\pi/100$ resolves this ambiguity in the phase estimate. The ambiguity in the phase estimate is because $\cos(\varphi - \theta) = \cos(2\pi - \varphi + 2\theta - \theta)$. For $\theta < \varphi$, we obtain two peaks centred at φ and $2\pi - \varphi + 2\theta$. To avoid this ambiguity the controlled phase should not be kept fixed after every measurement.

The calculations for Bayesian updating can be simplified considerably by writing the conditional probabilities of measurement results $P(u|\varphi, \theta)$ as a Fourier series [34], i.e.

$$P(u|\varphi, \theta) = \frac{1}{2\pi}\sum_j a_j e^{-ij\varphi}. \tag{1.27}$$

Since the probability distribution is real we have $a_j^* = a_{-j}$. The probability given in Eq. (1.23) can be written as a trivial three-term Fourier series with $a_0 = 1$, $a_1 = e^{-i\theta}/2$ and $a_{-1} = e^{i\theta}/2$, and $a_j = 0$ for all other j. The coefficients of the Fourier series after ℓ runs are updated according to

$$a_j^{(\ell)} = a_j^{(\ell-1)} + \frac{u_\ell}{2}\left(e^{i\theta}a_{j-1}^{(\ell-1)} + e^{-i\theta}a_{j+1}^{(\ell-1)}\right) \tag{1.28}$$

where u_ℓ is the measurement result of the ℓth run, and $\ell = 0$ corresponds to the initially flat probability distribution for which $a_0^{(0)} = 1$ and $a_j^{(0)} = 0$ for $j \neq 0$. In more general cases with different input states, the probability distribution for the detections can still be written as a

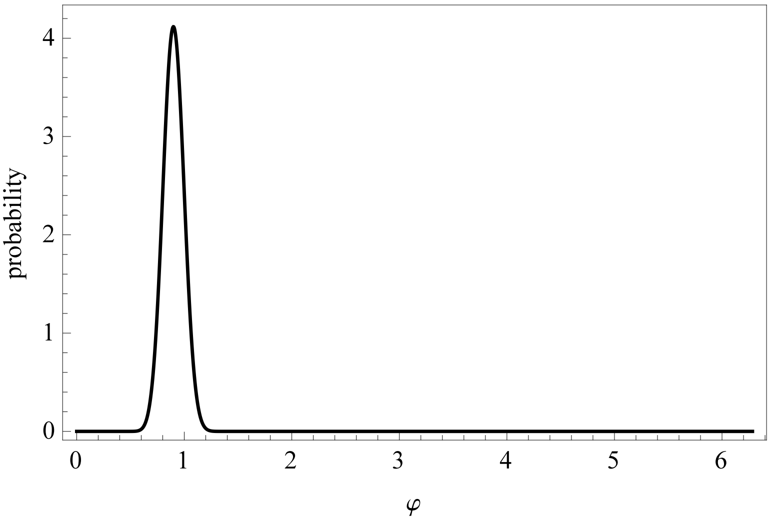


Figure 1.2: The probability distribution for the system phase after 100 measurements using single-photon states. This is an example for randomly generated measurement results, and the actual system phase is $\pi/4 \approx 0.8$.

Fourier series, and this approach still works. The unknown phase φ is estimated based on the updated probability for the phase $P(\varphi|u,\theta)$ according to

$$\check{\varphi} = \arg \int_0^{2\pi} e^{i\varphi} P(\varphi|u,\theta)\, d\varphi = \arg\left(a_{-1}\right) \tag{1.29}$$

where a_{-1} is the coefficient of the $e^{-i\varphi}$ term in the Fourier expansion. This estimate of the phase minimises the variance $V(\varphi)$ [35].

The variance of the estimate of φ can be estimated from the error propagation formula [38]

$$V(\varphi) = \frac{V\left(\hat{A}\right)}{\left|\partial\left\langle\hat{A}\right\rangle/\partial\varphi\right|^2}, \tag{1.30}$$

where $\hat{A}$ is the observable from which the estimate of the phase is obtained. According to this equation the variance of the phase can be decreased not only by decreasing the variance in $\hat{A}$, but also by increasing the gradient of $\left\langle\hat{A}\right\rangle$ with respect to phase. For the single photon case the observable $\hat{A}$ can be taken as $\hat{A} = |1,0\rangle\langle 0,1| + |0,1\rangle\langle 1,0|$ [39]. For this observable we obtain $\left\langle\hat{A}\right\rangle = P(1|\varphi,\theta) - P(0|\varphi,\theta) = \cos(\varphi-\theta)$, and therefore the gradient is obtained as $\partial\left\langle\hat{A}\right\rangle/\partial\varphi = \sin(\varphi-\theta)$.

The Fisher information, given in Eq. (1.15), can be written as

$$F = \sum_u \frac{1}{P(u|\varphi,\theta)} \left(\frac{\partial P(u|\varphi,\theta)}{\partial\varphi}\right)^2. \tag{1.31}$$

For the state $|1,0\rangle$, using the probability given in Eq. (1.23), we obtain $F = 1$. For N copies of the state $|1,0\rangle$, the additivity of the Fisher information gives $F = N$. We group together the N copies of the single photons and consider it as a single state with N photons. From the Cramér-Rao bound the variance of an unbiased estimate of the phase is lower bounded as

$$V(\varphi) \geq \frac{1}{N}. \tag{1.32}$$

Recall that in general it is not possible to have an unbiased estimate of the phase.

To find the quantum lower bound we calculate the quantum Fisher information. The state $|1, 0\rangle$ is a pure state and the phase shift action, $U = \exp\left(i\hat{b}_0^\dagger \hat{b}_0 \varphi\right)$, is a unitary operator where $\hat{b}_0^\dagger \hat{b}_0$ is the photon number operator in the upper arm of the interferometer. Therefore, as we discussed in Section 1.1.2, the quantum Fisher information can be obtained from the variance of the generator of the phase shift $\hat{b}_0^\dagger \hat{b}_0$ for the state just before the phase shift. The variance is easily calculated to be $V(\hat{b}_0^\dagger \hat{b}_0) = 1/4$. Therefore, we have $F_Q = 4V(\hat{b}_0^\dagger \hat{b}_0) = 1$, and for a state consisting of N copies of the single photon states, due to the additivity of the quantum Fisher information we obtain $F_Q = N$. In fact, this is just a special case of the general result that the quantum Fisher information is achieved by photon counting for pure states that are symmetric between the arms of the interferometer [40].

Since the quantum Fisher information is equal to the Fisher information, further optimisation over the measurement scheme cannot increase the Fisher information for this unentangled state. However, the variance will depend on the measurement scheme. For example adaptive schemes will improve over non-adaptive schemes. It is only the lower bound provided by the Cramér-Rao bound which is unchanged. This lower limit on the variance, is known as the standard quantum limit (SQL). It is the lowest variance possible with states that can be regarded as “quasi-classical”, in the sense that they have properties that can be described by classical physics. Single photons are often regarded as nonclassical, because they have a Wigner function that take negative values. However, for coherent states, which are usually regarded as quasi-classical, with mean photon number $\bar{n}$ in one of the input modes and vacuum in the other input mode we obtain the same scaling i.e. $1/\bar{n}$. In other words, the SQL scales with the inverse of the number of photons or mean photon number in the state.

It is clear that the lower limit on the variance can be decreased by increasing the resources, quantified by the total number of photons. Instead of increasing the resources, we aim to achieve more accurate measurements with the same resources by using more sophisticated techniques. This can be achieved by using input states with nonclassical properties such as entanglement.

1.2.2 NOON states

A well-known type of entangled states is NOON states. They were first considered in Ref. [41] in the study of decoherence on the Schrödinger cat states, and rediscovered in Refs. [39, 42] in the context of quantum imaging and quantum metrology. A NOON state is an equal superposition of N photons in one arm and 0 in the other arm and vice versa, i.e. $\left(|N, 0\rangle + |0, N\rangle\right)/\sqrt{2}$ [43]. Note that, the convention is that the NOON state is the state in the arms of the interferometer, rather than the state prior to the initial beam splitter. The phase shifts and the last beam splitter transform the NOON state according to

$$
\begin{aligned}
\frac{1}{\sqrt{2}}\left(|N,0\rangle + |0,N\rangle\right) &= \frac{1}{\sqrt{2N!}}\left[\left(\hat{b}_0^\dagger\right)^N |0,0\rangle + \left(\hat{b}_1^\dagger\right)^N |0,0\rangle\right] \\
&\mapsto \frac{1}{\sqrt{2N!}}\left[\left(\hat{c}_0^\dagger e^{i\theta}\right)^N |0,0\rangle + \left(\hat{c}_1^\dagger e^{i\varphi}\right)^N |0,0\rangle\right] \\
&\mapsto \frac{1}{\sqrt{2^{N+1}N!}}\left[e^{iN\theta}\left(\hat{d}_1^\dagger + i\hat{d}_0^\dagger\right)^N |0,0\rangle + e^{iN\varphi}\left(\hat{d}_0^\dagger + i\hat{d}_1^\dagger\right)^N |0,0\rangle\right] \\
&\mapsto \frac{1}{\sqrt{2^{N+1}N!}}\left[e^{iN\theta}\sum_{n_1=0}^{N}\begin{pmatrix} N \\ n_1 \end{pmatrix} i^{n_1}\left(\hat{d}_1^\dagger\right)^{N-n_1}\left(\hat{d}_0^\dagger\right)^{n_1}\right. \\
&\qquad\left. + e^{iN\varphi}\sum_{n_1=0}^{N}\begin{pmatrix} N \\ n_1 \end{pmatrix} i^{N-n_1}\left(\hat{d}_0^\dagger\right)^{n_1}\left(\hat{d}_1^\dagger\right)^{N-n_1}\right]|0,0\rangle \\
&\mapsto \frac{1}{\sqrt{2^{N+1}N!}}\sum_{n_1=0}^{N}\begin{pmatrix} N \\ n_1 \end{pmatrix}\left(\hat{d}_1^\dagger\right)^{n_2}\left(\hat{d}_0^\dagger\right)^{n_1} e^{iN\theta}\left[i^{n_1} + i^{n_2}e^{iN(\varphi-\theta)}\right]|0,0\rangle,
\end{aligned}
\tag{1.33}
$$

where n_1 and $n_2 = N - n_1$ correspond to the number of photons in the two output modes. The probability of detecting n_1 photons in one of the output modes and n_2 in the other is proportional to

$$
\begin{aligned}
P\left(n_1, n_2|\varphi\right) &\propto \left|i^{n_1} + i^{n_2}e^{iN(\varphi-\theta)}\right|^2 \\
&\propto 1 + \cos\left[N(\varphi-\theta) - \left(n_1 - n_2\right)\pi/2\right] \\
&\propto \begin{cases} 1 + (-1)^{(n_1-n_2)/2}\cos\left[N(\varphi-\theta)\right], & N \text{ even} \\ 1 - (-1)^{(n_1-n_2+1)/2}\sin\left[N(\varphi-\theta)\right], & N \text{ odd} \end{cases}
\end{aligned}
\tag{1.34}
$$

The expressions on the last line above are obtained by noting that for N even, $n_1 - n_2$ is also even and $(n_1 - n_2)\pi/2$ is an integer multiple of π. Similarly for N odd, $n_1 - n_2 = 2n_1 - N$ is also odd. For even N the power of (-1) is $(n_1 - n_2)/2$, and for odd N the power is $(n_1 - n_2 + 1)/2$. In each case this quantity is an integer, and the probability only depends on its parity. Note the N-fold phase enhancement in the output conditional probabilities. This makes the conditional probability vary N times faster, and therefore the gradient of the probability increases by a factor of N. This is shown in Fig. 1.3. This enhancement in the phase shift, known as super-resolution, results in $1/N$ enhancement in the precision. This can be seen from the error propagation formula given in Eq. (1.30). In this case, the observable can be taken as $\hat{A}_N = |N,0\rangle\langle 0,N| + |0,N\rangle\langle N,0|$ and the error propagation formula gives $V(\varphi) = 1/N^2$ [39].

Considering the proportionality factor in Eq. (1.34) and using Eq. (1.15), the Fisher information of a NOON state is obtained as $F = N^2$. A NOON state is a pure state, therefore its quantum Fisher information can be calculated easily by taking the variance of the photon number operator $\hat{b}_0^\dagger\hat{b}_0$. The quantum Fisher information of a NOON state is equal to the Fisher information, $F_Q = N^2$. Because the NOON state is path symmetric, the quantum Fisher information will be achieved with photon counting [40]. From the quantum Cramér-Rao bound, the variance of an unbiased phase estimate is obtained as $V(\varphi) \geq 1/N^2$. This lower bound is the ultimate bound allowed by quantum mechanics and is often called the Heisenberg limit [44]. The Heisenberg limit scales with the inverse of the square of the

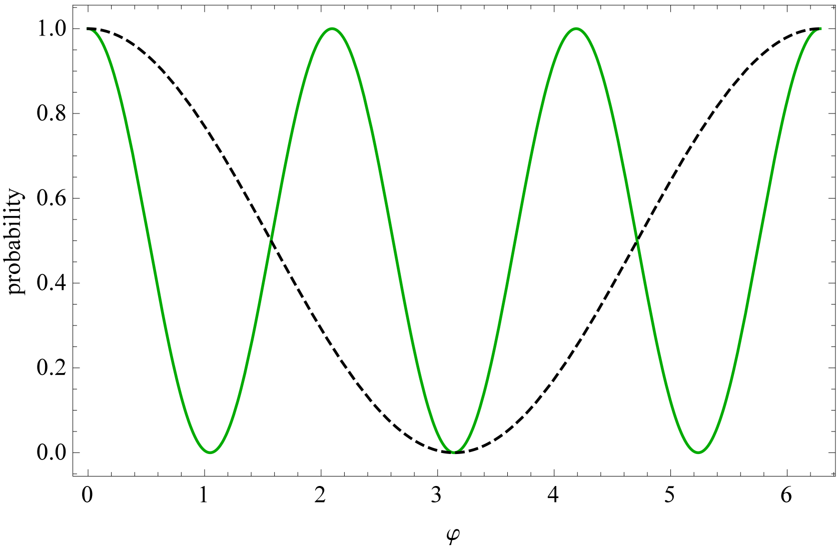


Figure 1.3: The output conditional probability. Green, solid line: three-photon NOON state. Black, dashed line: single photon state.

number of photons in the state. Note that, if we consider M repetitions of the state there will be an extra $1/M$ factor in the SQL and the Heisenberg limit.

Photon loss and NOON states

Although NOON states attain the Heisenberg scaling and have super-resolution, they are very sensitive to photon loss [45, 46]. Loss of even one photon from a NOON state destroys the entanglement and the phase information of the state; if a single photon is lost, detection of that photon would be sufficient to tell an observer which arm has all N photons. In other words, if the photon is lost from the upper arm, the NOON state collapses to $|N-1,0\rangle$ and if the photon is lost from the lower arm, the state collapses to $|0,N-1\rangle$. Neither of these states is sensitive to phase; for one state the phase shift only gives an undetectable global phase on that state and for the the other state there is no photon passing through the phase shift.

In reality, when a photon is lost there is no observer detecting the lost photon. However, the effect on the state is as if the photon were detected and the result, i.e. the knowledge of which arm the photon was in, were discarded. This means that a single photon lost gives an incoherent mixture of $|N-1,0\rangle$ and $|0,N-1\rangle$. In fact, it is not even known how many photons are lost. Therefore, the state will be a mixed state with different numbers of photons lost. The number of photons lost will only be known after detection, provided the initial number of photons is known.

Phase ambiguity

Another issue with NOON states is that the estimate of the phase obtained from NOON states is ambiguous. This is because the measurement results give $N\varphi \bmod 2\pi$ rather than $\varphi \bmod 2\pi$. In other words, if we obtain an estimate of the phase $\check{\varphi}$, the actual value of the phase could be $\varphi = \check{\varphi} + m\pi/(N-1)$ where m could be any integer between 0 and $N-1$. To remove this ambiguity we need to have additional knowledge of the phase. This information could be known prior to the measurement or obtained by extra measurement steps. A number of approaches have been proposed to resolve this ambiguity [47–50]. The quantum phase estimation algorithm discussed in the following section is an effective tool to remove this ambiguity.

1.3 Quantum phase estimation algorithm

The quantum phase estimation algorithm (QPEA) is an algorithm to measure the phase of an eigenvalue $e^{i\varphi}$ of a unitary operator U. Here, rather than using the inverse quantum Fourier transform, which requires entangling gates, we give the version that uses local measurements and feedback [51, 52]. The QPEA has widely been used in optical [26, 49, 50] and solid state systems [53–57] to measure a physical phase instead of an eigenvalue.

The circuit diagram for the QPEA is shown in Fig. 1.4. The algorithm uses an eigenstate of the unitary U, $|u\rangle$, and $K+1$ control qubits in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ which are measured in sequence. The algorithm estimates the phase φ, that can be expressed with $K+1$ binary digits, i.e. it can be expressed as the binary fraction

$$\varphi = 2\pi \sum_{k=0}^{K} \frac{u_k}{2^k} = 2\pi \times (0.u_0.u_1...u_K). \tag{1.35}$$

where u_k can take the values 0 or 1. Each qubit controls a power of U applied to the input state $|u\rangle$. In particular, the kth qubit, $|+\rangle_k$, controls the application of U^{2^k} on the input state. This introduces the phase shift $\exp(i2^k\varphi)$ on the $|1\rangle$ component of the kth qubit, i.e. evolves the qubit to

$$|+\rangle_k \to \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i2^k\varphi}|1\rangle\right). \tag{1.36}$$

The qubit is then measured in the X basis and based on the measurement result, the phase shift $R(\theta) = |0\rangle\langle 0| + e^{i\theta}|1\rangle\langle 1|$ is applied to the subsequent qubits. That is, there is no initial phase rotation, so none is indicated in the diagram in Fig. 1.4. The first measurement is performed on $|+\rangle_K$. If the phase φ can be exactly expressed as $2\pi \times (0.u_0...u_K)$, the measurement result gives u_K. Based on this measurement result, the feedback phase is set to $u_K\pi/2^K$ so that we have $\varphi - \theta = 2\pi \times (0.u_0...u_{K-1})$. The next measurement is on $|+\rangle_{K-1}$ and it gives u_{K-1}. This continues until all the digits of φ are determined.

The QPEA can be implemented in a Mach-Zehnder interferometer by using a sequence of NOON states starting from a 2^K-photon NOON state followed by a 2^{K-1}-photon NOON state down to a single photon state. The controlled application of the unitary U is simply the action of the unknown phase shift in one of the arms of the interferometer. The phase shift $R(\theta)$ corresponds to the controlled phase shift θ in the other arm of the interferometer. This phase is controlled for each NOON state based on the previous measurement results. In this interferometric version, the phase rotations on each qubit are all bundled together as one controlled phase.

If the phase can be expressed exactly with K bits, the phase estimate obtained from the QPEA corresponds to the actual phase. However, if the phase is unknown and not necessarily a binary fraction of finite length, the QPEA gives only SQL scaling for the variance of the phase estimate [49]. To achieve the Heisenberg scaling the QPEA needs to be generalized. In the generalized form of the QPEA each power of U, i.e. U^{2^k}, or equivalently each 2^k-photon NOON state, for $k = K, ..., 0$, needs to be repeated M times [49, 50]. Using Bayesian

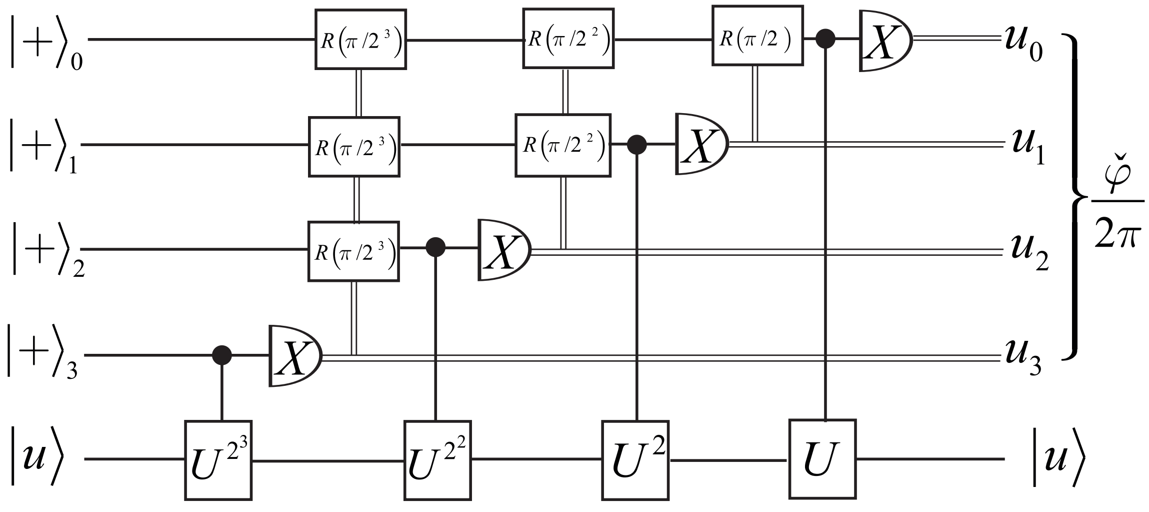


Figure 1.4: Quantum phase estimation algorithm diagram to estimate a phase with four binary digits. Double lines represent classical rails.

updating and adaptive feedback, this scheme can reach the Heisenberg scaling for constant M, provided $M \geq 4$ [49]. In this generalization, at each step the feedback phase is chosen to minimise the Holevo variance of the phase estimate after the next detection. To minimise the Holevo variance one needs to maximise $\left|\left\langle e^{i\check{\varphi}}\right\rangle\right|$ (see Eq. (1.7)). Note that, when measurements are performed with 2^k-photon NOON states, the phase is known modulo $2\pi/2^k$. Therefore, we instead need to maximise $\left|\left\langle e^{i2^k\check{\varphi}}\right\rangle\right|$.

In general, M can be taken to be a function of K and k. A variation commonly considered is linear scaling with k as $M(K,k) = G + F(K-k)$ for $F > 0$. As is clear, G is the number of applications of U^{2^K} or equivalently the number of repetitions of the 2^K-photon NOON state. As k is reduced the number of repetitions increases. The motivation for this variation is that measurements with smaller values of k distinguish between the multiple phase estimates obtained by larger values of k. Inaccuracy in distinguishing between these phase estimates, results in a large variance. Moreover, since smaller values of k require less resources, it is better to have larger number of measurements for small values of k to avoid large variances in the phase estimate.

For appropriately chosen values of G and F even a non-adaptive QPEA scheme reaches the Heisenberg scaling [50]. In the non-adaptive scheme, the controlled phase is incremented by $\pi/M(K,k)$ after measurement on each of the 2^k-photon NOON states, independent of the measurement results. In this case, the number of repetitions needs to be a function of k and in comparison to the adaptive scheme, the number of repetitions needs to be larger.

The measurement scheme we will use in Chapter 2 is inspired by the adaptive form of the QPEA. In Chapter 4 we use both adaptive and non-adaptive forms of the generalized QPEA.

1.4 Ramsey interferometry

The Ramsey interferometer is formally equivalent to the Mach-Zehnder interferometer. It is used for precise measurement of transition frequency in atoms [58] and to measure other physical quantities that affect the transition frequency such as magnetic field [59]. A diagram of the Ramsey interferometer is shown in Fig. 1.5. It consists of two $\pi/2$ Rabi pulses, which

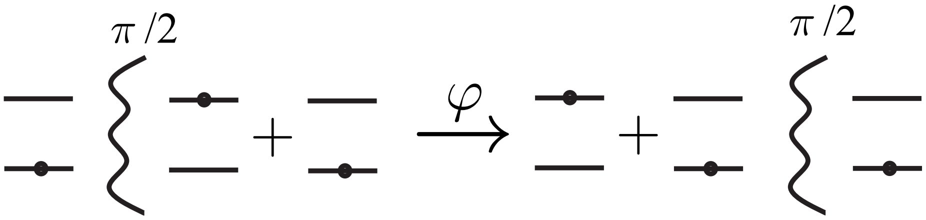


Figure 1.5: Ramsey interferometry. A $\pi/2$ pulse prepares an equal superposition of the ground and excited state of a two level atom, initially in the ground state. The atom then goes through a free evolution. This is followed by another $\pi/2$ pulse which projects the atom into one of the states depending on the phase φ imprinted on the superposition state.

are equivalent to the two beam splitters in the Mach-Zehnder interferometer, and a free evolution, equivalent to the phase shift in the Mach-Zehnder interferometer.

Consider a two-level atom initially prepared in its ground state. The first $\pi/2$ Rabi pulse prepares an equal superposition of the ground and excited state, $|0\rangle$ and $|1\rangle$, respectively. The system then undergoes a free evolution for a fixed time t_0 in which the relative phase shift $\Delta E t_0/\hbar = \omega t_0$ is induced between the two states, where ΔE is the energy difference between the two states, ω is the transition frequency, $\hbar$ is the reduced Planck constant and t_0 is the evolution time. The second $\pi/2$ pulse is then applied and the final state is measured. Depending on the relative phase shift, the atom is measured in either the ground or the excited state.

Treating the two-level atom as a qubit, the state evolution from the initial state to the final state can be described as

$$\begin{aligned}|0\rangle &\xrightarrow{R_y(\frac{\pi}{2})} \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)\\ &\xrightarrow{R_z(\varphi)} \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\varphi}|1\rangle\right)\\ &\xrightarrow{R_y(\frac{\pi}{2})} \begin{cases} p\left(0|\varphi\right) = \left(1+\cos\varphi\right)/2\\ p\left(1|\varphi\right) = \left(1-\cos\varphi\right)/2\end{cases}\end{aligned} \tag{1.37}$$

Here, $\varphi = \omega t_0$, $R_i(\varphi) = \exp(-i\varphi S_i)$ is the rotation operator on the Bloch sphere along the i axis, with S_i being the spin operator. If we consider the interaction time to be 2^k multiple of t_0, for an integer k, we obtain 2^k-fold phase shift. This is mathematically equivalent to a 2^k-photon NOON state in the Mach-Zehnder interferometer. The methods we discussed in the previous section can also be implemented in Ramsey interferometry to resolve ambiguity in the phase estimate and obtain the limit analogous to the Heisenberg limit for the variance of the phase estimate [54, 60, 61].

1.5 Dyne measurements

The phase of a beam in a single spatial mode can be estimated from measurements of quadratures of the field. Measurements of this type include homodyne (where just one quadrature is measured), heterodyne (where both quadratures are measured), and adaptive measurements that approximate homodyne measurements. In general, these measurements can be called dyne measurements [18, 62, 63]. A diagram of a dyne measurement is shown in Fig. 1.6. Consider a single-mode field such as a continuous coherent state $\left|\alpha = |\alpha|\, e^{i\varphi}\right\rangle$.

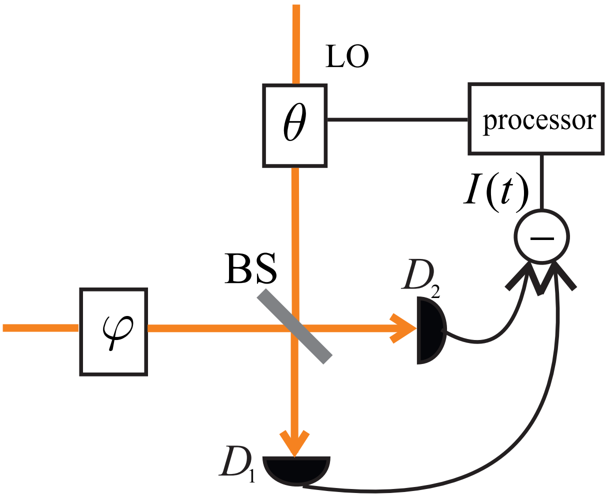


Figure 1.6: Adaptive measurement of the phase φ imposed on a single-mode beam. This scheme approximates the homodyne measurement. The single mode is combined with a strong local oscillator (LO) on a 50/50 beam splitter (BS). Based on the result of detectors D_1 and D_2 the processor adjusts the phase of the local oscillator θ.

The phase φ can be measured by measuring the electric field of the beam. The electric field operator is

$$\hat{\vec{E}} = \sqrt{\frac{\hbar\omega}{2\varepsilon_0 V}} \left(\vec{e}\,\hat{a}e^{i\left(\vec{k}.\vec{r}-\omega t\right)} + \vec{e}^{\,*}\,\hat{a}^{\dagger}e^{-i\left(\vec{k}.\vec{r}-\omega t\right)}\right) \tag{1.38}$$

where $\hbar$ is the reduced Planck constant, ε_0 is the vacuum permittivity, V is the quantization volume, ω is the angular frequency, $\vec{e}$ is the polarization vector, and a and $a^{\dagger}$ are annihilation and creation operators. This operator is proportional to the quadrature operator

$$\hat{X}_{\theta} = \hat{a}e^{-i\theta} + \hat{a}^{\dagger}e^{i\theta} \tag{1.39}$$

where $\theta = \omega t - \vec{k}.\vec{r}$.

To measure θ, the field is combined at a 50/50 beam splitter with a strong coherent state which can be treated classically, and is known as the local oscillator (LO). The LO acts as a phase reference. The annihilation operators of the two output ports of the beam splitter, $\hat{d}_0$ and $\hat{d}_1$ at time t can be written as [64]

$$\hat{d}_i(t) = \frac{1}{\sqrt{2}}\left(\hat{a} + (-1)^i\gamma\right), \qquad i = 0, 1 \tag{1.40}$$

where $\hat{a}$ is the annihilation operator of the coherent state, and γ is the amplitude of the LO. The photodetection rate in each of the outputs is given by

$$\left\langle \hat{d}_i^{\dagger}(t)\hat{d}_i(t)\right\rangle = \frac{1}{2}\left\langle\left(\hat{a}^{\dagger} + (-1)^i\gamma^*\right)\left(\hat{a} + (-1)^i\gamma\right)\right\rangle = \frac{1}{2}\left(|\alpha|^2 + |\gamma|^2 + (-1)^i 2\mathrm{Re}(\alpha\gamma^*)\right). \tag{1.41}$$

Denoting the number of photocounts in each output in the time interval $[t, t+\delta t)$ by δN_i, the difference photocurrent between the two output ports of the beam splitter can be be written as

$$I(t) = \lim_{\delta t\to 0}\lim_{|\gamma|\to\infty}\frac{\delta N_0 - \delta N_1}{|\gamma|\,\delta t} \tag{1.42}$$

The expectation values of the increments δN_i in the limit $\delta t \to 0$ can be written as

$$\langle dN_i\,(t)\rangle = \left\langle \hat{b}_i^{\dagger}\,(t)\,\hat{b}_i\,(t)\right\rangle dt. \tag{1.43}$$

For the continuous coherent state $|\alpha\rangle$ this gives [64]

$$I(t)dt = 2\mathrm{Re}\left(\alpha e^{-i\theta(t)}\right)dt + dW(t), \tag{1.44}$$

where $\theta(t)$ is the phase of γ and $dW(t)$ is a Wiener increment. Note that, $\langle I\,(t)\rangle = \left\langle \hat{X}_\theta\right\rangle$. Thus, photocurrent measurement is measurement of the θ quadrature. Based on the difference photocurrent an estimate of the phase can be obtained through Bayesian updating [65–67] or based on mathematical functions [19, 64, 68–71]. The technique we are following in this thesis is the latter.

In heterodyne measurements, the phase of the LO is varied linearly, i.e. it is set to $\theta(t) = \theta_0 + \Delta t$, where t is time, Δ is detuning and θ_0 is the phase at $t = 0$. The difference photocurrent has equal information about $X_{\theta=0}$ and $X_{\theta=\pi/2}$. The phase estimate $\breve{\varphi}(t)$ can be obtained from an appropriate functional of the measured photocurrent.

If the signal phase is known approximately, a lower phase variance can be obtained by setting the phase of the LO to $\theta = \varphi + \pi/2$ to measure the phase quadrature $X_{\varphi+\pi/2}$. This is known as homodyne measurement. If we don't have any information about the phase in advance, we can use adaptive homodyne measurement in which the phase of the LO is changed during the measurement. The result of the previous measurements is used to obtain an estimate of the phase $\breve{\varphi}$. The LO phase is then set to $\theta(t) = \breve{\varphi}(t) + \pi/2$. [68, 72].

For a phase that varies as a Wiener process, i.e.,

$$\frac{d\varphi}{dt} = \sqrt{\kappa}\xi(t), \tag{1.45}$$

where κ is the noise strength and ξ is a real Gaussian white noise defined as

$$\langle \xi(t)\xi(t')\rangle = \delta(t - t'), \tag{1.46}$$

using a coherent state in an adaptive homodyne scheme gives $N^{-1/2}/2$ for the variance of the phase [64, 70, 71]. A squeezed state, on the other hand, gives $N^{-2/3}$ scaling for the phase variance [64, 70, 71]. In Chapter 5 we show that using squeezed states in an adaptive homodyne scheme we can obtain Heisenberg scaling for a more general time varying phase.

1.6 Thesis layout

Having set the basis in this chapter, in the following chapters we address the following problems in quantum metrology.

In Chapter 2, we study the effect of photon loss in Mach-Zehnder interferometer. We propose a set of states that are resistant to photon loss and are obtained by processing the output from parametric down-conversion. We use sequences of such states in an adaptive measurement scheme in order to obtain an unambiguous phase measurement that beats the standard quantum limit.

In Chapter 3, we find optimal states for spectroscopy in an interferometric scheme. We consider a Mach-Zehnder interferometer with an ensemble of atoms in one of its arms. We show that at frequencies near the transition frequency of atoms, additional information about the transition frequency can be obtained from the phase. We consider the information about the transition frequency obtained from both the absorption and the phase shift, as quantified by the Fisher information. We examine the use of multiple single-photon states, NOON states, and numerically optimised states that are entangled and have multiple photons.

In Chapter 4, we move to Ramsey interferometry and magnetic field sensing using an NV centre. In this collaborative work with experimentalists we consider non-adaptive and a range of adaptive phase estimation algorithms to get an unambiguous estimate of an unknown time-independent magnetic field. We show that, although non-adaptive protocols reach the Heisenberg scaling, there is an optimised adaptive protocol that outperforms the non-adaptive protocol. We show that the enhancement obtained by the optimised adaptive protocol is significant if the overhead time, i.e. initialisation and measurement time are taken into account.

In Chapter 5, rather than single shot measurements considered up to that point in the thesis, we consider continuous measurements. Moreover, rather than entanglement, we exploit squeezing, another nonclassical property. The phase we consider in this chapter is a fluctuating phase with power law spectral density, imposed on a continuous squeezed state. We show that by using squeezed states in an adaptive homodyne measurement scheme we can reach the Heisenberg scaling for the variance of the phase estimate.

2

Loss resistant unambiguous phase measurement

2.1 Introduction

In the previous chapter we explained that although NOON states give the highest Fisher information per photon in phase estimation, they perform poorly in the presence of photon loss. However, it has been shown that there are other states that beat the SQL in the presence of loss. Examples of such states are low-photon number NOON states combined with single photon states [73], entangled coherent states [74], the combination of a coherent state and a squeezed vacuum state [75] corresponding to the original proposal of Caves [6], and the states that are optimised to be robust to photon loss [76–83]. Nevertheless, when there is photon loss the lower limit to the phase uncertainty is just a constant factor improvement over the SQL in the limit of large N [80, 81, 84, 85].

In considering optimised states for two-mode interferometry, they can be taken to be of the form $\sum_{k=0}^{N} \psi_k |N-k,k\rangle$. The total photon number can be taken to be a single value N; there is no advantage in using a superposition over different total photon numbers, because the detection process determines the number of photons and destroys such superposition. In the case of NOON states, one has ψ_k nonzero only for two values of k, whereas general optimised states would allow all ψ_k to be nonzero. This raises a challenging state engineering problem [86–89]. Such a general state could, for example, be produced by using N independent single photons [90].

The standard method to produce single photons is to use down-conversion, and postselect on detection of a single photon to obtain a single photon in the other output. This means that such a scheme would be very wasteful, because $2N$ photons would need to be produced to obtain a state with N photons. In this chapter, we first propose a scheme to produce states using all photons output by the down-conversion. We then show how to use sequences of such states with different total numbers of photons in order to obtain unambiguous phase estimates. This is done in much the same way as has been done for NOON states [26, 49, 91].

2.2 The state preparation scheme

We start with an M-port linear optical device (LOD) shown in Fig. 2.1. An LOD can be decomposed into a triangular array of beam splitters and phase shifters [92]. Any M port LOD can be represented by an $M \times M$ unitary matrix U with elements U_{ij}. It transforms input photon creation operators $\hat{a}_i^\dagger$ to the output creation operators $\hat{b}_i^\dagger$ according to

$$\hat{a}_i^\dagger \mapsto \sum_{j=1}^{M} U_{ij} \hat{b}_j^\dagger. \tag{2.1}$$

Our aim here is to propose an LOD scheme to generate two-mode multi-photon entangled states in the output which are resistant to photon loss. The most general two-mode N-photon pure state can be written as

$$|\psi\rangle_{\text{out}} = \sum_{k=0}^{N} \psi_k \, |N-k, k\rangle = \sum_{k=0}^{N} \psi_k \frac{1}{\sqrt{k!(N-k)!}} \left(\hat{b}_1^\dagger\right)^{N-k} \left(\hat{b}_2^\dagger\right)^{k} |0, 0\rangle, \tag{2.2}$$

where the quantities ψ_k are the probability amplitudes of the photon number components. When the loss is equal in both arms, as we consider here, optimal states for phase estimation are symmetric between the paths [77], so we aim to have $\psi_k = \psi_{N-k}$ in our output state. We postselect vacuum in the other $M-2$ output modes of the LOD in order to maximise the number of photons in the output.

Spontaneous parametric down-conversion (SPDC) is the most common source of photon pairs used in experiments. In order to make our state preparation scheme experimentally feasible, we consider the output from SPDC as the input to the LOD. The SPDC state can be written as [93]

$$|\psi\rangle_{\text{SPDC}} \propto |0\rangle_{\text{s}}|0\rangle_{\text{i}} + \xi|1\rangle_{\text{s}}|1\rangle_{\text{i}} + \xi^2|2\rangle_{\text{s}}|2\rangle_{\text{i}} + \xi^3|3\rangle_{\text{s}}|3\rangle_{\text{i}} + \ldots., \tag{2.3}$$

where the subscripts s and i stand for signal and idler; these will be omitted from this point on for brevity. We also use the proportionality symbol to indicate that a normalization constant has been omitted. The quantity ξ depends on the interaction time between the

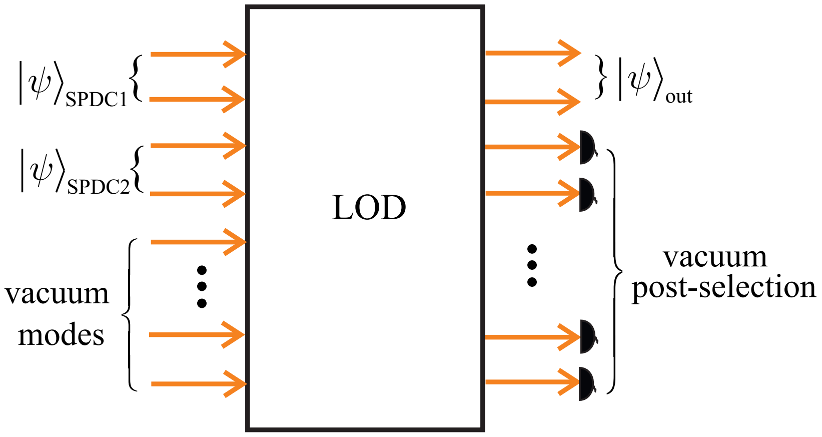

$|\psi\rangle_{\text{SPDC1}}$
$|\psi\rangle_{\text{SPDC2}}$
vacuum modes
LOD
$|\psi\rangle_{\text{out}}$
vacuum post-selection

Figure 2.1: A linear optical device (LOD) with two SPDC states and vacuum modes as input. $|\psi\rangle_{\text{out}}$ is the desired output state, and there is postselection on vacuum in the other modes. The two SPDC input states can be simplified to one.

optical nonlinear crystal and pump laser, the strength of the nonlinearity in the crystal and the power of the pump laser.

If we input the LOD with one SPDC source, recording a total of $N = 2n$ photons in the output postselects only the dual Fock state $|n, n\rangle$ as input

$$|n, n\rangle = \frac{1}{n!}\left(\hat{a}_1^\dagger\right)^n \left(\hat{a}_2^\dagger\right)^n |0, 0\rangle . \tag{2.4}$$

Note that here we are considering a simplified model where the value of n is known, whereas in practice it may not be known if there is loss. If vacuum is postselected in $M - 2$ of the output modes (so the $2n$ photons are detected in the first two modes), the output state can be represented by removing all the terms containing $\hat{b}_i^\dagger$ for $i \geq 3$. Transforming creation operators in Eq. (2.4) by Eq. (2.1), and omitting $\hat{b}_i^\dagger$ with $i \geq 3$ gives the output state in the form

$$|\psi\rangle_{\text{out}} = \frac{1}{n!}\left[\chi_1 \left(\hat{b}_1^\dagger\right)^2 + \chi_2 \hat{b}_1^\dagger \hat{b}_2^\dagger + \chi_3 \left(\hat{b}_2^\dagger\right)^2\right]^n |0, 0\rangle \tag{2.5}$$

where χ_1, χ_2, χ_3 are the factors determined by the unitary matrix U.

Alternatively, consider the case where there were two SPDC sources used, as in Fig. 2.1. Allowing different factors ξ_1 and ξ_2, the input state can be written as

$$\begin{aligned}
|\psi\rangle_{\text{in}} &\propto \left(|0, 0\rangle + \xi_1|1, 1\rangle + \xi_1^2|2, 2\rangle + \xi_1^3|3, 3\rangle + ...\right) \\
&\quad \otimes \left(|0, 0\rangle + \xi_2|1, 1\rangle + \xi_2^2|2, 2\rangle + \xi_2^3|3, 3\rangle + ...\right) \\
&\propto e^{\xi_1 \hat{a}_1^\dagger \hat{a}_2^\dagger + \xi_2 \hat{a}_3^\dagger \hat{a}_4^\dagger}|0, 0\rangle|0, 0\rangle \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\xi_1 \hat{a}_1^\dagger \hat{a}_2^\dagger + \xi_2 \hat{a}_3^\dagger \hat{a}_4^\dagger\right)^n |0, 0\rangle|0, 0\rangle.
\end{aligned} \tag{2.6}$$

Recording a total of $2n$ photons in the output modes of LOD postselects the state

$$|\psi\rangle_{\text{in}} \propto \frac{1}{n!}\left(\xi_1 \hat{a}_1^\dagger \hat{a}_2^\dagger + \xi_2 \hat{a}_3^\dagger \hat{a}_4^\dagger\right)^n |0, 0\rangle|0, 0\rangle \tag{2.7}$$

as the input state. If vacuum is postselected in all but two of the output modes, we again obtain a state of the form given in Eq. (2.5). Similarly, if we were to consider an arbitrary number of SPDC sources, we would again obtain the same form of state. Therefore there is no advantage in considering larger numbers of SPDC sources, and we consider an LOD fed with only one SPDC source and vacuum to the remaining modes. As we aim for a symmetric state we set $\chi_1 = \chi_3$. It is also convenient to denote $\chi = \chi_2/\chi_1$, which makes it clear that the states are parametrized by just one real number.

Thus, we consider an M-port LOD in which all but two of the input modes are vacuum and also all but two of the output modes are postselected as vacuum, but we are allowing a potentially large number of modes. The scheme can be simplified to a four-port LOD in the following way. Figure 2.2(a) shows a five-port LOD with three vacuum input modes and postselection of vacuum at three modes in the output, with the LOD simplified to a triangular array of beam splitters and phase shifts [92]. For simplicity, the beam splitters shown in the diagram also include phase shifts, rather than showing the phase shifts separately. It can be seen that there are six beam splitters that have vacuum input and output [those below the dashed line in Fig. 2.2(a)]. These beam splitters leave the field unchanged, and can be omitted. Therefore, the scheme can be simplified to the four-port LOD shown in Fig. 2.2(b). Similarly,

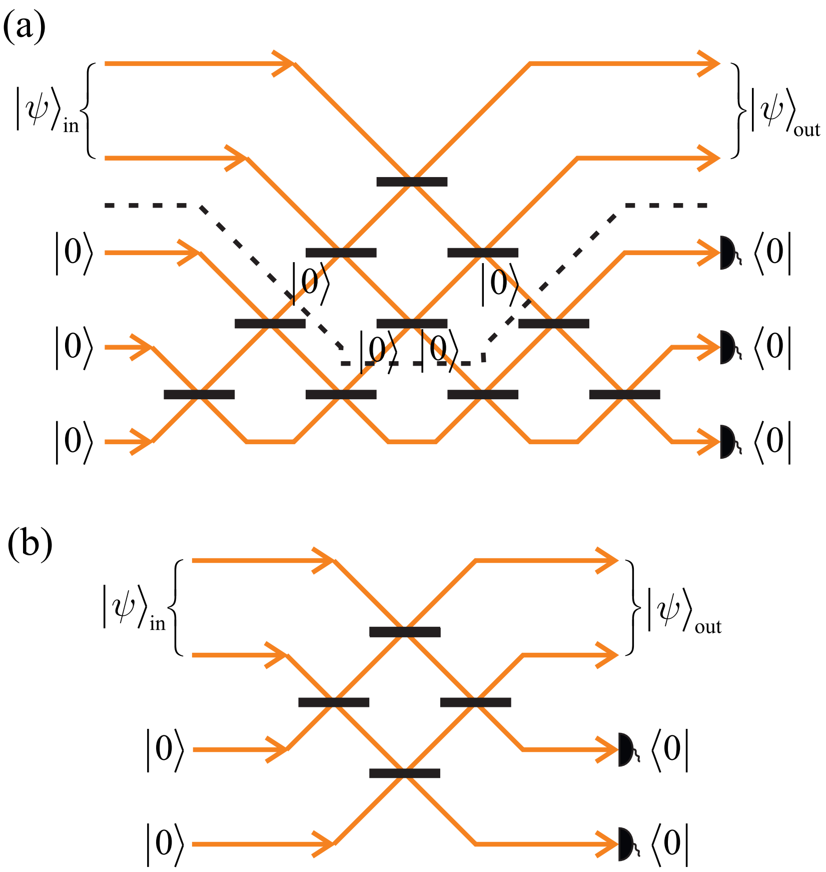

(a)
$|\psi\rangle_{\text{in}}$
$|\psi\rangle_{\text{out}}$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$|0\rangle$
$\langle 0|$
$\langle 0|$
$\langle 0|$
(b)
$|\psi\rangle_{\text{in}}$
$|\psi\rangle_{\text{out}}$
$|0\rangle$
$|0\rangle$
$\langle 0|$
$\langle 0|$

Figure 2.2: (a) A triangular array of beam splitters and phase shifts fed with a two-mode input state $|\psi\rangle_{\text{in}}$ and three vacuum modes, $|0\rangle$. The phase shifts are included with the beam splitters (shown as thick black lines) for simplicity. Output photons are detected in two of the modes, $|\psi\rangle_{\text{out}}$, and vacuum is postselected in three of the output modes. All the beams below the dashed line contain vacuum. (b) The five-port LOD can be simplified to a four-port one by keeping only the beam splitters above the dotted line.

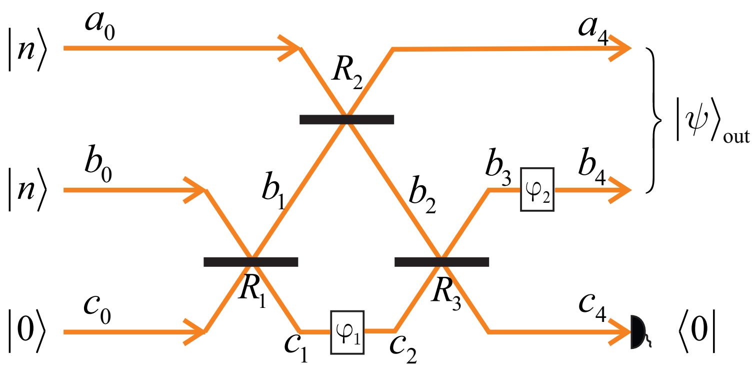

$|n\rangle$
a_0
a_4
R_2
$|\psi\rangle_{\text{out}}$
$|n\rangle$
b_0
b_1
b_2
b_3
φ_2
b_4
R_1
R_3
$|0\rangle$
c_0
c_1
φ_1
c_2
c_4
$\langle 0|$

Figure 2.3: The state preparation scheme for loss-resistant states. It is a three-port interferometer consisting of three beam splitters with reflectivities R_1, R_2, R_3 and two phase shifters φ_1, φ_2. The initial dual Fock state $|n\rangle\,|n\rangle$ gives the $2n$-photon loss-resistant state in the output, as indicated in Eq. (2.16).

if we started with an arbitrary number of modes (> 4), the scheme could be simplified to four modes in exactly the same way.

This shows that we can simplify to four modes, but we have found an even simpler scheme using three modes, as shown in Fig. 2.3. In this figure, R_1, R_2, R_3 are reflectivities of beam splitters and φ_1, φ_2 are phases of phase shifters. We use $\hat{a}_0^\dagger$, $\hat{b}_0^\dagger$, $\hat{c}_0^\dagger$ for the creation operators of the input modes. We denote the successive transformations by the subscripts, i.e. we use the subscript 1 after the first beam splitter, 2 after the second beam splitter and the phase shift φ_1, 3 after the third beam splitter and 4 for the output modes. We label the transmissivity of the ith beam splitter by T_i.

Consider the input state as

$$|\psi_0\rangle = |n, n, 0\rangle = \frac{1}{n!}\left(\hat{a}_0^\dagger\right)^n \left(\hat{b}_0^\dagger\right)^n |0, 0, 0\rangle \,. \tag{2.8}$$

Here, we assume that states with a given value of n can be produced on demand. This is a simplifying assumption for a source that is more controllable than SPDC. The first beam splitter leaves the first mode unaffected, $\hat{a}_0^\dagger = \hat{a}_1^\dagger$, and transforms the other two modes to

$$\hat{b}_0^\dagger = i\sqrt{R_1}\hat{b}_1^\dagger + \sqrt{T_1}\hat{c}_1^\dagger, \qquad \hat{c}_0^\dagger = i\sqrt{R_1}\hat{c}_1^\dagger + \sqrt{T_1}\hat{b}_1^\dagger. \tag{2.9}$$

Therefore, the state after the first beam splitter is

$$|\psi_1\rangle = \frac{1}{n!}\left(\hat{a}_1^\dagger\right)^n \left(i\sqrt{R_1}\hat{b}_1^\dagger + \sqrt{T_1}\hat{c}_1^\dagger\right)^n |0, 0, 0\rangle \,. \tag{2.10}$$

The second beam splitter transforms $\hat{a}_1^\dagger$ and $\hat{b}_1^\dagger$ to

$$\hat{a}_1^\dagger = i\sqrt{R_2}\hat{a}_2^\dagger + \sqrt{T_2}\hat{b}_2^\dagger, \qquad \hat{b}_1^\dagger = i\sqrt{R_2}\hat{b}_2^\dagger + \sqrt{T_2}\hat{a}_2^\dagger. \tag{2.11}$$

The phase shift gives $\hat{c}_1^\dagger = e^{i\varphi_1}\hat{c}_2^\dagger$. Thus we obtain the state

$$|\psi_2\rangle = \frac{1}{n!}\left(i\sqrt{R_2}\hat{a}_2^\dagger + \sqrt{1 - R_2}\hat{b}_2^\dagger\right)^n \left(i\sqrt{R_1}\left(i\sqrt{R_2}\hat{b}_2^\dagger + \sqrt{T_2}\hat{a}_2^\dagger\right) + \sqrt{T_1}e^{i\varphi_1}\hat{c}_2^\dagger\right)^n |0, 0, 0\rangle \,. \tag{2.12}$$

The final beam splitter leaves the first output mode unaffected, $\hat{a}_2^\dagger = \hat{a}_3^\dagger$, and for the other modes gives

$$\hat{b}_2^\dagger = i\sqrt{R_3}\hat{b}_3^\dagger + \sqrt{T_3}\hat{c}_3^\dagger, \qquad \hat{c}_2^\dagger = i\sqrt{R_3}\hat{c}_3^\dagger + \sqrt{T_3}\hat{b}_3^\dagger. \tag{2.13}$$

This gives the following state

$$\begin{aligned}|\psi_3\rangle = \frac{1}{n!}&\left(i\sqrt{R_2}\hat{a}_3^\dagger + \sqrt{T_2}\left(i\sqrt{R_3}\hat{b}_3^\dagger + \sqrt{T_3}\hat{c}_3^\dagger\right)\right)^n \\ &\times\left(i\sqrt{R_1}\left(i\sqrt{R_2}\left(i\sqrt{R_3}\hat{b}_3^\dagger + \sqrt{T_3}c_3^\dagger\right) + \sqrt{T_2}\hat{a}_3^\dagger\right) + \sqrt{T_1}e^{i\varphi_1}\left(i\sqrt{R_3}\hat{c}_3^\dagger + \sqrt{T_3}\hat{b}_3^\dagger\right)\right)^n \\ &\times|0, 0, 0\rangle \,.\end{aligned} \tag{2.14}$$

Acting the last phase shift and postselecting the vacuum in one of the output modes, replacing $\hat{c}_3^\dagger$ with zero, gives the unnormalized state

$$\begin{aligned}|\psi\rangle_{\text{out}} = \frac{1}{n!}&\left(i\sqrt{R_2}\hat{a}_4^\dagger + ie^{i\varphi_2}\sqrt{T_2R_3}\hat{b}_4^\dagger\right)^n \\ &\times\left(i\sqrt{R_1T_2}\hat{a}_4^\dagger + \hat{b}_4^\dagger e^{i\varphi_2}\left(e^{i\varphi_1}\sqrt{T_1T_3} - i\sqrt{R_1R_2R_3}\right)\right)^n |0, 0\rangle.\end{aligned} \tag{2.15}$$

To obtain the state

$$|\psi\rangle_{\text{out}} \propto \left[\left(\hat{b}_1^\dagger\right)^2 + \chi\hat{b}_1^\dagger\hat{b}_2^\dagger + \left(\hat{b}_2^\dagger\right)^2\right]^n |0, 0\rangle \tag{2.16}$$

with $\chi \in [0,2]$, we can use the reflectivities and phase shifts as

$$\varphi_1 = \arcsin\left(\frac{1}{2}(\chi-1)\sqrt{2+\chi}\right), \qquad \varphi_2 = \arccos\left(\frac{\chi}{2}\right),$$
$$R_1 = \frac{1}{1+\chi}, \quad R_2 = \frac{1}{2+\chi}, \quad R_3 = \frac{1}{1+\chi}. \tag{2.17}$$

For the two-photon state with $n = 1$, the output state as given in Eq. (2.16) can be expanded as

$$|\psi\rangle_{\text{out}} \propto \sqrt{2}\,|0,2\rangle + \chi\,|1,1\rangle + \sqrt{2}\,|2,0\rangle\,. \tag{2.18}$$

For $n = 2$, which is the four-photon state, we get the following expansion

$$|\psi\rangle_{\text{out}} \propto |0,4\rangle + \chi|1,3\rangle + \frac{2+\chi^2}{\sqrt{6}}|2,2\rangle + \chi|3,1\rangle + |4,0\rangle. \tag{2.19}$$

Unlike NOON states, the above states have terms other than $|N,0\rangle$ and $|0,N\rangle$. This makes these states more resilient to photon loss; in other words, loss of a single photon does not destroy the phase sensitivity of the states. Our proposed states still have some ambiguity in phase estimation, but by combining them with single photon states the ambiguity can be removed. By adjusting the reflectivities of the beam splitters and phases of the phase shifters we can choose the value of χ and optimise for phase measurement with loss.

In the case of two-photon states, the preparation technique is enough to obtain arbitrary symmetric states, and so is sufficient to obtain the optimal states. In contrast, for the four-photon states we would need two independent parameters. As there is only one parameter, χ, which can be varied, we cannot obtain the exactly optimal states of the form

$$|\psi\rangle_{\text{ex}} \propto |0,4\rangle + \chi_1'|1,3\rangle + \chi_2'|2,2\rangle + \chi_1'|3,1\rangle + |4,0\rangle, \tag{2.20}$$

where χ_1' and χ_2' are independent real variables. However, we will show that we can obtain results close to optimum. In the following sections we show the effect of photon loss on states of the form (2.16) when using an adaptive phase measurement scheme, as depicted in Fig. 2.4.

2.3 Photon loss

When using nonclassical states in optical interferometers, photon loss is the most important source of decoherence. A standard way to model photon loss in these systems is to introduce fictitious beam splitters in the arms of the interferometer [83]. Here, we consider the same amount of loss in both arms. The transmissivity, η, of the fictitious beam splitters determines the efficiency of the system. Note that, equal losses at any point in the interferometer are equivalent, so the loss in the arms of the interferometer is equivalent to having inefficient detectors with equal efficiency.

The loss-resistant states in Eq. (2.16) can be written in the standard form

$$|\psi\rangle_{\text{in}} = \sum_{k=0}^{N} \psi_k |N-k,k\rangle, \tag{2.21}$$

where the amplitudes of the photon number components, ψ_k, are no longer independent, but instead are all determined by the single parameter χ, introduced in the previous section. We

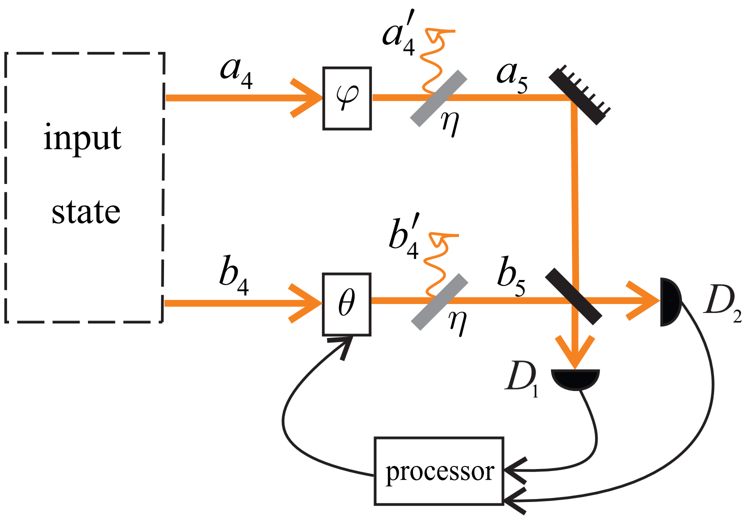


Figure 2.4: Adaptive phase measurement scheme. φ is the unknown phase, θ is the controlled phase and D_1 and D_2 are the photon detectors in the two outputs. The grey diagonal lines in the two arms are the fictitious beam splitters with transmissivity η modeling photon loss.

denote the action of the phase shift on the state by $\Psi_k = \psi_k e^{i(N-k)\varphi} e^{ik\theta}$. The fictitious beam splitters are depicted as being after the phase shifts (see Fig. 2.4). The overall effect on the state is the same regardless of whether the beam splitters are before or after the phase shifts, but we just adopt this convention for the calculations. The fictitious beam splitters transform the creation operators of the upper and lower arms, $\hat{a}_4^\dagger$ and $\hat{b}_4^\dagger$, according to

$$\hat{a}_4^\dagger = \sqrt{\eta}\hat{a}_5^\dagger + i\sqrt{1-\eta}\hat{a}_4'^\dagger, \qquad \hat{b}_4^\dagger = \sqrt{\eta}\hat{b}_5^\dagger + i\sqrt{1-\eta}\hat{b}_5'^\dagger. \tag{2.22}$$

where $\hat{a}_4'^\dagger$ and $\hat{b}_4'^\dagger$ are the creation operators of the loss modes. Thus, after loss the state $|N-k,k\rangle$ becomes

$$\sum_{n=0}^{N-k}\sum_{m=0}^{k} i^{n+m}\sqrt{C_n^{N-k}C_m^k}\,(\sqrt{\eta})^{N-n-m}\left(\sqrt{1-\eta}\right)^{n+m}|N-k-n,k-m,n,m\rangle, \tag{2.23}$$

where the third and fourth modes in the state are the loss modes and $C_m^k = \binom{k}{m}$ is the binomial coefficient. If we denote the total number of photons lost as $L = n + m$, and trace over loss modes, the reduced density operator can be written as

$$\rho = \sum_{L=0}^{N}\sum_{m=0}^{L}\sum_{r,s=0}^{N-L}\Psi_{r+m}\Psi_{s+m}^* A_{N,L,r,m}A_{N,L,s,m}^*|N-L-r,r\rangle\langle N-L-s,s|, \tag{2.24}$$

where

$$A_{N,L,r,m} \equiv \sqrt{\eta^{N-L}(1-\eta)^L C_{N-L-r}^{N-r-m}C_r^{r+m}}. \tag{2.25}$$

After loss, a 50/50 beam splitter acts on the state. Calculating the output density operator, its diagonal elements give the output detection probabilities

$$\begin{aligned} P_{L,k}(\varphi,\theta) = &\sum_{m=0}^{L}\sum_{r,s=0}^{N-L}\sum_{r_2=B_{kNLr}}^{\min(r,k)}\sum_{s_2=B_{kNLs}}^{\min(s,k)}\Psi_{r+m}\Psi_{s+m}^* A_{N,L,r,m}A_{N,L,s,m}^*\left(\frac{1}{2}\right)^{N-L}(-1)^{r_2+s_2} \\ &\times\frac{(N-L-k)!k!}{\sqrt{(N-L-r)!r!(N-L-s)!s!}}\binom{N-L-r}{k-r_2}\binom{r}{r_2}\binom{N-L-s}{k-s_2}\binom{s}{s_2}, \end{aligned} \tag{2.26}$$

where L is the total number of photons lost, $B_{kNL\ell} = \max(0, k - N + L + \ell)$ (for $\ell = r$ or s), and k is the number of photons detected in one of the output ports (which can be from 0 to $N - L$).

For the two-photon input state, Eq. (2.18), there are six different possible detection results:

- if no photons are lost ($L = 0$), then we can detect zero, one or two photons in output mode 2 ($k = 0, 1, 2$),
- if one photon is lost ($L = 1$), then we can detect zero or one photon in output mode 2 ($k = 0, 1$), and
- if both photons are lost ($L = 2$), then the only possible detection result is vacuum ($k = 0$).

Similarly for the four-photon state, Eq. (2.19), there are fifteen possible detection results. In Eq. (2.26), we have labeled possible detection results with the number of photons lost and the number of photons detected in mode 2, but the detections are of the number of photons in modes 1 and 2. The number of photons lost is determined from the initial number of photons and the total number of photons detected. The larger number of possible detection results in comparison to the lossless case makes the calculations more computationally difficult.

Because the probabilities depend on the Ψ_k coefficients, which in turn contain exponentials of the phase φ, the probabilities $P_{L,k}(\varphi, \theta)$ can be written as a Fourier series

$$P_{L,k}(\varphi, \theta) = \frac{1}{2\pi} \sum_j a_j e^{ij\varphi}. \tag{2.27}$$

When updating the phase probability by Bayes' theorem, explained in Section 1.2, the probability can be represented by a finite number of the Fourier coefficients.

In the following section we describe how our proposed states can be utilized in an adaptive measurement scheme to obtain phase variances less than the SQL.

2.4 The measurement scheme

The measurement scheme we propose is similar to the generalized quantum phase estimation algorithm (GQPEA), in the sense that a sequence of states is input to the interferometer. The states in this case are the loss-resistant states and they are input into the lossy interferometer. In the GQPEA the input states are NOON states, and the algorithm works based on the parity of the photon number difference in the output (see Eq. (1.34)). For either case, even or odd parity, the phase shift is $N\varphi$, N being the photon number. However, for the loss-resistant states proposed here, due to the presence of extra terms in the superposition, we have multiple outcomes. In other words, each of the possible detection results has a probability depending on the phase in a different way, and it is not possible to estimate the phase based on just the parity. Moreover, the possible combinations of detections is further increased by the loss. Here, in contrast to the GQPEA, we do not restrict ourselves to $M(K, k) = K + G(K - k)$ number of copies of each state. Instead, for a given total number of photons we find the number of copies of each state numerically. Moreover, we start the sequence with single photons rather than the highest photon number state.

We use the terminology "detection" for the measurement of an individual state, to contrast with the overall measurement of the phase, combining results of the individual detections.

We combine the results of detection of each of these successive states with single photon states to provide an overall measurement of the phase that is unambiguous. Moreover, we use feedback to adjust the controlled phase θ based on the previous detection results and controlled phases.

The globally optimal controlled phase is the one that minimises the final phase variance, but finding such a controlled phase requires a minimisation over an exponentially large number of variables. Here we adopt the approach of finding the locally optimal phase, that minimises the variance in the phase estimate after the next detection [26, 34]. This terminology is distinct from another type of terminology in optimisation, where locally optimal means that a parameter is optimal just for a region about that value, as opposed to optimal for all values. There are proposals to find the globally optimal phase in a more restricted sense that avoids needing an exponential number of variables [60, 94, 95], but those are still more computationally intensive than finding the locally optimal phase.

The measure we use for the variance of the phase is the Holevo phase variance, $V_{\mathrm{H}} = \mu^{-2} - 1$, explained in Section 1.1.1. Here, μ, called sharpness, is

$$\mu = \left|\left\langle e^{i(\check{\varphi}-\varphi)}\right\rangle\right|, \tag{2.28}$$

where the average is over the measurement results u_k. $\check{\varphi}$ is an unbiased estimator of the phase in the sense that $\varphi = \left\langle e^{i\check{\varphi}}\right\rangle$. Similar to Chapter 1 we use the vee accent $\check{\varphi}$ to indicate the phase estimator. The unbiased phase estimator with the smallest variance is $\check{\varphi} = \arg\left\langle e^{i\varphi}\right\rangle$ [34]. This average is over φ using the probability distribution for the phase given the measurement results [26]. As we explained in Section 1.1.1, the variance of the phase for an unbiased estimate is lower bounded by the inverse of the Fisher information according to the Cramér-Rao inequality

$$V(\varphi) \geq \frac{1}{F(\varphi,\theta)}, \tag{2.29}$$

where the Fisher information is

$$F(\varphi,\theta) = \sum_{u_k} \frac{1}{P(u_k|\varphi,\theta)} \left(\frac{\partial P(u_k|\varphi,\theta)}{\partial\varphi}\right)^2. \tag{2.30}$$

Note that this is for the usual variance, rather than the Holevo variance. The Fisher information effectively represents the amount of information about φ which is contained in the measurement result (though it is not quantified in bits as in the case of entropy). The probability $P(u_k|\varphi,\theta)$ is the probability of the measurement result u_k given the system and controlled phases φ and θ. It is the probability given in Eq. (2.26).

Optimal states for phase measurement with loss are typically evaluated via the Fisher information [76–81]. Often the quantum Fisher information is used, which gives a corresponding quantum Cramér-Rao bound [28]. The quantum Fisher information corresponds to a maximum of the Fisher information over all measurements. However, in practice the measurements are limited to photon counting [1]. For this reason we only consider the Fisher information. Using the Fisher information to evaluate our proposed four-photon loss-resistant states as given in Eq. (2.19), they are nearly as good as the general optimal states. As is shown in Fig. 2.5 the results are almost indistinguishable for $\eta < 0.7$.

However, the Fisher information only provides a lower bound on the variance. To provide a better test of our states, we consider the phase variance produced by measurements. The

---

[1] Recall that for *pure* path-symmetric states the quantum Fisher information is achieved with photon counting. Here we have path-symmetric states but they are not pure.

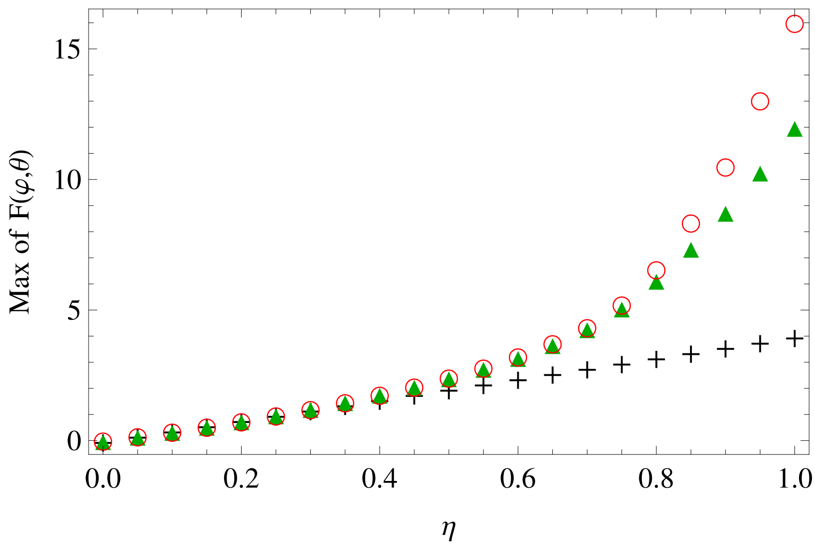


Figure 2.5: Maximum of classical Fisher information $F(\varphi,\theta)$ versus efficiency η. ○ : four-photon exact optimal state, Eq. (2.20), maximized over χ_1', χ_2', φ and θ. ▲: four-photon loss-resistant states, Eq. (2.19), maximized over χ, φ and θ. +: four single-photon states.

scheme we propose for phase estimation is as follows. We use a sequence of N_1 single-photon states followed by N_2 two- and N_4 four-photon loss-resistant states given in Eqs. (2.18) and (2.19). The value of χ must be determined for each kind of state for each amount of photon loss in the system. We denote the values of χ for the two- and four-photon loss-resistant states by $\chi^{(2)}$ and $\chi^{(4)}$, respectively. We have found the values of $\chi^{(2)}$ and $\chi^{(4)}$ by a numerical search for the values that minimise the phase variance, for a specific value of loss. For $\eta = 0.6$ we found the optimal value of $\chi^{(2)}$ to be 1.7 or 1.8, while we found $\chi^{(4)}$ to be 1.3 (for a total number of photons up to $\mathcal{N} = 30$).

Perhaps surprisingly, these values do not maximise the Fisher information given in Eq. (2.30); an example is shown in Fig. 2.6. In this figure, the Fisher information is shown for the two-photon loss-resistant state for the system phase $\varphi = \pi/4$ and controlled phase $\theta = 0$, and the maximum is for $\chi \approx 0.8$. We also show the phase variance for a sequence of seven single photons followed by one two-photon loss-resistant state. That is, there is one state dependent on the value of χ, and the remaining states are to resolve the phase ambiguity. In this case, the minimum is for $\chi \approx 1.7$, which is a radically different value to that which maximises the Fisher information.

We use Bayes' theorem to update the probability of the system phase given the measurement results u_m and controlled phases θ_m

$$P(\varphi|\vec{u}_m,\vec{\theta}_m) \propto P(u_m|\varphi,\theta_m)P(\varphi|\vec{u}_{m-1},\vec{\theta}_{m-1}) \tag{2.31}$$

where $\vec{u}_m = (u_1, u_2, \ldots, u_m)$ is the vector of successive measurement results and $\vec{\theta}_m = (\theta_1, \theta_2, \ldots, \theta_m)$ is the vector of the corresponding controlled phases (i.e. u_j is the measurement result with controlled phase θ_j). We also adopt the notation that $\vec{u}_0$ and $\vec{\theta}_0$ are the empty vectors. The proportionality factor is just a normalization constant, which is trivial to calculate. We assume that the phase is initially unknown, so the initial probability distribution is flat, and $P(\varphi|\vec{u}_0,\vec{\theta}_0) = 1/2\pi$. The probability $P(u_m|\varphi,\theta_m)$ is that given in Eq. (2.26). We set the first controlled phase to zero. This gives the same result as a random phase, because we

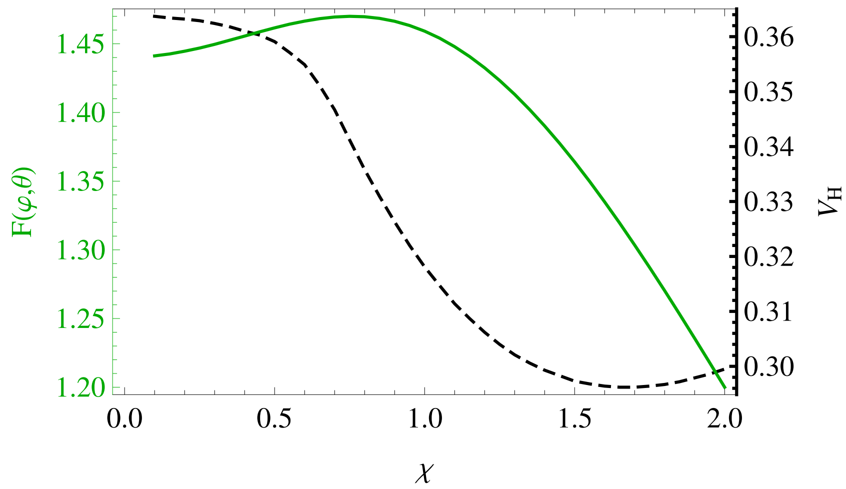


Figure 2.6: Fisher information, $F(\varphi, \theta)$ and Holevo phase variance, V_{H}, versus χ for $\eta = 0.6$. Solid green line: Fisher information for $\varphi = \pi/4$, $\theta = 0$ versus χ calculated using Eq. (2.30) for the two-photon state given in Eq. (2.18). Dashed black line: phase variance versus χ for the sequence of seven single photons followed by one two-photon loss-resistant state in the adaptive measurement protocol. The Fisher information gives a lower bound to the phase variance, rather than an exact phase variance, so the value of χ that maximises the Fisher information need not be the value that minimises the phase variance.

average over the system phase and only the relative phase between the arms is significant. The other controlled phases are obtained by maximising the sharpness after the next detection. In particular, the optimal θ_m is the one that maximises

$$\mu(\theta_m) = \frac{1}{2\pi} \sum_{u_m} \left| \int_0^{2\pi} d\varphi \, e^{i\varphi} \prod_{k=1}^{m} P(u_k|\varphi, \theta_k) \right|, \tag{2.32}$$

where the summation is over all the possible results for the m'th measurement, and the product over k corresponds to the updates of the probability distribution according to Bayes' rule [26]. Because the probabilities are represented as a Fourier series, as in Eq. (2.27), the integral over φ simply yields the coefficient of the $e^{-i\varphi}$ term, a_{-1}. Therefore this sum may be obtained by summing the predicted values of $|a_{-1}|$ after the next detection.

For the case of measurement with a single photon without loss, the formula for the controlled phase from Ref. [34] may be used. In the presence of loss the only extra detection result is where the photon is lost with a probability that is independent of the system and controlled phases. This just adds an extra constant to the sharpness and does not change the phases that maximise it. Therefore the formula from Ref. [34] may still be used. That is, we take one of the following three phases

$$\theta_0 = \arg\left(ba^* - c^*a\right), \quad \theta_\pm = \arg\left(\sqrt{\frac{c_2 \pm \sqrt{c_2^2 + |c_1|^2}}{c_1}}\right), \tag{2.33}$$

where

$$\begin{aligned} c_1 &= (a^*c)^2 - (ab^*)^2 + 4\left(|b|^2 - |c|^2\right) b^*c, \\ c_2 &= -2i \,\mathrm{Im}\left(a^2 b^* c^*\right), \end{aligned} \tag{2.34}$$

and a, b and c are functions of Fourier coefficients for the probability distribution: $a = a_{-1}$, $b = \frac{1}{2}a_{-2}$, $c = \frac{1}{2}a_0$. The optimal phase out of the above three possible phases is determined numerically.

The situation is more complicated with two- and four- photon loss-resistant states. For the simpler states considered in Ref. [1], it is possible to use the above formula in the two-photon case. However, here we have the complication that there is an additional $|1,1\rangle$ term in the state, and we also need to take account of the case where one photon is lost. This means that the formula no longer applies. For this reason we determined the optimal controlled phase numerically for the two- and four-photon cases.

We have calculated the exact phase variance by considering all the possible measurement results, of which there are $3^{N_1} \times 6^{N_2} \times 15^{N_4}$ (considering loss in all parts of the sequence). We sum over the sharpness for each sequence of measurement results as [26]

$$\mu = \frac{1}{2\pi}\sum_{\vec{u}_m}\left|\int_0^{2\pi} d\varphi\, e^{i\varphi}\prod_{k=1}^{m} P\left(u_k|\varphi,\theta_k\right)\right|. \tag{2.35}$$

To speed up the calculations it is useful to note that a measurement where n of the single photons were not lost is the same as one where there were n single photons without loss. This is because, if a single photon is lost there is no phase information and the probability distribution and controlled phase need not be updated. Using $\tilde{\mu}_n$ to denote the sharpness resulting from the sequence nN_2N_4 with no loss on the single photons, the actual sharpness of the sequence $N_1N_2N_4$, when there is loss in all parts, is given by

$$\mu = \sum_{n=0}^{N_1}\binom{N_1}{n}\eta^n(1-\eta)^{N_1-n}\tilde{\mu}_n. \tag{2.36}$$

In the above equation, $\binom{N_1}{n}\eta^n(1-\eta)^{N_1-n}$ is the probability of n photons remaining out of N_1 single photons. If we performed the calculation in the obvious way, where there are three possible measurement results for each single photon, the number of measurement results to sum over for the single photons would be 3^{N_1}. By performing the calculation in this way, the number needed is $1 + 2 + \ldots + 2^{N_1} = 2^{N_1+1} - 1$, which is considerably less.

It is possible to consider arbitrary sequences of states, where the one-, two-, and four-photon states are used in any order. To simplify the range of possible sequences to search over, we grouped together states of the same photon number. A similar approach was used in Ref. [1] for the lossless case, where numerical testing with small total numbers of photons found that this was optimal. Note that due to the adaptive nature of the measurement scheme the order of one, two- and four-photon states is important. In this case, we found the optimal values of N_1, N_2 and N_4 in the $N_1N_2N_4$ sequence; i.e., single photons were followed by two- then four-photon loss-resistant states.

We performed numerical searches over possible combinations of numbers of states $N_1N_2N_4$ to find the sequences which give the least variance for a range of total photon numbers $\mathcal{N} = N_1 + 2N_2 + 4N_4$. The sequence configurations for some of the total numbers of photons are given in the table below (see Appendix A.1 for the extended table of results).

$\mathcal{N}$	N_1	N_2	$\chi^{(2)}$	N_4	$\chi^{(4)}$
9	7	1	1.7	0	-
13	7	1	1.7	1	1.3
30	2	2	1.8	6	1.3

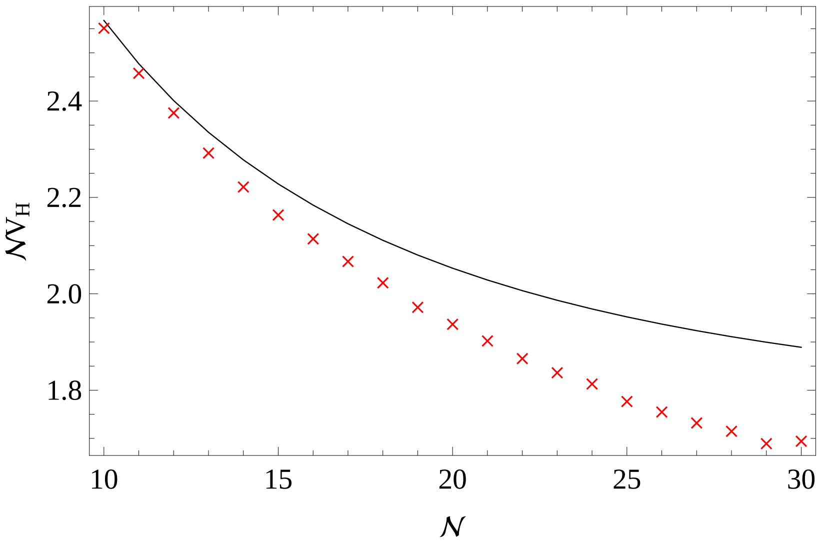


Figure 2.7: The scaled phase variance $\mathcal{N}V_H$ versus total number of photons $\mathcal{N}$ for $\eta = 0.6$ for: only single photon states (SQL) (solid line) and optimal sequence of single photon states combined with two- and four-photon loss-resistant states (shown by ×).

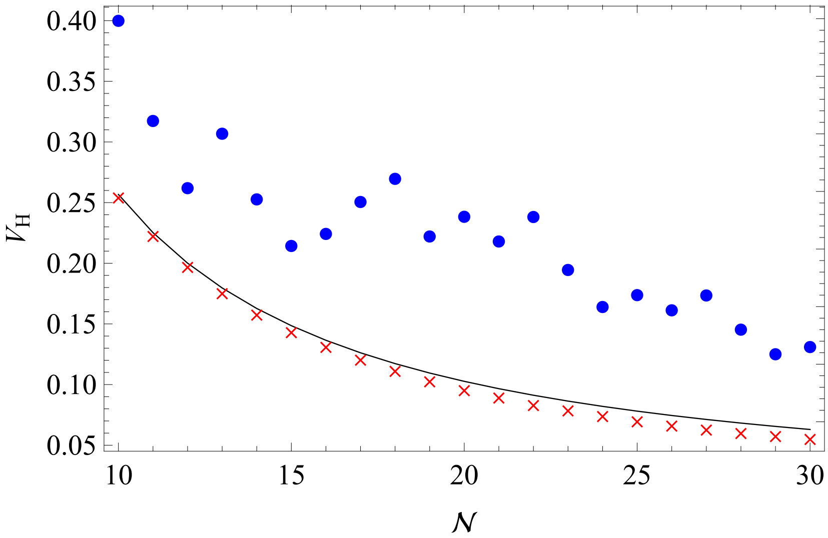


Figure 2.8: The phase variance V_H versus total number of photons $\mathcal{N}$ for $\eta = 0.6$ for three different input states. ×: optimal sequence of single photon states combined with two- and four-photon loss-resistant states. Solid line: only single photon states (SQL). ●: the scheme proposed in [1].

The results of the numerical optimization for $\eta = 0.6$ are plotted in Fig. 2.7. In this figure, the result for just single-photon states (with $\eta = 0.6$) is shown for comparison, and can be regarded as equivalent to the SQL. To show the difference between the optimal sequence of loss-resistant states and single photon states in this figure we have plotted the scaled phase variance $\mathcal{N}V_H$. Up to a total number of nine photons there is no advantage in using optimised multiphoton states, but after that the optimal sequence beats the SQL.

In Fig. 2.8 we have also shown the variance calculated using the states and sequences proposed in [1]. The two- and four-photon states proposed in [1] are obtained by setting $\chi = 0$ in Eqs. (2.18) and (2.19). These states give variances which are considerably larger than those obtained from single-photon states. This seems surprising, because using a sequence of states with different numbers of photons should be able to outperform a scheme limited to single-photon states. However, that requires choosing the values of N_1, N_2, and N_4 appropriately, and the order that the states are used. In this case, we have considered the scheme of [1] using the values of N_1, N_2, and N_4 that were chosen to minimise the variance *without* loss. It is clear that the optimal values of these quantities must be dependent on the loss; that is, the improvement over the scheme of [1] is primarily due to choosing the state sequences in such a way as to optimise the measurements for loss.

2.5 Conclusion

In this chapter we proposed an approach to generate multi-photon entangled states which are optimal for phase measurement in the presence of photon loss. In order to provide a technique that is experimentally feasible, we have considered methods of processing SPDC sources in order to provide improved loss tolerance. For two-photon states the method produced optimal states, but the technique is not able to exactly produce the optimal four-photon states. However, as shown in Fig. 2.5, the maximum of Fisher information for our proposed four-photon loss-resistant states is almost the same as the exact optimal states up to $\eta = 0.7$.

We proposed techniques of combining these loss-resistant states in order to provide an unambiguous measurement of the phase. Surprisingly, we find that the parameters that minimise the phase variance are not the same as those that maximise the Fisher information. This is likely because the lower bound set by the Fisher information is only achieved in the asymptotic case of large numbers of copies of the state, whereas we are considering a small number of copies.

By using the sequence of loss-resistant states, we showed it is possible to beat an SQL defined by the corresponding scheme with independent single photons. In comparison, if the measurement scheme of Ref. [1] is used, the phase variance is much greater. In order to obtain the best performance, the state sequence should be chosen based on the loss. Moreover optimizing the parameters for the loss-resistant states will provide an additional improvement.

In the next chapter we numerically find the optimal states of the form given in Eq. (2.21), for the task of spectroscopy of an ensemble of atoms in an interferometric system.

3

Quantum enhanced spectroscopy

3.1 Introduction

In Chapter 2 we explained that NOON states are very sensitive to loss, and there exist other quantum states that give better performance for phase measurement in the presence of absorption [76, 77, 79, 96]. Even with such states, the advantage over the standard quantum limit in phase estimation is reduced by loss. However, we can take advantage of the sensitivity of nonclassical properties of quantum states to absorption. The sensitivity of quantum coherence can be used efficiently to estimate absorption [97], and also estimate physical quantities that the absorption depends on.

In Ref. [98] a sub-shot-noise measurement of absorption is obtained using heralded single photons. In that work, a non-interferometric setup was used, therefore all the information is obtained from absorption, and the quantum enhancement results from sub-Poissonian statistics of single photons. According to the Kramers-Kronig relation, absorption is accompanied by a phase shift [99]. However, the information from the phase is only accessible if we take advantage of superposition and interference.

In Ref. [100] optimal states for simultaneous estimation of loss and phase are found. Such states are of the form $\sum_{k=0}^{N} \psi_k \left|N-k, k\right\rangle$ with a large weight on the loss mode to improve the estimation of loss. Here, we find the optimal states of similar form to estimate a parameter that both loss and phase depend on.

The system we are considering in this chapter is an ensemble of atoms. We are interested in measuring a transition frequency of the atoms. If this ensemble is probed by a beam of photons, the absorption of photons, and phase shift imposed on the probe, both depend on the transition frequency of atoms. We consider a Mach-Zehnder interferometer with the atomic ensemble placed in one of the arms of the interferometer. We optimise over the state in the arms of the interferometer and find the state from which we obtain maximum information about the atomic transition frequency.

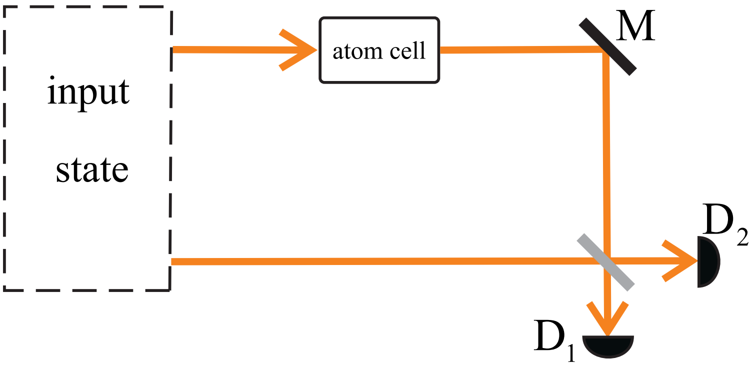


Figure 3.1: A Mach-Zehnder interferometer with an ensemble of atoms placed in the upper arm. D$_1$ and D$_2$ are photon number detectors in the output modes.

3.2 Interferometric scheme

Consider a Mach-Zehnder interferometer, shown in Fig. 3.1, with an ensemble of atoms placed in the upper arm of the interferometer. Here we consider an ensemble of identical two-level atoms in the absence of Doppler broadening and dipole dephasing. [1] This simple model gives a good qualitative description of the problem. Assuming that all atoms interact equally with the input quantum state and that there is no interaction between atoms, using the dipole and rotating-wave approximation the susceptibility of the ensemble is given by [101, 102]

$$\chi(\Delta) = \chi'(\Delta) + i\chi''(\Delta) = \frac{2\mathcal{N}_a\mu^2}{\hbar\varepsilon_0}\frac{\Delta + i\gamma_s}{\Delta^2 + \gamma_s^2}, \tag{3.1}$$

where $\Delta = \omega - \omega_0$ is the detuning between ω_0, the transition frequency of atoms, and ω, the frequency of input photons, γ_s is the spontaneous decay rate of the excited state, $\mathcal{N}_a$ is the number density of atoms, μ is the electric dipole moment, $\hbar$ is the reduced Planck constant and ε_0 is the vacuum permittivity. Details of the derivation of this susceptibility based on interaction of an ensemble of atoms with quantized light are given in Refs. [101, 102]. The imaginary and real parts of the susceptibility are plotted in Fig. 3.2. In this figure we have used data for the D1 transition line of Sodium from Ref. [103], i.e. $\mu = 0.704\times10^{-29}$ C$\cdot$ m and $\gamma_s = 61.354\times10^6$ s^{-1}. For the number density of atoms we have used $\mathcal{N}_a = 2.5\times10^{16}$ m^{-3}.

Knowing the susceptibility of the atomic medium, the effect of the atomic ensemble in the upper arm of the Mach-Zehnder interferometer can be modeled by the beam splitter model proposed in Refs. [104, 105]. Normally, one beam splitter is used to model loss in each of the arms of a Mach-Zehnder interferometer. However, here we consider a line of $n = L/\delta z$ beam splitters where each beam splitter represents one of the atoms in the ensemble (see Fig. 3.3). Here, L is the length of the ensemble, and δz is the spacing between two beam splitters. With the choice of a line of beam splitters the overall result will be the same as using a single beam splitter. However, with the line of beam splitters we obtain the transmissivity of the atom cell which needs to be conjectured if a single beam splitter was used. The kth beam splitter

---

[1] Doppler broadening and dipole dephasing are avoided to simplify the model. Doppler broadening means atoms traveling with different velocities with respect to the light beam have a modified detuning. In the presence of Doppler broadening each atom sees a different effect and the coherent effect washes away. This effect can be avoided using a cloud of trapped cold atoms. On the other hand, dipole dephasing can only change the coherence between two atomic levels. This adds an extra constant to the spontaneous decay rate γ_s in Eq. (3.1). This extra constant does not change the results.

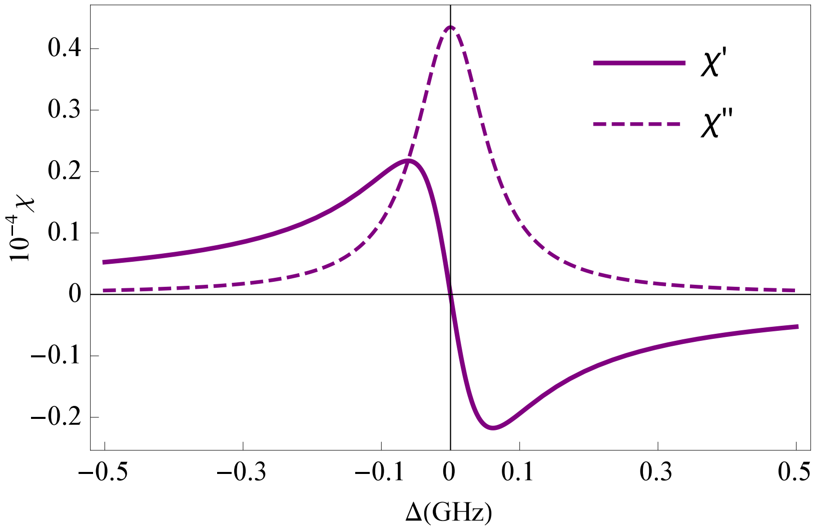


Figure 3.2: The real (solid line) and the imaginary (dashed line) parts of susceptibility, χ' and χ'' respectively, calculated using Eq. (3.1) for the D1 transition line of sodium.

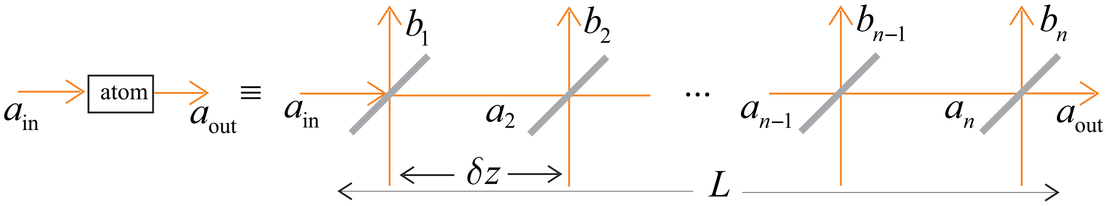


Figure 3.3: Beam splitter model to model the interaction of photons with the ensemble of atoms.

transforms the creation operator $\hat{a}_k^\dagger$ according to

$$\hat{a}_k^\dagger = \sqrt{t(\omega)}\hat{a}_{k+1}^\dagger + \sqrt{r(\omega)}\hat{b}_k^\dagger, \tag{3.2}$$

where $t(\omega)$ and $r(\omega)$ are the transmissivity and reflectivity of the beam splitter, ω is the frequency of input photons, and b_k is the loss mode of the kth beam splitter. After passing through n beam splitters, the creation operator of the input mode $\hat{a}_{\text{in}}^\dagger$ is transformed to

$$\hat{a}_{\text{in}}^\dagger = \left(\sqrt{t(\omega)}\right)^n \hat{a}_{\text{out}}^\dagger + \sqrt{r(\omega)} \sum_{k=1}^{n} \left(\sqrt{t(\omega)}\right)^{n-k} \hat{b}_k^\dagger. \tag{3.3}$$

The effect of the atomic ensemble is obtained by taking the limits

$$n = L/\delta z \to \infty, \qquad \delta z \to 0, \qquad r(\omega) \to 0, \tag{3.4}$$

such that the attenuation coefficient defined by $K(\omega) = |r(\omega)|/\delta z$ remains finite. We can write

$$|t(\omega)|^n = (1 - |r(\omega)|)^n = (1 - K(\omega)L/n)^n \to e^{-K(\omega)L}, \tag{3.5}$$

where the arrow is for the limit $n \to \infty$. Choosing the phase of the output creation operator, $\hat{a}_{\text{out}}^\dagger$, in Eq. (3.3) such that it produces the conventional propagation phase, $\exp\left(i\eta(\omega)\,\omega L/c\right)$, expressed in terms of refractive index $\eta(\omega)$, we have

$$\left(\sqrt{t(\omega)}\right)^n = e^{(i\eta(\omega)\omega/c - K(\omega)/2)L}. \tag{3.6}$$

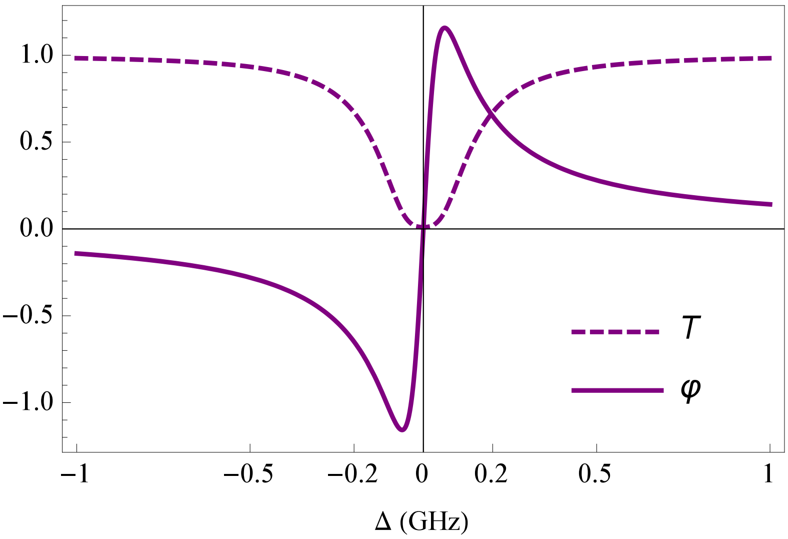


Figure 3.4: The transmissivity T (dashed line) and the phase shift φ (solid line) versus detuning Δ for the D1 transition line of sodium, $L = 1$ cm and $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3}.

The attenuation coefficient and the index of refraction are determined by the susceptibility of the medium according to [105]

$$\eta(\omega) + i\frac{c}{2\omega}K(\omega) = \sqrt{1+\chi} \approx 1 + \frac{1}{2}(\chi' + i\chi''), \tag{3.7}$$

where c is the speed of light, χ' and χ'' are the real and imaginary parts of the susceptibility. Here, we have omitted the dependence of χ on ω for brevity, and the approximation is due to χ being small. The real part of susceptibility, χ', describes dispersion and the imaginary part, χ'', describes absorption by the medium.

In the limits of Eq. (3.4), the summation is converted to integration, and the discrete loss modes $\hat{b}_k^\dagger$ are converted to continuous mode variable $\hat{b}^\dagger(z)$ according to

$$\sum_{k=1}^{n} \rightarrow \frac{1}{\delta z}\int_0^L dz, \qquad \hat{b}_k^\dagger \rightarrow \sqrt{\delta z}\hat{b}^\dagger(z). \tag{3.8}$$

Therefore, Eq. (3.3) can be written as

$$\hat{a}_{\text{in}}^\dagger = \hat{a}_{\text{out}}^\dagger e^{-i\frac{\omega L}{c}\sqrt{1+\chi}} - i\sqrt{\frac{\omega}{c}\chi''}\int_0^L e^{-i\frac{\omega}{c}(L-z)\sqrt{1+\chi}}\hat{b}^\dagger(z)dz. \tag{3.9}$$

According to Eqs. (3.3), (3.6) and (3.7), the transmissivity of the ensemble, T, and the phase shift imposed on the state from the ensemble, φ, can be written in terms of imaginary and real parts of the susceptibility

$$T = e^{-\chi''\omega L/c}, \qquad \varphi = -\frac{\chi'\omega L}{2c}. \tag{3.10}$$

The quantities T and φ are plotted in Fig. 3.4. This figure is plotted for $\omega_0 = 2\pi(508.33)$ THz, which is the D1 transition line of sodium [103], and $L = 1$ cm. As can be seen from Fig. 3.4, for detunings close to zero, φ has the highest slope. However, in this region, T is very small.

In the following section we find the optimal states to measure the transition frequency of the atoms i.e. Δ in this scheme.

3.3 Optimised states

We consider the general form of the state in the arms of the interferometer to be

$$|\psi\rangle = \sum_{k=0}^{N} \psi_k \, |N-k, k\rangle, \tag{3.11}$$

i.e. a pure state with the total photon number of N. We use Fisher information as the measure to quantify the metrological value of the states. According to the Cramér-Rao bound [28] the variance in estimating a parameter, Δ in this case, using an unbiased estimate, is lower bounded by the inverse of the Fisher information $F(\Delta)$

$$\mathrm{var}(\Delta) \geq 1/F(\Delta). \tag{3.12}$$

Here, we are considering photon number detection in the output modes, thus we are using classical rather than quantum Fisher information. The Fisher information represents the amount of information about Δ contained in the measurement results. It is given as

$$F(\Delta) = \sum_{n_1,n_2} \frac{1}{P_{n_1,n_2}(\Delta)} \left(\frac{\partial P_{n_1,n_2}(\Delta)}{\partial \Delta} \right)^2, \tag{3.13}$$

where $P_{n_1,n_2}(\Delta)$ is the probability of detecting n_1 and n_2 photons in the output ports.

Considering the state given in Eq. (3.11), acting the atom cell transformation given in Eq. (3.9) on the first mode, and the last 50/50 beam splitter of the interferometer on both modes, we obtain

$$\begin{aligned} P_{n_1,n_2}(\omega) = {} & \sum_{k=0}^{n_1+n_2} \sum_{k'=0}^{n_1+n_2} \sum_{u=n_2-k}^{n_2} \sum_{v=n_2-k'}^{n_2} \psi_k \psi_{k'}^{*} \frac{n_1! n_2! (N-n_1-n_2)!}{\sqrt{k! k'! (N-k)! (N-k')!}} \left(\frac{1}{2}\right)^{n_1+n_2} (-1)^{k-n_2} \\ & \times \binom{N-k'}{N-n_1-n_2} \binom{N-k}{N-n_1-n_2} \binom{n_1+n_2-k}{u} \binom{n_1+n_2-k'}{v} \binom{k}{k+u-n_2} \binom{k'}{k'+v-n_2} \\ & \times \left(1 - e^{-\omega L \chi''/c}\right)^{N-n_1-n_2} e^{i\omega\chi' L(k-k')/(2c)} e^{-\omega L \chi''(2n_1+2n_2-k-k')/(2c)}, \end{aligned} \tag{3.14}$$

where $\omega = \Delta + \omega_0$ is the frequency of the photons in the input state. Note that we have only considered the loss due to the atomic ensemble. More generally there might be additional loss in the system which we don't consider here. Because we are quantifying the metrological value of the states via the Fisher information, we regard the optimal states to be those which maximise the Fisher information. We have found the optimal values of ψ_k numerically using the particle swarm optimization (PSO) algorithm explaind in Appendix A.2. In this optimization problem a swarm of particles searches the space of ψ_k coefficients for those that maximize the Fisher information. Here, in our simulations we used $\chi = 0.729$, $c_l = c_g = 2.05$ with 10 particles and 100 iterations.

We have found that the optimal state for a specific type of atoms, only depends on the number density of atoms $\mathcal{N}_a$ or the length of the ensemble L. This can be explained in the following way. In Eqs. (3.10) and (3.14), we have dependence on the parameters of the system via $\omega\chi' L$ and $\omega\chi'' L$, which can be written as

$$\omega\chi L = \omega(\chi' + \chi'')L = \frac{2\mathcal{N}_a \mu^2 L \omega_0}{\hbar \varepsilon_0 \gamma_s} \frac{(\Delta/\gamma_s + i)\,(1 + \Delta/\gamma_s)}{1 + (\Delta/\gamma_s)^2}. \tag{3.15}$$

For a given type of atom the multiplying factor at the front can only be varied via $\mathcal{N}_a$ or L . The other parameters, μ, ω_0 and γ_s can be varied by changing the type of atom. These parameters affect the variation of $\omega\chi$ in three ways:

1. They change the multiplicative factor at the front. As that factor can also be changed by varying $\mathcal{N}_a$ or L, that does not give any qualitatively different results than simply changing $\mathcal{N}_a$ or L.

2. The parameter γ_s appears in the ratio Δ/γ_s, and therefore provides a scaling to the variation of $\omega\chi L$ with Δ. It therefore does not qualitatively change the results.

3. The parameter ω_0 appears in the factor $(1 + \Delta/\omega_0)$. This factor affects the variation very little, because we consider a parameter regime where $\Delta/\omega_0 \ll 1$.

In the following we take L constant at 1 cm and discuss the two cases: $\mathcal{N}_a > 10^{17}$ m^{-3} (large $\mathcal{N}_a$) and $\mathcal{N}_a < 10^{17}$ m^{-3} (small $\mathcal{N}_a$).

3.3.1 Large $\mathcal{N}_a$

For $\mathcal{N}_a > 10^{17}$ m^{-3}, we have found that numerically optimised states of the form given in Eq. (3.11) perform better than NOON states and independent single photons. In Fig. 3.5 we have compared the Fisher information of the N-photon optimal state, N independent single photon states $|1,0\rangle^{\otimes N}$, N copies of a single-photon NOON state $(|1,0\rangle + |0,1\rangle)/\sqrt{2}$, and an N-photon NOON state $(|N,0\rangle + |0,N\rangle)/\sqrt{2}$. This figure is plotted for $N = 2$ (upper row) and $N = 10$ (lower row). In this figure, we have used $\omega_0 = 2\pi(508.332)$ THz, which is the transition frequency of the D1 line of Sodium [103], and an atom density of $\mathcal{N}_a = 2.5 \times 10^{17}$m^{-3}.

Figure 3.5 shows that, even for $N = 2$, the enhancement obtained by optimal states is significant. For larger photon numbers, as is shown in the graphs for $N = 10$, there is no further significant improvement in the enhancement of the optimal states. Moreover, the optimal states with high photon numbers are not experimentally achievable with current technology. On the other hand, the optimal states for $N = 2$ can be generated with a scheme similar to the one proposed in Chapter 2.

Note that, close to resonance, for copies of single-photon NOON states the maximum peak is higher than independent single photons and N-photon NOON states. This is as would be expected, since single-photon NOON states are the least sensitive NOON states to loss. From Fig. 3.5(c), we see that ten-photon NOON states perform worse than independent single photons close to resonance. However, far from resonance, their Fisher information is even higher than the numerically obtained optimal states. The reason why this is possible is that the optimal states are only optimal in the sense of giving the largest maximum Fisher information, but it is possible for other states to have larger Fisher information for detunings where the optimal states do not give their maximum Fisher information. On the other hand, as can be seen in Fig. 3.5(a), two-photon NOON states are less sensitive to loss (compared to ten-photon NOON states), and close to resonance they perform better than independent single photons.

The other thing to note from Fig. 3.5 is that to be able to work in the region with maximum Fisher information we need to have prior knowledge of the detuning. This is because the peaks of maximum Fisher information are quite narrow. In other words this scheme could be used to measure hyperfine splitting of atomic levels, or measuring external effects, such as magnetic field, on the transition frequency of atoms.

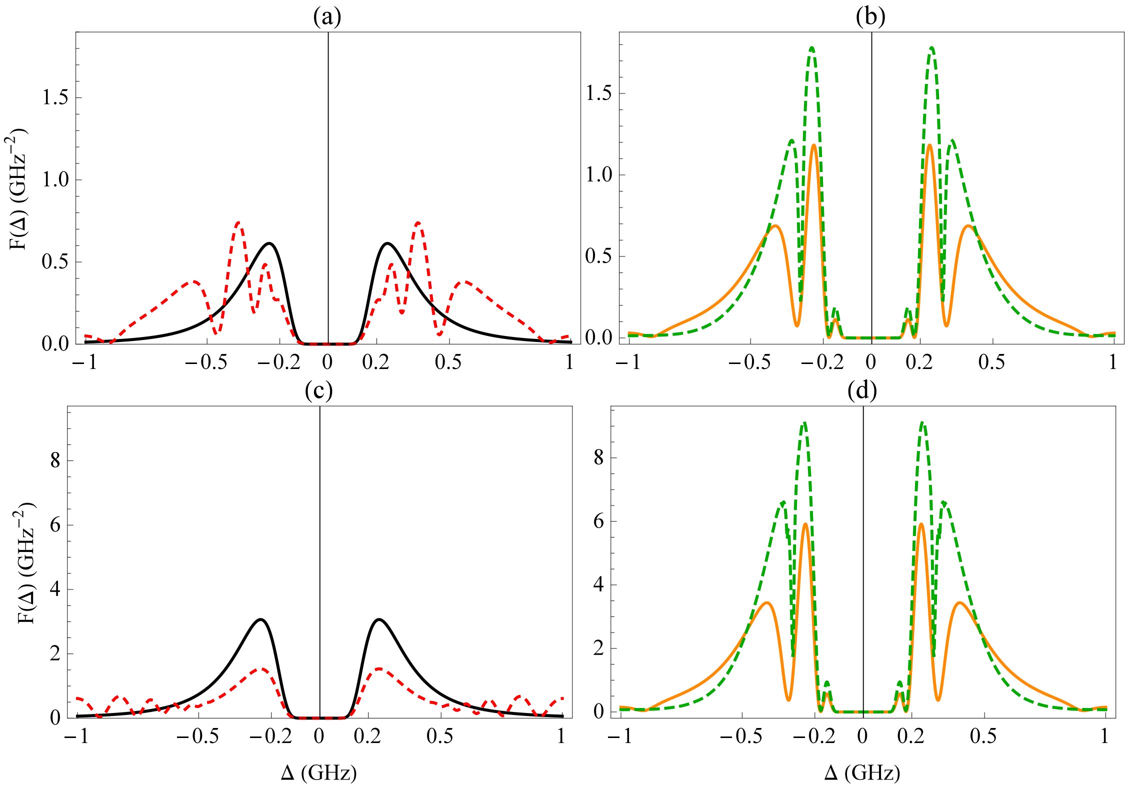


Figure 3.5: Fisher information $F(\Delta)$ versus detuning Δ for $N = 2$ photons [(a) and (b)] and $N = 10$ [(c) and (d)]. In (a) and (c), the solid black lines are for N independent single photons $|1, 0\rangle^{\otimes N}$, and the dashed red lines are for N-photon NOON states $(|N, 0\rangle + |0, N\rangle)/\sqrt{2}$. In (b) and (d) the dashed green lines are for N-photon optimal states, and the solid orange lines are for N copies of single-photon NOON states $(|1, 0\rangle + |0, 1\rangle)^{\otimes N}/\sqrt{2^N}$.

In Fig. 3.6 we have plotted the values of the coefficients of the optimal states, ψ_k in the superposition (3.11), for a range of photon numbers from $N = 2$ to $N = 10$. This figure shows

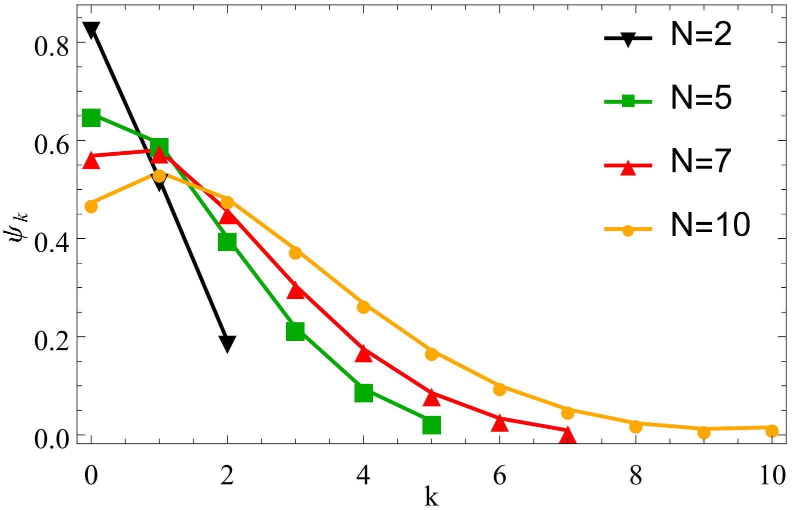


Figure 3.6: Coefficients ψ_k of the optimal states for number density of atoms $\mathcal{N} = 2.5 \times 10^{17}$ m^{-3} for four values of the total photon number N.

that the optimal states have higher amplitudes for the terms with higher photon numbers in the arm that contains the atomic ensemble. This variation is what would be expected because when there are more photons in the arm with the ensemble, they are more likely to be lost, giving more information about Δ.

3.3.2 Small $\mathcal{N}_a$

For smaller values of $\mathcal{N}_a$ than considered in the previous subsection, the range of the phase shift is small (see Fig. 3.7). In this case, the optimal state is N independent single photons, $|1,0\rangle^{\otimes N}$. Having all the photons in the upper arm, only the loss is being probed, and no information is being obtained from the phase shift. The phase shift must be significant so that we can take advantage of interferometric schemes in spectroscopy. Surprisingly for $\mathcal{N}_a = 2.5\times10^{16}\ \mathrm{m}^{-3}$ the maximum of the Fisher information for N independent single photons is even higher than the maximum of the Fisher information for the N-photon numerically optimised states with a larger number density of atoms which were considered in the previous subsection (see Fig. 3.8).

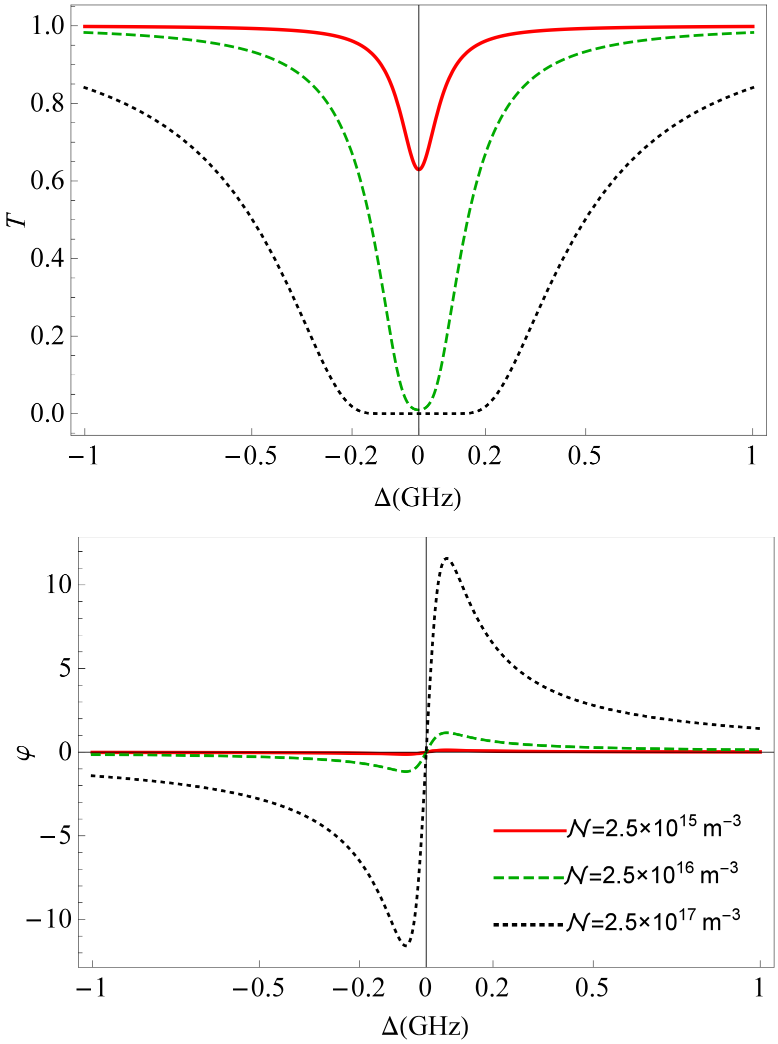


Figure 3.7: Transmissivity T and phase shift φ versus Δ, for a range of values of number density of atoms $\mathcal{N}$. Solid-red line: $\mathcal{N} = 2.5\times10^{15}\ \mathrm{m}^{-3}$. Dashed-green line: $\mathcal{N} = 2.5\times10^{16}\ \mathrm{m}^{-3}$. Dotted-black line: $\mathcal{N} = 2.5\times10^{17}\ \mathrm{m}^{-3}$.

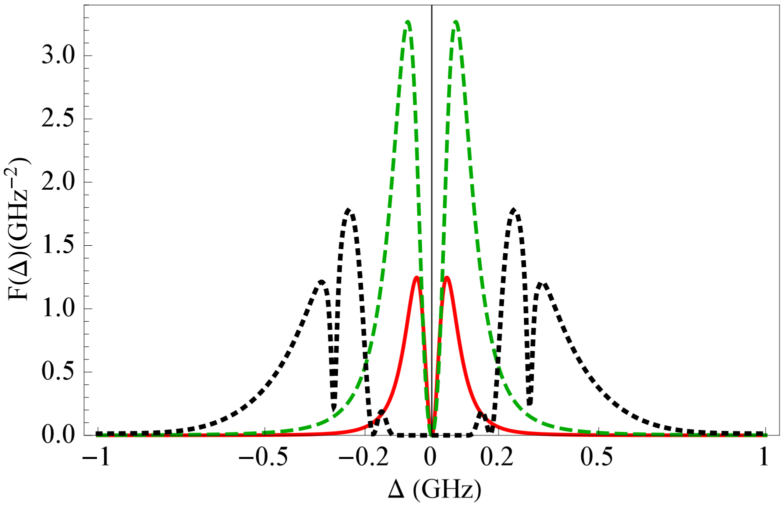


Figure 3.8: Fisher information versus Δ for total number of photons $N = 2$, for the optimal states of a range of number densities of atoms $\mathcal{N}$. Solid-red line: N independent single photons with $\mathcal{N} = 2.5 \times 10^{15}$ m^{-3}. Dashed-green line: N independent single photons with $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3}. Dotted-black line: numerically optimized state with $\mathcal{N} = 2.5 \times 10^{17}$ m^{-3}.

This could be understood from the variation of the transmissivity T and phase shift φ with $\mathcal{N}_a$, shown in Fig. 3.7. For smaller values of $\mathcal{N}_a$ the range of the phase shift is also smaller, which eliminates the advantage in using entangled states. In this case, the Fisher information is coming from the variation in the absorption. As $\mathcal{N}_a$ is decreased further, the dip in the absorption is reduced which results in a smaller Fisher information. For the higher densities, there is a larger phase shift, but it is in a region where the absorption is very high.

3.4 Conclusion

In this work we found optimal multi-photon states for measurement of the transition frequency of atoms. The scheme proposed here is an interferometric scheme with photon number detection in the output. In order to find the best states for measurement of the transition frequency, we numerically optimized for the states that provide the largest Fisher information.

For the number density of atoms we considered initially, the imposed phase on the probe is large, and it is advantageous to using information from both the absorption and the phase shift for measuring the transition frequency. In this case, the optimal state is an entangled multi-photon state. This optimal state has a large weighting on the state with all photons in the arm with the atomic ensemble. On the other hand, for a smaller number density of atoms, the phase shift imposed on the probe is small and therefore the information from the phase shift is not significant enough to give any advantage. In this case, the optimal state is independent single photons. In other words, it is advantageous to pass all the photons through the atom cell and obtain all the information from absorption.

Surprisingly there is a value of the number density, $\mathcal{N} = 2.5 \times 10^{16}$ m^{-3} for which N independent single photons have the highest Fisher information. This Fisher information is even higher than the Fisher information for the N-photon numerically optimized states with a larger number density of atoms. Therefore, if we have control over the number density of atoms it is better to choose this number density and probe the ensemble with independent single photons.

More generally, it would be possible to consider loss in both arms of the interferometer in addition to the atom cell. For different amount of loss the optimal states needs to be found which is left for future work.

4

Magnetometry with an NV centre

4.1 Introduction

In the past two chapters, we investigated phase estimation in optical interferometers. Another area of interferometry is Ramsey interferometry. Ramsey interferometry with atoms and solid-state systems has been used extensively in frequency measurement [58], and measurement of physical quantities that affect the frequency, such as magnetic field [59]. In solid-state systems, of particular interest is the electron spin of the nitrogen-vacancy (NV) centre, because it has long coherence time [106, 107] and it can be optically initialised and read out [108, 109]. The electron spin of NV centre has been used to measure a range of physical quantities such as electric field [110], magnetic field [55, 56, 111, 112] and temperature [113].

In this chapter, we give the theoretical background of the experiment done in Ref. [61]. I contributed to the numerical and theoretical part of this collaborative work with experimentalists. In this experiment, a single electron spin in diamond is used to measure a time-independent magnetic field using single-shot readout and real-time feedback in the system. We first review Ramsey interferometry with an NV centre. We then use a non-adaptive and a range of adaptive phase estimation algorithms to get an unambiguous estimate of an unknown magnetic field. We show that, although non-adaptive protocols reach the Heisenberg-like scaling, there is an optimised adaptive protocol that outperforms the non-adaptive protocol.

4.2 Ramsey interferometry

The nitrogen-vacancy (NV) centre is a defect in the diamond lattice that consists of a substitutional nitrogen atom and a vacancy at its adjacent lattice position (see Fig. 4.1(a)). The experiment in Ref. [61] is performed with a negatively charged NV centre, denoted by NV$^-$. The electronic energy level structure of NV$^-$ has a spin triplet ($S = 1$) ground ($^3A_2$) and excited states (^3E), shown in Fig. 4.1(b) [114]. The labeling of the energy levels is based on the standard notation for $C_{3\mathrm{V}}$ group symmetry operations [115]. In the ground state, the $m_s = \pm 1$ is lifted from the $m_s = 0$ state by 2.87 GHz due to spin-spin interaction. In the

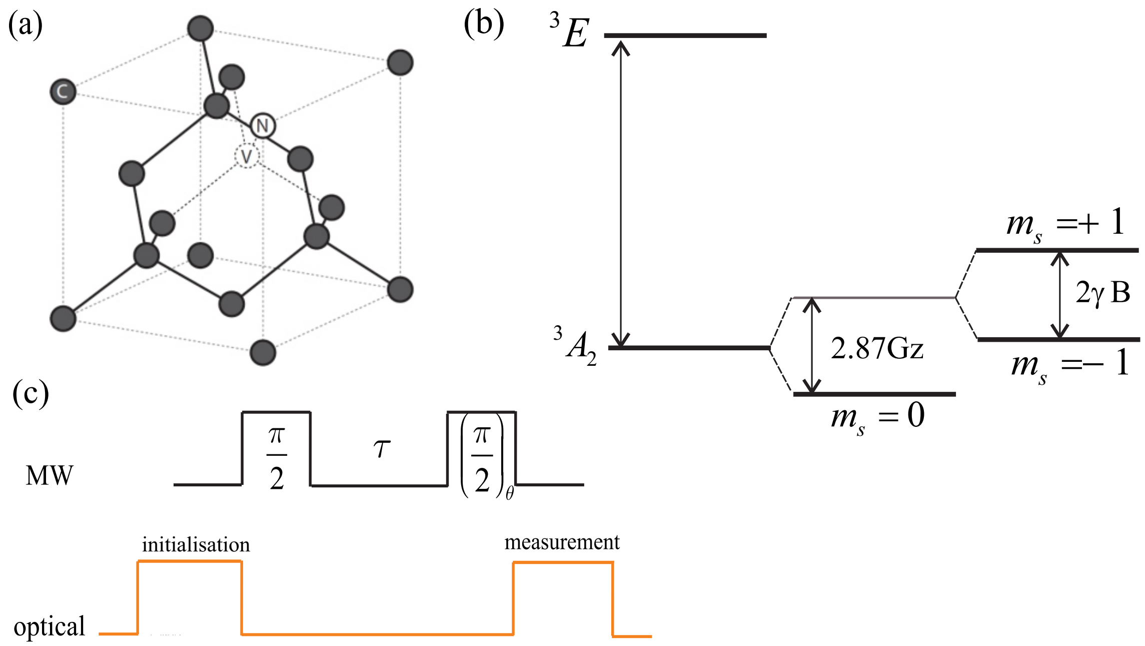


Figure 4.1: (a) An NV centre is formed by a substitutional nitrogen atom with a vacancy in its adjacent lattice position in a diamond lattice. This figure is adapted from Ref. [2] (b) The energy level structure of a negatively charged NV centre. (c) The pulse sequence used to estimate the magnetic field. The black line represents the microwave (MW) field used for coherent control of the spin. The phase θ of the second $\pi/2$ pulse is the controlled phase which determines the measurement basis. The orange line represents the optical field used for initialisation and measurement.

presence of an external magnetic field $\vec{B}$ the $m_s = \pm 1$ state splits due to the Zeeman effect. The Hamiltonian of the interaction with the magnetic field $\vec{B}$ is $\mathcal{H}_{\text{int}} = \gamma \vec{B}.\vec{S}$, where $\vec{S}$ is the spin 1 operator, and $\gamma = 28$ MHz$\cdot$ mT^{-1} is the gyromagnetic ratio.

In Ramsey interferometry, first a $\pi/2$ microwave (MW) pulse resonant with $|m_s = 0\rangle$ and $|m_s = -1\rangle$, denoted by $|0\rangle$ and $|1\rangle$, respectively, prepares the superposition state $(|0\rangle + |1\rangle)/\sqrt{2}$. This state then evolves under the Hamiltonian $\mathcal{H}_{\text{int}}$ for time τ to $\big(|0\rangle + e^{-i\varphi}|1\rangle\big)/\sqrt{2}$, where $\varphi = \gamma B \tau$. The phase φ can be estimated by measuring the spin in a suitable basis, by adjusting the phase θ of a second $\pi/2$ pulse. For magnetic fields past a certain value, the phase will wrap around and give the same measurement results, so magnetic fields past a certain value cannot be distinguished unambiguously. To have φ in the range of $(-\pi, \pi]$, the magnetic field must be in the range $(-B_{\text{max}}, B_{\text{max}}]$, where $B_{\text{max}} = \pi/(\gamma\tau)$.

In a Ramsey interferometer, if all the measurements have the same interaction time τ, the lower bound of the uncertainty in the phase estimate, for the total interaction time $T = N\tau$, scales as $\sigma_\varphi \sim \sqrt{\tau/T}$ [112]. This corresponds to the uncertainty of the field scaling as

$$\sigma_B \sim \frac{1}{\gamma\sqrt{\tau T}} \tag{4.1}$$

which we call standard measurement sensitivity (SMS). The uncertainty can be decreased by increasing τ, but this also reduces the magnetic field range. The dynamic range, defined as the ratio of the maximum detectable magnetic field to the uncertainty in the field, is then bounded as

$$\frac{B_{\text{max}}}{\sigma_B} \leq \pi\sqrt{T/\tau}. \tag{4.2}$$

However, using multiple interaction times in an estimation sequence, with the minimum interaction time τ and the total interaction time T, the lower bound on the phase uncertainty scales as $1/T$ [54], which corresponds to an enhancement in the dynamic range by a square factor, i.e.

$$\frac{B_{\max}}{\sigma_B} \lesssim T/\tau. \tag{4.3}$$

Note that there still needs to be some minimum interaction time τ, because that is what limits $B_{\max}$. The enhancement is analogous to the square enhancement obtained by using NOON states in optical interferometers [60]. We will call this scaling "Heisenberg-like", because it is analogous, but distinct from the enhancement obtained with multiple entangled systems.

The ultimate goal in magnetometry is to obtain the highest dynamic range, i.e. the highest precision over the largest range. This is a phase estimation problem for which the methods of optical interferometers can be used. A significant difference between measurements in optical and Ramsey interferometers is that Ramsey measurements have lower visibility. The initial visibility is low and it also reduces exponentially with interaction time. Prior to this work, there was speculation that adaptive schemes can outperform non-adaptive schemes when the visibility is low; simpler adaptive schemes did not outperform the non-adaptive measurements [54, 57]. This is surprising since typically adaptive schemes give better measurements, and in many cases far better measurements [26]. A sufficiently general adaptive scheme should always be able to match a non-adaptive scheme, because the non-adaptive scheme is just a special case. In the following section, we show that there is an optimised adaptive protocol that outperforms the non-adaptive protocol even for low visibilities.

4.3 Phase estimation

We convert the task of estimating the unknown time-independent magnetic field B to the task of estimating the unknown phase φ, imprinted on the spin state, due to the interaction with the magnetic field. Similar to Chapter 2, we use the terminology “detection” for individual measurements as opposed to the overall measurement combining the results of the sequence.

We use the generalized quantum phase estimation algorithm, explained in Section 1.3, to estimate the system phase φ. We perform a sequence of detections that consists of $K+1$ different interaction times with the shortest interaction time being τ. We start with the longest possible interaction time $t = 2^K\tau$, and perform a sequence of measurements with interaction times reduced by successive factors of 2, i.e. the sequence of interaction times is $2^K\tau, 2^{K-1}\tau, \ldots, 2^0\tau$ [54]. For each interaction time we perform some number of repetitions. More specifically, if the longest interaction time is $2^K\tau$, we perform $M(K,k)$ repetitions for the interaction time $t_k = 2^k\tau$ where

$$M(K,k) = G + F(K-k). \tag{4.4}$$

Here, G is the number of detections for the longest interaction time, $2^K\tau$. The number of detections increases by F when the interaction time is divided by 2. The motivation for performing the largest number of detections for the shortest interaction time is that the shortest interaction time makes the largest distinction in phase and therefore errors are most detrimental. Using a larger number of repetitions for shorter interaction times provides a larger reduction in the error for less cost.

We update the probability of the system phase given the detection results according to Bayes' rule:

$$P(\varphi|\vec{u}_\ell) \propto P(u_\ell|\varphi)P(\varphi|\vec{u}_{\ell-1}). \tag{4.5}$$

Here, $\varphi = \gamma B \tau$, $\vec{u}_\ell$ is the vector of ℓ measurement results u_1, u_2, ..., u_ℓ where $u_\ell = 0, 1$ represents measuring $|m_s = 0\rangle$ and $|m_s = -1\rangle$ in the ℓth detection, respectively. The proportional to symbol is due to omitting a term in Bayes' rule that does not depend on φ. Because it does not depend on φ it only gives a normalisation to the probability distribution, which can be calculated at the end. $P(u_\ell|\varphi)$ is the conditional probability of the measurement result u_ℓ given the system phase φ, and is given by [116]

$$\begin{aligned} P(u_\ell = 0|\varphi) &= \frac{1 + f_0 - f_1}{2} + \frac{f_0 + f_1 - 1}{2} e^{-(2^k \tau / T_2^*)^2} \cos\left(2^k \varphi - \theta\right), \\ P(u_\ell = 1|\varphi) &= 1 - P(u_\ell = 0|\varphi). \end{aligned} \tag{4.6}$$

Here, θ is the controlled phase, T_2^* is the decay rate of the spin coherence, τ is the shortest interaction time, and f_0 and f_1 are the maximum probabilities of getting $u_\ell = 0$ and $u_\ell = 1$, respectively. The Gaussian decay factor $e^{-(2^k \tau / T_2^*)^2}$ accounts for the bath-induced dephasing [116].

We assume the only knowledge we have about the system phase φ is that it is confined in the interval $(-\pi, \pi]$, thus $P(\varphi|u_0) = 1/(2\pi)$. In other words, we assume that the initial probability distribution for the magnetic field is flat over the possible range of values, which then translates to the probability distribution for the phase. It is worth mentioning that the flat distribution is not as natural for magnetic fields as it is for phase measurements. Different initial distributions might be more appropriate for specific applications, but in the absence of another proposed distribution we use the flat distribution to represent minimal initial information.

The total measurement interaction time is

$$T = \sum_{k=0}^{K} 2^k \tau M_k = \tau \left[G\left(2^{K+1} - 1\right) + F\left(2^{K+1} - 2 - K\right) \right]. \tag{4.7}$$

We quantify the resources by the total interaction time. In our simulations we used $\tau = 20$ ns which is the value used in the experiment. We use the Holevo variance to estimate the accuracy of the phase estimate

$$V_{\mathrm{H}} = \left|\left\langle e^{i\breve{\varphi}} \right\rangle\right|^{-2} - 1, \tag{4.8}$$

where $\breve{\varphi}$ is an unbiased estimate of the phase. The Holevo variance can be calculated efficiently if we represent the probability distribution for the phase in a compact form using a Fourier series;

$$P(u_\ell|\varphi) = \sum_j a_j e^{ij\varphi}. \tag{4.9}$$

Because Eq. (4.6) can be written with three Fourier terms, Bayes' rule gives the following updating rule for the Fourier coefficients

$$\begin{aligned} a_j^{(\ell)} = & \frac{1 + (-1)^{u_\ell}(f_0 - f_1)}{2} a_j^{(\ell-1)} \\ & + \frac{f_0 + f_1 - 1}{4} e^{-(2^k \tau / T_2^*)^2} \left(e^{i(u_l \pi + \theta_l)} a_{j-2^k}^{(\ell-1)} + e^{-i(u_\ell \pi + \theta_\ell)} a_{j+2^k}^{(\ell-1)} \right). \end{aligned} \tag{4.10}$$

In the following, we describe a range of measurement protocols. We first start with a non-adaptive protocol in which the controlled phase is incremented independent of the detection results. This is followed by adaptive protocols in which the controlled phase is determined based on previous detection results.

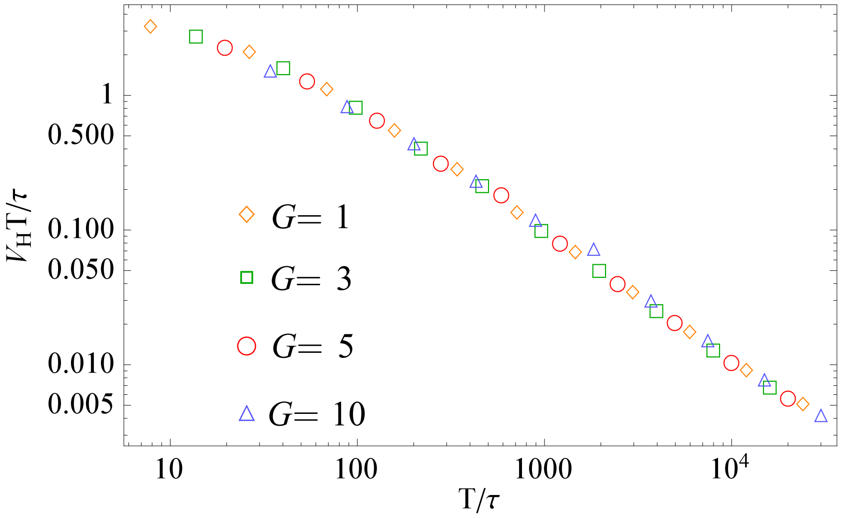


Figure 4.2: The scaled variance for a non-adaptive protocol where the number of repetitions for each interaction time varies as in Eq. (4.4), with $F = 5$ for a range of values of G.

4.3.1 Non-adaptive protocol

In the non-adaptive protocol [54], the controlled phase is not updated based on previous detection results, but is swept between 0 and π for detections with the same interaction time, according to predefined values. If we perform M_k detections for the interaction time t_k, the controlled phase is incremented by a step of π/M_k after each detection. In other words, the mth controlled phase θ_m for the interaction time t_k is set to

$$\theta_m = \frac{m\pi}{M_k}, \qquad m = 1, \ldots, M_k. \tag{4.11}$$

This protocol was experimentally demonstrated in Refs. [55] and [56].

We seek to find the most accurate non-adaptive protocol. We will then compare the performance of this optimised non-adaptive protocol with adaptive protocols. To find the optimised non-adaptive protocol we find the values of G and F that are optimal, in that they yield the smallest variance. By extensive numerical search over a range of G from 1 to 20 and range of F from 0 to 20, we have found that the best value of F is 5. For this value of F, the scaled variance does not change significantly for different values of G. This is shown in Fig. 4.2 in which the scaled variance $V_{\mathrm{H}}T/\tau$ is plotted versus T/τ for $F = 5$ and a range of values of G. In other words, there is no optimal value of G, so it is convenient to take $G = 1$ to minimise the number of detections and therefore the interaction time.

In Fig. 4.3 the scaled variance $V_{\mathrm{H}}T/\tau$ is shown for $G = 1$ (top) and $G = 5$ (bottom) for a range of F from 0 to 5. Note that, once the longest interaction time $2^K\tau$ reaches T_2^* ($T_2^* = 96$ μs) there is no further improvement in sensitivity, and the scaled variance starts increasing.

4.3.2 Adaptive protocols

For adaptive protocols we consider three methods to update the controlled phase: limited-adaptive, full-adaptive, and optimised-adaptive protocols. The details of each protocol are given below.

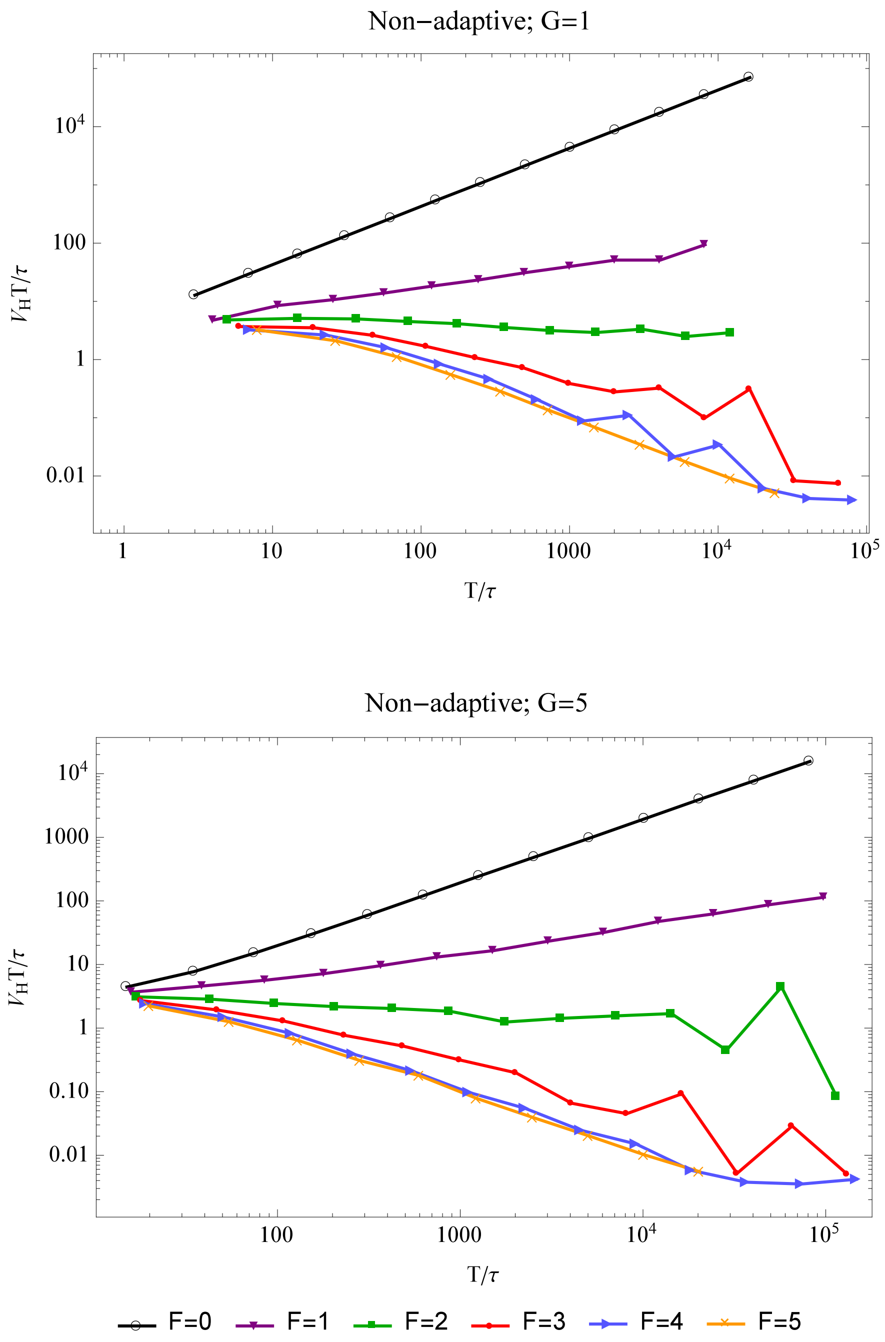


Figure 4.3: The scaled variance, $V_H T/\tau$, versus T/τ for the non-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F. Here we have used the experimental parameters $f_0 = 0.88$, $f_1 = 0.993$ and $T_2^* = 96\mu$s.

Limited-adaptive protocol

In this protocol, proposed by Cappellaro [117], the controlled phase is updated only when the interaction time is changed. In other words, when the interaction time is changed from t_k to t_{k-1}, the controlled phase is updated by

$$\theta = \frac{1}{2} \arg\left(a_{-2^k}\right). \tag{4.12}$$

We set the controlled phases in the interaction time $2^K\tau$ to zero. When the interaction time is changed from $2^K\tau$ to $2^{K-1}\tau$ we calculate the controlled phase using the above equation with $k = K$ and use this phase for all the M_{K-1} detections. We do not recalculate θ for every single detection in this interaction time, the same phase as the first detection is being used. In a similar manner we calculate all the other controlled phases for different interaction times. This controlled phase maximises the sharpness after the first detection with a given interaction time. This is similar to the strategy we used in Chapter 2.

In Fig. 4.4 we have plotted the simulations for this protocol for $G = 1$ (top) and $G = 5$

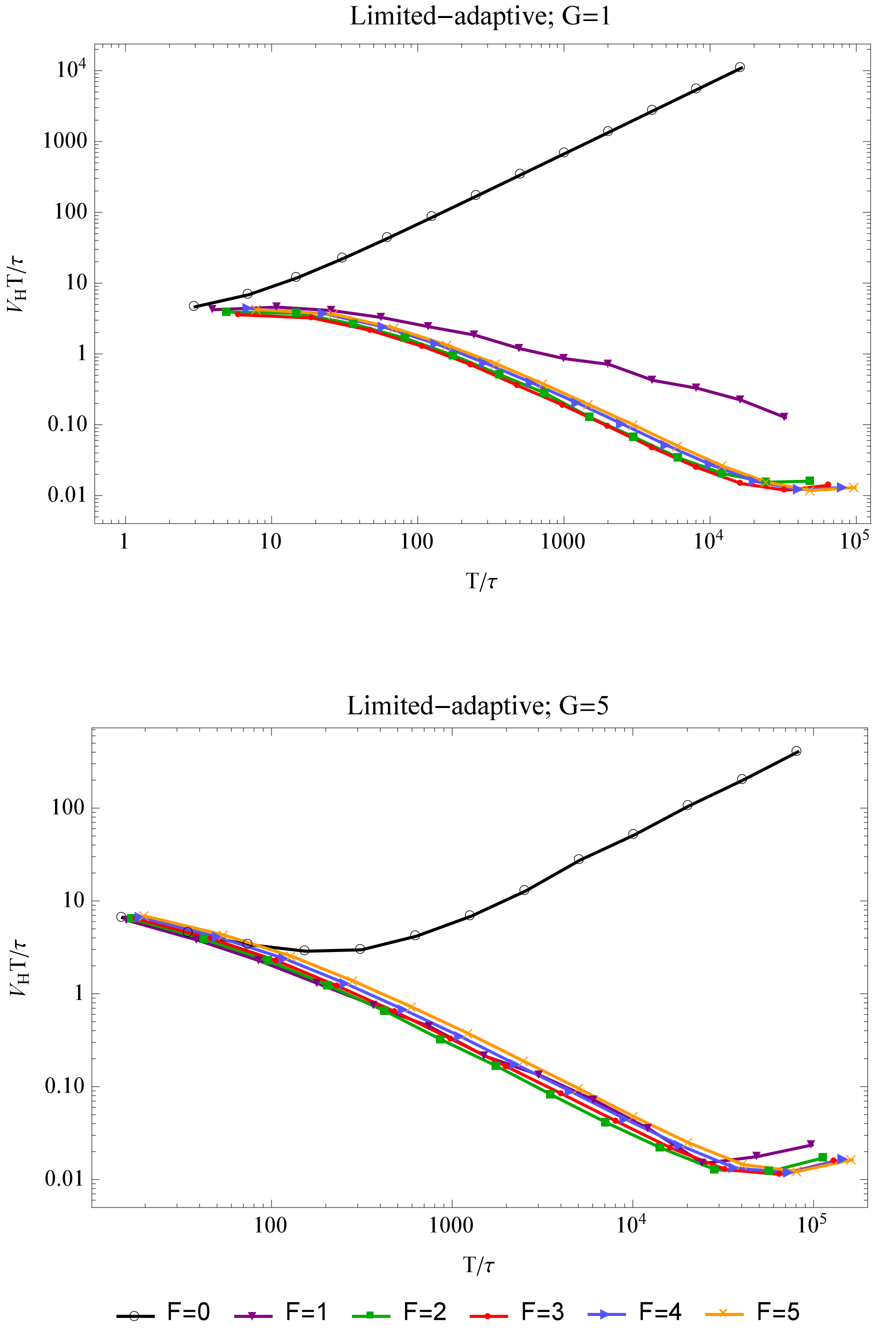


Figure 4.4: The scaled variance, $V_{\mathrm{H}}T/\tau$, versus T/τ for the limited-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F. This figure is plotted for $f_0 = 0.88$, $f_1 = 0.993$ and $T_2^* = 96\mu$s.

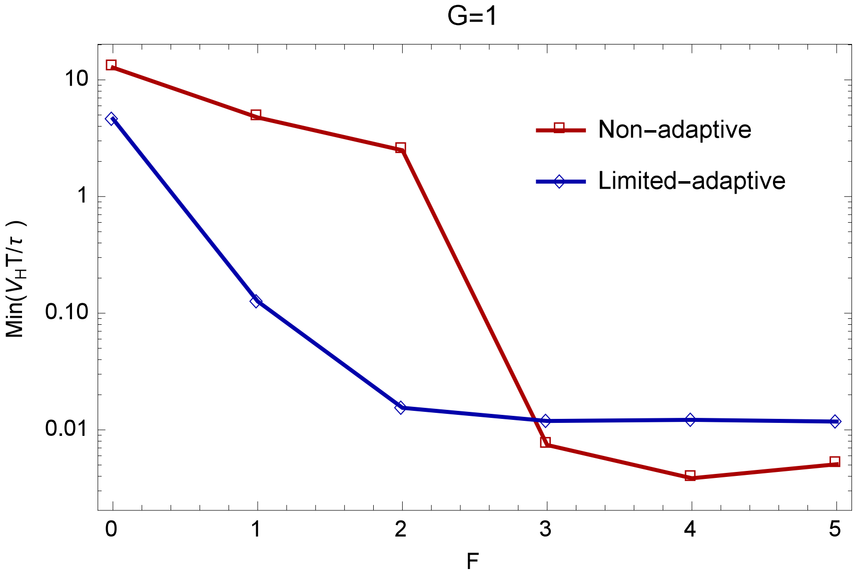


Figure 4.5: Minimum of the scaled variance versus F for two protocols; limited-adaptive and non-adaptive.

(bottom) for a range of values of F. Compared to the non-adaptive protocol, the limited-adaptive protocol reaches the Heisenberg-like scaling for $F = 2$ (if $G = 1$) and $F = 1$ (if $G = 5$). It is easier to compare two protocols if we plot the minimum of the scaled variance versus F. This is shown in Fig. 4.5. It can be seen that up to $F = 3$ the limited-adaptive protocol outperforms the non-adaptive protocol. However, for $F > 3$ as is shown in Fig. 4.3, the non-adaptive protocol reaches the Heisenberg-like scaling and it gives lower variance than the limited adaptive protocol.

The number of detections for a sequence of $K + 1$ interaction times, i.e. the sequence $2^K\tau, 2^{K-1}\tau, ..., \tau$, is

$$\sum_{k=0}^{K} [G + F(K - k)] = \frac{1}{2}(K + 1)(KF + 2G)\,. \tag{4.13}$$

Each detection corresponds to a Ramsey measurement. Each Ramsey measurement requires time for the intialisation and detection, in addition to the interaction time. In the experiment in Ref. [61], the initialisation time was 40 μs and the measurement time was 200μs. We call this extra 240 μs time for each detection “overhead time”. The total overhead time in μs is therefore 240 times the total number of detections.

$$T_{\text{overhead}} = 240 \times \frac{1}{2}(K + 1)(KF + 2G)\,. \tag{4.14}$$

In all the plots, except in Fig. 4.10, we have only included the interaction time, and haven’t considered the overhead time. In practice, it is important to minimise the total time of the sequence, including the overhead time. Therefore, the sequence that reaches the Heisenberg-like scaling with smaller F is a better sequence. Moreover, non-adaptive protocols can always be obtained as a special case of a sufficiently general adaptive protocol. Because the limited-adaptive protocol can be outperformed by the non-adaptive protocol, there must be other adaptive protocols with better performance.

Full-adaptive protocol

In this case the controlled phase is not only updated when the interaction time is changed, but also updated after each detection for a given interaction time $t_k = 2^k\tau$. In this protocol, after each detection with interaction time t_k, the phase is updated by

$$\theta = \frac{1}{2}\arg\left(a_{-2^{k+1}}\right). \tag{4.15}$$

This is essentially continuing to use the same formula for each detection with interaction time t_k as for the first detection after the interaction time is changed. Note that, in this protocol θ is recalculated after each detection, which is distinct from the limited-adaptive protocol in which θ is calculated for the first detection with interaction time t_k, and kept fixed for all the remaining detections with this interaction time.

This controlled phase is basically the best estimate of the phase modulo $\pi/2^k$ instead of $\pi/2^{k-1}$. This is surprising, because normally in adaptive schemes the best estimate is used to approximate a homodyne measurement. The reason is that a π phase ambiguity does not matter for the controlled phase θ. In other words, adding π to the controlled phase only exchanges the detection results $u_\ell = +1$ and $u_\ell = 0$. If we add π to the controlled phase and we get $u_\ell = 0$ as the detection result, all the following calculations are the same as if we hadn't added π and the detection result was $u_\ell = +1$. This means that the final phase variance does not change. This choice of controlled phase is different from the scheme we used in chapter 2. There, the controlled phase was chosen to maximise the average sharpness after the next detection result. We have seen through numerical simulations that such a scheme does not work well here.

Figure 4.6 shows the scaled variance obtained by this protocol for $G = 1$ and $G = 5$ for a range of values of F. The improvement achieved by the full-adaptive protocol over the limited adaptive protocol is shown in Fig. 4.7. However, as can be seen in this figure, the non-adaptive protocol for $F = 5$ still has better performance. The minimum of the scaled phase variance for the non-, limited- and full-adaptive protocols is plotted in Fig. 4.8. This figure shows the enhancement obtained by updating the phase after each detection. Moreover, it shows that the non-adaptive protocol performs better than even the full adaptive protocol for values of $F > 3$. We will now show that there is an optimised adaptive protocol which performs better than the non-adaptive protocol for all values of F.

Optimised adaptive protocol

In this protocol, we add phase increments θ^{incr} to the controlled phases based on the full-adaptive protocol. We found that this gives better results than adding the increments to the limited-adaptive or non-adaptive protocols. The increments in the phase are obtained by numerically minimising the final phase variance through the particle swarm optimization (PSO) algorithm [60], explained in Appendix A.2. The search space of the PSO algorithm for this case is the phase increments after each detection. Each increment only depends on the last detection result and the number of previous detections. This is not the most general possible scheme. The most general possible is prohibitively difficult to calculate because the number of possible increments is exponential in the number of detections. In our simulations for particle swarm optimization, we used 10 particles and 400 iterations. For the constants we used $c_g = c_l = 2.05$ and $\chi = 0.729$.

Figure 4.9 shows that the optimised adaptive protocol for $G = 1$, $F = 2$ performs as well as the best non-adaptive protocol ($G = 1$, $F = 5$). Moreover, the optimised adaptive protocol

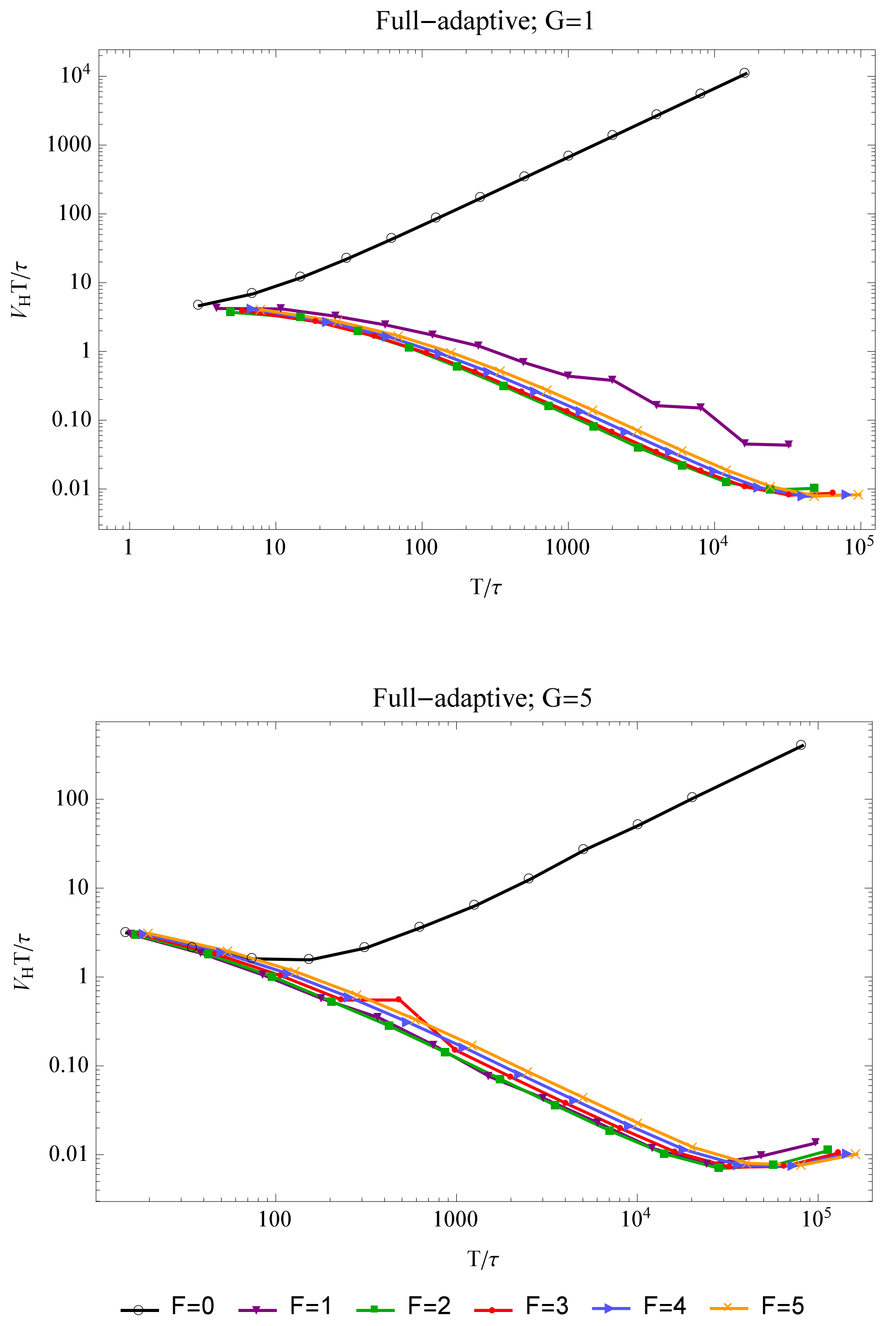


Figure 4.6: The scaled variance, $V_H T/\tau$, versus T/τ for the full-adaptive protocol for $G = 1$ (top) and $G = 5$ (bottom) for a range of values of F.

for $G = 1$, $F = 3$, and also $G = 5$, $F = 2$ outperforms the optimised non-adaptive protocol. If we consider the overhead time, the enhancement obtained by the optimised adaptive protocol is significant. This is shown in Fig. 4.10. As would be expected the optimised adaptive protocol for $G = 1$, $F = 1$ has the best performance when overhead time is considered. As can be seen in Fig. 4.9, $F = 1$ gives Heisenberg-like scaling. Moreover, because it has the smallest number of detections, it has the lowest overhead time.

To compare all the protocols we have shown the minimum of the scaled phase variance for the non-, limited-, full-, and optimised adaptive protocols in Fig. 4.11. This figure shows that the optimised adaptive protocol outperforms the non-adaptive protocol for all values of F.

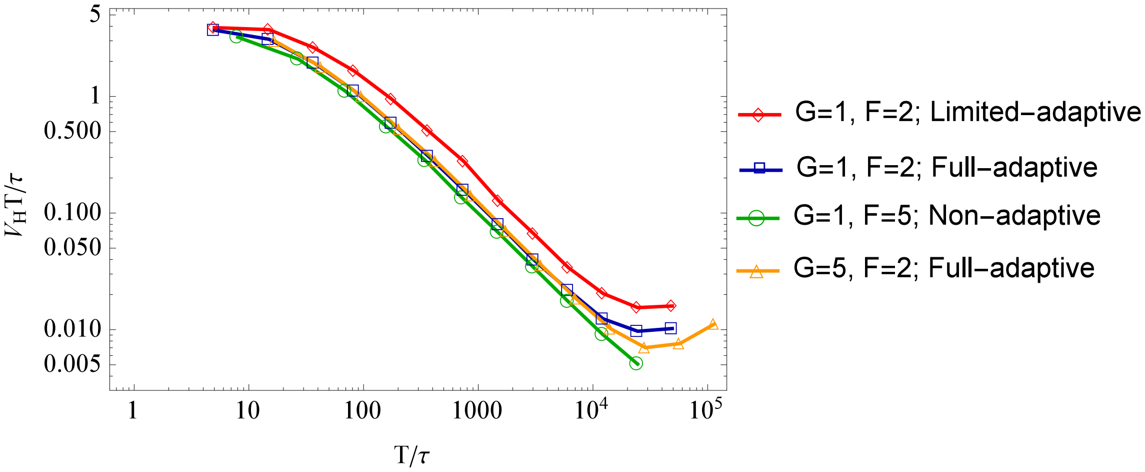


Figure 4.7: Scaled variance of the optimised non-adaptive protocol ($G = 1$, $F = 5$) compared to the best limited- and full-adaptive protocols.

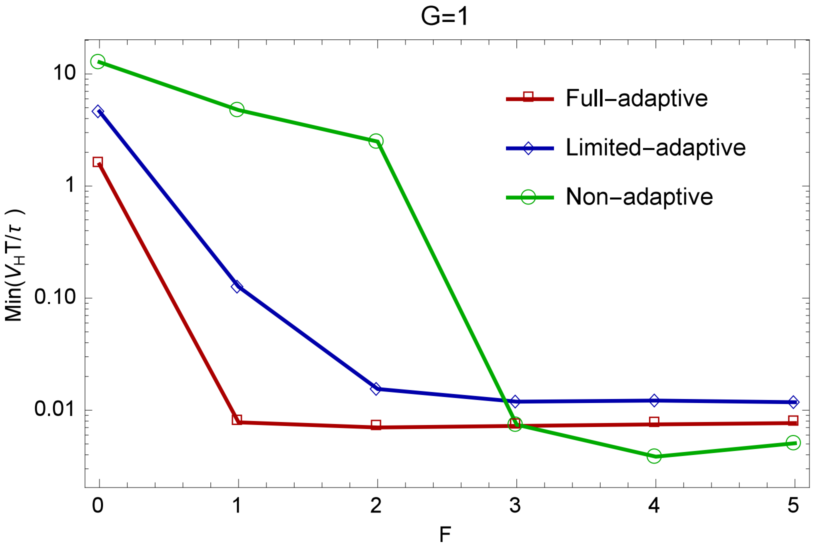


Figure 4.8: Minimum of the scaled variance versus F for full-, limited- and non-adaptive protocols.

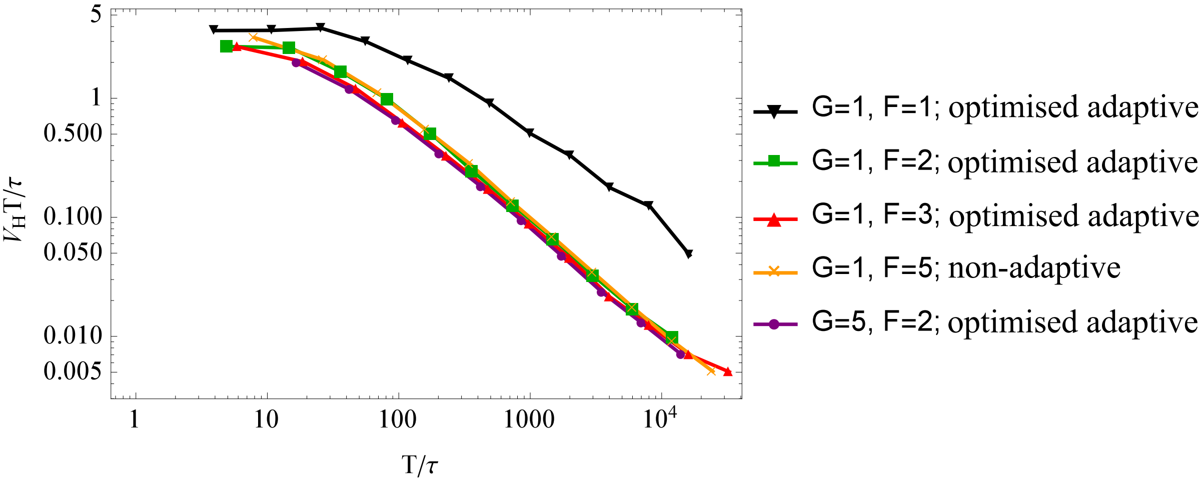


Figure 4.9: The scaled variance $V_H T/\tau$ versus T/τ_0 comparing the optimised adaptive protocol with a range of values of G and F with the optimised non-adaptive protocol $G = 1$, $F = 5$.

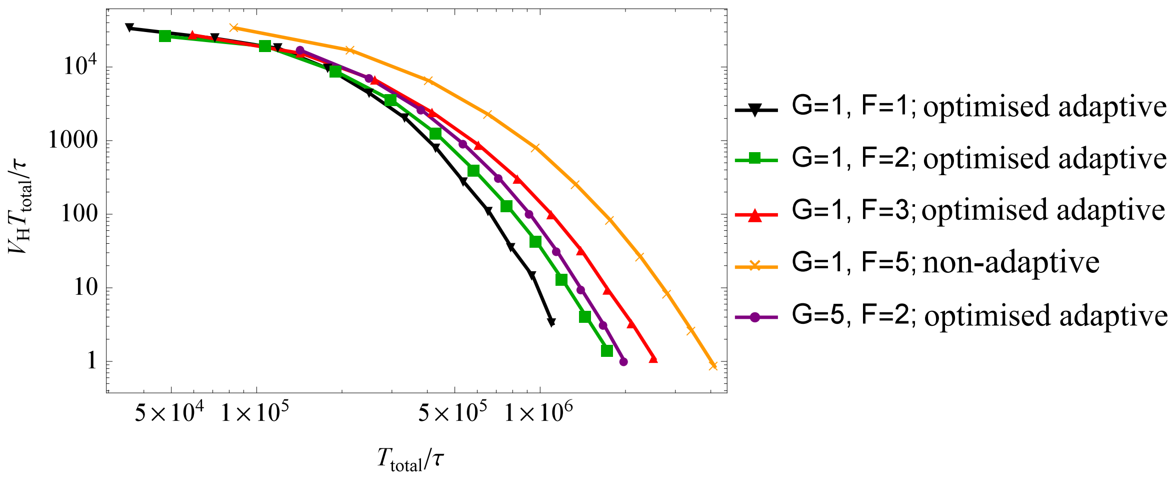


Figure 4.10: The scaled variance including the overhead time; $V_{\mathrm{H}}T_{\mathrm{total}}/\tau$ versus T_{total}/τ where $T_{\mathrm{total}} = T + T_{\mathrm{overhead}}$ with T_{overhead} given in Eq. 4.14.

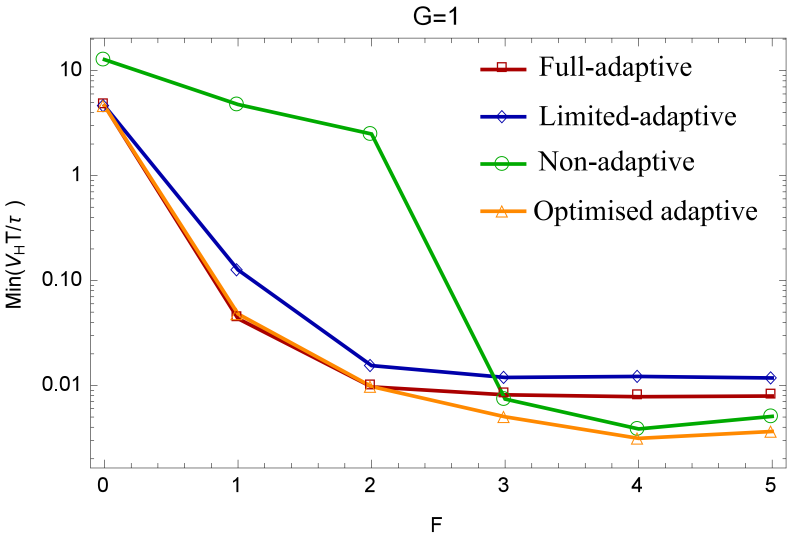


Figure 4.11: The minimum of the scaled variance versus F for optimised, full-, limited- and non-adaptive protocols.

4.4 Conclusion

In this chapter, we gave simulations for different protocols in estimation of a time-independent magnetic field using a single electron spin in an NV centre. We showed that non-adaptive protocols reach Heisenberg-like scaling. However, there is an optimised adaptive protocol which requires a smaller number of detections and outperforms the optimised non-adaptive protocol. The enhancement obtained by the optimised adaptive protocol is significant if the overhead time is considered. In this optimised adaptive protocol, the phase is updated after each detection via increments found by the PSO algorithm. The numerical simulations given here are confirmed in the experiment done in Ref. [61]. This optimised adaptive protocol is probably not the ultimate optimal adaptive protocol. Further improvements might be obtained by taking into account the full measurement history.

5
Estimation of a time varying phase

5.1 Introduction

In the previous chapters we considered estimation of a constant phase in interferometric schemes. For a constant phase the SQL in the phase variance scales as $1/N$ and the Heisenberg limit scales as $1/N^2$, N being the total number of photons. However, in many applications the phase varies with time. Examples of such are a signal with a time varying phase to transmit information, and phase estimation in gravitational wave detectors. Another scenario is phase estimation in the presence of phase diffusion, considered in Refs. [118, 119]. This is distinct from what we are considering in this chapter.

We consider phase variation that is stationary and has Gaussian statistics. Such phase variation can be characterised just by its spectral density, which we take to have power law scaling, i.e. $\kappa^{p-1}/|\omega|^p$ for large frequency ω. The constant κ is a constant of proportionality, which is given with a power of $p-1$ so that it has units of frequency. For such a varying phase the SQL scales as $\left(\kappa/\bar{N}\right)^{(p-1)/p}$ and the Heisenberg limit scales as $\left(\kappa/\bar{N}\right)^{2(p-1)/(p+1)}$, $\bar{N}$ being the photon flux of the beam [120, 121]. For $p \to \infty$ this time varying phase is analogous to a constant phase, and the stochastic limits in phase estimation approach the constant phase limit.

The stochastic Heisenberg limit can be obtained, up to a constant factor, by sampling regularly spaced sequence of pulses each measured by a canonical phase measurement [121]. However, it is not possible to deterministically realise canonical measurements with linear optical elements. Similar scaling for a constant phase can be achieved using adaptive homodyne measurements. Moreover, the sampling approach requires the pulses to be very short so that the phase is effectively constant during each sample. It therefore requires very fast adaptive measurements where a controlled phase is varied orders of magnitude faster than the system phase [121]. It is much more practical to consider a continuous beam, where the controlled phase need only be varied slowly.

In prior work it has been shown that, for phase variation with Wiener statistics, continuous squeezed states in an adaptive homodyne scheme give Heisenberg scaling for the variance of the phase estimate [64, 70, 122]. For the Wiener process the spectrum has power law scaling

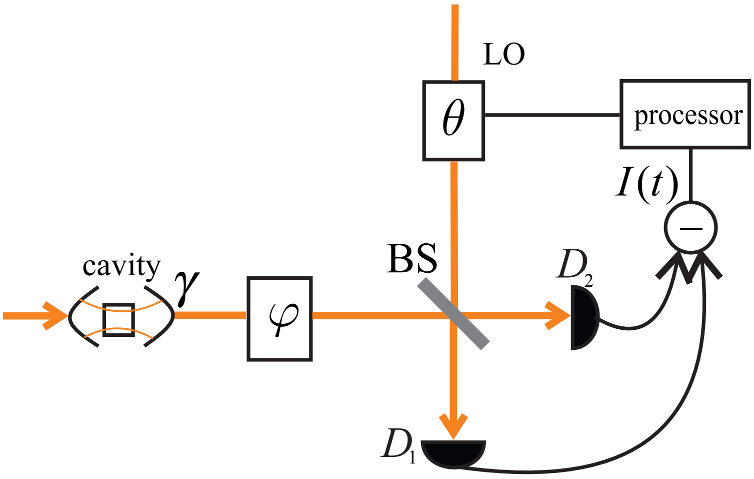


Figure 5.1: The scheme for adaptive homodyne measurement of the phase φ imposed on a squeezed coherent state generated by a cavity with decay constant γ. D_1 and D_2 are the photodetectors. $I(t)$ is the difference photocurrent between the two outputs of the 50/50 beam splitter (BS). The processor adjusts the phase of the local oscillator (LO) labeled by θ based on $I(t)$.

with $p = 2$. In this chapter, we show that with such an adaptive scheme we can also obtain Heisenberg scaling for $p > 1$.

5.2 Adaptive homodyne measurement with squeezed states

A diagram of an adaptive homodyne scheme is shown in Fig. 5.1. As we explained in Chapter 1, in adaptive homodyne measurements the field, which we here take to be a squeezed coherent state, is combined with a strong local oscillator (LO) on a 50/50 beam splitter. The difference photocurrent in the outputs of the beam splitter is then used to adjust the phase of the LO for the following measurements [68, 72]. The continuous squeezed coherent beam is produced in an optical parametric oscillator (OPO) [123]. In this method a nonlinear medium inside a cavity is pumped with a coherent beam. The light leaked out of the cavity provides the continuous beam. The decay constant of the cavity is denoted by γ.

The output photon flux from the cavity can be written in terms of the quadrature operators just outside the cavity and before the phase, denoted by $\hat{X}$ and $\hat{Y}$, as [71]

$$4\bar{\mathcal{N}} = \left\langle \hat{X} \right\rangle^2 + \left\langle \hat{Y} \right\rangle^2 + \left\langle : \Delta\hat{X}^2 + \Delta\hat{Y}^2 : \right\rangle. \tag{5.1}$$

Here, $\left\langle \hat{X} \right\rangle = 0$ and $\left\langle \hat{Y} \right\rangle = E$, where E is the coherent amplitude of the field. The normally ordered variances of the quadratures are [124, 125]

$$\begin{aligned} \left\langle : \Delta\hat{X}^2 : \right\rangle = \left\langle : \hat{X}^2 : \right\rangle - \left\langle : \hat{X} : \right\rangle^2 = -\frac{\gamma\varepsilon}{1+\varepsilon}, \\ \left\langle : \Delta\hat{Y}^2 : \right\rangle = \frac{\gamma\varepsilon}{1-\varepsilon}, \end{aligned} \tag{5.2}$$

where ε is a parameter related to the squeezing parameter r according to

$$\varepsilon = \frac{e^r - 1}{e^r + 1}. \tag{5.3}$$

This gives

$$\bar{N} = \frac{E^2}{4} + \frac{\gamma}{2}\sinh^2\left(\frac{r}{2}\right). \tag{5.4}$$

We denote the quadrature operators of the field inside the cavity by $\hat{x}$ and $\hat{y}$. The Heisenberg equation of motion for these quadrature operators can be written as [70, 126]

$$\frac{d\hat{x}}{dt} = -\hat{x}\gamma(1+\varepsilon)/2 + \sqrt{\gamma}\hat{\xi}, \tag{5.5}$$

$$\frac{d\hat{y}}{dt} = -\hat{y}\gamma\left(1-\varepsilon\right)/2 + \sqrt{\gamma}\hat{\eta}. \tag{5.6}$$

Here, $\hat{\xi}$ and $\hat{\eta}$ are the quadrature noise operators, and we have considered the squeezed quadrature to be $\hat{x}$. The phase φ is imposed on the squeezed state before it combines on a 50/50 beam splitter with a LO which has phase θ. The output quadrature at angle $\theta-\varphi$ is obtained as [70, 126]

$$\hat{I} = \cos\left(\theta-\varphi\right)\left(\sqrt{\gamma}\hat{x}-\hat{\xi}\right) + \sin\left(\theta-\varphi\right)\left(\sqrt{\gamma}\hat{y}+E-\hat{\eta}\right). \tag{5.7}$$

This corresponds to the measured difference photocurrent in the output modes.

Equations (5.5), (5.6) and (5.7) can be solved by replacing the quadrature operators with the corresponding quadrature variables for the Wigner distribution, and replacing $\hat{I}$, the output quadrature, with the difference photocurrent in the output [70]. Therefore, we can write

$$\frac{dx}{dt} = -x\gamma(1+\varepsilon)/2 + \sqrt{\gamma}\xi \tag{5.8}$$

$$\frac{dy}{dt} = -y\gamma\left(1-\varepsilon\right)/2 + \sqrt{\gamma}\eta \tag{5.9}$$

$$I = \cos\left(\theta-\varphi\right)\left(\sqrt{\gamma}x-\xi\right) + \sin\left(\theta-\varphi\right)\left(\sqrt{\gamma}y+E-\eta\right). \tag{5.10}$$

Here, ξ and η are Gaussian increments satisfying $\langle\xi\left(t\right)\xi\left(t'\right)\rangle = \langle\eta\left(t\right)\eta\left(t'\right)\rangle = \delta\left(t-t'\right)$. One way to numerically integrate these equations is to directly discretise the equations over time steps of length Δt [70]. The method we describe here is to instead integrate the differential equations over a time step of length Δt. This method is still not exact because we assume that the system and controlled phases are constant over these time intervals. That is, the remaining approximation in the discretisation is now in taking the phases to be constant over the time intervals. Provided the time intervals are short, the approximation will be accurate, and it will be more accurate than the approximation without the integrals.

Integrating Eqs. (5.8) and (5.9) we obtain

$$x\left(t\right) = e^{\gamma(1+\varepsilon)(t_0-t)/2}x_0 + \sqrt{\gamma}\int_{t_0}^{t} du\, e^{\gamma(1+\varepsilon)(u-t)/2}\xi(u), \tag{5.11}$$

$$y\left(t\right) = e^{\gamma(1-\varepsilon)(t_0-t)/2}x_0 + \sqrt{\gamma}\int_{t_0}^{t} du\, e^{\gamma(1-\varepsilon)(u-t)/2}\eta(u), \tag{5.12}$$

where x_0 and y_0 are the values of x and y at $t=0$. To obtain the effect of a step from time t_0 to t_1 we integrate I over this interval. Therefore, we need to integrate $\sqrt{\gamma}x-\xi$ and $\sqrt{\gamma}y-\eta$.

We have

$$\begin{aligned}\sqrt{\gamma}\int_{t_0}^{t_1} dt x &= \sqrt{\gamma}x_0\int_{t_0}^{t_1} dt e^{\gamma(1+\varepsilon)(t_0-t)/2} + \gamma\int_{t_0}^{t_1} dt\int_{t_0}^{t} du\, e^{\gamma(1+\varepsilon)(u-t)/2}\xi(u)\\ &= x_0\frac{2\sqrt{\gamma}}{\gamma\,(1+\varepsilon)}\left(1-e^{-\gamma(1+\varepsilon)\Delta t/2}\right) + \gamma\int_{t_0}^{t_1} du\,\xi(u)\int_{u}^{t_1} dt\, e^{\gamma(1+\varepsilon)(u-t)/2}\\ &= x_0\frac{e^{-r}+1}{\sqrt{\gamma}}\left(1-e^{-\gamma(1+\varepsilon)\Delta t/2}\right) + (e^{-r}+1)\int_{t_0}^{t} du\,\xi(u)\left(1-e^{\gamma(1+\varepsilon)(u-t_1)/2}\right),\end{aligned} \tag{5.13}$$

which gives

$$\begin{aligned}\int_{t_0}^{t_1} dt(\sqrt{\gamma}x-\xi(t)) &= x_0\frac{e^{-r}+1}{\sqrt{\gamma}}\left(1-e^{-\gamma(1+\varepsilon)\Delta t/2}\right)\\ &\quad -(e^{-r}+1)\int_{t_0}^{t} du\,\xi(u)e^{\gamma(1+\varepsilon)(u-t_1)/2} + \int_{t_0}^{t} e^{-r}\xi(u)\,du.\end{aligned} \tag{5.14}$$

In the above equations we have defined $\Delta t = t_1 - t_0$. Similarly for $\sqrt{\gamma}y - \eta$ we obtain

$$\begin{aligned}\int_{t_0}^{t_1} dt\,(\sqrt{\gamma}y-\eta(t)) &= y_0\frac{e^{r}+1}{\sqrt{\gamma}}\left(1-e^{-\gamma(1-\varepsilon)\Delta t/2}\right)\\ &\quad -(e^{r}+1)\int_{t_0}^{t} du\,\eta(u)e^{\gamma(1-\varepsilon)(u-t_1)/2} + \int_{t_0}^{t} e^{r}\eta(u)\,du.\end{aligned} \tag{5.15}$$

We define

$$\chi_x \equiv \int_{t_0}^{t} du\,\xi(u)e^{\gamma(1+\varepsilon)(u-t_1)/2}, \qquad \chi_y \equiv \int_{t_0}^{t} du\,\eta(u)e^{\gamma(1-\varepsilon)(u-t_1)/2} \tag{5.16}$$

$$\psi_x \equiv \int_{t_0}^{t} du\,\xi(u)e^{-r}, \qquad \psi_y \equiv \int_{t_0}^{t} du\,\eta(u)e^{r}. \tag{5.17}$$

In terms of these new variables the integral of I can be written as

$$\begin{aligned}\int_{t_0}^{t_1} dt\,I &= \cos{(\theta-\varphi)}\left[x_0(e^{-r}+1)\left(1-e^{-\gamma(1+\varepsilon)\Delta t/2}\right)/\sqrt{\gamma}-(e^{-r}+1)\,\chi_x+\psi_x\right]\\ &\quad+\sin{(\theta-\varphi)}\left[E\Delta t+y_0(e^{r}+1)\left(1-e^{-\gamma(1-\varepsilon)\Delta t/2}\right)/\sqrt{\gamma}-(e^{r}+1)\,\chi_y+\psi_y\right].\end{aligned} \tag{5.18}$$

The expectation values of χ_ℓ and ψ_ℓ for both $\ell = x, y$ are zero because ξ and η both have mean zero. Therefore the variances are as the following

$$\left\langle\chi_x^2\right\rangle = \int_{t_0}^{t} du\,e^{\gamma(1+\varepsilon)(u-t_1)} = (e^{-r}+1)\left(1-e^{-\gamma(1+\varepsilon)\Delta t}\right)/2\gamma, \tag{5.19}$$

$$\left\langle\chi_y^2\right\rangle = \int_{t_0}^{t} du\,e^{\gamma(1-\varepsilon)(u-t_1)} = (e^{r}+1)\left(1-e^{-\gamma(1-\varepsilon)\Delta t}\right)/2\gamma, \tag{5.20}$$

$$\left\langle\psi_x^2\right\rangle = e^{-2r}\Delta t, \qquad \left\langle\psi_y^2\right\rangle = e^{2r}\Delta t, \tag{5.21}$$

and the covariances are

$$\langle \chi_x \psi_x \rangle = \int_{t_0}^{t} du\, e^{\gamma(1+\varepsilon)(u-t_1)/2} e^{-r} = e^{-r}\left(e^{-r}+1\right)\left(1 - e^{-\gamma(1+\varepsilon)\Delta t/2}\right)/\gamma, \quad (5.22)$$

$$\left\langle \chi_y \psi_y \right\rangle = \int_{t_0}^{t} du\, e^{\gamma(1+\varepsilon)(u-t_1)/2} e^{r} = e^{r}\left(e^{r}+1\right)\left(1 - e^{-\gamma(1-\varepsilon)\Delta t/2}\right)/\gamma. \quad (5.23)$$

We also define

$$m_x^{(1)} \equiv \left(e^{-r}+1\right)\left(1 - e^{-\gamma(1+\varepsilon)\Delta t/2}\right)/\sqrt{\gamma}, \quad (5.24)$$

$$m_y^{(1)} \equiv \left(e^{r}+1\right)\left(1 - e^{-\gamma(1-\varepsilon)\Delta t/2}\right)/\sqrt{\gamma}, \quad (5.25)$$

and $\Omega_x = \psi_x - \lambda_x \chi_x$ and $\Omega_y = \psi_y - \lambda_y \chi_y$ in such a way that the covariances $\langle \Omega_x \chi_x \rangle$ and $\left\langle \Omega_y \chi_y \right\rangle$ are zero. This could be obtained by having

$$\lambda_x = \frac{\langle \chi_x \psi_x \rangle}{\left\langle \chi_x^2 \right\rangle}, \qquad \lambda_y = \frac{\left\langle \chi_y \psi_y \right\rangle}{\left\langle \chi_y^2 \right\rangle}. \quad (5.26)$$

In terms of these scaling factors we can write

$$x\left(t_1\right) = e^{-\gamma(1+\varepsilon)\Delta t/2} x_0 + \sqrt{\gamma}\chi_x, \quad (5.27)$$

$$y\left(t_1\right) = e^{-\gamma(1+\varepsilon)\Delta t/2} y_0 + \sqrt{\gamma}\chi_y, \quad (5.28)$$

$$\begin{aligned} I(t_1) = I(t_0) &+ \cos\left(\theta - \varphi\right)\left(m_x^{(1)} x_0 + \Omega_x + m_x^{(4)} \chi_x\right) \\ &+ \sin\left(\theta - \varphi\right)\left(m_y^{(1)} y_0 + E\Delta t + \Omega_y + m_y^{(4)} \chi_y\right), \end{aligned} \quad (5.29)$$

where $m_x^{(4)} = \lambda_x - e^{-r} - 1$ and $m_y^{(4)} = \lambda_y - e^{r} - 1$.

In the numerical simulations we use these formulae to improve the accuracy, and determine the controlled phase θ from the photocurrent. Before explaining how this is done, we first give the details for the time variation of the system phase φ.

5.3 System phase time variation

We consider a time-varying system phase φ which is stationary and has Gaussian statistics. Here, we mean stationary in the wide sense [127]. For a wide-sense stationary process, the mean value of the phase, $\langle \varphi(t) \rangle$ is independent of time, and its autocorrelation function, $\Sigma(t_1, t_2) = \langle \varphi(t_1)\varphi(t_2) \rangle$ is only a function of $t_1 - t_2$ [127]. Moreover, we assume the spectral density of the process, defined as the Fourier transform of the autocorrelation function,

$$\tilde{\Sigma}(\omega) = \int_{-\infty}^{\infty} \Sigma(t)\, e^{-i\omega t}\, dt, \quad (5.30)$$

has power law scaling for large ω, i.e. $\tilde{\Sigma}(\omega) \sim \kappa^{p-1}/|\omega|^p$. The multiplicative factor κ is a constant with units of frequency, and is the inverse of the characteristic time of the spreading of the process. To ensure that the spectrum is limited at $\omega = 0$, we consider the spectral density to be [121, 128]

$$\tilde{\Sigma}(\omega) = \frac{\kappa^{p-1}}{\omega^p + \Gamma^p}. \quad (5.31)$$

Γ is a constant and is the characteristic time for the relaxation of the phase towards zero [121].

Now we show how a time-varying phase with such spectral density can be generated. If we take the Fourier transform of the phase $\varphi(t)$ and calculate the two-frequency expectation value we obtain

$$
\begin{aligned}
\langle \tilde{\varphi}(\omega_1)\tilde{\varphi}(\omega_2)\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt_1\, dt_2 \langle \varphi(t_1)\varphi(t_2)\rangle e^{-i(\omega_1 t_2 - \omega_2 t_2)} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt_1\, dt_2\, \Sigma(t_1 - t_2) e^{-i\left[\frac{1}{2}(\omega_1+\omega_2)(t_1-t_2)+\frac{1}{2}(\omega_1-\omega_2)(t_1+t_2)\right]} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dT\, d\Delta\, \Sigma(\Delta) e^{-i\left[\frac{1}{2}(\omega_1+\omega_2)\Delta+(\omega_1-\omega_2)T\right]} \\
&= 2\pi\delta(\omega_1-\omega_2)\int_{-\infty}^{\infty} dT\, \Sigma(\Delta) e^{-i\left[\frac{1}{2}(\omega_1+\omega_2)\Delta\right]} \\
&= 2\pi\delta(\omega_1-\omega_2)\tilde{\Sigma}(\omega_1).
\end{aligned} \tag{5.32}
$$

Here, we have used the change of variables $T = (t_1 + t_2)/2$, $\Delta = t_1 - t_2$, and in the last line we have replaced $(\omega_1 + \omega_2)/2$ by ω_1 because of the delta function $\delta(\omega_1 - \omega_2)$. Note also that, because the phase $\varphi(t)$ is real, $\tilde{\varphi}(-\omega) = \tilde{\varphi}^*(\omega)$. As a result, we can write the Fourier transform of the phase in the form

$$
\tilde{\varphi}(\omega) = \sqrt{2\pi\tilde{\Sigma}(\omega)}\xi(\omega), \tag{5.33}
$$

where $\xi(\omega)$ has the correlations

$$
\langle \xi(\omega_1)\xi^*(\omega_2)\rangle = \langle \xi(\omega_1)\xi(-\omega_2)\rangle = \delta(\omega_1-\omega_2). \tag{5.34}
$$

Taking the inverse Fourier transform of $\tilde{\varphi}(\omega)$ we obtain

$$
\varphi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, \sqrt{2\pi\tilde{\Sigma}(\omega)}\xi(\omega)e^{i\omega t}. \tag{5.35}
$$

Calculating the correlation function we obtain

$$
\begin{aligned}
\langle \varphi\left(t+\tau\right)\varphi\left(t\right)\rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \sqrt{\tilde{\Sigma}\left(\omega_1\right)\tilde{\Sigma}\left(\omega_2\right)}\,\langle \xi\left(\omega_1\right)\xi\left(\omega_2\right)\rangle\, e^{i\omega_1(t+\tau)} e^{i\omega_2 t} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, \tilde{\Sigma}\left(\omega\right) e^{i\omega\tau}.
\end{aligned} \tag{5.36}
$$

This confirms that $\varphi(t)$ has power spectral density $\tilde{\Sigma}(\omega)$.

To generate this phase in our simulations we generate discretised complex white noise and use a discretised Fourier transform. We take $\xi(\omega)$ to be approximated by

$$
\xi(\omega_k) \approx (z_{k,1} + iz_{k,2})/\sqrt{2\delta\omega}, \tag{5.37}
$$

where $z_{k,1}$ and $z_{k,2}$ are normally distributed random numbers with $z_{k,1} = z_{-k,1}$, $z_{k,2} = -z_{-k,2}$. We approximate the integral in Eq. (5.35) by

$$
\varphi(t_n) \approx \frac{1}{\sqrt{4\pi}}\sum_k \sqrt{\delta\omega}\sqrt{\tilde{\Sigma}\left(\omega_k\right)}\left(z_{k,1} + iz_{k,2}\right) e^{i\omega_k t_n}. \tag{5.38}
$$

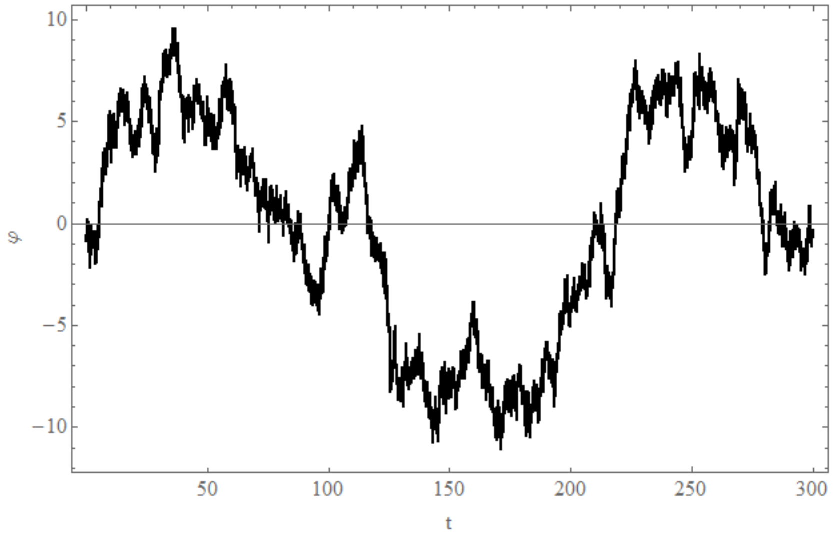


Figure 5.2: A Gaussian random process with power law spectral density $1/(\omega^p + \Gamma^p)$, with $p = 2$, $\Gamma = 10^{-3}$, and $\delta t = 10^{-3}$.

Taking $t_n = n\delta t$, $\omega_k = k\delta\omega$ and $\delta\omega\delta t = 2\pi/N$ the above equation becomes

$$\begin{aligned}
\varphi(t_n) &\approx \frac{1}{\sqrt{4\pi}} \sum_k \sqrt{\delta\omega}\sqrt{\tilde{\Sigma}(\omega_k)(z_{k,1} + iz_{k,2})}e^{i2\pi nk/N} \\
&= \frac{1}{\sqrt{2N\delta t}} \sum_k \sqrt{\tilde{\Sigma}\left(\omega_k\right)}\left(z_{k,1} + iz_{k,2}\right) e^{i2\pi nk/N} \\
&\approx \frac{1}{\sqrt{2N\delta t}} \left[\sum_{k=0}^{N-1} \sqrt{\tilde{\Sigma}(\omega_k)}\left(z'_{k,1} + iz_{k,2}\right) e^{i2\pi nk/N} + \sum_{k=0}^{N-1} \sqrt{\tilde{\Sigma}(\omega_k)}\left(z'_{k,1} - iz_{k,2}\right) e^{-i2\pi nk/N}\right] \\
&= \sqrt{\frac{2}{N\delta t}} \left[\mathrm{Re}\left(\sum_{k=0}^{N-1} \sqrt{\tilde{\Sigma}(\omega_k)}\left(z'_{k,1}\right) e^{-i2\pi nk/N}\right) - \mathrm{Im}\left(\sum_{k=0}^{N-1} \sqrt{\tilde{\Sigma}(\omega_k)}\left(z_{k,2}\right) e^{-i2\pi nk/N}\right)\right]
\end{aligned} \tag{5.39}$$

where $z'_{k,1} = z_{k,1}$ for $k \neq 0$ and $z'_{0,1} = z_{0,1}/2$. Figure 5.2 shows the generated noise for $p = 2$ and $\Gamma = 10^{-3}$. As p is increased the time variation of the phase becomes slower.

5.4 Feedback phase

To estimate this time-varying phase, we change the LO phase, θ, based on the difference photocurrent given in Eq. (5.29) during the course of the measurement. As we mentioned in Section 1.5, the LO phase could be updated by Bayesian updating [65–67] or based on the functionals of the photocurrent record [64, 70]. The Bayesian updating is highly numerically intensive for this problem. Moreover, it is shown in Ref. [70] that Bayesian updating gives only a few percent enhancement over the other method. Therefore, we follow Refs. [64, 70] and update the phase based on the functionals of the difference photocurrent.

The relevant information from the measurement record can be formulated in the following quantities [64, 70]

$$A(t) = \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\theta} I(u)\, du, \tag{5.40}$$

$$B(t) = -\int_{-\infty}^{t} e^{\chi(u-t)} e^{2i\theta}\, du, \tag{5.41}$$

where χ is a scaling parameter which scales the weight $e^{\chi(u-t)}$ given to the difference photocurrent at time u, $I(u)$. The phase estimate at time t, $\breve{\varphi}(t)$, is a function of the two functionals $A(t)$ and $B(t)$ according to

$$\breve{\varphi}(t) = \arg\left(C(t)\right), \qquad C(t) = A(t) + \chi B(t) A^*(t). \tag{5.42}$$

However, it is found that using this phase estimate as the LO phase gives poor results [64, 70]. This is because for very good estimates of the phase in the feedback the results do not distinguish easily between the system phase and system phase plus π. Therefore, many of the results are off by π which results in a large phase variance [18]. Thus, following the technique of previous works [64, 70] we set the LO phase to

$$\theta(t) = \arg\left(C^{1-\delta}(t) A^{\delta}(t)\right) + \pi/2, \tag{5.43}$$

and find the optimal value of δ numerically. Recall that, in Section 4.3.2 we explained that changing the controlled phase by π does not make any difference to the final phase variance. If we were only attempting to measure the phase at a final time, then it would not matter if there were errors of π in the phase estimate at intermediate times, because errors of π in the controlled phase do not adversely affect the results. This means that it would be reasonable to use $\arg\left(C(t)\right)$ as the phase estimate at intermediate times. However, because we require accurate estimates of the phase at all times, we must be able to resolve the π ambiguity at all times, and it is better to use the LO phase given in Eq. (5.43).

5.5 Numerical results

We find the minimum variance for each value of $\bar{\mathcal{N}}$ numerically. If we scale the time by κ, we obtain the dimensionless parameters $\bar{\mathcal{N}}/\kappa$, γ/κ, and χ/κ. In addition to these three parameters, the other parameters that we can vary are the dimensionless parameters r and δ. Here, we consider arbitrary squeezing and do not consider any limitations for the squeezing parameter, r. This is because we want to find the ultimate scaling obtained with this scheme regardless of current technological status of sources of squeezing. For each value of $\bar{\mathcal{N}}/\kappa$ we have found the minimum variance by a numerical search for the four parameters γ/κ, χ/κ, δ, and e^r. In order to do this, we systematically incremented the value of each parameter in turn to find the values that give the minimum variance.

For $\bar{\mathcal{N}}/\kappa < 5 \times 10^7$ we used the Holevo variance,

$$\mathrm{Re}\left[\frac{1}{M}\sum_{i=1}^{M} e^{i(\breve{\varphi}_i - \varphi)}\right]^{-2} - 1. \tag{5.44}$$

For $\bar{\mathcal{N}}/\kappa \geq 5 \times 10^7$ the formula for the standard mean-square error given as

$$\frac{1}{M}\sum_{i=1}^{M} (\breve{\varphi}_i - \varphi)^2, \tag{5.45}$$

approximates the Holevo variance more accurately than using the formula for the Holevo variance. This is due to roundoff error using the formula for the Holevo variance. To give the system of equations time to reach its steady state, we did not sample the error up to time $100/\chi$ in our simulations. We then estimated the variance by sampling the error for every time step up to $300/\chi$. Even though the error was sampled every time step, the samples are strongly correlated for times below $1/\chi$. Therefore the number of independent samples is effectively the multiple of $1/\chi$ used for the time. We performed 2^6 independent integrations from time 0 to $300/\chi$. Therefore, the effective number of independent samples of the error was 200×2^6. This includes those from different times within one integration.

To calculate the integrals (5.40) and (5.41) we use time steps of $\delta t = 1/\left(10^3\chi\right)$ and approximate the integrals with

$$A(t+\delta t) \approx (1-\chi\delta t)A(t) - I(t)e^{i\theta}\delta t, \tag{5.46}$$

$$B(t+\delta t) \approx (1-\chi\delta t)B(t) + e^{2i\theta}\delta t, \tag{5.47}$$

where we have used the approximation $e^{\chi\delta t} \approx 1-\chi\delta t$ and assumed I does not change in the interval $[t, t+\delta t)$.

We found the values of the parameters e^r, χ, γ and δ that give the minimum MSE for a range of values of p. We then used linear regression to find the scaling of each parameter as a function of $\bar{N}/\kappa$ for a general value of p and found

$$\begin{gathered} e^r \sim \left(\bar{N}/\kappa\right)^{(p-1)/(2p+2)}, \quad \chi/\kappa \sim \left(\bar{N}/\kappa\right)^{2/(p+1)}, \\ \gamma/\kappa \sim \left(\bar{N}/\kappa\right)^{(p+3)/(2p+2)}, \quad \delta \sim \left(\kappa/\bar{N}\right)^{(p-1)/(p+2)}, \\ \mathrm{MSE} \sim \left(\kappa/\bar{N}\right)^{2(p-1)/(p+1)}. \end{gathered} \tag{5.48}$$

The optimal values of e^r, χ/κ, γ/κ and δ, are shown in Figs. 5.3 and 5.4. The minimum MSE is shown in Fig. 5.5 for a range of values of p. For the case of $p = 4$, it was not possible to push $\bar{N}/\kappa$ to large values. This is due to the rapid decrease of the MSE for large values of p and the resulting roundoff error in the simulations.

The scaling obtained for the MSE with the adaptive homodyne scheme using squeezed state is of the Heisenberg scaling. In Fig. 5.6 we have compared the scaling constant of the proposed scheme here with the Heisenberg limit, and the pulsed measurement proposed in Ref. [121]. For large values of $\bar{N}$ the scaling constant for the Heisenberg limit is [121]

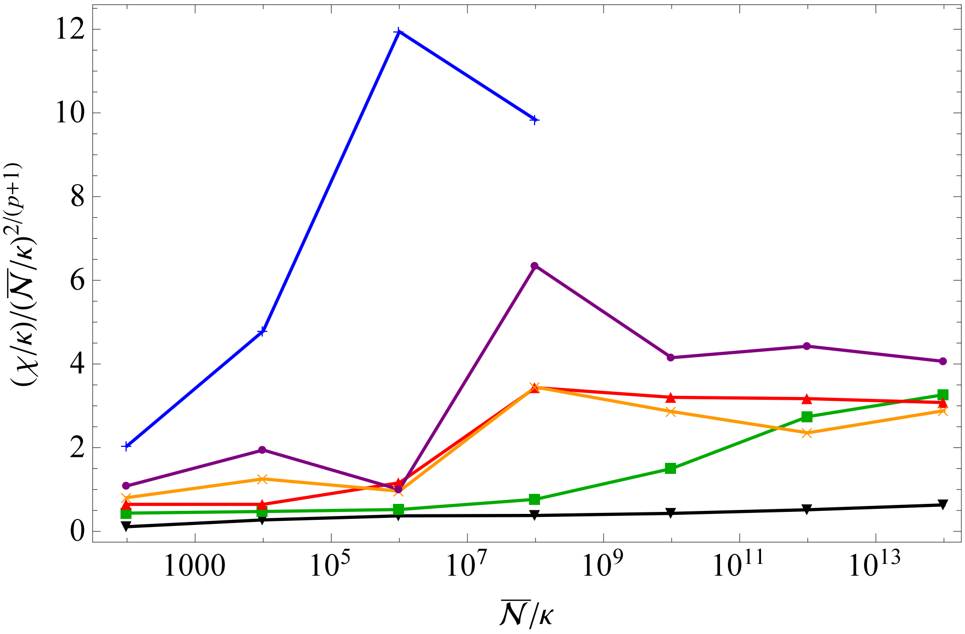
$$c_z = \frac{11}{420}\left(\frac{p_3}{4}\right)^{2/(p+1)}\left(\frac{1}{4\pi\lambda}\right)^{2(p-1)/(p+1)}, \tag{5.49}$$

with $\lambda \approx 0.7246$ and $p_3 = (p+1)(p+2)(p+3)$. The scaling constant for the MSE achievable by the pulsed method of Ref. [121] is

$$c_A = \frac{p+1}{p-1}\left(4|z_A|^3/27\right)^{(p-1)/(p+1)}\pi^{2p/(p+1)}, \tag{5.50}$$

with $z_A \approx -2.338$. As can be seen from this figure, the pulsed method (with assumed ideal phase measurements) performs better than the continuous squeezing method (with adaptive homodyne measurements) for large p.

In Fig. 5.7 we have plotted the system phase and the estimated phase obtained based on Eq. (5.42) for $p = 1.5$, and $p = 3$. The initial period of transience of the phase estimate can be seen in this figure. The phase estimate is initially far from the system phase but as we obtain more information from the measurements it locks into the system phase and follows its fluctuations quite well.



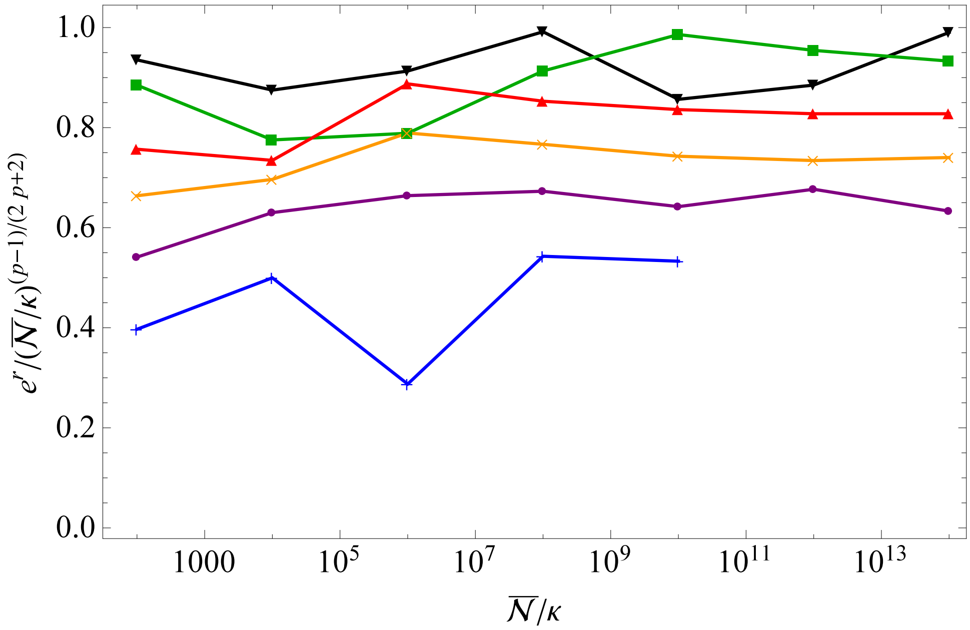


Figure 5.3: The optimal values of the parameters χ and e^r versus $\bar{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$.

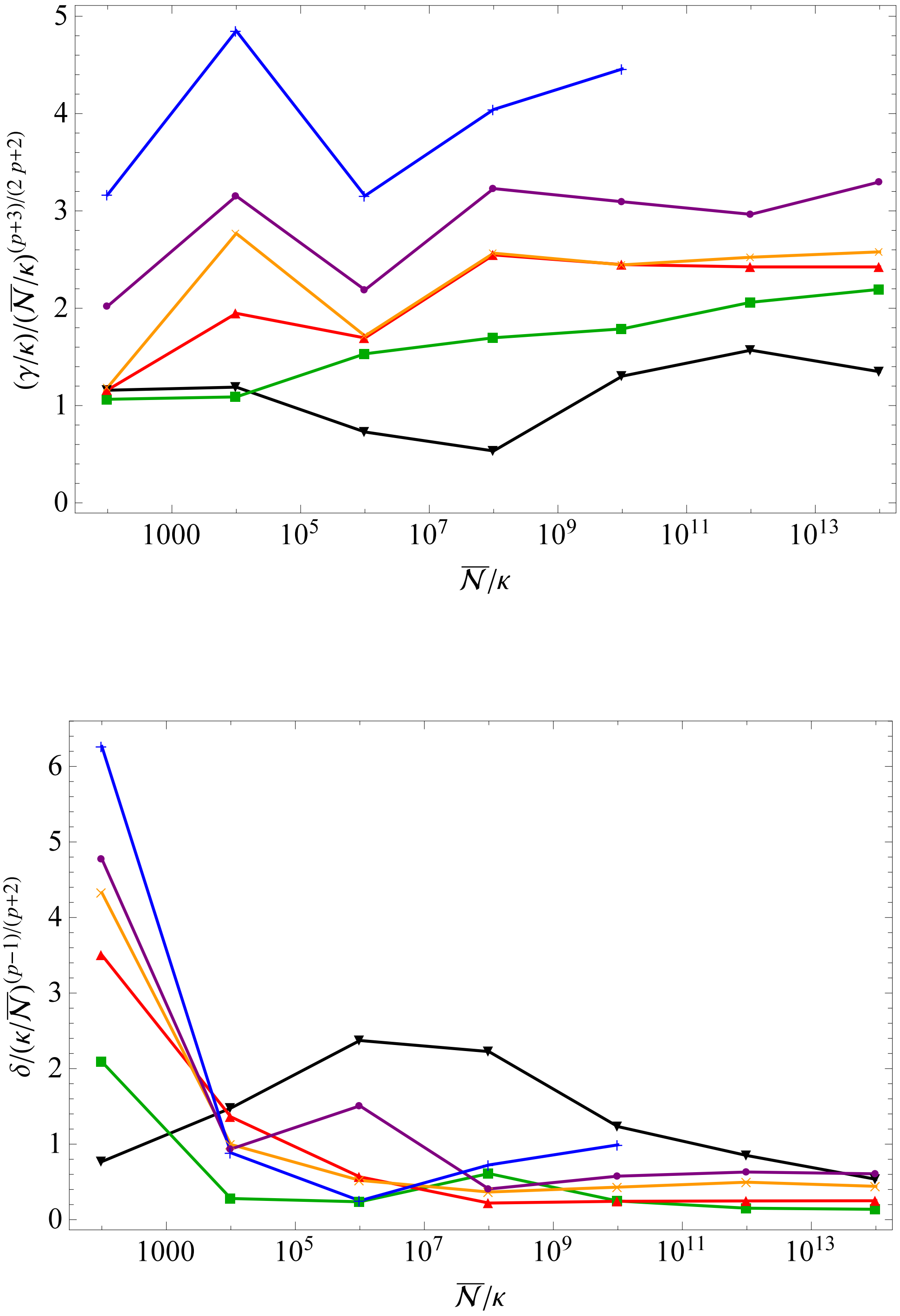


Figure 5.4: The optimal values of the parameters γ and δ versus $\bar{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$.

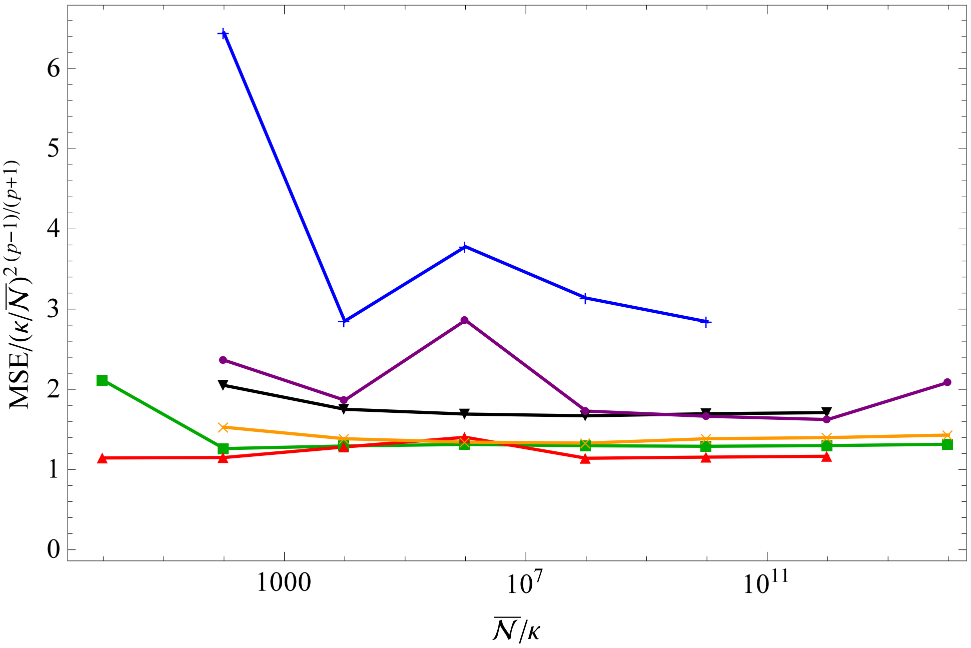


Figure 5.5: The scaled MSE versus $\bar{N}/\kappa$ for a range of values of p. Black line: $p = 1.25$, green line: $p = 1.5$, red line: $p = 2$, yellow line: $p = 2.5$, purple line: $p = 3$, and blue line: $p = 4$.

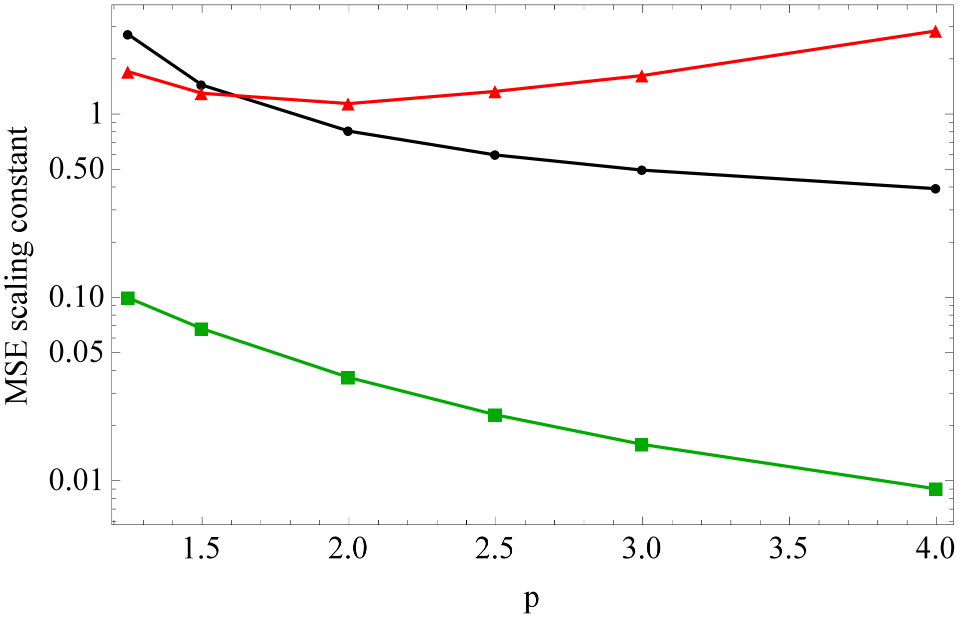


Figure 5.6: The scaling constant of the MSE for the Heisenberg limit (green line), c_z given in Eq. (5.49), the pulsed measurement (black line), c_A given in Eq. (5.50), and the homodyne scheme with a continuous squeezed state (red line).

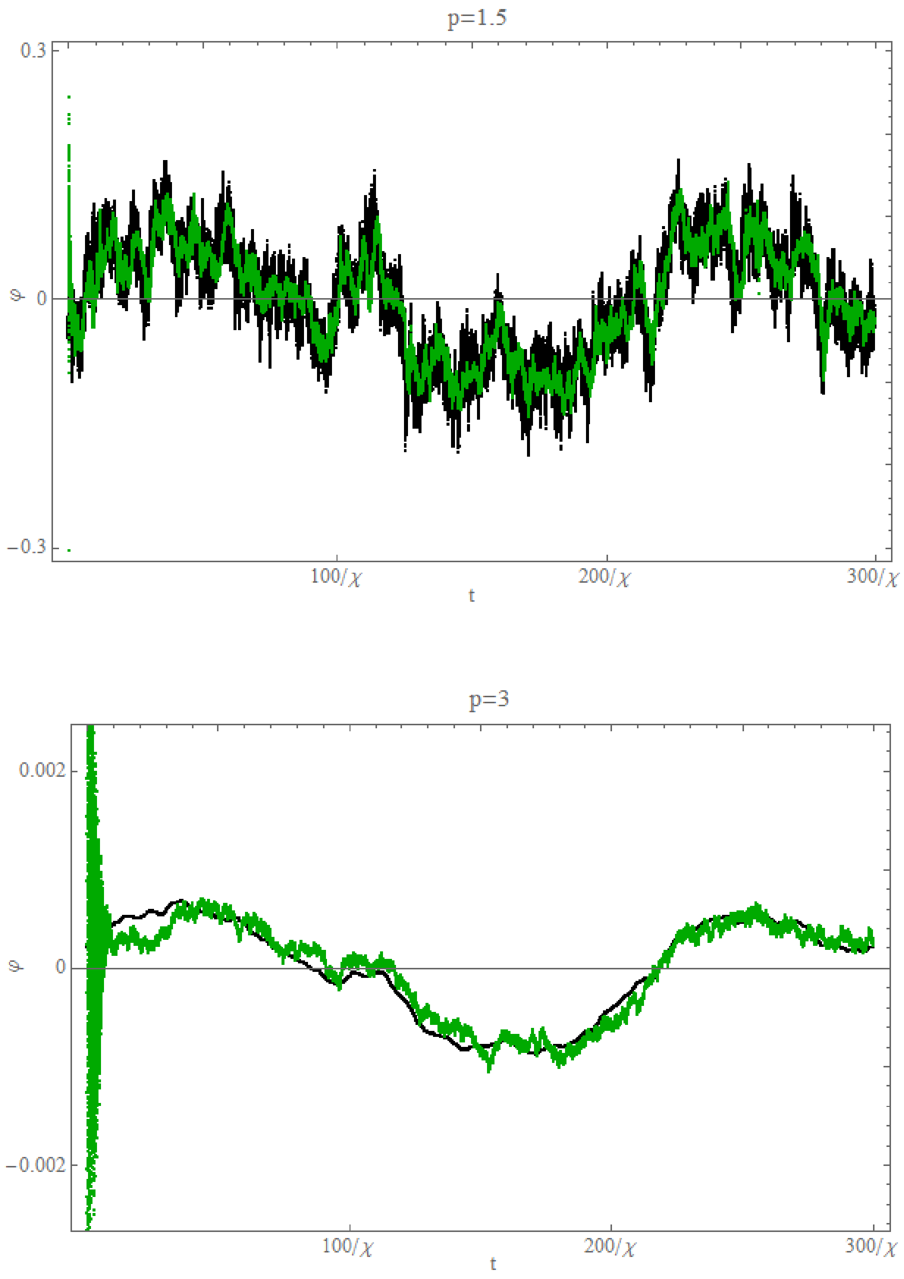


Figure 5.7: System phase (black line) and the estimated phase (green line) for $\bar{N}/\kappa = 10^8$. Top: $p = 1.5$, bottom: $p = 3$.

5.6 Conclusion

In this chapter, we investigated estimation of a time-varying phase in an adaptive homodyne scheme using a continuous squeezed state. We considered a phase with time-invariant Gaussian statistics and power law spectral density. We showed that assuming it is possible to achieve arbitrarily high squeezing, this scheme gives Heisenberg scaling for the variance of the phase. Moreover, we showed that the scaling constant obtained with the adaptive method is larger than the scaling obtained with the sampling method proposed in Ref. [121].

An interesting topic for future work is estimation of a time varying phase with entangled states in an interferometric scheme. As we discussed in Chapters 2 and 3 for the case of a constant phase, not all the entangled states can beat the SQL. The entanglement useful for estimating a constant phase can be characterized by the Fisher information [129]. It would be interesting to characterize the entanglement useful for estimation of a time-varying phase. Other possible areas for further study are to consider non-Gaussian correlations in the phase variation and also the effect of loss in the system.

6

Conclusions

In this thesis we investigated methods to achieve improved precision in measurement of noisy quantum systems. The focus of this work was on adaptive measurement schemes. We considered phase estimation in optical and solid-state systems. In the optical case, we considered estimation of a phase in a Mach-Zehnder interferometer in the presence of photon loss, and estimation of a time-varying phase imposed on a continuous beam in a single spatial mode. The solid-state system we considered was an NV centre. In such a system, the noise results in low visibility.

We started with phase estimation in a Mach-Zehnder interferometer. We proposed a scheme to generate multi-photon entangled states which are close to optimal for phase measurement in the presence of photon loss. We considered loss in both arms of the interferometer, and to simplify the calculations we considered the same amount of loss in both arms. In order to provide a technique that is experimentally feasible, we proposed a scheme that uses an SPDC source to produce loss-resistant states. The scheme is able to produce $2n$-photon states ($n = 1, 2, \dots$), which are loss-resistant. By loss-resistant we mean that the state is still useful for measurement in the presence of loss and ideally is the state which yields the most accurate measurements with that loss.

We proposed an adaptive technique which uses sequences of these loss-resistant states together with single-photon states in order to provide an unambiguous estimate of an unknown phase. This adaptive technique is inspired by the generalized form of the quantum phase estimation algorithm. We found the optimal sequence by a numerical search over the range of possible sequences. We found that for small total photon numbers it is optimal to use only single-photon states. As the number of total photons gets larger, we obtain an enhancement in the precision by using two-photon optimal states, and for larger total photon numbers by using four-photon optimal states in the sequence, and so forth. We showed that by using the optimal sequence of states, we are able to obtain variances smaller than the SQL, defined by the corresponding scheme with independent single photons.

We then considered a special case in which both loss and the phase shift depend on the parameter we wish to measure, which is the transition frequency of atoms. We considered an ensemble of two-level atoms in one of the arms of a Mach-Zehnder interferometer. The interaction of the probe state in the arms of the interferometer with the atomic ensemble results in the photon loss, and imposes a phase shift on the state. We formulated this interaction by the beam splitter model. In this case, both the loss and the phase shift depend on the transition frequency of the atoms. We considered a weighted superposition of dual Fock states in the arms of the interferometer. We found the optimal state for estimating the transition frequency of the atoms by a numerical search for the coefficients that provide the largest Fisher information. In other parts of the thesis we have considered unambiguous measurements, whereas for this case, due to the complexity of this scheme, we have not considered techniques to obtain an unambiguous measurement, and just used the Fisher information.

We found that the loss and the phase shift strongly depend on the number density of atoms. For large number densities, the phase imposed on the probe is large. Therefore, there is an advantage in using both the phase shift and loss to measure the transition frequency of atoms. In this case, the optimal state is an entangled multi-photon state, which has a large weighting on the state with all photons in the arm with the atomic ensemble. However, for small number densities, the phase shift imposed on the probe is not large enough to obtain significant information from the phase shift and optimal states obtain all the information from absorption. This results in unentangled Fock state as the optimal state which is equivalent to sending independent single photons through the atomic ensemble. In other words, the optimal state is that with all photons in the arm with the atomic ensemble. Therefore, there is no advantage in using an interferometric scheme in this case. Surprisingly, there is a number density of atoms for which independent single photons have the highest Fisher information, even higher than the Fisher information for the numerically optimised states with a larger number density. Comparing the plots of the imposed phase shift on the state and the absorption for a range of number densities clarified this. This showed that the optimal number density is the one for which the absorption covers the full range of 0 to 1, but is only close to 1 over a narrow range. This is in contrast to situations where there is a wide region where the absorption is very close to 1. Those will give larger phase shifts but because the absorption is so high the large phase shift is not useful for measurement.

Another consideration is that the Fisher information takes its maximum value over a relatively narrow range of frequencies, and otherwise it is close to zero. Thus, an approach for eliminating ambiguities as has been considered for NOON states would not work. Therefore, we need to have prior information about the transition frequency. In other words, this scheme can be used to measure small changes in the transition frequency.

Next, we studied measurement of an unknown time-independent magnetic field with the electron spin of an NV centre. We investigated the performance of a non-adaptive protocol and various adaptive protocols. The estimation of the magnetic field in this case is equivalent to phase estimation in a Mach-Zehnder interferometer. However, in the case of magnetometry not only is the variance of the estimated magnetic field important but also the range of the detectable magnetic field. Therefore, we aimed to maximise the dynamic range, which is defined as the ratio of the maximum detectable magnetic field to the uncertainty in the field.

We used the generalized form of the quantum phase estimation algorithm to obtain an unambiguous estimate of the magnetic field. We used a sequence of Ramsey measurements with multiple interaction times. The sequence starts with the longest interaction time, and the interaction time is then decreased by successive powers of 2 until the shortest interaction time is reached. For each interaction time, there are a number of detections before the next interaction time is used. The number of detections for each interaction time is increased as the interaction time is decreased. This is necessary to resolve the ambiguity accurately. We found the best non-adaptive scheme numerically, which gives variance for the measurement scaling as the inverse of the square of the total interaction time. This scaling is analogous to the Heisenberg limit in optical interferometry. Despite the speculations about the performance of adaptive schemes in magnetometry with an NV centre, we showed that there is an adaptive scheme that reaches the Heisenberg-like scaling. Compared to the non-adaptive protocol, this protocol requires a lower number of detections, and therefore shorter total interaction time, and gives lower variance for the estimate of the magnetic field. For a general measurement scheme there would be a different controlled phase based on each possible measurement record. Instead we consider controlled phases that are determined by increments based on the most recent detection result and the number of the detections in the sequence. The adaptive scheme that is optimal for this set of possible schemes is then found by particle swarm optimisation. The enhancement obtained by this adaptive protocol over the non-adaptive protocol is significant when the overhead time, consisting of initialisation and measurement time, is taken into account.

Finally, we looked into tracking a time-varying phase imposed on a continuous beam in a single spatial mode. This case is very important because time-varying signals are frequently encountered in practice, for example in gravitational wave detection or communication. Moreover, a continuous beam is easier to produce in practice. For a time-varying phase, the SQL and Heisenberg limit are different from the corresponding limits for a constant phase [121]. We considered the phase as a stochastic process with power law spectral density, i.e. $1/\omega^p$, for $p > 1$. The Wiener process is the case for which $p = 2$. We considered using squeezed states with adaptive feedback based on the photocurrent difference in the output modes. We showed that by numerically optimising over the squeezing and other experimental parameters we can reach the Heisenberg scaling in tracking such a time-varying phase.

A

Appendices

A.1 Optimal sequence for loss resistant phase estimation

In this appendix we give the optimal sequence for estimation of an unknown phase in a lossy interferometer with efficiency η in both arms (see Fig. 2.4). The optimal sequence given here is found for $\eta = 0.6$. The sequence starts with single photons followed by two-photon and four-photon optimal states given in Eqs. (2.18) and (2.19). In the table given below, $\mathcal{N} = N_1 + 2N_2 + 4N_4$ is the total number of photons in the sequence, where N_1, N_2, and N_4 are the numbers of single-photon, two-photon, and four-photon states in the sequence. For $\mathcal{N} < 9$ the sequence only has single-photon states and there is no advantage in using two- and four-photon states. The values given for the parameters $\chi^{(2)}$, and $\chi^{(4)}$ are the values that minimize the Holevo variance for the given amount of loss.

$\mathcal{N}$	N_1	N_2	$\chi^{(2)}$	N_4	$\chi^{(4)}$
9	7	1	1.7	0	-
10	8	1	1.7	0	-
11	7	2	1.7	0	-
12	8	2	1.7	0	-
13	7	1	1.7	1	1.3
15	9	1	1.7	1	1.3
16	8	2	1.7	1	1.3
17	7	3	1.7	1	1.3
18	6	4	1.7	1	1.3
19	5	3	1.7	2	1.3
20	6	3	1.7	2	1.3

A.2 Particle swarm optimisation

In this appendix we give details of the particle swarm optimisation (PSO) algorithm we used in Chapters 3 and 4. The PSO is a powerful algorithm to find the global minimum or maximum when we have little knowledge about the underlying function and we don't know if the candidate solutions are near or far from the global optimums. The PSO algorithm is a population based search technique. That is, it represents the state of the algorithm by a population which is iteratively modified until a termination criterion is satisfied. First introduced by Eberhart and Kennedy in 1995 [130, 131], this algorithm is inspired by the behaviour of groups of some species of animals such as flocks of birds and schools of fish, where a desired position such as a food source is located by team work.

Suppose we have a d-dimensional function $f(\vec{x}) = f(x_1, x_2, ..., x_d)$ and we want to find the global minimum or maximum of this function. In the PSO algorithm a group of particles, called a swarm is considered in the search space of the problem. Here, particles are conceptual entities. Each particle evaluates the function at its current location. Each particle then determines its movement through the search space based on its own current and previous best position and the best position of other members of the swarm. To each particle three d-dimensional vectors are assigned: its position $\vec{x} = (x_1, x_2, ..., x_d)$, the best position it has found $\vec{x_\ell} = (x_{\ell 1}, x_{\ell 2}, ..., x_{\ell d})$ and its velocity $\vec{v} = (v_1, ..., v_d)$. The initial position of the particles and the initial velocity are chosen randomly. The velocity $\vec{v}$ and position $\vec{x}$ of each particle is updated to $\vec{v}'$ and $\vec{x}'$ based on its previous best position $\vec{x_\ell}$, and the best position found so far in the entire swarm $\vec{x_g}$ as [132]

$$\begin{aligned} \vec{x}' &= \vec{x} + \vec{v}', \\ \vec{v}' &= \chi \left[\vec{v} + c_g r_g \left(\vec{x_g} - \vec{x} \right) + c_\ell r_\ell \left(\vec{x_\ell} - \vec{x} \right) \right]. \end{aligned} \tag{A.1}$$

Here r_g and r_l are uniform random numbers in the interval [0, 1] generated at each step. The constant χ is known as a constriction factor and is determined from the other two constants c_g and c_ℓ according to [132]

$$\chi = \frac{2}{\left| 2 - c - \sqrt{c^2 - 4c} \right|}, \qquad c = c_g + c_\ell. \tag{A.2}$$

To ensure quick convergence of the algorithm the constant c is taken to be 4.1, and for simplicity c_g and c_ℓ are normally taken to be equal and therefore $c_\ell = c_g = 2.05$ [132]. After some number of repetitions the whole swarm is collectively merging to the optimum of the function, in a similar way as a school of fish merge for food. The number of repetitions is either set by the user or it is determined by a termination criterion.

Each of the terms in Eq. (A.1) plays a role in the PSO algorithm. The first term in Eq. (A.1) keeps the particle moving in the same direction as it was moving. The second term makes the particle to move towards the region of the search space in which it has found its best value. The third term moves the particle towards the region of the space in which the swarm has found its best value. The random values r_g and r_ℓ gives a stochastic nature to the velocity updating formula. This causes the particle to move semi-randomly towards the particles previous best position and the swarm's best position.

To stop the particles from going beyond the search space, the maximum of the velocity needs to be restricted. For a search space bounded by $[x_{\min}, x_{\max}]$, normally $v_{\max}$ is taken to be $v_{\max} = k(x_{\max} - x_{\min})/2$ with $0.1 \leq k \leq 1$. The number of particles is usually chosen between 10 and 50. In our algorithms we used 10 particles and we did not achieve enhancement by increasing the number of particles up to 40.

The PSO algorithm has been used extensively in solving optimisation problems in a range of fields including parameter estimation and quantum metrology [60, 61, 94, 95, 133].

References

[1] G. L. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde. *Entanglement-enhanced measurement of a completely unknown optical phase*. Nat. Photon. **5**, 43 (2011).

[2] H. Bernien. *Control, measurement and entanglement of remote quantum spin registers in diamond*. PhD thesis, Delft, University of Technology (2014).

[3] M. J. Holland and K. Burnett. *Interferometric detection of optical phase shifts at the Heisenberg limit*. Phys. Rev. Lett. **71**, 1355 (1993).

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

[5] V. Giovannetti, S. Lloyd, and L. Maccone. *Advances in quantum metrology*. Nat. Photon. **5**, 222 (2011).

[6] C. M. Caves. *Quantum-mechanical noise in an interferometer*. Phys. Rev. D **23**, 1693 (1981).

[7] J. Aasi *et al. Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light*. Nat. Photon. **7**, 613 (2013).

[8] The LIGO Scientific Collaboration. *A gravitational wave observatory operating beyond the quantum shot-noise limit*. Nat. Phys. **7**, 962 (2011).

[9] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker. *LIGO: The laser interferometer gravitational-wave observatory*. Science **256**, 325 (1992).

[10] K. Goda, O. Miyakawa, E. E. Mikhailov, S. Saraf, R. Adhikari, K. McKenzie, R. Ward, S. Vass, A. J. Weinstein, and N. Mavalvala. *A quantum-enhanced prototype gravitational-wave detector*. Nat. Phys. **4**, 472 (2008).

[11] The LIGO Scientific Collaboration. *Observation of gravitational waves from a binary black hole merger*. Phys. Rev. Lett. **116**, 061102 (2016).

[12] R. X. Adhikari. *Gravitational radiation detection with laser interferometry*. Rev. Mod. Phys. **86**, 121 (2014).

[13] M. G. A. Paris. *Canonical quantum phase variable*. Il Nuovo Cimento B **111**, 1151 (1996).

[14] M. G. A. Paris. *Quantum estimation for quantum technology*. International Journal of Quantum Information **7**, 125 (2009).

[15] L. Pezzé and A. Smerzi. *Quantum theory of phase estimation*. Proc. Int. School of Physics Enrico Fermi, Course 188, Varenna ed G. M. Tino and M. A. Kasevich (Amsterdam: IOS) p. 691 (2014).

[16] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński. *Quantum limits in optical interferometry*. Progress in Optics **60**, 345 (2015).

[17] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).

[18] D. W. Berry. PhD thesis, The University of Queensland, e-print quant-ph/0202136 (2001).

[19] H. M. Wiseman and R. B. Killip. *Adaptive single-shot phase measurements: A semiclassical approach*. Phys. Rev. A **56**, 944 (1997).

[20] G. W. Forbes and M. A. Alonso. *Measures of spread for periodic distributions and the associated uncertainty relations*. Am. J. Phys. **69**, 340 (2001).

[21] M. J. Collet. *Phase noise in squeezed states*. Phys. Scr. **T48**, 124 (1993).

[22] A. S. Holevo. *Covariant measurements and imprimitivity systems*. Lect. Notes Math. **1055**, 153 (1984).

[23] D. W. Berry, M. J. W. Hall, M. Zwierz, and H. M. Wiseman. *Optimal Heisenberg-style bounds for the average performance of arbitrary phase estimates*. Phys. Rev. A **86**, 053813 (2012).

[24] G. A. Durkin and J. P. Dowling. *Local and global distinguishability in quantum interferometry*. Phys. Rev. Lett. **99**, 070801 (2007).

[25] Y. R. Zhang, G. R. Jin, J. P. Cao, W. M. Liu, and H. Fan. *Unbounded quantum fisher information in two-path interferometry with finite photon number*. J. Phys. A **46**, 035302 (2013).

[26] D. Berry, B. L. Higgins, S. D. Bartlett, M. Mitchell, G. J. Pryde, and H. M. Wiseman. *How to perform the most accurate possible phase measurements*. Phys. Rev. A **80**, 052114 (2009).

[27] Y. C. Eldar. *Uniformly improving the Cramér-Rao bound and maximum-likelihood estimation*. IEEE Trans. Signal Process. **54**, 2943 (2006).

[28] S. L. Braunstein and C. M. Caves. *Statistical distance and the geometry of quantum states*. Phys. Rev. Lett. **72**, 3439 (1984).

[29] S. L. Braunstein, C. M. Caves, and G. J. Milburn. *Generalized uncertainty relations: Theory, examples, and Lorentz invariance*. Annals of Physics **247**, 135 (1996).

[30] W. Heisenberg. *Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik*. Zeitschrift für Physik **43**, 172 (1927).

[31] J. A. Wheeler and H. Zurek. *Quantum theory and measurement* (Princeton University Press, 1983).

[32] H. P. Robertson. *The uncertainty principle*. Phys. Rev. **34**, 163 (1929).

[33] L. Mandelstam and I. Tamm. *The uncertainty relation between energy and time in nonrelativistic quantum mechanics*. J. Phys. USSR **9**, 249 (1945).

[34] D. W. Berry, H. M. Wiseman, and J. K. Breslin. *Optimal input states and feedback for interferometric phase estimation*. Phys. Rev. A **63**, 053804 (2001).

[35] D. W. Berry and H. M. Wiseman. *Optimal states and almost optimal adaptive measurements for quantum interferometry*. Phys. Rev. Lett. **85**, 5098 (2000).

[36] T. Bayes and R. Price. *An essay towards solving a problem in the doctrine of chances*. Phil. Trans. R. Soc. Lond. **53**, 370 (1763).

[37] L. L. Cam. *Asymptotic Methods in Statistical Decision Theory* (Springer, Berlin, 1986).

[38] C. W. Helstrom. *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[39] H. Lee, P. Kok, and J. P. Dowling. *A quantum Rosetta stone for interferometry*. J. Mod. Opt. **49**, 2325 (2002).

[40] H. F. Hofmann. *All path-symmetric pure states achieve their maximal phase sensitivity in conventional two-path interferometry*. Phys. Rev. A **79**, 033822 (2009).

[41] B. C. Sanders. *Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement*. Phys. Rev. A **40**, 2417 (1989).

[42] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling. *Quantum interferometric optical lithography: Exploiting entanglement to beat the diffraction limit*. Phys. Rev. Lett. **85**, 2733 (2000).

[43] J. P. Dowling. *Quantum optical metrology*. Contemp. Phys. **49**, 125 (2008).

[44] V. Giovannetti, S. Lloyd, and L. Maccone. *Quantum metrology*. Phys. Rev. Lett. **96**, 010401 (2006).

[45] M. A. Rubin and S. Kaushik. *Loss-induced limits to phase measurement precision with maximally entangled states*. Phys. Rev. A **75**, 053805 (2007).

[46] G. Gilbert and Y. S. Weinstein. *Aspects of practical remote quantum sensing*. J. Mod. Opt. **55**, 3283 (2008).

[47] M. W. Mitchell. *Metrology with entangled states*. Proc. SPIE **5893**, 589310 (2005).

[48] L. Pezzé and A. Smerzi. *Sub shot-noise interferometric phase sensitivity with beryllium ions Schrödinger cat states*. Europhys. Lett. **78**, 30004 (2007).

[49] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde. *Entanglement-free Heisenberg-limited phase estimation*. Nature **450**, 393 (2007).

[50] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde. *Demonstrating Heisenberg-limited unambiguous phase estimation without adaptive measurements*. New. J. Phys. **11**, 073023 (2009).

[51] R. Griffiths and C.-S. Niu. *Semiclassical Fourier transform for quantum computation*. Phys. Rev. Lett. **76**, 3228 (1996).

[52] B. L. Higgins. *Quantum measurement: Concepts, algorithms, and experiments in photonic quantum information science*. PhD thesis, Griffiths University (2001).

[53] G. Giedke, J. M. Taylor, D. DAlessandro, M. D. Lukin, and A. Imamoğlu. *Quantum measurement of a mesoscopic spin ensemble*. Phys. Rev. A **74**, 032316 (2006).

[54] R. S. Said, D. W. Berry, and J. Twamley. *Nanoscale magnetometry using a single-spin system in diamond*. Phys. Rev. B **83**, 125410 (2011).

[55] G. Waldherr, J. Beck, P. Neumann, R. S. Said, M. Nitsche, M. L. Markham, D. J. Twitchen, J. Twamley, F. Jelezko, and J. Wrachtrup. *High-dynamic-range magnetometry with a single nuclear spin in diamond*. Nat. Nanotech. **7**, 105 (2012).

[56] N. M. Nusran, M. U. Momeen, and M. V. G. Dutt. *High-dynamic-range magnetometry with a single electronic spin in diamond*. Nat. Nanotech. **7**, 109 (2012).

[57] N. M. Nusran and M. V. G. Dutt. *Optimizing phase estimation algorithm for diamond spin magnetometry*. Phys. Rev. B **90**, 024422 (2014).

[58] S. A. Diddams, J. C. Bergquist, S. R. Jefferts, and C. W. Oates. *Standards of time and frequency at the outset of the 21st century*. Science **306**, 1318 (2004).

[59] D. Budker and M. Romalis. *Optical magnetometry*. Nat. Phys. **3**, 227 (2007).

[60] A. J. F. Hayes and D. W. Berry. *Swarm optimization for adaptive phase measurements with low visibility*. Phys. Rev. A **89**, 013838 (2014).

[61] C. Bonato, M. S. Block, H. T. Dinani, D. W. Berry, L. Markham, D. J. Twichen, and R. Hanson. *Optimized quantum sensing with a single electron spin using real-time adaptive measurements*. Nat. Nanotech. **11**, 247 (2016).

[62] M. G. Genoni, S. Mancini, and A. Serafini. *General-dyne unravelling of a thermal master equation*. Russ. J. Math. Phys. **21**, 329 (2014).

[63] H. M. Wiseman and G. J. Milburn. *Quantum measurement and control* (Cambridge University Press, 2010).

[64] D. W. Berry and H. M. Wiseman. *Adaptive quantum measurements of a continuously varying phase*. Phys. Rev. A **65**, 043803 (2002).

[65] D. T. Pope, H. M. Wiseman, and N. K. Langford. *Adaptive phase estimation is more accurate than nonadaptive phase estimation for continuous beams of light*. Phys. Rev. A **70**, 043812 (2004).

[66] S. Olivares and M. G. A. Paris. *Bayesian estimation in homodyne interferometry*. J. Phys. B **42**, 055506 (2009).

[67] A. A. Berni, T. Gehring, M. G. A. Paris, B. M. Nielsen, V. Handchen, and M. G. A. Paris. *Ab initio quantum-enhanced optical phase estimation using real time feedback control*. Nat. Photon. **9**, 577 (2015).

[68] H. W. Wiseman. *Adaptive phase measurements of optical modes: Going beyond the marginal Q distribution*. Phys. Rev. Lett. **75**, 4587 (1995).

[69] H. M. Wiseman and R. B. Killip. *Adaptive single-shot phase measurements: The full quantum theory*. Phys. Rev. A **57**, 2169 (1998).

[70] D. W. Berry and H. M. Wiseman. *Adaptive phase measurements for narrowband squeezed beams*. Phys. Rev. A **73**, 063824 (2006).

[71] D. W. Berry and H. M. Wiseman. *Erratum: Adaptive phase measurements for narrowband squeezed beams [Phys. Rev. A 73, 063824 (2006)]*. Phys. Rev. A **87**, 019901(E) (2013).

[72] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi. *Adaptive homodyne measurement of optical phase*. Phys. Rev. Lett. **89**, 133602 (2002).

[73] N. Gkortsilas, J. J. Cooper, and J. A. Dunningham. *Measuring a completely unknown phase with sub-shot-noise precision in the presence of loss*. Phys. Rev. A **85**, 063827 (2012).

[74] J. Joo, W. J. Munro, and T. P. Spiller. *Quantum metrology with entangled coherent states*. Phys. Rev. Lett. **107**, 083601 (2012).

[75] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel. *Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector geo 600*. Phys. Rev. A **88**, 041802(R) (2013).

[76] U. Dorner, R. Demkowicz-Dobrzański, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley. *Optimal quantum phase estimation*. Phys. Rev. Lett. **80**, 013825 (2009).

[77] R. Demkowicz-Dobrzański, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley. *Quantum phase estimation with lossy interferometers*. Phys. Rev. A **80**, 013825 (2009).

[78] P. A. Knott, W. J. Munro, and J. A. Dunningham. *Attaining subclassical metrology in lossy systems with entangled coherent states*. Phys. Rev. A **89**, 053812 (2014).

[79] M. Kacprowicz, R. Demkowicz-Dobrzański, W. Wasilewski, K. Banaszek, and I. A. Walmsley. *Experimental quantum-enhanced estimation of a lossy phase shift*. Nat. Photon. **4**, 357 (2010).

[80] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin. *Scaling laws for precision in quantum interferometry and the bifurcation landscape of the optimal state*. Phys. Rev. A **83**, 021804(R) (2011).

[81] S. Knysh, E. H. Chen, and G. A. Durkin. *True limits to precision via unique quantum probe*. arXiv:1402.0495v1 .

[82] K. Jiang, C. J. Brignac, Y. Weng, M. B. Kim, H. Lee, and J. P. Dowling. *Strategies for choosing path-entangled number states for optimal robust quantum-optical metrology in the presence of loss*. Phys. Rev. A **86**, 013826 (2012).

[83] T. W. Lee, S. D. Huver, H. Lee, L. Kaplan, C. M. S. B. McCracken, D. B. Uskov, C. F. Wildfeuer, G. Veronis, and J. P. Dowling. *Optimization of quantum interferometric metrological sensors in the presence of photon loss*. Phys. Rev. A **80**, 063803 (2009).

[84] B. M. Escher, R. L. de Matos Filho, and L. Davidovich. *General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology*. Nat. Phys. **7**, 406 (2011).

[85] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guta. *The elusive Heisenberg limit in quantum-enhanced metrology*. Nat. Commun. **3**, 1063 (2012).

[86] H. Lee, P. Kok, N. J. Cerf, and J. P. Dowling. *Linear optics and projective measurements alone suffice to create large-photon-number path entanglement*. Phys. Rev. A **65**, 030101(R) (2002).

[87] J. Fiurášek. *Conditional generation of N-photon entangled states of light*. Phys. Rev. A **65**, 053818 (2002).

[88] G. J. Pryde and A. G. White. *Creation of maximally entangled photon-number states using optical fiber multiports*. Phys. Rev. A **68**, 052315 (2003).

[89] N. M. VanMeter, P. Lougovski, D. B. Uskov, K. Kieling, J. Eisert, and J. P. Dowling. *General linear-optical quantum state generation scheme: Applications to maximally path-entangled states*. Phys. Rev. A **76**, 063808 (2007).

[90] X. Zou, K. Pahlke, and W. Mathis. *Generation of entangled states of two traveling modes for fixed number of photon*. arXiv:quant-ph/0110149v1 .

[91] L. Pezzé and A. Smerzi. *Sub shot-noise interferometric phase sensitivity with beryllium ions Schrödinger cat states*. Euro. Phys. Lett. **78**, 30004 (2007).

[92] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. *Experimental realization of any discrete unitary operator*. Phys. Rev. Lett. **73**, 58 (1994).

[93] Z. Y. J. Ou. *Multi-Photon Quantum Interference* (Springer, New York, 2007).

[94] A. Hentschel and B. C. Sanders. *Machine learning for precise quantum measurement*. Phys. Rev. Lett. **104**, 063603 (2010).

[95] A. Hentschel and B. C. Sanders. *Efficient algorithm for optimizing adaptive quantum metrology processes*. Phys. Rev. Lett. **107**, 233601 (2011).

[96] H. T. Dinani and D. W. Berry. *Loss-resistant unambiguous phase measurement*. Phys. Rev. A **90**, 023856 (2014).

[97] S. Scheel. *Decoherence of nonclassical states of light can be very efficiently used in optical sensor technology*. J. Mod. Opt. **50**, 1327 (2003).

[98] R. Whittaker, C. Erven, M. B. A. Neville, J. L. O'Brien, H. Cable, and J. C. F. Matthews. *Quantum enhanced absorption spectroscopy*. arXiv:1508.00849 (2015).

[99] J. D. Jackson. *Classical electrodynamics* (Wiley, 1999).

[100] P. J. D. Crowley, A. Datta, M. Barbieri, and I. A. Walmsley. *Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry*. Phys. Rev. A **89**, 023845 (2014).

[101] P. Kok and B. W. Lovett. *Introduction to optical quantum information processing* (Cambridge University Press, 2010).

[102] R. G. Beausoleil, W. J. Munro, and T. P. Spiller. *Applications of coherent population transfer to quantum information processing*. J. Mod. Opt. **51**, 1559 (2004).

[103] D. A. Steck. *Sodium d line data http://steck.us/alkalidata/sodiumnumbers.pdf* (2001).

[104] J. R. Jeffers, N. Imoto, and R. Loudon. *Quantum optics of traveling-wave attenuators and amplifiers*. Phys. Rev. A **47**, 3346 (1993).

[105] R. Loudon. *The quantum theory of light* (Oxford University Press, 2001).

[106] C. A. Ryan, J. S. Hodges, and D. G. Cory. *Robust decoupling techniques to extend quantum coherence in diamond*. Phys. Rev. Lett. **105**, 200402 (2010).

[107] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. M. J. Isoya, J. Achard, J. Beck, J. Tissler, and V. Jacques. *Ultralong spin coherence time in isotopically engineered diamond*. Nat. Mater. **8**, 383 (2009).

[108] C. Kurtsiefer, S. Mayer, P. Zarda, and H. Weinfurter. *Stable solid-state source of single photons*. Phys. Rev. Lett. **85**, 290 (2000).

[109] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup. *Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate*. Phys. Rev. Lett. **93**, 130501 (2004).

[110] F. Dolde, H. Fedder, M. W. Doherty, T. Nobauer, F. Rempp, G. Balasubramanian, T. Wolf, F. Reinhard, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup. *Electric-field sensing using single diamond spins*. Nat. Phys. **7**, 459 (2011).

[111] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. G. Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin. *Nanoscale magnetic sensing with an individual electronic spin in diamond*. Nature **455**, 644 (2008).

[112] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin. *High-sensitivity diamond magnetometer with nanoscale resolution*. Nat. Phys. **4**, 810 (2008).

[113] V. M. Acosta, E. Bauch, M. P. Ledbetter, A. Waxman, L.-S. Bouchard, and D. Budker. *Temperature dependence of the nitrogen-vacancy magnetic resonance in diamond*. Phys. Rev. Lett. **104**, 070801 (2010).

[114] M. W. Doherty, N. B. Mansonb, P. Delaney, F. Jelezko, J. Wrachtrupe, and L. C. L. Hollenberg. *The nitrogen-vacancy colour centre in diamond*. Phys. Rep. **528**, 1 (2013).

[115] N. B. Manson, J. P. Harrison, and M. J. Sellars. *Nitrogen-vacancy center in diamond: Model of the electronic structure and associated dynamics*. Phys. Rev. B **74**, 104303 (2006).

[116] J. R. Maze, A. Dreau, V. Waselowski, H. Duarte, J.-F. Roch, and V. Jacques. *Free induction decay of single spins in diamond*. New J. Phys. **14**, 103041 (2012).

[117] P. Cappellaro. *Spin-bath narrowing with adaptive parameter estimation*. Phys. Rev. A **85**, 030301 (2012).

[118] M. G. Genoni, S. Olivares, and M. G. A. Paris. *Optical phase estimation in the presence of phase diffusion*. Phys. Rev. Lett. **106**, 153603 (2011).

[119] M. D. Vidrighin, G. Donati, M. G. Genoni, X. M. Jin, W. S. Kolthammer, M. Kim, A. Datta, M. Barbieri, and I. A. Walmsley. *Joint estimation of phase and phase diffusion for quantum metrology*. Nat. Commun. **5**, 3532 (2014).

[120] D. W. Berry, M. J. W. Hall, and H. M. Wiseman. *Stochastic Heisenberg limit: Optimal estimation of a fluctuating phase*. Phys. Rev. Lett. **111**, 113601 (2013).

[121] D. W. Berry, M. Tsang, M. J. Hall, and H. M. Wiseman. *Quantum Bell-Ziv-Zakai bounds and Heisenberg limits for waveform estimation*. Phys. Rev. X **5**, 031018 (2015).

[122] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. Berry, T. C. Ralph, H. M.Wiseman, E. H. Huntington, and A. Furusawa. *Quantum-enhanced optical-phase tracking*. Science **337**, 1514 (2012).

[123] D. F. Walls and G. J. Milburn. *Quantum optics* (Springer, 2008).

[124] M. J. Collett and C. W. Gardiner. *Squeezing of intracavity and traveling wave light fields produced in parametric amplification*. Phys. Rev. A **30**, 1386 (1984).

[125] C. W. Gardiner and P. Zoller. *Quantum noise* (Springer Berlin, 2000).

[126] P. D. Drummond and Z. Fizek. *Quantum squeezing* (Springer Verlag Berlin, 2004).

[127] N. J. Kasdin. *Discrete simulation of colored noise and stochastic processes and* $1/f^{\alpha}$ *power law noise generation*. Proceedings of the IEEE **83**, 802 (1995).

[128] H. L. V. Trees. *Detection, Estimation, and Modulation Theory, Part II: Nonlinear Modulation Theory* (John Wiley and Sons New York, 2003).

[129] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi. *Fisher information and multipartite entanglement*. Phys. Rev. Lett. **85**, 022321 (2012).

[130] J. Kennedy and R. Eberhart. *Particle swarm optimization*. Proceedings of the 1995 IEEE International Conference on Neural Networks (Perth, Australia): IEEE Service Center, Piscataway, NJ, IV p. 1942.

[131] R. Eberhart and J. Kennedy. *A new optimizer using particle swarm theory*. Proceedings of the Sixth International Symposium on Micro Machine and Human Science (Nagoya, Japan): IEEE Service Center, Piscataway, NJ p. 39.

[132] D. Bratton and J. Kennedy. *Defining a standard for particle swarm optimization*. Proceedings of the 2007 IEEE Swarm Intelligence Symposium p. 120 (2007).

[133] A. Sergeevich and S. D. Bartlett. *Optimizing qubit Hamiltonian parameter estimation algorithm using PSO*. Proceedings of 2012 IEEE Conference on Evolutionary Computation (CEC), arXiv:1206.3830 .