ADAPTIVE PHASE MEASUREMENTS

By

Dominic William Berry

A thesis submitted to the University of Queensland
for the degree of Doctor of Philosophy
Department of physics
December 2001

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.

Dominic William Berry

Acknowledgments

Primary acknowledgments go to my PhD supervisor, Dr. Howard Wiseman, who took me on as a student and continued to come to the university for weekly discussions, despite having moved to Griffith University. He originated the topic of this project, as well as directing me to the most fruitful areas for study. He also originated the general ideas behind some of the approaches I used in this project. I am grateful for his untiring help with the drafts and revisions of my papers, as well as this thesis. Also his advice on publishing and applying for postdoctoral work was invaluable.

I am grateful to Zhong-Xi Zhang and John Breslin, who contributed valuable insight to the method of solution of some of the problems considered here. Thanks also go to the Australian Research Council, that supported this research financially through the Australian Postgraduate Awards. I give thanks to our system administrator, Ian Mortimer, for enabling me to use the department's computing resources, as well as Prof. Peter Drummond for allowing me some computing time on the quantum optics cluster.

List of Publications

In this chronological list of publications, those on which this thesis is based are asterisked.

1. D. W. Berry, N. R. Heckenberg and H. Rubinsztein-Dunlop, "Effects associated with bubble formation in optical trapping" J. Mod. Opt. **47**, 1575 (2000).

2. * D. W. Berry, H. M. Wiseman and Zhong-Xi Zhang, "Heterodyne and adaptive phase measurements on states of fixed mean photon number" Phys. Rev. A **60**, 2458 (1999).

3. * D. W. Berry and H. M. Wiseman, "Phase measurements at the theoretical limit" Phys. Rev. A **63**, 013813 (2000).

4. * D. W. Berry and H. M. Wiseman, "The effects of time delays in adaptive phase measurements" J. Mod. Opt. **48**, 797 (2001).

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

6. * 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).

7. * D. W. Berry and H. M. Wiseman, "Adaptive quantum measurements of a continuously-varying phase" Phys. Rev. A (to be published).

Abstract

In this thesis I consider the general problem of how to make the best possible phase measurements using feedback. Both the optimum input state and optimum feedback are considered for both single-mode dyne[1] measurements and two-mode interferometric measurements. I derive the optimum input states under general dyne measurements when the mean photon number is fixed, both for general states and squeezed states. I propose a new feedback scheme that introduces far less phase uncertainty than mark II feedback, and is very close to the theoretical limit. I also derive results for the phase variance when there is a time delay in the feedback loop, showing that there is a lower limit to the introduced phase variance, and this is approached quite accurately under some conditions. I derive the optimum input states for interferometry, showing that the phase uncertainty scales as N^{-1} for all the common measures of uncertainty. This is contrasted with the $|j0\rangle_z$ state, which does not scale as N^{-1} for all measures of phase uncertainty. I introduce an adaptive feedback scheme that is very close to optimum, and can give scaling very close to N^{-1} for the uncertainty. Lastly I consider the case of continuous measurements, for both the dyne and interferometric cases.

[1] Dyne measurements are those based on continuous measurement of field quadratures, including heterodyne, homodyne and more general adaptive measurements.

List of Symbols

The following list is neither exhaustive nor exclusive, but may be helpful. It generally contains only those symbols used frequently or in more than one chapter.

a	annihilation operator for an optical mode
$a^\dagger$	creation operator for an optical mode
$\|n\rangle$	a Fock number state with n photons
φ	the system phase
ϕ	a phase estimate
$F(\ldots)$	the POM for $\ldots$
H_{nm}	the matrix giving the POM as in Eq. (1.33)
$V(\phi)$	the Holevo phase variance
$V_\varphi(\phi)$	a modified Holevo phase variance given by Eq. (1.49)
$V'(\phi)$	an alternative measure of phase variance given by Eq. (2.3)
$h(n)$	$1 - H_{n,n+1}$
Φ	the local oscillator phase
$\hat{X}_\Phi$	the Φ quadrature operator
t	unscaled time, in the interval $[0, \infty)$
$I(t)$	the unscaled difference photocurrent
v	scaled time, in the interval $[0, 1)$
$I(v)$	the scaled difference photocurrent
A_v	defined in Eq. (1.65)
B_v	defined in Eq. (1.66)
C_v	defined in Eq. (1.81)
A, B, C	the values of these variables at $v = 1$
dW	a Wiener stochastic increment
$\hat{\varphi}$	an intermediate phase estimate used for feedback (for dyne measurements)
$\hat{\varphi}$	the optimal phase estimate (for interferometry)
c, p	the constants in the asymptotic scaling for $h(n)$
α	coherent amplitude
η	photodetector efficiency
$\overline{n}$	mean photon number
z_1	the first zero of the Airy function, about -2.338
ζ	the squeezing parameter for squeezed states
n_0	given by $\overline{n}e^{2\zeta}$
Δ	a constant for optimised squeezed states, about 2.427
Δ'	an alternative constant, about 0.927
α^{P}	the coherent amplitude for the squeezed state in the POM
ζ^{P}	the squeezing parameter for the squeezed state in the POM
ξ^{P}	the squeezing parameter relative to the phase of α^{P}
$\overline{n}^{\mathrm{P}}$	the mean photon number for the squeezed state in the POM
$Q(A, B)$	the ostensible probability distribution for A and B
ε	defined in Eq. (3.30)
γ	the local oscillator amplitude
$A_t^{\mathrm{S}}, B_t^{\mathrm{S}}$	squeezing parameters defined analogously to A_v and B_v by Eq. (3.109)

α_v^{P}	an intermediate value of α^{P}, defined in Eq. (3.200)
ζ_v^{P}	an intermediate value of ζ^{P}, defined in Eq. (3.201)
ξ_v^{P}	the intermediate squeezing parameter relative to the phase of α_v^{P}
λ	the variable controlling when the corrections for dyne measurements are used
α^f	the value of α found by fitting to the data
τ	the feedback time delay
$\hat{J}_x, \hat{J}_y, \hat{J}_z, \hat{J}^2$	operators in the Schwinger representation
$\|j\mu\rangle_z$	the common eigenstate of $\hat{J}_x$ and $\hat{J}^2$
$\|j\mu\rangle_y$	the common eigenstate of $\hat{J}_y$ and $\hat{J}^2$
$I^j_{\mu,\nu}(\pi/2)$	interferometer matrix elements
$P_n^{(\alpha,\beta)}$	Jacobi polynomials
H_n	Hermite polynomials
$V_\pi(\phi)$	the modified Holevo variance for phase modulo π
$\Delta\phi$	the square root of the standard variance
$\Delta\phi_{\mathrm{H}}$	the square root of the Holevo variance
L_{H}	the entropic length
L_{F}	the Fisher length
L_C	the $C \times 100\%$ confidence interval
L_{rp}	the reciprocal-of-peak-value
L_{S}	the Süssman measure
u_m	the result of the mth detection for interferometry
n_m	the sequence of m detection results $u_m \dots u_2 u_1$
Φ_m	the feedback phase before the mth detection for interferometry
$\psi_{\mu,m,k}(n_m)$	the coefficients for the system state after m detections
$\theta(t)$	the phase estimate based on the infinitesimal data in the interval $[t, t+dt)$
$\Theta(t)$	the phase estimate based on the data up till time t
κ	the diffusion coefficient for the system phase
χ	the scaling constant for the exponential weighting
K, X	the scaled values of κ and χ

List of Abbreviations

POM	Probability operator measure
POVM	Positive-operator-valued measure
BS	Beam splitter
PD	Photodetector
EOM	Electro-optic phase modulator
SDE	Stochastic differential equation

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this study I consider the general problem of how to most efficiently measure phase. This is a very important subject, as many high precision measurements are based upon measurements of phase. In particular, the current search for gravitational waves requires extremely accurate phase measurements, where new approaches that surpass the standard quantum limit may be necessary to obtain useful results [1].

Unlike most other quantities that we would wish to measure, it is not possible to measure phase directly. We must measure phase indirectly, and this almost always introduces an extra uncertainty beyond the intrinsic uncertainty in the phase of the mode. In general, it is possible to improve measurements by introducing an auxiliary phase shift. In the case of homodyne measurements, this phase shift would be based upon the previous knowledge about the phase. In this study I consider the case that the phase is unknown, and instead the auxiliary phase is adjusted based on the data obtained during the measurement. I investigate schemes for adjusting the auxiliary phase so as to introduce the minimum possible phase uncertainty.

Another aspect of the problem of efficiently measuring the phase is the state itself. Every state with finite energy will have some intrinsic uncertainty in the phase. We would wish to optimise the state so that it has the minimum possible intrinsic phase uncertainty, or alternatively so that it gives the minimum phase uncertainty for some specific phase measurement scheme.

Before I discuss these problems, I will briefly review the theory behind the description of phase and phase measurements.

1.1 The Description of Phase

Classically there is no ambiguity in the definition of phase. A general propagating sinusoidal wave can be described by

$$\psi(x,t) = A\sin(kx - \omega t + \phi), \tag{1.1}$$

where A is the amplitude and ϕ is the phase. The phase is not a quantity that can be measured directly; however, it is just a real number and can be determined unambiguously from the variation of $\psi(x,t)$ (provided the amplitude is constant and nonzero).

For a monochromatic electromagnetic field propagating in the $\mathbf{k}$ direction, we can express the electric field as

$$\mathbf{E}(\mathbf{R},t) = \mathbf{E}_0\sin(\mathbf{k}\cdot\mathbf{R} - \omega t + \phi). \tag{1.2}$$

In quantising the electric field in the Heisenberg picture, the operator for the electric field is

$$\hat{\mathbf{E}}(\mathbf{R},t) = \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}}\left(\varepsilon e^{i(\mathbf{k}\cdot\mathbf{R}-\omega t)}a + \varepsilon^* e^{-i(\mathbf{k}\cdot\mathbf{R}-\omega t)}a^\dagger\right), \tag{1.3}$$

where $\hbar$ is Planck's constant divided by 2π, ϵ_0 is the dielectric permittivity for a vacuum, V is the quantisation volume, ε is the polarisation vector and a and $a^\dagger$ are the annihilation and creation

operators respectively. The variation of the electric field in time and space, as well as the polarisation direction are contained in this field operator, but the amplitude and phase of the field are contained in the state.

The annihilation and creation operators have the commutation relations

$$\left[a, a^\dagger\right] = 1, \tag{1.4}$$

$$[a, a] = \left[a^\dagger, a^\dagger\right] = 0. \tag{1.5}$$

I will generally use the hat over variables to indicate operators when they are used in both an operator and non-operator sense (like $\mathbf{E}$), but not when the variable is used only as an operator (like a).

The quantum mechanical decription of phase was first considered by London [2, 3] and Dirac [4]. In general, in quantum mechanics we wish to represent physical quantities by Hermitian operators. Unfortunately, this approach produces difficulties when applied to phase. For example, consider the result for a coherent state

$$\langle a \rangle = e^{i\phi}\sqrt{N}. \tag{1.6}$$

From this, an obvious way of defining an operator for the phase is by

$$a = e^{i\hat{\phi}}\sqrt{\hat{N}}, \tag{1.7}$$

where $\hat{N}$ is the usual operator for the photon number,

$$\hat{N} = a^\dagger a. \tag{1.8}$$

The phase operator defined in this way is equivalent to that considered by Dirac [4]. It is easily seen that the phase operator defined in this way is not Hermitian.

This definition of the phase operator produces other problems. For example, Dirac derives the commutation relation

$$[\hat{N}, \hat{\phi}] = i. \tag{1.9}$$

This would seem to imply the uncertainty relation [5]

$$\Delta N \Delta\phi \geq \tfrac{1}{2}. \tag{1.10}$$

This uncertainty relation does not make sense, because the uncertainty in N can become arbitrarily small, whereas the uncertainty in ϕ can not be over π. In fact, this uncertainty relation is not necessarily implied by Eq. (1.9), because the phase operator is not Hermitian. Another problem is that, for a number state, the expectation value of the commutator should be zero [6], rather than i.

One resolution of this problem is that although it is not possible to define a Hermitian phase operator, it is possible to define Hermitian sine and cosine operators that give valid uncertainty relations [6, 7]. These Susskind-Glogower operators satisfy the commutation relations

$$\left[\widehat{\cos\phi}, \hat{N}\right] = i\,\widehat{\sin\phi} \tag{1.11}$$

$$\left[\widehat{\sin\phi}, \hat{N}\right] = -i\,\widehat{\cos\phi}. \tag{1.12}$$

The corresponding uncertainty relations are

$$\Delta N \Delta\cos\phi \geq \tfrac{1}{2}\left|\left\langle\widehat{\sin\phi}\right\rangle\right| \tag{1.13}$$

$$\Delta N \Delta\sin\phi \geq \tfrac{1}{2}\left|\left\langle\widehat{\cos\phi}\right\rangle\right|. \tag{1.14}$$

These uncertainty relations make sense, because in the limit of small number uncertainty the expectation values on the right hand side also become small, so these uncertainty relations do not imply ridiculously large values of $\Delta\cos\phi$ or $\Delta\sin\phi$.

Unfortunately there are also problems with these operators. For example, we find that for the vacuum state $\langle \cos^2\theta \rangle = \frac{1}{4}$ [8]. As the vacuum state should have a uniform phase distribution, we would expect that $\langle \cos^2\theta \rangle$ is equal to $\frac{1}{2}$. In addition, using these operators we find $\langle \exp(ip\theta) \rangle = 0$ for all integers $p > 0$. This implies a uniform phase distribution, which is what we would expect, but is inconsistent with the result obtained for $\langle \cos^2\theta \rangle$.

Another approach to finding a Hermitian phase operator is the Pegg-Barnett formalism [9, 10, 11]. The basis of this formalism is to put an upper limit s on the photon number, then take the limit as s tends to infinity. The reference phase states are taken to be

$$|\theta_m\rangle_s = (s+1)^{-1/2} \sum_{n=0}^{s} \exp(in\theta_m)|n\rangle, \tag{1.15}$$

where

$$\theta_m = \theta_0 + \frac{2m\pi}{s+1}, \quad m = 0, 1, \ldots, s \tag{1.16}$$

and θ_0 is an arbitrary constant.

Then a Hermitian phase operator is defined by

$$\hat{\phi}_s = \sum_{m=0}^{s} \theta_m |\theta_m\rangle_{ss}\langle\theta_m|. \tag{1.17}$$

Note that for this operator the phase states $|\theta_m\rangle_s$ are clearly eigenstates with eigenvalues of θ_m. In terms of the number basis, the operator is

$$\hat{\phi}_s = \theta_0 + \frac{s\pi}{s+1} + \frac{2\pi}{s+1} \sum_{j\neq k}^{s} \frac{\exp\left[i(j-k)\theta_0\right] |j\rangle\langle k|}{\exp\left[i(j-k)2\pi/(s+1)\right] - 1}. \tag{1.18}$$

In the Pegg-Barnett formalism, the limit of $s \to \infty$ is taken after expectation values have been determined.

It is also possible to take the limit $s \to \infty$ of Eq. (1.18), which gives

$$\hat{\phi}_\infty = \theta_0 + \pi + \sum_{j\neq k}^{\infty} \frac{\exp\left[i(j-k)\theta_0\right] |j\rangle\langle k|}{i(j-k)}. \tag{1.19}$$

This operator was also considered before the Pegg-Barnett formalism was developed [12, 13, 14]. This would appear to be a Hermitian operator that can be used to describe phase; however, there are problems with this operator. This operator leads to different expectation values than given by the Pegg-Barnett formalism (where the expectation values are taken before the limit $s \to \infty$).

The problem is that (1.18) converges to (1.19) only weakly [15]. The weak limits of operators do not preserve the operator algebra, for example, the weak limit of $\hat{\phi}_s^2$ is not $\hat{\phi}_\infty^2$. One result of this is that, for the vacuum state, $\langle \hat{\phi}_\infty^2 \rangle = \pi^2/6$ [16, 17]. For a uniform phase distribution, the result should be $\pi^2/3$, which is that obtained from the limit of $\langle \hat{\phi}_s^2 \rangle$.

Fortunately, in this study we do not require an explicit phase operator, and usually all we need to know is the phase variance, or at most the probability distribution. For this what we want is the probability operator measure, or POM.

1.2 Probability Operator Measures

In quantum mechanical systems, the most general way of obtaining the probability of some measurement result E is by the expectation value of an operator $F(E)$, i.e.

$$P(E) = \mathrm{Tr}[\rho F(E)], \tag{1.20}$$

where ρ is the state matrix for the system. If the set of all possible measurement results is Ω, it is evident that $P(\Omega) = 1$ for all ρ, which implies that $F(\Omega) = 1$. Thus $F(E)$ can be called a probability

operator, and the mapping $E \mapsto F$ defines a probability operator measure (POM), sometimes also called a positive-operator-valued measure (POVM), on Ω [18, 19]. This method does not require a specific operator to represent the quantity that is being measured.

This is quite different to the simple method for pure states, where the probability is given by the square of the inner product between the initial and final states. To see the similarity between the two methods, recall that in the simple method we represent the physical quantity being measured by a Hermitian operator, for example $\hat{R}$. This has associated eigenvalues and eigenvectors r and $|\psi_r\rangle$. After a measurement yielding the result r the system is in state $|\psi_r\rangle$, and the probability for this result is $|\langle\psi_r|\psi\rangle|^2$.

Alternatively we can define the projection operators

$$\Pi_r = |\psi_r\rangle\langle\psi_r|. \tag{1.21}$$

The probability of obtaining the result r is then

$$\begin{aligned} P_r &= |\langle\psi_r|\psi\rangle|^2 \\ &= \langle\psi|\psi_r\rangle\langle\psi_r|\psi\rangle \\ &= \langle\psi|\Pi_r|\psi\rangle. \end{aligned} \tag{1.22}$$

The normalised state after the measurement is then

$$|\psi_r\rangle = \frac{\Pi_r|\psi\rangle}{\sqrt{P_r}}. \tag{1.23}$$

For more general measurements, we can replace the projection operator with a more general operator Ω_r, called the measurement operator. Then the probability is given by

$$P_r = \langle\psi|\Omega_r^\dagger\Omega_r|\psi\rangle. \tag{1.24}$$

Note that in this case $\Omega_r^\dagger\Omega_r$ does not necessarily simplify to Ω_r, as is the case for projection operators. The normalised state after the measurement is given by

$$|\psi_r\rangle = \frac{\Omega_r|\psi\rangle}{\sqrt{P_r}}. \tag{1.25}$$

For the most general case we replace $\Omega_r^\dagger\Omega_r$ in Eq. (1.24) [or Π_r in Eq. (1.22)] with a general operator F_r. The probability is then determined using

$$P_r = \langle\psi|F_r|\psi\rangle. \tag{1.26}$$

Clearly, the generalisation to this for mixed states is Eq. (1.20).

For phase measurements, the probability distribution for the measurement result can be determined using the POM $F(\phi)$. This approach was first considered be Helstrom [19], and is also considered in Refs [20, 21]. If the phase measurement treats all phases equally, it should be invariant under a phase translation

$$R(\theta)F(\phi)R(-\theta) = F(\phi+\theta) \tag{1.27}$$

where $R(\theta) = \exp(ia^\dagger a\theta)$ is the phase translation operator. Now $F(\phi)$ has the general expansion

$$F(\phi) = \sum_{n,m=0}^{\infty} |n\rangle\langle m|F_{nm}(\phi). \tag{1.28}$$

The phase shift invariance condition gives

$$F(\phi+\theta) = \sum_{n,m=0}^{\infty} e^{ia^\dagger a\theta}|n\rangle\langle m|e^{-ia^\dagger a\theta}F_{nm}(\phi) \tag{1.29}$$

$$\sum_{n,m=0}^{\infty} |n\rangle\langle m|F_{nm}(\phi+\theta) = \sum_{n,m=0}^{\infty} e^{in\theta}|n\rangle\langle m|e^{-im\theta}F_{nm}(\phi) \tag{1.30}$$

This means that we must have

$$F_{nm}(\phi+\theta) = e^{i(n-m)\theta}F_{nm}(\phi). \tag{1.31}$$

This implies that $F_{nm}(\phi)$ must have the form

$$F_{mn}(\phi) = \frac{1}{2\pi}e^{i(n-m)\phi}H_{nm}. \tag{1.32}$$

A factor of $1/(2\pi)$ has been added for normalisation. Therefore the general form of $F(\phi)$ for a shift invariant phase measurement is

$$F(\phi) = \frac{1}{2\pi}\sum_{n,m=0}^{\infty}|n\rangle\langle m|e^{i(n-m)\phi}H_{nm}. \tag{1.33}$$

For the integral of the probabilities to equal 1, we must have

$$\int_{-\pi}^{\pi}F(\phi)d\phi = 1. \tag{1.34}$$

Applying this to Eq. (1.33) above we find that

$$\sum_{n=0}^{\infty}|n\rangle\langle n|H_{nn} = 1. \tag{1.35}$$

This means that the diagonal elements H_{nn} must all be equal to 1.

In addition there is the condition that the probability given by $\mathrm{Tr}[\rho F(\phi)]$ always be real and positive. This, together with the above result means that all of the H_{nm} must have absolute values between 0 and 1. Usually these are all assumed to be real and positive; however, we will see in Sec. 6.7 that they need not be.

In Ref. [22] it is shown that the additional condition that a number shifter does not alter the phase distribution gives $H_{nm} = 1$, corresponding to the POM

$$F^{\mathrm{can}}(\phi) = \frac{1}{2\pi}\sum_{n,m=0}^{\infty}e^{i(n-m)\phi}|n\rangle\langle m|. \tag{1.36}$$

This POM corresponds to a canonical phase measurement [22]. Ref. [20] also derives this POM using the maximum likelihood approach. In general, real measurements will give smaller values of H_{nm}, and the closer these are to 1 the better the phase measurement is. This POM is the best possible for the main states that I will be considering in thesis, but not for every possible input state. The best POM as derived in [19] is actually dependent on the input state; this is discussed further in Sec. 6.7.

Note that we may express the POM (1.36) in the form

$$F^{\mathrm{can}}(\phi) = |\phi\rangle\langle\phi|, \tag{1.37}$$

where

$$|\phi\rangle = \frac{1}{(2\pi)^{1/2}}\sum_{n=0}^{\infty}e^{in\phi}|n\rangle. \tag{1.38}$$

These states are eigenstates of the Susskind-Glogower operators $\widehat{\cos}\,\phi$ and $\widehat{\sin}\,\phi$, and may be interpreted as phase states. We therefore see that the POM of Eq. (1.36) is consistent with the Susskind-Glogower formalism. It can also be shown [23] that identical results are obtained using this POM as using the Pegg-Barnett formalism. In addition, London's treatment of phase [2, 3] is also equivalent to this.

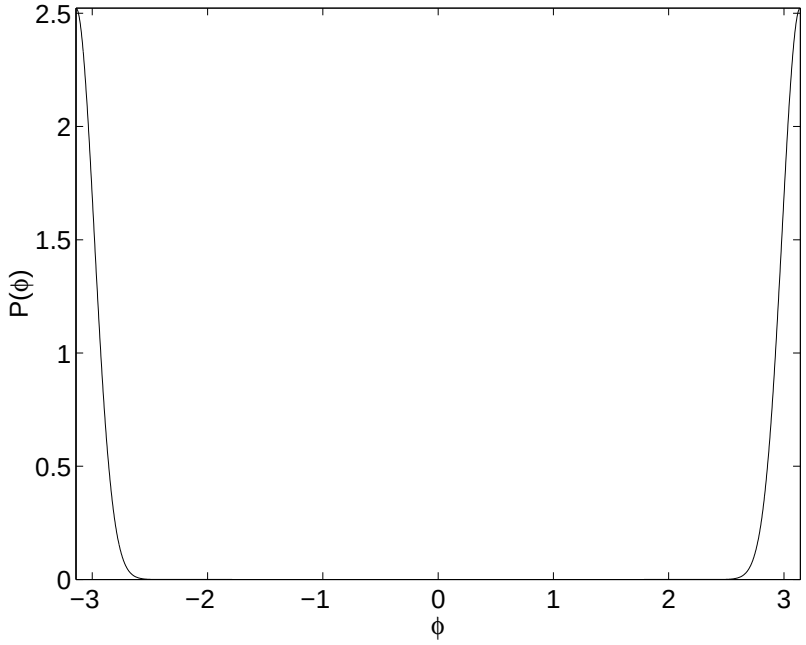

Figure 1.1: A phase probability distribution that is narrowly peaked at $\pm\pi$.

The fact that these different approaches to phase give equivalent results is a compelling reason to consider this to be an accurate description of phase. Nevertheless, there is another class of descriptions of phase which is not equivalent to this: those based on an 'operational' approach [24, 25, 26, 27, 28, 29]. Here the phase is defined as the quantity measured by a particular experiment. The disadvantage of this approach is that the description of phase is dependent on the experiment. For further discussion of the problems involved in the description of phase, see Refs [30, 31].

1.3 Phase Variance

In this work we are primarily interested in the phase variance, and not the total phase distribution. Since phase is a cyclic variable the usual definition of phase variance, as given by

$$\mathrm{var}(\phi) = \left\langle \phi^2 \right\rangle - \left\langle \phi \right\rangle^2, \tag{1.39}$$

does not work well. For example, we would usually expect the variance to go to infinity in the limit of a flat distribution. For phase, as the distribution is limited to a region of length 2π, the variance will be finite for a flat distribution. This means that it is not possible to give an uncertainty relation in the usual way. In addition, if the mean of the distribution is at one bound of the phase, the phase variance obtained from this definition will be artificially large. For example, the distribution shown in Fig. 1.1 is narrowly peaked, but Eq. (1.39) will give a very large variance.

These problems can be solved by using the Holevo phase variance [32]

$$V(\phi) = \left|\left\langle e^{i\phi} \right\rangle\right|^{-2} - 1. \tag{1.40}$$

This definition is naturally modulo 2π, and in addition if the distribution is flat then $\left\langle e^{i\phi} \right\rangle$ will be zero, so the variance will be infinite. For this definition of the variance, there is the uncertainty relation

$$V(\phi)\Delta N^2 \geq \tfrac{1}{4}. \tag{1.41}$$

In contrast, using Eq. (1.39) for the phase variance there is no uncertainty relation of this form because, for example, a number state has zero number variance but a finite variance under Eq. (1.39).

Also, for a distribution that is sharply peaked, we find

$$\begin{aligned} V(\phi) &= \left|\left\langle e^{i\phi}\right\rangle\right|^{-2} - 1 \\ &= \left|e^{i\langle\phi\rangle}\left\langle e^{i(\phi-\langle\phi\rangle)}\right\rangle\right|^{-2} - 1 \\ &= \left|\left\langle e^{i(\phi-\langle\phi\rangle)}\right\rangle\right|^{-2} - 1 \\ &\approx \left|\left\langle 1 + i(\phi-\langle\phi\rangle) - \tfrac{1}{2}(\phi-\langle\phi\rangle)^2\right\rangle\right|^{-2} - 1 \\ &= \left|1 - \tfrac{1}{2}\left\langle(\phi-\langle\phi\rangle)^2\right\rangle\right|^{-2} - 1 \\ &= \left[1 - \tfrac{1}{2}\left\langle(\phi-\langle\phi\rangle)^2\right\rangle\right]^{-2} - 1 \\ &\approx 1 + \left\langle(\phi-\langle\phi\rangle)^2\right\rangle - 1 \\ &= \left\langle(\phi-\langle\phi\rangle)^2\right\rangle. \end{aligned} \tag{1.42}$$

Therefore this measure of the phase variance is approximately the same as the usual estimate of the variance for sharply peaked distributions.

The definition of the Holevo variance (1.40) can be used when the exact phase distribution is known. In obtaining numerical results, it is is generally the case that the exact distribution is unknown, so the Holevo phase variance must be estimated from a set of samples. In the case of the standard variance, it is a standard result in statistics that the variance calculated from the samples by

$$\frac{1}{M}\sum_{i=1}^{M}(\phi_i - \langle\phi\rangle)^2, \tag{1.43}$$

where M is the number of samples, is a biased estimator for the variance of the distribution. The unbiased estimator is where the dividing factor is $M-1$, i.e. the unbiased estimator is

$$\frac{1}{M-1}\sum_{i=1}^{M}(\phi_i - \langle\phi\rangle)^2. \tag{1.44}$$

This is because taking the average of the data removes one degree of freedom. If the mean of the distribution is known, then an unbiased estimator for the variance is

$$\frac{1}{M}\sum_{i=1}^{M}(\phi_i - \bar{\phi})^2, \tag{1.45}$$

where $\bar{\phi}$ is the known mean for the distribution. If the measurements are unbiased, then the mean for the distribution will be the same as the actual phase, φ. If we use the actual phase, then the estimator becomes

$$\frac{1}{M}\sum_{i=1}^{M}(\phi_i - \varphi)^2. \tag{1.46}$$

The situation is analogous for the Holevo phase variance. If the measurements are unbiased, we can estimate the Holevo variance from the samples by

$$\left[\mathrm{Re}\left(\frac{1}{M}\sum_{i=1}^{M}e^{i(\phi_i-\varphi)}\right)\right]^{-2} - 1. \tag{1.47}$$

It is easy to see that when the phase distribution is narrowly peaked, this simplifies to Eq. (1.46). On the other hand, if the measurements are *biased*, then this will not be an estimator for the Holevo variance, and Eq. (1.46) will not be an estimator for the standard variance.

For measurements that may be biased the variance is not a good measure of the accuracy of the measurement. An arbitrary bias may be added to the distribution without altering the variance, as

the deviation is measured from the average of the distribution, rather than the actual phase. For a biased measurement scheme, it is more appropriate to define the standard variance as

$$\mathrm{var}_{\varphi}(\phi) = \left\langle (\phi - \varphi)^2 \right\rangle. \tag{1.48}$$

Here the subscript φ has been used to distinguish this variance from the usual definition. For this definition, biased distributions will give larger variances. Also Eq. (1.46) will be an unbiased estimator for this modified variance.

The analogous definition for the Holevo variance is

$$V_{\varphi}(\phi) = \left[\mathrm{Re}\left\langle e^{i(\phi-\varphi)} \right\rangle\right]^{-2} - 1. \tag{1.49}$$

In the case that $\varphi = 0$, or ϕ is the deviation from the system phase, this definition of the variance simplifies to

$$V_0(\phi) = \left[\mathrm{Re}\langle e^{i\phi}\rangle\right]^{-2} - 1. \tag{1.50}$$

Similarly to the case for the standard variance, this definition will give larger variances for biased distributions, whereas the standard definition does not distinguish biased and unbiased distributions. If the measurements are biased, then Eq. (1.47) will be an estimator for this variance, but not the Holevo variance as given by Eq. (1.40).

In this thesis I usually consider phase distributions that are unbiased, so this modified definition of the Holevo phase variance is not required. For the problem of finding optimal phase estimates, however, this altered definition is necessary in order to eliminate the possibility of biased phase estimates.

Now if we consider arbitrary measurements on an arbitrary pure state $|\psi\rangle$, the probability distribution for the phase is given by

$$\begin{aligned} P(\phi) &= \mathrm{Tr}[\rho F(\phi)] \\ &= \frac{1}{2\pi} \sum_{n,m=0}^{\infty} \langle\psi|n\rangle\langle m|\psi\rangle e^{i(n-m)\phi} H_{nm}. \end{aligned} \tag{1.51}$$

In order to determine the Holevo phase variance we must determine $\langle e^{i\phi}\rangle$. Evaluating this gives

$$\begin{aligned} \langle e^{i\phi}\rangle &= \int_{-\pi}^{\pi} P(\phi) e^{i\phi} d\phi \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{n,m=0}^{\infty} \langle\psi|n\rangle\langle m|\psi\rangle e^{i(n+1-m)\phi} H_{nm} d\phi \\ &= \sum_{n,m=0}^{\infty} \langle\psi|n\rangle\langle m|\psi\rangle \delta_{n+1,m} H_{nm} \\ &= \sum_{n=0}^{\infty} \langle\psi|n\rangle\langle n+1|\psi\rangle H_{n,n+1}. \end{aligned} \tag{1.52}$$

Therefore the phase variance does not depend on all the elements H_{nm}, but only on the off-diagonal elements $H_{n,n+1}$. As we are often only interested in the phase variance resulting from a measurement scheme, this greatly reduces the number of variables required to describe the measurement. Note that the condition $|H_{nm}| \leq 1$ means that this expression gives $|\langle e^{i\phi}\rangle| \leq 1$, as we would expect.

For canonical measurements we have $H^{\mathrm{can}}_{nm} = 1$, so this simplifies to

$$\left\langle e^{i\phi}\right\rangle_{\mathrm{can}} = \sum_{n=0}^{\infty} \langle\psi|n\rangle\langle n+1|\psi\rangle. \tag{1.53}$$

For more arbitrary measurements, we will have $H_{n,n+1} \leq 1$, so it is convenient to define the vector $h(n)$ by

$$h(n) = 1 - H_{n,n+1}. \tag{1.54}$$

In terms of this we obtain

$$\begin{aligned}\left\langle e^{i\phi}\right\rangle &= \sum_{n=0}^{\infty}\langle\psi|n\rangle\langle n+1|\psi\rangle(1-h(n)) \\ &= \sum_{n=0}^{\infty}\langle\psi|n\rangle\langle n+1|\psi\rangle - \sum_{n=0}^{\infty}\langle\psi|n\rangle\langle n+1|\psi\rangle h(n) \\ &= \left\langle e^{i\phi}\right\rangle_{\text{can}} - \sum_{n=0}^{\infty}\langle\psi|n\rangle\langle n+1|\psi\rangle h(n).\end{aligned} \tag{1.55}$$

If the photon number distribution is reasonably sharply peaked then we can typically replace $h(n)$ by its value for the average photon number. Then we find

$$\left\langle e^{i\phi}\right\rangle \approx \left\langle e^{i\phi}\right\rangle_{\text{can}}(1-h(\bar{n})). \tag{1.56}$$

For large mean photon number, we find that $h(\overline{n}) \ll 1$ for most measurement schemes. Using this approximation, the Holevo phase variance is

$$\begin{aligned}V(\phi) &\approx \left|\left\langle e^{i\phi}\right\rangle_{\text{can}}(1-h(\bar{n}))\right|^{-2} - 1 \\ &= \left|\left\langle e^{i\phi}\right\rangle_{\text{can}}\right|^{-2}(1-h(\bar{n}))^{-2} - 1 \\ &\approx \left|\left\langle e^{i\phi}\right\rangle_{\text{can}}\right|^{-2}(1+2h(\bar{n})) - 1 \\ &\approx \left|\left\langle e^{i\phi}\right\rangle_{\text{can}}\right|^{-2} + 2h(\bar{n}) - 1 \\ &= V_{\text{can}}(\phi) + 2h(\bar{n}).\end{aligned} \tag{1.57}$$

Thus we see that to a first approximation the phase variance is equal to the intrinsic phase variance plus twice the value of $h(n)$ for the average photon number. This means that the introduced phase variance is fairly independent of the input state (apart from the photon number).

For many types of real measurements, the value of $h(n)$ decreases as some power of n for large n. We can therefore approximate it by

$$h(n) \approx cn^{-p}. \tag{1.58}$$

When this is the case, we can describe the measurement by just the two variables c and p, rather than the entire vector $h(n)$, or the matrix H_{nm}.

1.4 Real Measurements

In practice, it is not possible to directly measure phase. What we must do is infer the phase from measurements of some other quantity. The electric field, with the operator given by Eq. (1.3), is directly measurable. Ignoring the field direction, this operator is proportional to

$$\hat{X}_{\Phi} = ae^{-i\Phi} + a^{\dagger}e^{i\Phi}, \tag{1.59}$$

where

$$\Phi = \omega t - \mathbf{k}\cdot\mathbf{R}. \tag{1.60}$$

This is called the Φ quadrature operator. For high frequency light Φ varies far too rapidly for convenience, so we wish to measure the quadrature in a more convenient way.

This can be done by combining the light with a strong local oscillator field at a beam splitter, as in Fig. 1.2. The local oscillator field provides a reference phase. The difference between the photon

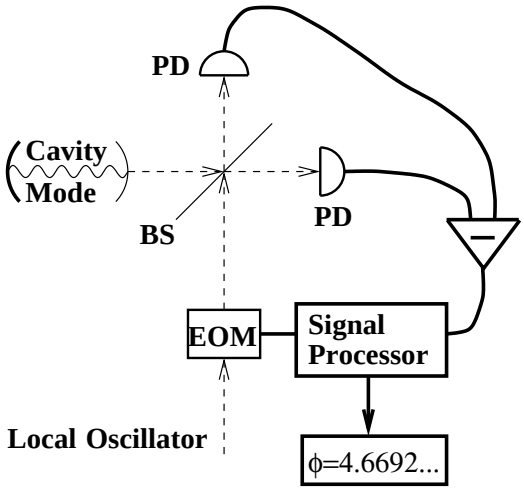

Figure 1.2: Diagram of the apparatus for making a dyne phase measurement. The signal from the cavity and the local oscillator field are combined at a 50/50 beam splitter (BS) and the outputs are detected by photodetectors (PD). For adaptive measurements the signals from these photodetectors are processed by the digital signal processor, which determines a phase estimate and adjusts the electro-optic phase modulator (EOM) accordingly.

numbers at the two photodetectors gives a measurement of the quadrature $\hat{X}_\Phi$, where Φ is the phase of the local oscillator field.

In this study I consider measurements that take an appreciable time, so rather than just considering the total photon counts at the detectors, it is possible to consider the instantaneous photocurrent. For this extended measurement we can alter the phase of the local oscillator during the measurement based on the data obtained so far.

The standard technique for measuring a completely unknown phase is heterodyne detection, where the phase of the local oscillator is varied linearly. This is simple to do, as it simply means that the local oscillator has a slightly different frequency than the signal. This method suffers from the drawback that it introduces a fairly large phase uncertainty. It is possible to obtain more accurate measurements if we use a constant feedback phase Φ, that is close to the actual phase plus $\pi/2$. This is called the homodyne technique.

The major problem with the homodyne technique is that it requires knowledge of the system phase beforehand. The guiding principle behind adaptive phase measurements is that we wish to approximate homodyne phase measurements by varying the local oscillator phase during the measurement based on data obtained so far. We will see that it is not quite so simple, however, as we do not use the best estimate of the phase, and in rare cases will not use a phase estimate at all. I will use the general term "dyne" for heterodyne and homodyne phase measurements, as well as these more complicated adaptive phase measurements.

In order to describe the adaptive measurements I will first introduce the notation. The numbers of detections at the two photodetectors in the time interval δt are denoted δN_+ and δN_-. The complex amplitude of the local oscillator will be denoted as γ. The difference photocurrent $I(t)$ is then defined in terms of the noncommuting limits

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_+ - \delta N_-}{|\gamma|\, \delta t}. \tag{1.61}$$

This simplifies to

$$I(t)dt = 2\mathrm{Re}(\langle a \rangle\, e^{-i\Phi})dt + dW(t), \tag{1.62}$$

where $dW(t)$ is an infinitesimal Wiener increment with variance dt. For a constant amplitude coherent state we have $\langle a \rangle = \alpha$. Note that $\langle I(t) \rangle = \langle \hat{X}_\Phi \rangle$, so this is essentially a measurement of the Φ quadrature.

For more general states, $\langle a \rangle$ may have a systematic variation given by

$$\langle a \rangle = \alpha_t \propto \sqrt{u(t)}, \tag{1.63}$$

where $u(t)$ is a mode function. In this case we alter the definition of $I(t)$, and scale α_t and the time to obtain

$$I(v)dv = 2\mathrm{Re}(\alpha_v e^{-i\Phi})dv + dW(v), \tag{1.64}$$

similarly to the case for constant amplitude. The detail of how this is done is given in Ch. 3.

We then define the quantities A_v and B_v by

$$A_v = \int_0^v e^{i\Phi} I(u)du, \tag{1.65}$$

$$B_v = -\int_0^v e^{2i\Phi} du, \tag{1.66}$$

The time is scaled to the unit interval, and for the values of A_v and B_v at $v = 1$ the subscripts are omitted. It is shown in Ref. [33] that the only relevant information from the measurement record is contained in these two variables. (B_v does not explicitly depend on the measurement record, but it does depend on it implicitly if the feedback phase is varied based on the measurement record.)

To see why this is so, firstly recall that the way the normalised state varies for measurement operator Ω_r is

$$|\psi_r\rangle = \frac{\Omega_r|\psi\rangle}{\sqrt{P_r}}. \tag{1.67}$$

Alternatively we can consider an unnormalised state $|\tilde{\psi}_r\rangle$, given by

$$|\tilde{\psi}_r\rangle = \frac{\Omega_r|\psi\rangle}{\sqrt{\Lambda_r}}, \tag{1.68}$$

where the Λ_r are called the ostensible probabilities. The actual probability is then given by

$$P_r = \Lambda_r \langle\tilde{\psi}_r|\tilde{\psi}_r\rangle. \tag{1.69}$$

This method is used in Ref. [33] to consider the evolution of the signal state under dyne measurements, with the ostensible probabilites chosen as those for a vacuum.

It is found that the evolution of the unnormalised state depends on the photocurrent record and feedback phases only through the variables A_v and B_v. This means that the probability distribution for the photocurrent record up to time v, $\mathbf{I}_{[0,v)}$, is given by

$$P(\mathbf{I}_{[0,v)}) = P_0(\mathbf{I}_{[0,v)})\langle\tilde{\psi}_v(A_v, B_v)|\tilde{\psi}_v(A_v, B_v)\rangle, \tag{1.70}$$

where $P_0(\mathbf{I}_{[0,v)})$ is the ostensible probability distribution. Note that this does not prove that the probability distribution for the photocurrent record only depends on A_v and B_v, as the ostensible probability distribution may depend on the photocurrent record in some more complicated way. The ostensible distribution, however, is for a vacuum, and does not contain any information about the state. This means that all the information about the state from the photocurrent record $\mathbf{I}_{[0,v)}$ is contained in A_v and B_v.

To show this in a more rigorous way, consider some arbitrary parameter of the input state, x. This could, for example, be the phase or the photon number. Using Bayes' theorem we find that the probability distribution for this quantity given the photocurrent record $\mathbf{I}_{[0,v)}$ is

$$P(x|\mathbf{I}_{[0,v)}) = \frac{P(x)P(\mathbf{I}_{[0,v)}|x)}{P(\mathbf{I}_{[0,v)})}. \tag{1.71}$$

Here $P(x)$ is the probability distribution for this quantity at the start of the measurement. It will be flat, as we are assuming that there is no knowledge about this parameter before the start of the measurement. The probability in the denominator is independent of x, so the probability distribution is therefore

$$P(x|\mathbf{I}_{[0,v)}) \propto P(\mathbf{I}_{[0,v)}|x). \tag{1.72}$$

The ostensible probability distribution $P_0(\mathbf{I}_{[0,v)})$ contains no information about the state, and will therefore be independent of x. Thus we find

$$P(x|\mathbf{I}_{[0,v)}) \propto \langle\tilde{\psi}_v(A_v, B_v)|\tilde{\psi}_v(A_v, B_v)\rangle. \tag{1.73}$$

This means that the probability distribution for any parameter of the signal state only depends on the photocurrent record through A_v and B_v. This means, for example, that estimators for the phase should be functions of A_v and B_v. They can be functions of other variables, for example the initial coherent amplitude, if that is known, but they should depend on the photocurrent record only through A_v and B_v.

1.5 Adaptive Phase Measurements

For a coherent state with $\alpha = |\alpha|\, e^{i\varphi}$, Eq. (1.64) becomes

$$I(v)dv = 2\,|\alpha|\cos(\varphi - \Phi)dv + dW(v). \tag{1.74}$$

In order for the measurement to be close to an ideal measurement of the phase, the cos function should be as close as possible to zero. This implies that its argument should be close to $\pm\pi/2$. Therefore it is best to take the local oscillator phase to be $\hat{\varphi} + \pi/2$, where $\hat{\varphi}$ is some estimate of the system phase. In that case

$$I(v)dv = 2\,|\alpha|\sin(\varphi - \hat{\varphi})dv + dW(v). \tag{1.75}$$

In the limit that $|\varphi - \hat{\varphi}| \ll 1$, this becomes

$$I(v)dv \approx 2\,|\alpha|\,(\varphi - \hat{\varphi})dv + dW(v), \tag{1.76}$$

so the measurement is very close to an ideal measurement of the phase. The reason why there is always an excess phase uncertainty is because the sin function is not completely linear, so this is not exactly equivalent to a direct measurement of the phase.

The basis of homodyne measurements is to use an estimate of the phase that is known before the measurement begins. For adaptive measurements, we only have information about the phase once the measurement has begun, and we use a phase estimate based on the data obtained so far. To see what we can use as a phase estimator, we can expand the variable A_v to give

$$A_v = v\alpha - \alpha^* B_v + i\sigma_v, \tag{1.77}$$

where

$$\sigma_v = \int_0^v e^{i\Phi(u) - i\pi/2} dW(u). \tag{1.78}$$

From this result it is simple to show that

$$vA_v + B_v A_v^* = \alpha(v^2 - |B_v|^2) + i(v\sigma_v - B_v\sigma_v^*). \tag{1.79}$$

Taking the expectation value gives

$$\langle vA_v + B_v A_v^*\rangle \approx \alpha\left\langle v^2 - |B_v|^2\right\rangle. \tag{1.80}$$

Table 1.1: The values of c and p for heterodyne, mark I and mark II measurements.

	c	p
heterodyne	1/8	1
mark I	1/8	0.5
mark II	1/16	1.5

If the local oscillator phase is independent of the photocurrent record, then this is exact. In the case of feedback, B_v may be correlated with σ_v, but this result should still be approximately true. This means that the phase of $vA_v + B_v A_v^*$ should be close to the phase of the signal. It is convenient to define the new variable

$$C_v = vA_v + B_v A_v^*. \tag{1.81}$$

This should not be confused with the variable C used in Ref. [34], which is defined differently.

For heterodyne measurements, the local oscillator phase $\Phi(v)$ is varying linearly. This means that the average value of $e^{2i\Phi(v)}$ will be zero, so B_v will be close to zero. This means that $C_v \approx A_v$, so $\arg A_v$ will also be a good estimator of the phase. It is simple to show from this [35] that the phase variance for heterodyne measurements on coherent states is approximately $\frac{1}{2}|\alpha|^{-2}$. As the intrinsic phase variance for coherent states is $\frac{1}{4}|\alpha|^{-2}$, this means that there is an introduced phase variance of $\frac{1}{4}\overline{n}^{-1}$.

For more general measurements where $B_v \neq 0$, $\arg A_v$ will not be as good an estimator of the phase; however, it can still be used as an estimator. If $\hat{\varphi} = \arg A_v$ is used as the estimator for the phase in feedback to the local oscillator phase, we have a measurement scheme that can be analysed analytically. In the scheme considered in Ref. [36], $\arg A$ was the phase estimate used at the end of the measurement as well.

It was shown in Ref. [36] that these measurements give the canonical result if the system has at most one photon. For systems with large numbers of photons, however, the introduced phase variance scales as $\frac{1}{4}\overline{n}^{-1/2}$. This means that for large photon numbers this measurement scheme gives far higher phase variances than heterodyne measurements.

It is possible to greatly improve on this result by simply using the best phase estimate at the end of the measurement, $\arg C$. This is called the mark II adaptive phase measurement scheme, whereas the scheme where $\arg A$ is used at the end of the measurement is called mark I. This measurement scheme introduces a phase uncertainty of $\frac{1}{8}\overline{n}^{-3/2}$ [35]. This is a vast improvement over the mark I scheme, and also improves on the heterodyne scheme.

Recall from Eq. (1.57) that the introduced phase variance is approximately $2h(\overline{n})$. This means that the variation of $h(n)$ for the heterodyne, mark I and mark II measurement schemes is approximately

$$h_{\rm het}(n) \approx \tfrac{1}{8}n^{-1}, \tag{1.82}$$

$$h_{\rm I}(n) \approx \tfrac{1}{8}n^{-0.5}, \tag{1.83}$$

$$h_{\rm II}(n) \approx \tfrac{1}{16}n^{-1.5}. \tag{1.84}$$

To summarise this, the values of c and p for these three measurement schemes are as given in Table 1.1.

The mark II measurement scheme still leaves questions, however. It has been shown [37] that optimum states have intrinsic phase variances that scale as $1.89 \times \overline{n}^{-2}$. For these states the phase uncertainty that is introduced by mark II measurements will be far greater than the intrinsic phase uncertainty of the state. In addition it has been shown [34] that the theoretical limit to the variance that is introduced by dyne phase measurements is $\frac{1}{4}\log\overline{n} \times \overline{n}^{-2}$. This is not quite as good as the scaling for the intrinsic phase variance, but it is far better than the scaling for mark II measurements.

Another issue is that the intermediate phase estimates used for mark II measurements are still $\arg A_v$. As evidenced by the poor scaling for mark I measurements, these phase estimates are far worse than $\arg C_v$. This raises the question of whether phase measurements can be made closer to

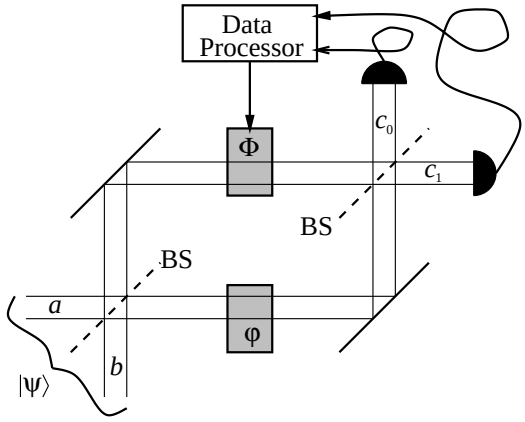

Figure 1.3: The Mach-Zehnder interferometer, with the addition of a controllable phase Φ in one arm. The unknown phase to be estimated is φ. Both beam splitters (BS) are 50/50.

the theoretical limit by using better intermediate phase estimates. This is considered in Ch. 3, and it is shown that measurements can be made very close to the theoretical limit.

Another factor in attempting to make the most accurate possible phase measurements is the input state. Rather than just considering states that are optimised for minimum intrinsic phase variance, the fairest way to evaluate the various measurement schemes is to consider states that are optimised for minimum phase variance under that particular phase measurement scheme. This is considered in Ch. 2.

It is necessary to have a constraint on the state being optimised in order to avoid obtaining a state with indefinitely large photon number. The two main ways of constraining the state that is being optimised are to put an upper limit on the photon number and to specify the mean photon number. A third alternative is to only consider squeezed states. These states are far more convenient to work with, both theoretically and experimentally, and in addition they give results that are extremely close to those for the general optimisation problem for general measurements. All these alternatives are considered in Ch. 2.

1.6 Interferometry

The major alternative to dyne measurements is interferometric phase measurements. Here, rather than measuring the phase of a single mode, and assuming that the local oscillator field is sufficiently intense that it can be treated classically, we are measuring the phase difference between two modes, both of which are treated quantum mechanically.

The most convenient way of considering this is via a Mach-Zehnder interferometer, as in Fig. 1.3. Two input modes are combined at a beam splitter, after which each of the modes is subjected to a phase shift, and the two modes are recombined at a second beam splitter. Usually we would consider measurement of the phase difference between the two arms. Here for simplicity I consider the phase shift to be measured, φ, to be in one arm, and add a controllable phase shift, Φ, in the second arm. This allows us to make adaptive measurements analogous to the case for dyne measurements.

Note that if the first beam splitter is omitted, the configuration is identical to the configuration used to make dyne measurements of a single mode. The arm of the interferometer subjected to the phase shift Φ is equivalent to the local oscillator mode, and the arm subjected to the phase shift φ is equivalent to the signal mode in the dyne case. The difference here is that we are now treating both modes quantum mechanically.

The main role of the first beam splitter is that it creates quantum correlations between the two modes, which generally improves the phase properties of the state. For example, it is not possible to measure a phase difference between a number state and a vacuum state, but the two mode state produced after the beam splitter has good phase properties.

In general it is much easier to produce input states without quantum correlations between the modes, so the initial beam splitter is very useful. In this study, I consider two-mode input states with arbitrary correlations between the modes. In this case, the initial beam splitter is fairly superfluous, as it merely transforms the input state into another correlated state that could have been considered directly. I still include the initial beam splitter in this study, for consistency with previous work.

It is well known that if we feed a state with a mean photon number of $\overline{n}$ into one arm, it is possible to obtain a phase variance scaling as $\overline{n}^{-1}$. This is analogous to the case of coherent states for dyne measurements, which are the easiest to produce and also give $\overline{n}^{-1}$ scaling. There have been several proposals for reducing the phase variance to $\overline{n}^{-2}$. The first of these is that by Caves [1].

Caves considers an interferometer with a coherent state in one arm, but in the other arm, rather than a vacuum state, a squeezed vacuum is used. The measurements considered are equivalent to homodyne measurements, where the phase difference is very close to 0. For the right squeezing parameter these measurements have a variance scaling as $\overline{n}^{-2}$. Squeezed states were also considered by Bondurant and Shapiro [38].

Yurke *et al.* [39] consider an input state that is a combination of input number states,

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|j\rangle_a |j\rangle_b + |j+1\rangle_a |j-1\rangle_b \right), \tag{1.85}$$

where the subscripts a and b indicate number states in ports a and b respectively. According to their analysis this state should give a phase uncertainty scaling close to N^{-1}, where $N = 2j$ is the fixed total photon number. Similarly to Ref. [1] the phase must be very close to 0 to obtain this scaling. Yurke *et al.* also consider active interferometers, characterised by SU(1,1) rather than SU(2), as is the conventional Mach-Zehnder interferometer considered in this study.

Holland and Burnett [40] also considered the case where the total photon number is fixed. They considered the state with equal photon numbers in both input ports, $|j\rangle_a |j\rangle_b$. For this state the phase uncertainty scales as N^{-1}, but again only for phases very close to 0. This state also has the additional problem that it gives results at 0 and $\pm\pi$ with equal probability, and therefore must be considered modulo π.

Sanders and Milburn [41, 42] considered "optimal" measurements, for which the phase uncertainty is independent of the system phase. Unfortunately these measurements are derived from projections onto the phase states, rather than a physical measurement scheme involving photodetectors. In Ch. 6 it is shown that it is not possible to implement these measurements using photodetectors, even allowing feedback. Therefore I will generally call these measurements ideal or canonical rather than optimal, and reserve the term optimal for the best possible physically realisable scheme.

Sanders and Milburn considered the same input state as Holland and Burnett, $|j\rangle_a |j\rangle_b$. In Ref. [42] they show that the phase uncertainty for this state scales as N^{-1} according to two common measures of the uncertainty. In Ch. 5 I show that this state has very poor scaling if we consider the phase variance. The optimal input states as evaluated using the Holevo phase variance are derived, and it is shown that all the common measures of the phase uncertainty scale as N^{-1} for these states. Both the $|j\rangle_a |j\rangle_b$ and $\frac{1}{\sqrt{2}} \left(|j\rangle_a |j\rangle_b + |j+1\rangle_a |j-1\rangle_b \right)$ states are rough approximations of this state.

1.7 Experimental Imperfections

The majority of the work considered in this thesis ignores all experimental imperfections. This is because the main motivation for this work is to lay a theoretical foundation for how far phase measurements can, in principle, be improved, rather than to examine the limits to phase measurements using current technology.

Firstly there are a number of problems that are very specific to the apparatus. For example an inaccurate calibration in the phase shifter producing the phase Φ will result in a corresponding error in the measured phase. These types of problems are relatively simple to analyse, and as they are very specific to the equipment used they will not be discussed here.

One problem that is common to any phase measurement is inefficient photodetectors. No current photodetectors will register every photon. There are two main types of photodetectors with different

efficiencies. Firstly there are large amplitude photodetectors that do not need to distinguish the exact photon number. These are used for dyne measurements, where the photon counts are treated in the continuous approximation. These can be made very efficient, and the best current photodetectors have efficiencies around 98% [43].

The case of dyne measurements with inefficient detectors was considered in [35]. For the case of coherent states, the analysis is simple, as an efficiency η can be treated by changing the mean photon number from $\overline{n}$ to $\eta\overline{n}$. This means that, for mark II measurements, the phase variance to first order is

$$\begin{aligned} V(\phi) &= \frac{1}{4\eta\alpha^2} + O(\alpha^{-3}) \\ &= \frac{1}{4\alpha^2} + \frac{1-\eta}{4\eta\alpha^2} + O(\alpha^{-3}) \end{aligned} \tag{1.86}$$

We will get the same result (to first order) for any other reasonably good phase feedback scheme where the introduced phase variance is of higher order than the intrinsic phase variance. The phase variance is in the form of the intrinsic phase variance plus an extra term due to the inefficient photodetectors. As explained previously, the introduced phase variance is generally independent of the input state (to first order). This means that

$$\Delta V(\phi) = \frac{1-\eta}{4\eta\overline{n}} \tag{1.87}$$

will be the introduced phase variance due to the inefficient photodetectors for other states with reduced phase uncertainty.

The second type of photodetectors is photon counters, that count photons one by one. This is the type required for interferometric measurements, where we wish to alter the feedback phase after every detection. These generally have far lower efficiencies, less than 90% [44]. This is a serious problem, as the analysis breaks down if even a single photon is missed. Taking account of a fixed probability of missing a photon would greatly complicate the analysis, and was not attempted in this study.

Another problem, that is unique to phase measurements involving feedback, is the time delay in the feedback loop. This is arguably an even more fundamental problem than inefficient photodetectors, because the feedback loop cannot operate faster than the time it takes light to reach the phase modulator. This is a very serious problem for short pulses. For example, if the pulse is shorter than the distance between the phase modulator and the photodetector it is not possible to perform any sort of adaptive measurement. Adaptive measurements rely on the pulse being sufficiently long that the feedback phase can be altered during the passage of the signal.

Even if the time delay is a small fraction of the length of the signal, there will be an appreciable increase in the phase variance. This was investigated by a highly simplified theory in Ref. [35], and in Ch. 4 the introduced phase variance is estimated in a far more rigorous way.

1.8 Structure of the Thesis

This thesis is structured around the problems of determining the optimum input states and measurement schemes for dyne measurements and interferometry. The optimum input states for dyne measurements are derived in Ch. 2, and the problem of performing optimum dyne measurements is discussed in Ch. 3. Optimum input states for interferometry are discussed in Ch. 5, and optimum interferometric measurements are discussed in Ch. 6. These chapters form the main theme of this thesis. Some additional problems considered are time delays in Ch. 4 and continuous (rather than pulsed) measurements in Ch. 7.

As mentioned above there are three main alternatives when considering optimum input states for dyne measurements:

1. An upper limit is placed on the photon number.
2. The mean photon number is fixed, but the state is otherwise arbitrary.

3. The state is a squeezed state, and the mean photon number is fixed.
These alternatives are considered in Secs 2.1, 2.2 and 2.3 respectively of Ch. 2. In each case both canonical measurements and more general dyne measurements where $h(n) \approx cn^{-p}$ are considered.

The analytic results for canonical measurements and for general dyne measurements with an upper limit on the photon number were previously derived in Refs [37, 35, 45]. These derivations are summarised here, as the method involved sheds light on the method of solution for the more complicated cases of general dyne measurements for fixed mean photon number and for squeezed states. The accuracy of all these results is evaluated by extensive numerical calculations.

The problem of performing dyne phase measurements at the theoretical limit is too complicated to solve analytically, and Ch. 3 therefore relies upon numerical methods. The various feedback schemes are evaluated numerically by solving the stochastic differential equations (SDEs) for a large number of samples. This is made simpler by using squeezed states, for which only the two squeezing parameters need be kept track of. The problem of deriving the SDEs for the squeezing parameters, as well as much of the background theory of the evolution of the state, is described in Sec. 3.3.

A series of different phase feedback schemes that give results progressively closer to the theoretical limit are described in Secs 3.5 to 3.7. The last of these is a corrected feedback scheme that gives variances within about 5% of the theoretical limit. Lastly in this chapter, the possibilities of surpassing the theoretical limit are considered.

The perturbation approach for time delays that is considered in Ref. [35] is repeated in a more rigorous way in Ch. 4. It is shown that the same result is obtained for the final value of the intermediate phase estimate, but this approach fails when it is performed for mark II measurements. In Sec. 4.3 an alternative approach based on the POM is considered that gives a result that is valid not only for the mark II measurements, but for all other feedback schemes. These results are all backed by extensive numerical calculations in Sec. 4.4.

In Ch. 5 the optimum input states for interferometry are derived, as evaluated using the Holevo variance for ideal measurements. The phase variance for these states scales as N^{-2}, so the uncertainty scales as N^{-1}. The Holevo variance for three alternative input states is calculated, and it is shown that all of these have poorer scaling. The optimum input states are then evaluated under several alternative measures of phase uncertainty, and are shown to have N^{-1} scaling for all of these.

In Ch. 6 a feedback scheme is introduced that approximates the ideal measurements very closely. For optimal input states, the scaling in the phase variance under this measurement scheme is very close to N^{-2}. For photon numbers above 5, this measurement scheme is not exactly optimal, but it is possible to solve numerically for the optimal feedback scheme (for small photon numbers). Although it is not always possible to obtain variances as small as canonical, using this feedback scheme it is possible to obtain variances smaller than canonical for some states. In Ch. 6 the resolution of this apparent contradiction is discussed.

The last area that is considered in this thesis is that of continuous measurements. If we wish to transmit information via the phase, a pulsed signal with a single phase is not very useful, as only a single real number is transmitted. In order to transmit a significant amount of information we can either send a whole series of pulses with different phases, or produce a constant beam with a fluctuating phase. This is the alternative that is considered in Ch. 7.

In this chapter both the cases of dyne measurements and interferometry are considered. Unfortunately the non-classical states with reduced phase uncertainty do not necessarily have equivalents in the continuous case. This is because the reduced phase uncertainty is due to the back-action of the measurement on the state. If the state does not change, then it is not possible to get the reduced variance.

In the case of dyne measurements it is possible to consider squeezed states by altering the statistics for the detections in time dt, but keeping α constant. Therefore both the coherent and squeezed state cases are considered in Ch. 7. It does not seem to be possible to use this method for interferometry, as we cannot alter the statistics for individual detections. Therefore only the case with all photons in one port is considered in this chapter.

Chapter 2

Optimal Input States for Dyne Measurements

There are, in general, two areas for improvement in the use of dyne measurements for phase estimation. Firstly there is the input state, and secondly there is the measurement technique that is used. In this chapter I will be focusing on optimising the input states for minimum phase variance under various types of dyne measurements, and in the next chapter I discuss how to perform the best possible dyne measurements.

In optimising the input states, there must be some constraint placed on the state, because we can reduce the phase variance indefinitely by using larger and larger photon numbers. The limit of this is phase eigenstates, which have zero intrinsic phase variance, but infinite mean photon number. There are two main ways in which we can constrain the input states in order to avoid this problem. One way is to place an upper limit on the photon number that the state can have contributions from, and another way is to fix the mean photon number, but allow the state to have contributions from indefinitely large photon numbers.

A third alternative is to consider squeezed states rather than arbitrary states. For these states we can optimise the squeezing parameter while keeping the mean photon number fixed. The reason for considering this case is that it is rather more realistic than considering arbitrary states, as it is not possible to produce an arbitrary state experimentally, whereas it is possible to produce a squeezed state. In addition squeezed states are easily treated numerically, as only the two squeezing parameters need be considered, rather than the full state. A third reason is that squeezed states give results extremely close to general optimised states for realistic dyne measurements.

I will firstly consider the case where there is an upper limit placed on the photon number.

2.1 Upper Limit on Photon Number

This case is far simpler than the case of a fixed mean photon number, and the analytic results were derived before this study commenced. I will summarise the derivations here, however, as they are necessary to understand the numerical results that will be presented. In addition, these derivations cast light on how to derive the more complicated results with fixed mean photon number.

2.1.1 Canonical Measurements

Firstly I will consider the optimum states for canonical measurements. This is the simplest case, and in fact the only case which is exactly soluble. It is solved, for example in Refs [37] and [35]. As discussed in the introduction, the measure of the phase variance that will be used throughout this thesis is the Holevo phase variance,

$$V(\phi) = \left|\left\langle e^{i\phi}\right\rangle\right|^{-2} - 1. \tag{2.1}$$

For simplicity, the mean phase of the optimised state can be taken to be zero. This means that $\langle e^{i\phi} \rangle$ is real. Therefore, rather than using Eq. (2.1), we can use

$$\begin{aligned} V(\phi) &= \left| \tfrac{1}{2} \left(\langle e^{i\phi} \rangle + \langle e^{-i\phi} \rangle \right) \right|^{-2} - 1 \\ &= \langle \cos\phi \rangle^{-2} - 1. \end{aligned} \tag{2.2}$$

This means that minimising the Holevo phase variance is equivalent to maximising $\langle \cos\phi \rangle$. Thus minimising the Holevo phase variance is equivalent to minimising the measure of the variance

$$V'(\phi) = 2(1 - \langle \cos\phi \rangle). \tag{2.3}$$

In addition, this measure of the phase variance has the same value in the limit of small variance as the Holevo phase variance. This measure is not exactly equal to the Holevo phase variance, and as is shown later it differs in the higher order terms. This is significant for general dyne measurements where we wish to obtain higher order terms in the expressions for the phase variance.

Now recall from Eq. (1.52) that

$$\langle e^{i\phi} \rangle = \sum_{n=0}^{\infty} \langle \psi | n \rangle \langle n+1 | \psi \rangle H_{n,n+1}. \tag{2.4}$$

This means that we can represent $e^{i\phi}$ as the operator

$$\widehat{\exp(i\phi)} = \sum_{n=0}^{\infty} |n\rangle\langle n+1| H_{n,n+1}, \tag{2.5}$$

and we can represent $\cos\phi$ by the operator

$$2\widehat{\cos\phi} = \sum_{n=0}^{\infty} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n| \right] H_{n,n+1}. \tag{2.6}$$

When we put an upper limit of N on the photon number, we can replace this with

$$2\widehat{\cos\phi} = \sum_{n=0}^{N-1} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n| \right] H_{n,n+1}. \tag{2.7}$$

For canonical measurements $H_{n,n+1} = 1$, so this becomes

$$2\widehat{\cos\phi} = \sum_{n=0}^{N-1} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n| \right]. \tag{2.8}$$

In general, when we wish to maximise (or minimise) the expectation value of some Hermitian operator $\hat{A}$ while keeping the expectation values of other Hermitian operators $\hat{B}$, $\hat{C}\ldots$, constant, we use the method of undetermined multipliers. For this, we require that the matrix elements of all of these operators, $\langle n|\hat{X}|m\rangle$, be real, so the corresponding matrices are symmetric.

The state will be expressed as $|\psi\rangle = \sum_n \psi_n |n\rangle$, where the coefficients ψ_n are real. It is not possible to obtain any smaller phase variance using complex ψ_n. To see this, note that Eq. (2.7) gives

$$\begin{aligned} 2\left\langle \widehat{\cos\phi} \right\rangle &= \sum_{n=0}^{N-1} \left[\psi_n^* \psi_{n+1} + \psi_{n+1}^* \psi_n \right] H_{n,n+1} \\ &= \sum_{n=0}^{N-1} 2\mathrm{Re}\left[\psi_n^* \psi_{n+1} \right] H_{n,n+1}. \end{aligned} \tag{2.9}$$

From this it is clear that the maximal value of $\langle\widehat{\cos\phi}\rangle$ is obtained for real $\psi_n^*\psi_{n+1}$. This in turn implies that the minimum phase variance will be when $\arg(\psi_n)$ is independent of n. Without loss of generality $\arg(\psi_n)$ can be taken to be zero.

Therefore, taking the ψ_n and increments $d\psi_n$ to be real, we obtain

$$\begin{aligned} d\langle\psi|\hat{X}|\psi\rangle &= (d\langle\psi|)\,\hat{X}|\psi\rangle + \langle\psi|\hat{X}\,(d|\psi\rangle) \\ &= \sum_n d\psi_n \left(\langle n|\hat{X}|\psi\rangle + \langle\psi|\hat{X}|n\rangle\right) \\ &= 2\,(d\langle\psi|)\,\hat{X}|\psi\rangle. \end{aligned} \tag{2.10}$$

If we have a maximum of $\langle\hat{A}\rangle$ while $\langle\hat{B}\rangle$, $\langle\hat{C}\rangle\ldots$ are kept constant, then for any increment $(d|\psi\rangle)$ which does not change the expectation values of $\hat{B}$, $\hat{C}\ldots$, the increment in $\langle\hat{A}\rangle$ must be zero also. This means that if

$$(d\langle\psi|)\hat{X}|\psi\rangle = 0, \tag{2.11}$$

for $\hat{X} = \hat{B}$, $\hat{C}\ldots$, then $(d\langle\psi|)\hat{A}|\psi\rangle$ must be equal to zero. This means that $\hat{A}|\psi\rangle$ cannot be linearly independent of $\hat{B}|\psi\rangle$, $\hat{C}|\psi\rangle\ldots$, so it must be possible to write

$$\left(\alpha\hat{A} + \beta\hat{B} + \gamma\hat{C} + \ldots\right)|\psi\rangle = 0, \tag{2.12}$$

for some combination of constants α, β, $\gamma\ldots$.

We wish to maximise the expectation value of the above operator with the single constraint that the state is normalised. Using the method of undetermined multipliers gives

$$\left[\alpha\left(2\widehat{\cos\phi}\right) + \beta\hat{1}\right]|\psi\rangle = 0. \tag{2.13}$$

Rearranging this gives the eigenvalue equation

$$\left(2\widehat{\cos\phi}\right)|\psi\rangle = \nu|\psi\rangle. \tag{2.14}$$

When we expand the state in terms of the photon number states $|\psi\rangle = \sum_{n=0}^{N}\psi_n|n\rangle$, we find

$$\sum_{n=0}^{N-1}\left(\psi_{n+1}|n\rangle + \psi_n|n+1\rangle\right) = \sum_{n=0}^{N}\nu\psi_n|n\rangle. \tag{2.15}$$

Rearranging this gives

$$\sum_{n=0}^{N-1}\psi_{n+1}|n\rangle + \sum_{n=1}^{N}\psi_{n-1}|n\rangle = \sum_{n=0}^{N}\nu\psi_n|n\rangle. \tag{2.16}$$

This equation is equivalent to the recurrence relation

$$\psi_{n+1} = \nu\psi_n - \psi_{n-1}, \tag{2.17}$$

for $0 < n < N$, with the boundary conditions

$$\begin{aligned} \psi_1 &= \nu\psi_0, \\ \psi_{N-1} &= \nu\psi_N. \end{aligned} \tag{2.18}$$

The recurrence relation (2.17) is satisfied by exponentials of the form

$$\psi_n = e^{\pm in\arccos\frac{\nu}{2}}. \tag{2.19}$$

To obtain real solutions, we can use

$$\psi_n = \mathrm{Re}\left(Ae^{in\arccos\frac{\nu}{2}}\right). \tag{2.20}$$

The boundary conditions can be more conveniently expressed by extending the range of the recurrence relation to $0 \leq n \leq N$, and taking

$$\begin{aligned} \psi_{-1} &= 0, \\ \psi_{N+1} &= 0. \end{aligned} \tag{2.21}$$

The first boundary condition then gives the phase of A as

$$\arg A = \operatorname{acos}\frac{\nu}{2} - \frac{\pi}{2}. \tag{2.22}$$

This boundary condition does not give the magnitude, which must be found by normalisation. The last boundary condition gives

$$(N+2)\operatorname{acos}\frac{\nu}{2} = k\pi, \tag{2.23}$$

where k is an arbitrary integer. The eigenvalues are therefore

$$\nu_k = 2\cos\left(\frac{k\pi}{N+2}\right), \tag{2.24}$$

and the coefficients for the corresponding (unnormalised) eigenvectors are

$$\begin{aligned} \psi_n &= \operatorname{Re}\left(-ie^{\frac{i(n+1)k\pi}{N+2}}\right) \\ &= \sin\left(\frac{(n+1)k\pi}{N+2}\right). \end{aligned} \tag{2.25}$$

It is clear from Eq. (2.14) that

$$\left\langle \widehat{2\cos\phi} \right\rangle = \nu, \tag{2.26}$$

so

$$V(\phi) = (\nu/2)^{-2} - 1. \tag{2.27}$$

Therefore the eigenvalue that minimises the Holevo phase variance is the maximum eigenvalue,

$$\nu_1 = 2\cos\left(\frac{\pi}{N+2}\right). \tag{2.28}$$

For this eigenvalue the exact Holevo phase variance is

$$\begin{aligned} V(\phi) &= \left(\cos\left(\frac{\pi}{N+2}\right)\right)^{-2} - 1 \\ &= \tan^2\left(\frac{\pi}{N+2}\right). \end{aligned} \tag{2.29}$$

The state is given by

$$|\psi\rangle = \frac{1}{\sqrt{N/2+1}} \sum_{n=0}^{N} \sin\left(\frac{(n+1)\pi}{N+2}\right)|n\rangle. \tag{2.30}$$

Here the factor of $1/\sqrt{N/2+1}$ is required for the state to be normalised.

Note that for large photon number we have

$$\begin{aligned} V(\phi) &= \tan^2\left(\frac{\pi}{N+2}\right) \\ &\approx \left(\frac{\pi}{N+2}\right)^2 \\ &\approx \frac{\pi^2}{N^2}. \end{aligned} \tag{2.31}$$

This means that the minimum intrinsic phase variance scales down as rapidly as N^{-2}.

2.1.2 General Dyne Measurements

The method of solution for canonical measurements can be generalised to the case of more general dyne measurements, where instead of $H_{n,n+1} = 1$, we have

$$\begin{aligned} H_{n,n+1} &= 1 - h(n) \\ &\approx 1 - cn^{-p}. \end{aligned} \tag{2.32}$$

This case was also considered in Ref. [35]. Again the eigenvalue equation is

$$\left(2\widehat{\cos\phi}\right)|\psi\rangle = \nu|\psi\rangle, \tag{2.33}$$

except this time

$$\begin{aligned} 2\widehat{\cos\phi} &= \sum_{n=0}^{N-1} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n|\right] H_{n,n+1} \\ &= \sum_{n=0}^{N-1} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n|\right] (1 - h(n)). \end{aligned} \tag{2.34}$$

Expanding this out gives

$$\begin{aligned} \sum_{n=0}^{N-1} \psi_{n+1}|n\rangle + \sum_{n=1}^{N} \psi_{n-1}|n\rangle - \sum_{n=0}^{N-1} \left[h(n)\psi_{n+1}|n\rangle + h(n)\psi_n|n+1\rangle\right] &= \sum_{n=0}^{N} \nu\psi_n|n\rangle, \\ \sum_{n=0}^{N-1} \psi_{n+1}|n\rangle + \sum_{n=1}^{N} \psi_{n-1}|n\rangle - \sum_{n=0}^{N-1} h(n)\psi_{n+1}|n\rangle - \sum_{n=1}^{N} h(n-1)\psi_{n-1}|n\rangle &= \sum_{n=0}^{N} \nu\psi_n|n\rangle. \end{aligned} \tag{2.35}$$

This gives the recursion relation for $0 < n < N$,

$$\psi_{n+1} + \psi_{n-1} - h(n)\psi_{n+1} - h(n-1)\psi_{n-1} = \nu\psi_n. \tag{2.36}$$

In order to solve this, in Ref. [35] the photon number n is treated as a continuous variable, and the coefficients ψ_n are replaced with the function $\psi(n)$, which is assumed to be twice differentiable. Then the approximations are made

$$\psi(n+1) + \psi(n-1) \approx \left[2 + \frac{\partial^2}{\partial n^2}\right]\psi(n), \tag{2.37}$$

$$h(n)\psi(n+1) + h(n-1)\psi(n-1) \approx 2h(n)\psi(n). \tag{2.38}$$

The second approximation (2.38) is based upon using $h(n) \approx h(n-1)$ and $\psi(n+1)+\psi(n-1) \approx 2\psi(n)$. The second derivative is not used, because this term is already much smaller than $\psi(n)$.

With these approximations, the eigenvalue equation becomes

$$\left(2 + \frac{\partial^2}{\partial n^2} - 2cn^{-p}\right)\psi(n) = \nu\psi(n). \tag{2.39}$$

In Ref. [35], the term $2cn^{-p}$ is linearised about $n = N$, and the variables are changed to $y = 1 - N^{-1}n$. The equation is then

$$\left(-\frac{\partial^2}{\partial y^2} + by\right)\psi(y) = a_k\psi(y), \tag{2.40}$$

where

$$\begin{aligned} a_k &= N^2(2 - \nu_k - 2cN^{-p}), \\ b &= 2cpN^{2-p}. \end{aligned} \tag{2.41}$$

The boundary condition is $\psi(0) = 0$, and the boundary condition at $y = 1$ is ignored.

Note that, similarly to the case for canonical measurements, the exact boundary conditions should be $\psi_{-1} = \psi_{N+1} = 0$. In terms of y, this means we should have $\psi(y) = 0$ for $y = -1/N$ and $1 + 1/N$. The $1/N$ terms in these boundary conditions are ignored, as they merely give higher order corrections to the results.

Using these approximations, the solutions are

$$\psi_k(y) \propto \mathrm{Ai}(z_k + b^{1/3} y), \tag{2.42}$$

where Ai is the Airy function and z_k is the kth real zero of the Airy function satisfying $0 > z_1 > z_2 > \ldots$. The eigenvalues are

$$\nu_k = 2 - \left(2cN^{-p} + |z_k| \, (2cp)^{2/3} N^{-2(1+p)/3}\right), \tag{2.43}$$

and so it is obvious that the maximum eigenvalue is for $k = 1$. The corresponding Holevo phase variance is

$$\begin{aligned} V(\phi) &= (\nu_k/2)^{-2} - 1 \\ &= \left[1 - \left(cN^{-p} + \tfrac{1}{2} \, |z_1| \, (2cp)^{2/3} N^{-2(1+p)/3}\right)\right]^{-2} - 1 \\ &\approx \left[1 + \left(2cN^{-p} + |z_1| \, (2cp)^{2/3} N^{-2(1+p)/3}\right) + 3(cN^{-p})^2\right] - 1 \\ &= 2cN^{-p} + |z_1| \, (2cp)^{2/3} N^{-2(1+p)/3} + 3c^2 N^{-2p}. \end{aligned} \tag{2.44}$$

This method should be used rather than using the approximation $V(\phi) \approx 2 - v_k$ (as is used in [35]), as the third term here will be of the same order as the second for mark I measurements.

The value of the first zero of the Airy function is $z_1 \approx -2.33810741045976$. Using these results, and the values of c and p for heterodyne, mark I and mark II measurements given in Table 1.1, the variances for these measurements should be given by

$$V(\phi_{\mathrm{het}}) \approx \tfrac{1}{4} N^{-1} + 0.927879 \times N^{-4/3}, \tag{2.45}$$

$$V(\phi_{\mathrm{I}}) \approx \tfrac{1}{4} N^{-1/2} + 0.631402 \times N^{-1}, \tag{2.46}$$

$$V(\phi_{\mathrm{II}}) \approx \tfrac{1}{8} N^{-3/2} + 0.765947 \times N^{-5/3}. \tag{2.47}$$

The second term for mark I measurements here includes the third term from Eq. (2.44). This term has been omitted in the other two cases, as it is of higher order. As is discussed in the next section, $h_{\mathrm{I}}(n)$ has an extra term of order N^{-1}. This term would also need to be taken into account in order for the second term for mark I measurements to be accurate.

2.1.3 Numerical Results

These analytic results have been verified numerically by calculating the optimised states for heterodyne measurements and adaptive mark I and II measurements. For moderate maximum photon numbers the calculations were exact. For heterodyne measurements the vector $h_{\mathrm{het}}(n)$ is given exactly by [34]

$$h_{\mathrm{het}}(n) = 1 - \frac{\Gamma\left(n + 3/2\right)}{\sqrt{\Gamma\left(n+1\right)\Gamma\left(n+2\right)}}. \tag{2.48}$$

This form of the equation can not be used for large n due to roundoff error. For $n \geq 10$ the first twelve terms of an asymptotic expansion were used, giving results as accurate as or more accurate than the exact expression (2.48). These terms were determined from the asymptotic expansion for $\log\Gamma\left(n\right)$ [46], and are given below:

$$\begin{aligned} h_{\mathrm{het}}(n) \sim{}& \frac{1}{8(n+1)} - \frac{1}{2^7(n+1)^2} - \frac{5}{2^{10}(n+1)^3} + \frac{21}{2^{15}(n+1)^4} + \frac{399}{2^{18}(n+1)^5} - \frac{869}{2^{22}(n+1)^6} \\ &- \frac{39325}{2^{25}(n+1)^7} + \frac{334477}{2^{31}(n+1)^8} + \frac{28717403}{2^{34}(n+1)^9} - \frac{59697183}{2^{38}(n+1)^{10}} - \frac{8400372435}{2^{41}(n+1)^{11}} + \frac{34429291905}{2^{46}(n+1)^{12}} + \ldots \end{aligned} \tag{2.49}$$

The exact expression for mark I measurements is slightly more complicated. From [34] the exact expression for H_{mn} for mark I measurements is

$$H^{\mathrm{I}}_{mn} = \sum_{p=0}^{\lfloor m/2 \rfloor} \sum_{q=0}^{\lfloor n/2 \rfloor} \gamma_{mp}\gamma_{nq} M^{p,q}, \tag{2.50}$$

where $\lfloor m/2 \rfloor$ is the integer part of $m/2$,

$$\gamma_{mp} = \frac{\sqrt{m!}}{2^p (m-2p)! p!}, \tag{2.51}$$

and $M^{p,q}$ are calculated using the recursion relation

$$M^{n,m} = \frac{n M^{n-1,m} + m M^{n,m-1}}{2(n-m)^2 + n + m}, \tag{2.52}$$

with the boundary values

$$M^{n,0} = M^{0,n} = \frac{1}{(2n+1)(2n-1)\ldots 1} = \frac{1}{(2n+1)!!}. \tag{2.53}$$

The values of $h_{\mathrm{I}}(n) = 1 - H^{\mathrm{I}}_{n,n+1}$ were calculated up to $n = 3000$ using this expression. Further values were extrapolated by fitting an asymptotic expansion to the results below 3000. The first term, $1/(8\sqrt{n})$, was assumed, and three further terms were obtained by fitting techniques. The terms found were

$$h_{\mathrm{I}}(n) \approx \frac{1}{8n^{1/2}} - \frac{0.101561734071163}{n} - \frac{0.0508542807551548}{n^{3/2}} + \frac{0.0959147066823866}{n^2}. \tag{2.54}$$

These coefficients were the exact values used in calculations, but the digits given do not reflect the accuracy of the fit. (I have given all digits used so that it is possible to accurately reproduce the results.)

The exact expression for mark II measurements is even more complicated than that for mark I measurements. From [34] the expression is

$$H^{\mathrm{II}}_{mn} = \sum_{p=0}^{\lfloor m/2 \rfloor} \sum_{q=0}^{\lfloor n/2 \rfloor} \gamma_{mp}\gamma_{nq} \left\langle \left(\frac{1+D}{1+D^*} \right)^{(n-m)/2} D^p (D^*)^q \right\rangle_Q. \tag{2.55}$$

where $D = BA^*/A$, and the subscript Q indicates that the expectation value is for the "ostensible" or vacuum distribution. This notation differs slightly from that in [34], where the symbol C was used, rather than D. In Ref. [34] it is shown that

$$\langle D^n D^{*m} \rangle_Q = M^{n,m}, \tag{2.56}$$

where $M^{n,m}$ is calculated as above. In order to obtain $h_{\mathrm{II}}(n)$ we require $H^{\mathrm{II}}_{n,n+1}$. In order to calculate these we can expand Eq. (2.55) to obtain

$$H^{\mathrm{II}}_{n,n+1} = \sum_{p=0}^{\lfloor n/2 \rfloor} \sum_{q=0}^{\lfloor (n+1)/2 \rfloor} \gamma_{np}\gamma_{n+1,q} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(3/2-k)_k}{k!} \frac{(1/2-l)_l}{l!} M^{k+p,l+q}, \tag{2.57}$$

where

$$(\alpha)_n = \alpha(\alpha+1)\ldots(\alpha+n-1). \tag{2.58}$$

The infinite sum in (2.57) can be determined to within double precision (15 digits) by summing about the first 100 terms. The values of $h_{\mathrm{II}}(n)$ were calculated up to $n = 1000$ in this way. I attempted to obtain the higher order terms using this data set in a similar way as for the mark

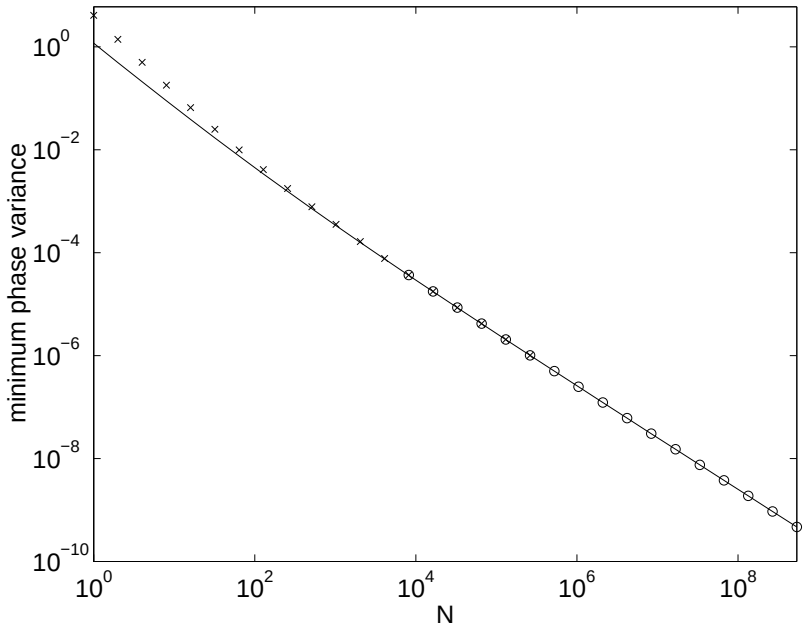

Figure 2.1: The minimum phase variance for heterodyne measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic analytic expression as the continuous line.

I case, but unfortunately it was not found to be possible to consistently obtain any terms beyond $\frac{1}{16}n^{-1.5}$. Therefore all elements beyond $n = 1000$ were determined using $h_{\mathrm{II}}(n) \approx \frac{1}{16}n^{-1.5}$.

Using the above methods for calculating $h(n)$, the minimum phase variance was determined by numerically determining the eigenvalues for heterodyne, mark I and mark II measurements, for photon numbers up to 2^{18}. These calculations were exact except for the above approximations to $h(n)$ for mark I and II measurements.

For larger photon numbers it was not feasible to solve the exact eigenvalue problem, but an approximate solution was obtained by using the continuous approximation of the eigenvalue problem and discretising it. In order to reduce the number of intervals required in the discretised equation, the equation was solved for three different numbers (512, 1024 and 2048) of intervals. The result for the continuous case was then estimated by projecting to zero step size assuming the error is quadratic in the step size. The approximations that were *not* made (that were made in order to derive the analytic result) were the linear approximation of $h(n)$ and the omission of the boundary condition at $y = 1$.

The results for heterodyne measurements are shown in Fig. 2.1. The results for the exact calculations agree extremely well with the results for the continuous approximation over the region where both values have been calculated. This indicates that the continuous approximation is a very good approximation of the exact eigenvalue problem.

The results also agree well with the asymptotic expression (2.44), with good agreement for photon numbers above 100. In order to better see the difference between the numeric results and the analytic expression, and in particular to see how accurate the second term in the analytic expression is, it is convenient to define the parameter z by

$$z = \frac{V(\phi) - 2cN^{-p}}{(2cp)^{2/3}N^{-2(1+p)/3}}. \tag{2.59}$$

From Eq. (2.44), provided the third term $3c^2N^{-2p}$ can be ignored, this should converge to $|z_1| \approx 2.338107$. The values of z for heterodyne measurements are plotted in Fig. 2.2. Again there is extremely good agreement between the values calculated exactly and those calculated using the

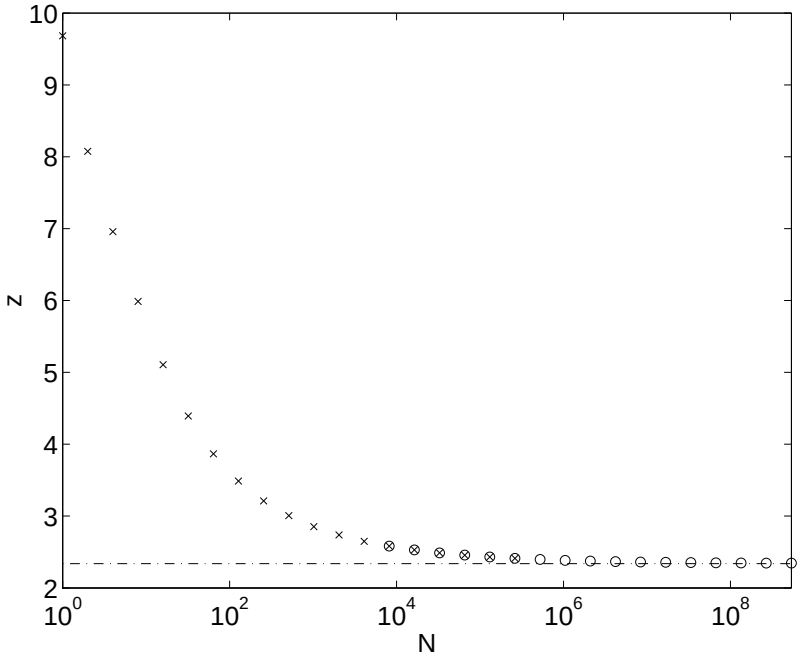

Figure 2.2: The value of z for heterodyne measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the theoretical asymptotic value of $|z_1| \approx 2.338107$ is shown as the dash-dotted line.

continuous approximation. The value of z does not converge closely to $|z_1|$ until a photon number of around 10^6. In fact, we do not have 1% agreement until a photon number of 8×10^6.

The results for mark I measurements are plotted in Fig. 2.3. The asymptotic expression plotted here is Eq. (2.46), which includes the extra $3c^2N^{-2p}$ term. There is good agreement between the results and this asymptotic expression for photon numbers above about 10. There is also good agreement between the values calculated exactly and those calculated using the continuous approximation.

For mark I measurements the values of z cannot be expected to converge to $|z_1|$ due to the extra $3c^2N^{-2p}$ term. From the asymptotic expression (2.46), the asymptotic value including this term should be 2.525607. As can be seen in Fig. 2.4, however, z does not converge to this value. The reason for this is that the second term of the asymptotic expression for $V(\phi_{\mathrm{I}})$ is of the same order as the second term in the asymptotic expression for $h_{\mathrm{I}}(n)$ in Eq. (2.54). Taking account of this term, $V(\phi_{\mathrm{I}})$ should be

$$V(\phi_{\mathrm{I}}) \approx \tfrac{1}{4}N^{-1/2} + 0.428278 \times N^{-1}, \tag{2.60}$$

and z should converge to approximately 1.713114. This value is also plotted in Fig. 2.4. As can be seen, the results converge quite accurately to this value. The results for the largest photon numbers do not agree accurately; however, this is just due to poor convergence of the numerical technique.

The results for mark II measurements are plotted in Fig. 2.5. In this case there is good agreement between the results calculated using the two different methods, and the results appear to be close to the asymptotic result for photon numbers above about 10^5. When we look at the values of z (see Fig. 2.6), we see that there is again excellent agreement between the results calculated using the two different methods, but there is good agreement with the asymptotic value only for photon numbers above 10^{10}. In fact, we require a photon number above 5×10^{14} in order to have better than 1% agreement.

Note that in [35] it is claimed that there should be good convergence to the asymptotic values

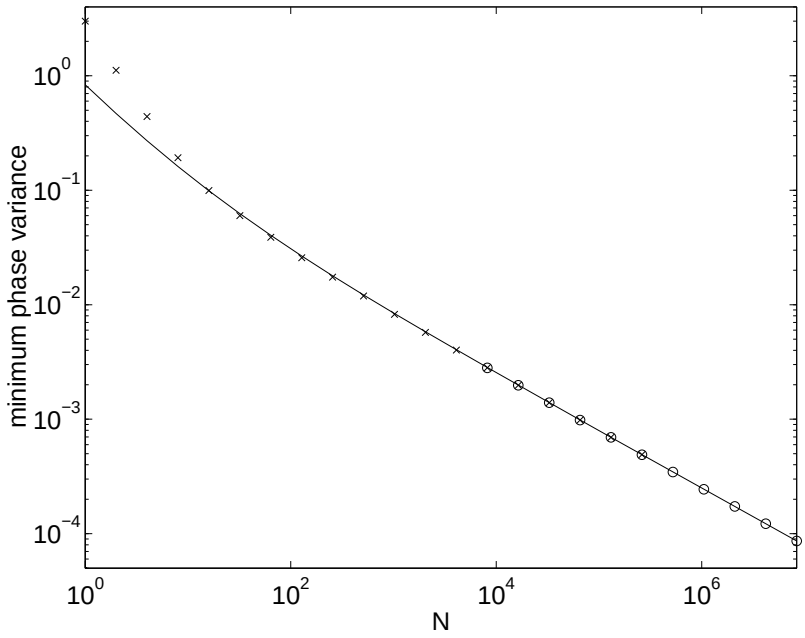

Figure 2.3: The minimum phase variance for mark I measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic analytic expression as the continuous line.

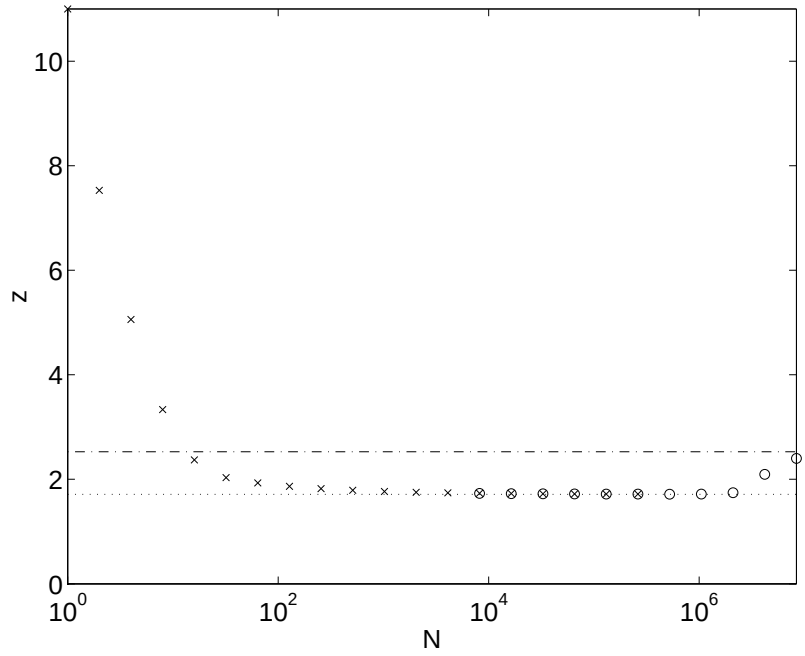

Figure 2.4: The value of z for mark I measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the theoretical asymptotic value of 2.525607 is shown as the dash-dotted line. The asymptotic value taking into account the second term for $h_{\mathrm{I}}(n)$ is shown as the dotted line.

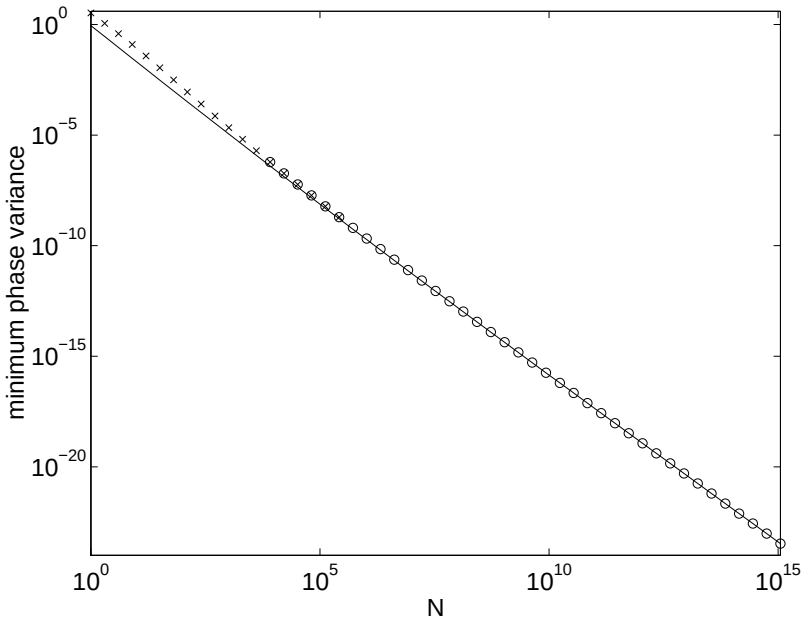

Figure 2.5: The minimum phase variance for mark II measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic analytic expression as the continuous line.

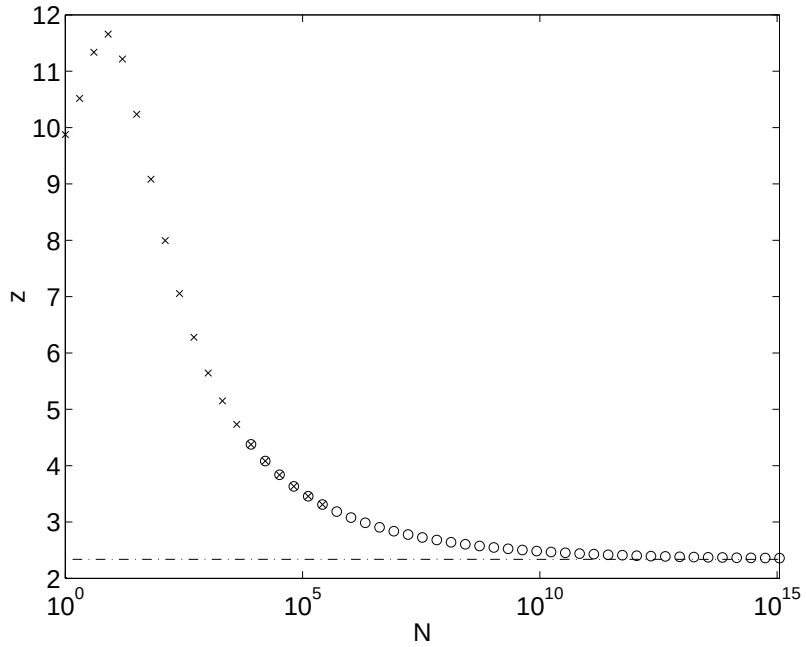

Figure 2.6: The value of z for mark II measurements on states with an upper limit N on the photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the theoretical asymptotic value of $|z_1| \approx 2.338107$ is shown as the dash-dotted line.

for

$$N \geq \left(\frac{10^3}{2cp}\right)^{1/(1-p)}. \tag{2.61}$$

This means that there should be good agreement with the asymptotic values for photon numbers above about 4000, 400 and 3×10^7 for heterodyne, mark I and mark II measurements respectively. The numerical results obtained here show that, in order to have better than 1% agreement with the asymptotic results, the maximum photon numbers should be above about 8×10^6, 4000 and 5×10^{14} respectively.

These apparent discrepancies of many orders of magnitude are not so great when considered in terms of the difference in z rather than the photon number. For a photon number close to 4000 for heterodyne measurements, z is about 13% from the asymptotic value, which is not very poor agreement. Similarly for a photon number around 400 for mark I measurements the value of z is 4% from the asymptotic value, and for a photon number around 3×10^7 for mark II measurements the value of z is 17% from the asymptotic value. Nevertheless, these discrepancies are still larger than would be expected.

To understand the reason for this, note that the criterion used in [35] was that 99.5% of the solution $\psi(y)$ [as given by Eq. (2.42)] be confined within the lower half of the interval $[0, 1]$. This criterion means that the approximation that we can ignore the boundary condition at $y = 1$ is very good, but it is a fairly weak criterion for the approximation that we can linearise $h(n)$ about $n = N$ (equivalent to $y = 0$). A better criterion for this approximation to be accurate is that 99.5% of the solution be confined within $y < 0.04$. This criterion means that we require $5b^{-1/3} < 0.04$, where $b = 2cpN^{2-p}$, and gives

$$N > \left(\frac{10^6}{cp}\right)^{1/(2-p)}. \tag{2.62}$$

With this criterion the minimum photon numbers for heterodyne, mark I and mark II measurements are 8×10^6, 6×10^4 and 10^{14} respectively. With the exception of the mark I result, these values are much closer to the photon numbers required for 1% agreement with the asymptotic value of z.

2.2 Fixed Mean Photon Number

The next case that I consider is that where, rather than an upper limit being put on the photon number, the mean photon number is fixed. This case is rather more complicated, as the method of undetermined multipliers gives two undetermined constants, so the problem is more complicated than a simple eigenvalue problem.

2.2.1 Canonical Measurements

The simpler case of canonical measurements with a fixed mean photon number was solved by Summy and Pegg [37]. I will give the derivation here, as it is simple and indicates the method of solution for the case of general dyne measurements. I am giving a different but equivalent version of the derivation for consistency with the other derivations given in this chapter.

We wish to maximise the expectation value of $2\widehat{\cos}\,\phi$ while keeping the state normalised and the photon number constant. Using the method of undetermined multipliers gives the equation

$$\left[\alpha\left(2\widehat{\cos}\,\phi\right) + \beta\hat{1} + \gamma\hat{N}\right]|\psi\rangle = 0. \tag{2.63}$$

Rearranging this gives

$$\left[\left(2\widehat{\cos}\,\phi\right) - \mu\hat{N}\right]|\psi\rangle = \nu|\psi\rangle. \tag{2.64}$$

This has two unknown constants, and therefore cannot be solved as a simple eigenvalue problem. Instead what we do is solve it as an eigenvalue equation for ν with a fixed value of μ, and the eigenstate corresponding to the maximum eigenvalue is an optimised state. The mean photon

number can then be found from this state. We can obtain a range of mean photon numbers by adjusting μ.

Expanding the equation out in terms of the number coefficients of $|\psi\rangle$ gives

$$\sum_{n=0}^{\infty}\psi_{n+1}|n\rangle+\sum_{n=1}^{\infty}\psi_{n-1}|n\rangle-\sum_{n=0}^{\infty}\mu\psi_n n|n\rangle=\sum_{n=0}^{\infty}\nu\psi_n|n\rangle, \tag{2.65}$$

which implies the recurrence relation

$$\psi_{n+1}+\psi_{n-1}-\mu\psi_n n=\nu\psi_n. \tag{2.66}$$

Similarly to the derivation for general dyne measurements when there is an upper limit on the photon number, we can use the continuous approximation where we replace ψ_n with $\psi(n)$, and use

$$\psi(n+1)+\psi(n-1)\approx\left[2+\frac{\partial^2}{\partial n^2}\right]\psi(n). \tag{2.67}$$

With this approximation the recurrence relation becomes the differential equation

$$\left(-\frac{\partial^2}{\partial n^2}+\mu n\right)\psi(n)=(2-\nu)\psi(n). \tag{2.68}$$

This equation is exactly equivalent to the differential equation obtained in the maximum photon number case (2.40), and it therefore has the solutions

$$\psi_k(n)\propto\mathrm{Ai}(z_k+\mu^{1/3}n), \tag{2.69}$$

with the eigenvalues

$$\nu_k=2-|z_k|\,\mu^{2/3}. \tag{2.70}$$

It is clear that the solution that maximises ν is again that with $k=1$. In Ref. [37] the authors say that the solution must have a zero for $n=-\epsilon$, where $0<\epsilon<1$. The boundary condition for the discrete equation is $\psi_{-1}=0$, which implies that the solution should have a zero for $n=-1$. I will take $\psi(0)=0$ here, as the difference only gives higher order terms to the solution.

It can be shown numerically that

$$\left\langle z_1+\mu^{1/3}n\right\rangle\approx-0.779369136819922. \tag{2.71}$$

For brevity I will call this constant $\langle X\rangle$ (for consistency with [37]). Rearranging this gives

$$\mu^{1/3}=\frac{\langle X\rangle-z_1}{\langle n\rangle}. \tag{2.72}$$

This gives the relation between the mean photon number and μ. The eigenvalue is then

$$\nu_1=2-\frac{|z_1|\,(\langle X\rangle-z_1)^2}{\langle n\rangle^2}. \tag{2.73}$$

From Eq. (2.64) it is clear that

$$\left\langle 2\widehat{\cos\phi}\right\rangle-\mu\,\langle n\rangle=\nu. \tag{2.74}$$

Substituting the values for μ and ν from above gives

$$\left\langle 2\widehat{\cos\phi}\right\rangle-\frac{(\langle X\rangle-z_1)^3}{\langle n\rangle^2}=2-\frac{|z_1|\,(\langle X\rangle-z_1)^2}{\langle n\rangle^2}. \tag{2.75}$$

This simplifies to

$$2-\left\langle 2\widehat{\cos\phi}\right\rangle=\frac{-\langle X\rangle\,(\langle X\rangle-z_1)^2}{\langle n\rangle^2} \tag{2.76}$$

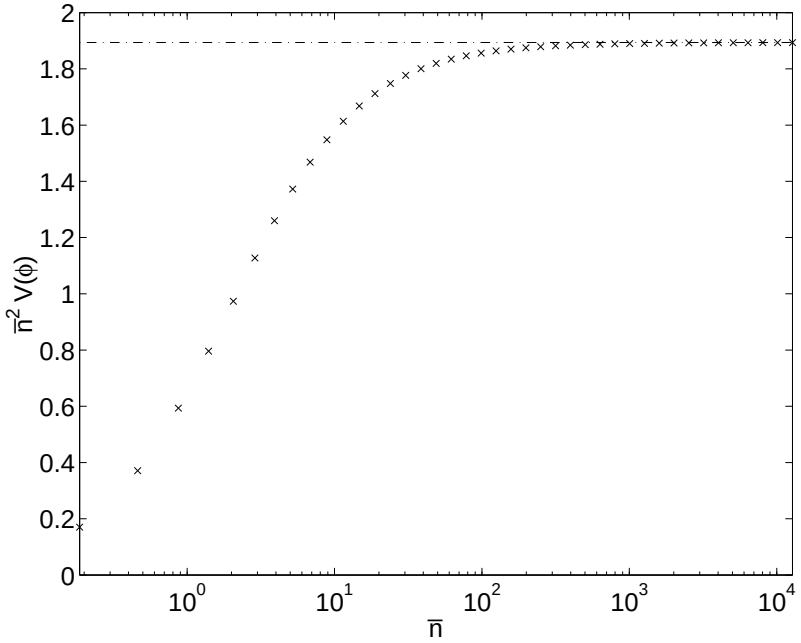

Figure 2.7: The phase variance multiplied by the square of the mean photon number for states optimised for minimum phase variance with the constraint of fixed mean photon number. The theoretical asymptotic value of 1.893606 is shown as the dash-dotted line.

Therefore the phase variance is

$$V(\phi) = \frac{1.89360591826155}{\overline{n}^2}. \tag{2.77}$$

This is the result obtained in [37]. Therefore we see that the phase variance for optimum states scales as $\bar{n}^{-2}$, the same as for the case with an upper limit on the photon number. Note that it is accurate to approximate the Holevo phase variance by $2 - \left\langle 2\widehat{\cos\phi} \right\rangle$ here, as the differences will be of order $\overline{n}^{-4}$.

This analytic result has been verified by determining the optimum states numerically. The eigenvalue problem was solved with various values of μ, and for each value of μ the maximum eigenvalue was chosen, and the mean photon number was determined from the corresponding eigenstate. In order to make the problem finite, the state coefficients were only considered for photon numbers up to around 11 times the mean photon number of the state. At this point the state coefficients had fallen to around 10^{-16}.

The results multiplied by $\overline{n}^2$ are plotted in Fig. 2.7. The results converge rapidly to the asymptotic value, with agreement within 1% of the asymptotic value for mean photon numbers above about 250. Note that it does not make sense to perform calculations with the continuous version of the eigenvalue equation, as this case was solved exactly.

2.2.2 General Dyne Measurements

Now I will consider the case of optimising for minimum phase variance with a fixed mean photon number for the case of more general dyne measurements. The complete derivation is original to this study, and is based on a partial derivation by Zhong-Xi Zhang (personal communication). This derivation was published in a brief form in [47].

Similarly to the case of canonical measurements, the method of undetermined multipliers gives the equation

$$\left[\left(2\widehat{\cos\phi}\right) - \mu\hat{N}\right]|\psi\rangle = \nu|\psi\rangle. \tag{2.78}$$

For the case of more general dyne measurements, we have

$$2\widehat{\cos\phi} = \sum_{n=0}^{\infty} \left[|n\rangle\langle n+1| + |n+1\rangle\langle n|\right](1-h(n)). \tag{2.79}$$

If the state is expressed in the number states basis

$$|\psi\rangle = \sum_{n=0}^{\infty} \psi_n |n\rangle, \tag{2.80}$$

then Eq. (2.78) can be expanded to obtain

$$\sum_{n=0}^{\infty} \psi_{n+1}|n\rangle + \sum_{n=1}^{\infty} \psi_{n-1}|n\rangle - \sum_{n=0}^{\infty} h(n)\psi_{n+1}|n\rangle - \sum_{n=1}^{\infty} h(n-1)\psi_{n-1}|n\rangle - \sum_{n=0}^{\infty} \mu\psi_n n|n\rangle = \sum_{n=0}^{\infty} \nu\psi_n|n\rangle. \tag{2.81}$$

This gives the recurrence relation

$$\psi_{n+1} + \psi_{n-1} - h(n)\psi_{n+1} - h(n-1)\psi_{n-1} - \mu\psi_n n = \nu\psi_n. \tag{2.82}$$

Similarly to the derivation for an upper limit on the photon number, we can treat n as a continuous variable, and make the approximations

$$\begin{aligned} \psi(n+1) + \psi(n-1) &\approx \left[2 + \frac{\partial^2}{\partial n^2}\right]\psi(n) \\ h(n)\psi(n+1) + h(n-1)\psi(n-1) &\approx 2h(n)\psi(n). \end{aligned} \tag{2.83}$$

Using these approximations, Eq. (2.82) becomes

$$-\frac{d^2\psi(n)}{dn^2} + \psi(n)\left[2h(n) + \nu + \mu n - 2\right] \approx 0. \tag{2.84}$$

Now define

$$f(n) = 2h(n) + \nu + \mu n - 2. \tag{2.85}$$

To solve Eq. (2.84), it is convenient to expand $f(n)$ in a Taylor series around

$$n_0 = \left(\frac{\mu}{2cp}\right)^{-1/(p+1)}. \tag{2.86}$$

The derivatives of $f(n)$ are

$$f(n_0) = 2h(n_0) + \nu + \mu n_0 - 2, \tag{2.87}$$
$$f'(n_0) = 2h'(n_0) + \mu = 0, \tag{2.88}$$
$$f''(n_0) = 2h''(n_0) = 2cp(p+1)n_0^{-p-2}, \tag{2.89}$$
$$f'''(n_0) = -2cp(p+1)(p+2)n_0^{-p-3}. \tag{2.90}$$

Note that this technique requires that the number distribution has its maximum near n_0. This is justified in the derivation in Appendix A.1.

Using the Taylor series for $f(n)$, and defining $f_0 = f(n_0)$, $f_2 = f''(n_0)/2$ and $f_3 = f'''(n_0)/6$, Eq. (2.84) becomes

$$\left\{-\frac{d^2}{dn^2} + \left[(n-n_0)^2 f_2 + (n-n_0)^3 f_3\right]\right\}\psi(n) \approx -f_0\psi(n). \tag{2.91}$$

Note that $-f_0 = 2 - (2h(n_0) + \nu + \mu n_0)$, so the above equation is equivalent to solving (2.78) as an eigenvalue equation for ν with a fixed value of μ. Now Eq. (2.91) is equivalent to the time-independent Schrödinger's equation with energy eigenvalue

$$E = -f_0, \tag{2.92}$$

for a perturbed harmonic Hamiltonian $\hat{H}$. We can apply perturbation theory with

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \tag{2.93}$$

$$\hat{H}_0 = -\frac{d^2}{dn^2} + \left[f_2(n-n_0)^2\right], \tag{2.94}$$

$$\hat{H}_1 = f_3(n-n_0)^3. \tag{2.95}$$

This perturbation theory derivation is fairly lengthy, and the details are contained in Appendix A.1. It is shown that the energy eigenvalue of the unperturbed ground state is $\sqrt{f_2}$, and that the mean photon number for the perturbed state is

$$\bar{n} \approx n_0 + \frac{p+2}{4\sqrt{cp(p+1)}} n_0^{p/2}. \tag{2.96}$$

Using this we can find the minimum phase variance based on

$$\langle 2\cos\phi \rangle_{\max} = (\nu + \mu\bar{n})_{\max}. \tag{2.97}$$

This can be evaluated as

$$\nu + \mu\overline{n} = 2 - 2h(\overline{n}) + f(\overline{n}). \tag{2.98}$$

Using a Taylor expansion for $f(\overline{n})$ gives

$$\nu + \mu\overline{n} = 2 - 2h(\overline{n}) + f_0 + (\overline{n} - n_0)^2 f_2. \tag{2.99}$$

Using the result for $\overline{n}$ in Eq. (2.96), the last term is of order $\overline{n}^{-2}$, which is small enough to be omitted here.

It is shown in Appendix A.1 that the energy eigenvalue of the state corresponding to the smallest phase variance is $\sqrt{f_2}$. The correction found by perturbation theory is of order $\overline{n}^{-2}$, and can be omitted. As $-f_0$ is the energy eigenvalue, we get

$$\begin{aligned} \nu + \mu\bar{n} &= 2 - \left[2h(\overline{n}) + \sqrt{f_2}\right] \\ &= 2 - \left[2c\bar{n}^{-p} + \sqrt{cp(p+1)}n_0^{-p/2-1}\right]. \end{aligned} \tag{2.100}$$

The Holevo phase variance is then given by

$$\begin{aligned} V(\phi) &= [(\nu + \mu\bar{n})/2]^{-2} - 1 \\ &\approx 2c\bar{n}^{-p} + \sqrt{cp(p+1)}\bar{n}^{-p/2-1} + 3c^2\bar{n}^{-2p}. \end{aligned} \tag{2.101}$$

Note that the first term here is the same as the result when an upper limit is put on the photon number, but the second term scales as a different power of $\bar{n}$. Similarly to the case where there is an upper limit on the photon number, the phase variance should be found in this way rather than using $V(\phi) \approx 2 - (\nu + \mu\bar{n})$. This method gives the extra term $3c^2\bar{n}^{-2p}$, which is of lower order than the second term for mark I measurements. This term can be ignored in the other two cases, as it is of higher order.

Using this relation, the Holevo phase variance for the three cases, heterodyne, mark I and mark II, should scale as

$$V(\phi_{\text{het}}) \approx \tfrac{1}{4}\bar{n}^{-1} + \tfrac{1}{2}\bar{n}^{-3/2}, \tag{2.102}$$

$$V(\phi_{\text{I}}) \approx \tfrac{1}{4}\bar{n}^{-1/2} + \frac{3}{64}\overline{n}^{-1} + \frac{\sqrt{3}}{4\sqrt{2}}\bar{n}^{-5/4}, \tag{2.103}$$

$$V(\phi_{\text{II}}) \approx \tfrac{1}{8}\bar{n}^{-3/2} + \frac{\sqrt{15}}{8}\bar{n}^{-7/4}. \tag{2.104}$$

For mark I measurements there will also be a term of order $\overline{n}^{-1}$ due to the second term in the expansion for $h_{\mathrm{I}}(n)$. Taking this term into account the variance should be

$$V(\phi_{\mathrm{I}}) \approx \tfrac{1}{4}\bar{n}^{-1/2} - 0.156248 \times \overline{n}^{-1} + \frac{\sqrt{3}}{4\sqrt{2}}\bar{n}^{-5/4}. \tag{2.105}$$

The case of heterodyne detection is of particular interest because it differs radically from the result claimed by D'Ariano and Paris [48] of

$$V(\phi_{\mathrm{het}}) = \frac{1.00 \pm 0.02}{\bar{n}^{1.30 \pm 0.02}}. \tag{2.106}$$

As the quoted errors suggest, this result was obtained entirely numerically, in contrast to the analytical result obtained here. Judging from the graphs given in [48], the numeric fit was performed for relatively small photon numbers, only up to about 100. As was found in previous sections, it generally takes very large photon numbers for the asymptotic scaling to become evident. In Sec. 2.2.3 I present numerical results that show that the analytic result in Eq. (2.103) is a far better fit than the power law of D'Ariano and Paris.

2.2.3 Numerical Results

These analytic results have been verified by numerically calculating the optimum states for dyne measurements with a fixed mean photon number. In these calculations the values of $h(n)$ for heterodyne, mark I and mark II measurements were calculated using the formulae described in Sec. 2.1.3. The phase variances and photon numbers were determined in a similar way as in Sec. 2.2.1. Fixed values of μ were used, and for each value of μ the maximum eigenvalue was found, and the mean photon number was determined from the corresponding eigenvector.

Similarly to Sec. 2.2.1 a cutoff was used at a photon number sufficiently above the mean photon number that the state had fallen to around 10^{-16}. In addition, calculations were performed using the continuous approximation in order to obtain results for very large photon numbers. The method used was similar to that used in Sec. 2.1.3. The phase variances and photon numbers were determined for 512, 1024 and 2048 intervals, and the result for the continuous case was then estimated by projecting to zero step size.

The results for the exact case and continuous approximation for heterodyne measurements are shown in Fig. 2.8. The power law claimed by D'Ariano and Paris is also shown in this figure. As can be seen, there is good agreement between the numerically calculated values and the asymptotic analytic expression for photon numbers above 100. There is also excellent agreement between the values calculated exactly and those calculated using the continuous approximation.

Note that the power law of D'Ariano and Paris gives good agreement for photon numbers below 100, better than the asymptotic analytic expression obtained here. For photon numbers above 100, however, this power law differs greatly from the numerically calculated values, demonstrating that this power law is not the correct asymptotic scaling.

In order to see how accurately the numerical results agree with second term in the asymptotic limit, we can define the parameter z in a similar way as was defined in Sec. 2.1.3:

$$z = \frac{V(\phi) - 2c\bar{n}^{-p}}{\bar{n}^{-p/2-1}}. \tag{2.107}$$

Provided the term $3c^2\overline{n}^{-2p}$ can be ignored, the asymptotic value of z should be $\sqrt{cp(p+1)}$. This parameter is plotted in Fig. 2.9. Again there is excellent agreement between the values calculated exactly and the continuous approximation. There is still a small but significant difference between the numeric values and the asymptotic limit for a photon number of 100, but for photon numbers above 10^4 there is better than 1% agreement with the asymptotic value.

The calculated results for mark I measurements are shown in Fig. 2.10. In this case the algorithm for calculating the continuous approximation did not give convergent results, so only the exact results are shown. There is good agreement with the asymptotic expression (2.103) for photon numbers

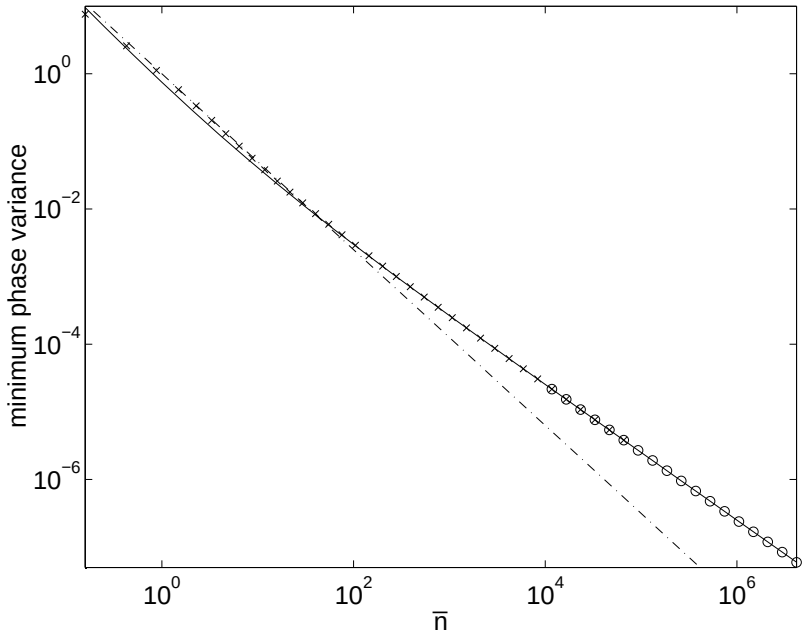

Figure 2.8: The minimum phase variance for heterodyne measurements on states with a fixed mean photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic analytic expression as the continuous line. The power law claimed by D'Ariano and Paris for heterodyne detection is also plotted (dash-dotted line).

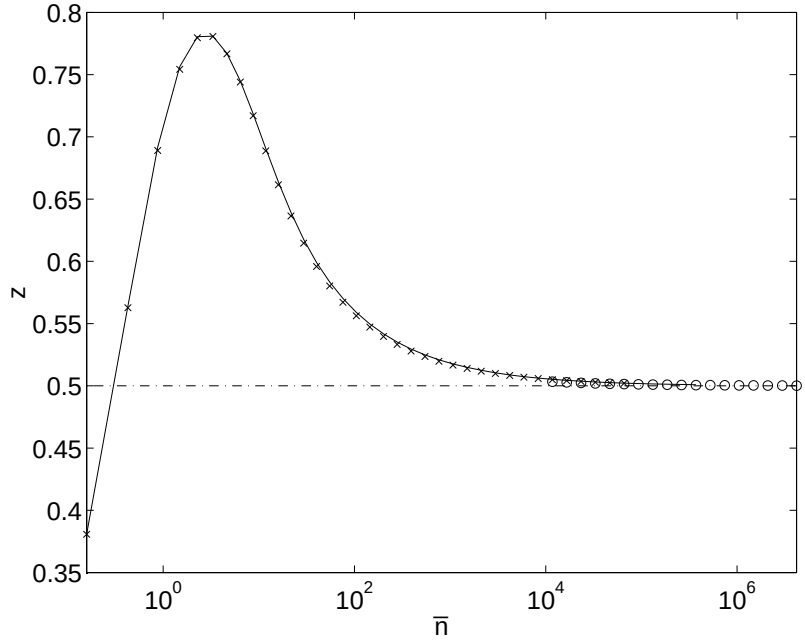

Figure 2.9: The value of z for heterodyne measurements on states with a fixed mean photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic value of 1/2 as the dash-dotted line. The results for squeezed states are shown as the continuous line.

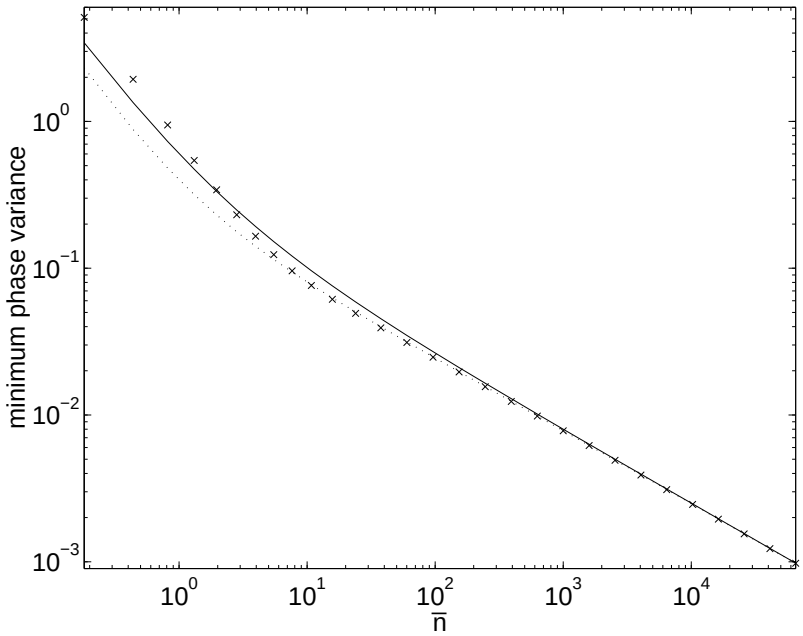

Figure 2.10: The minimum phase variance for mark I measurements on states with a fixed mean photon number. The exact calculations are shown as the crosses and the asymptotic analytic expression as the continuous line. The analytic expression taking account of the second term for $h_{\mathrm{I}}(n)$ is shown as the dotted line.

above about 100. The asymptotic analytic expression taking account of the second term for $h_{\mathrm{I}}(n)$ is also shown in this figure. As can be seen, the numerical results agree with this expression even more accurately, with good agreement for photon numbers above about 10.

For mark I measurements there is a term of order $\overline{n}^{-1}$, so we can not expect z to converge to $\sqrt{cp(p+1)}$. Using Eq. (2.105), we see that z should converge to

$$\sqrt{cp(p+1)} - 0.156248 \times \overline{n}^{1/4}. \tag{2.108}$$

The value of z is plotted in Fig. 2.11, and as can be seen it converges reasonably accurately to this value, but not to $\sqrt{cp(p+1)}$.

The results for mark II measurements are shown in Fig. 2.12. There is very good agreement with the asymptotic analytic expression, and between the results calculated using the exact method and continuous approximation. If we plot z (Fig. 2.13), we find that the second term does not agree well until very large photon numbers. In fact, a photon number greater than 7×10^8 is required to obtain better than 1% agreement. The agreement is again poor for the largest photon numbers; however, this is just due to poor convergence of the numerical technique.

2.3 Optimised Squeezed States

As an alternative to optimising completely general states with a fixed mean photon number, we can restrict our attention to squeezed states. As mentioned in the introduction to this chapter, there are three reasons for this:
1. Squeezed states are relatively easily generated in the laboratory, whereas there is no known way of producing general optimised states experimentally.
2. Squeezed states can be treated numerically far more easily than general optimised states.
3. It has been found numerically (see Sec. 2.3.4) that the phase uncertainties of optimised squeezed states are very close to those of optimised general states, and a partial theoretical explanation can be obtained by the following analysis.

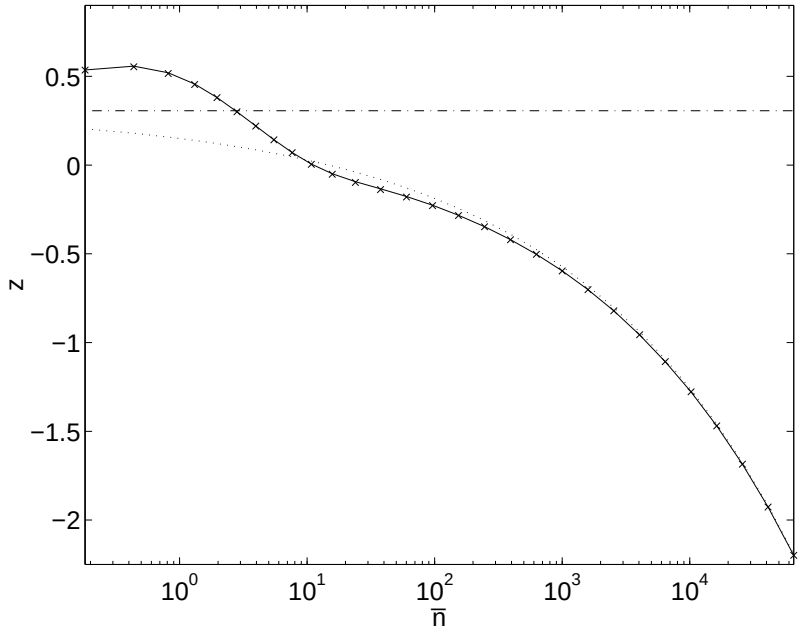

Figure 2.11: The value of z for mark I measurements on states with a fixed mean photon number. The exact calculations are shown as crosses. The asymptotic value ignoring the second term in Eq. (2.105) is shown as the dash-dotted line, and the asymptotic expression taking this term into account is plotted as the dotted line. The results for squeezed states are shown as the continuous line.

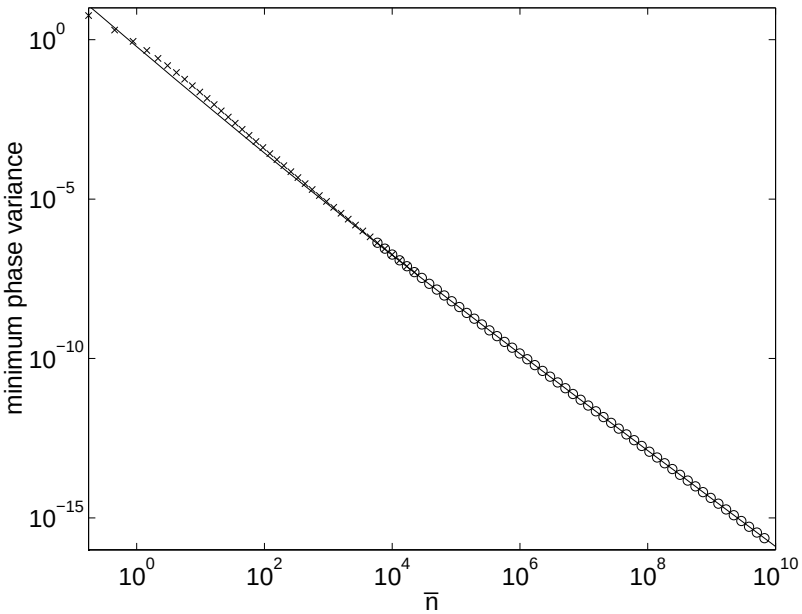

Figure 2.12: The minimum phase variance for mark II measurements on states with a fixed mean photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic analytic expression as the continuous line.

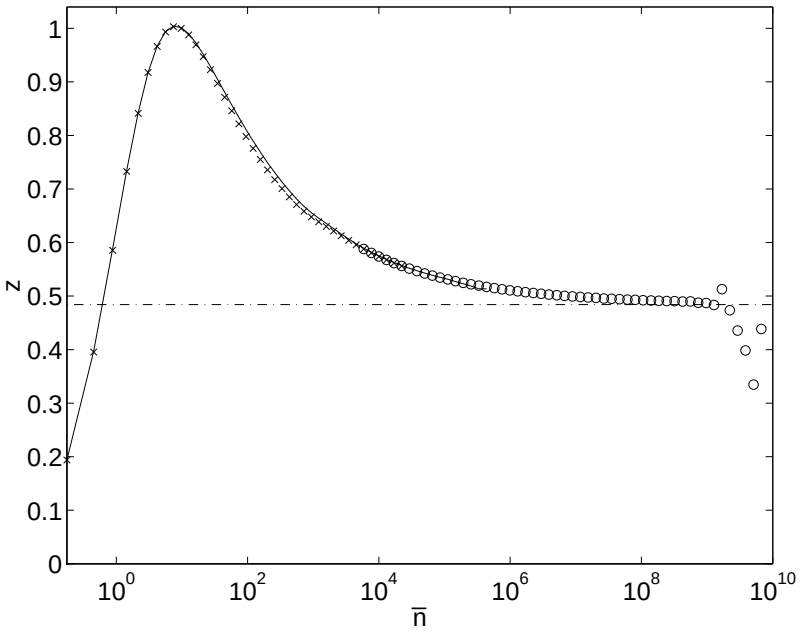

Figure 2.13: The value of z for mark II measurements on states with a fixed mean photon number. The exact calculations are shown as the crosses, the continuous approximation as the circles, and the asymptotic value of $\sqrt{15}/8$ as the dash-dotted line. The results for squeezed states are shown as the continuous line.

Squeezed states can be described by just two parameters, the coherent amplitude and the squeezing parameter. When the mean photon number is fixed there is only one independent parameter, and the optimisation problem is reduced to function minimisation in one dimension.

2.3.1 Canonical Measurements

Again the simplest case is that of canonical measurements. This case was solved by Collett [45], and I will outline the derivation here. The squeezed states considered are of the form

$$|\alpha, \zeta\rangle = \exp(\alpha a^\dagger - \alpha^* a) \exp[(\zeta^* a^2 - \zeta {a^\dagger}^2)/2] |0\rangle. \tag{2.109}$$

The mean photon number for this state is given by

$$\bar{n} = |\alpha|^2 + \sinh^2 |\zeta| \,. \tag{2.110}$$

Using this relation we can take a fixed value of $\bar{n}$, and vary ζ. The value of α can then be determined from $\bar{n}$ and ζ.

One complication is the phases of α and ζ. Only relative phase is important, so the phase of ζ can be considered relative to α. Specifically, if the phase of α is rotated by θ, an equivalent state is obtained by rotating the phase of ζ by 2θ. We can therefore take the phase of α to be zero without loss of generality.

All of the following analysis relies on ζ being real also. We can see that ζ should be real if we consider the phase quadrature diagram for the state. As in the introduction, the general quadrature $\hat{X}_\Phi$ is given by

$$\hat{X}_\Phi = a e^{-i\Phi} + a^\dagger e^{i\Phi}. \tag{2.111}$$

The 0 and $\pi/2$ quadratures are therefore

$$\hat{X}_0 = (a + a^\dagger), \tag{2.112}$$

$$\hat{X}_{\pi/2} = -i(a - a^\dagger). \tag{2.113}$$

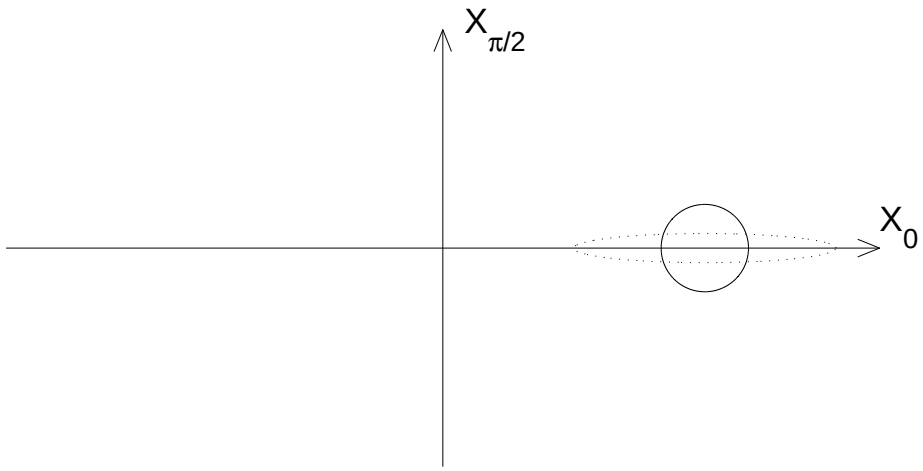

Figure 2.14: Contours of the probability distribution for X_0 and $X_{\pi/2}$. The contour for a coherent state is shown as the continuous line, and a contour for a squeezed state with $\phi_\zeta = \pi$ is shown as the dotted line.

For squeezed states the expectation values of these operators are

$$\left\langle \hat{X}_0 \right\rangle = 2\,|\alpha| \cos\varphi, \tag{2.114}$$

$$\left\langle \hat{X}_{\pi/2} \right\rangle = 2\,|\alpha| \sin\varphi, \tag{2.115}$$

where φ is the phase of α. In addition, the uncertainties of these quadrature operators are

$$\left\langle \Delta\hat{X}_0^2 \right\rangle = e^{2r} \sin^2 \tfrac{1}{2}\phi_\zeta + e^{-2r} \cos^2 \tfrac{1}{2}\phi_\zeta, \tag{2.116}$$

$$\left\langle \Delta\hat{X}_{\pi/2}^2 \right\rangle = e^{2r} \cos^2 \tfrac{1}{2}\phi_\zeta + e^{-2r} \sin^2 \tfrac{1}{2}\phi_\zeta, \tag{2.117}$$

where r and ϕ_ζ are the magnitude and phase of ζ.

A common way of representing coherent and squeezed states is by a contour of the probability distribution for the measured values X_0 and $X_{\pi/2}$, as in Fig. 2.14. The contour for a coherent state is a circle; however, the contour for a squeezed state is an ellipse, with the major axis at an angle $\frac{1}{2}(\phi_\zeta + \pi)$.

If we are considering small deviations from zero phase,

$$\left\langle \hat{X}_0 \right\rangle \approx 2\,|\alpha|\,, \tag{2.118}$$

$$\left\langle \hat{X}_{\pi/2} \right\rangle \approx 2\,|\alpha|\,\varphi, \tag{2.119}$$

so measuring X_0 gives information about the photon number, whereas measuring $X_{\pi/2}$ gives information about the phase. If the actual phase is zero, then ϕ_ζ should be equal to π in order for $X_{\pi/2}$, and therefore the phase, to have a reduced uncertainty. This means that ζ should be real and negative in order to have a phase squeezed state.

A complex value of ζ will mean that the major axis of the ellipse is at an angle from the horizontal, and this should give a larger phase uncertainty than if the ellipse is horizontal. This is not a rigorous derivation, and I have found numerically (see Fig. 2.18) that for very small $\overline{n}$ the optimum ζ is positive. Even for these exceptional values of $\overline{n}$, complex values of ζ were never found to give a better result. It is therefore reasonable to assume a real value of ζ in the following analysis.

In order to determine the phase uncertainty, we wish to determine

$$\left\langle \widehat{\cos\phi} \right\rangle = \tfrac{1}{2} \sum_{n=0}^{\infty} \left(\langle \alpha, \zeta | n \rangle \langle n+1 | \alpha, \zeta \rangle + \langle \alpha, \zeta | n+1 \rangle \langle n | \alpha, \zeta \rangle \right). \tag{2.120}$$

Collett determines this sum using the number state representation of squeezed states as given by [49]

$$\langle n|\alpha,\zeta\rangle = (n!\mu)^{-1/2}(\nu/2\mu)^{n/2}H_n[\beta(2\mu\nu)^{-1/2}]\exp[-|\beta|^2/2 + (\nu^*/2\mu)\beta^2], \tag{2.121}$$

where

$$\mu = \cosh r, \tag{2.122}$$
$$\nu = e^{i\phi_\zeta}\sinh r, \tag{2.123}$$
$$\beta = \alpha\mu + \alpha^*\nu. \tag{2.124}$$

The H_n are Hermite polynomials and are given by [46]

$$H_n(x) = n!\sum_{m=0}^{\lfloor n/2\rfloor}\frac{(-1)^m(2x)^{n-2m}}{m!(n-2m)!}. \tag{2.125}$$

Note that this formula is for the general case of complex α and ζ.

Collett evaluated this for real α and real, negative ζ, and found

$$\left\langle\widehat{\cos\phi}\right\rangle \approx \mathrm{erf}(\sqrt{2n_1})\left[1 - \frac{(\coth r - 1)^2}{32n_1\sqrt{2e^{-r}\sinh r}}\left(1+\frac{1}{2n_1}\right)\right], \tag{2.126}$$

where

$$n_1 = \tfrac{1}{2}\alpha^2(1-\tanh r). \tag{2.127}$$

This in turn gives the phase variance as

$$V(\phi) \approx \frac{n_0+1}{4\bar{n}^2} + 2\mathrm{erfc}(\sqrt{2n_0}), \tag{2.128}$$

where $n_0 = \bar{n}e^{-2r}$. Taking the derivative with respect to n_0 gives

$$\frac{\partial}{\partial n_0}V(\phi) \approx \frac{1}{4\bar{n}^2} - \sqrt{\frac{8}{\pi n_0}}e^{-2n_0}. \tag{2.129}$$

This is zero for

$$\frac{1}{4\bar{n}^2} = \sqrt{\frac{8}{\pi n_0}}e^{-2n_0}, \tag{2.130}$$

or

$$n_0 = \log(4\bar{n}) - \tfrac{1}{4}\log(2\pi n_0). \tag{2.131}$$

In Ref. [45] the factor of n_0 is omitted in the second term on the right hand side. We can make a slightly better approximation by using $n_0 \approx \log(4\bar{n})$ on the right hand side, so

$$n_0 \approx \log(4\bar{n}) - \tfrac{1}{4}\log\log(4\bar{n}) - \tfrac{1}{4}\log(2\pi). \tag{2.132}$$

We can also include an arbitrary number of additional terms with iterated logs; however, this expression gives a very accurate solution, as shown in Fig. 2.17.

Using the asymptotic expansion of the complementary error function

$$\mathrm{erfc}(x) \sim \frac{1}{x\sqrt{\pi}}e^{-x^2}\sum_{k=0}^{\infty}(-1)^k\frac{1\cdot 3\cdots(2k-1)}{(2x^2)^k}, \tag{2.133}$$

we have, for large n_0,

$$\begin{aligned} 2\mathrm{erfc}(\sqrt{2n_0}) &\approx \sqrt{\frac{2}{\pi n_0}}e^{-2n_0} \\ &= \frac{1}{8\bar{n}^2}. \end{aligned} \tag{2.134}$$

Using this result and substituting (2.132) into (2.128) gives the minimum phase variance as

$$V(\phi) \approx \frac{\log \bar{n} - \frac{1}{4}\log\log(4\bar{n}) + \Delta}{4\bar{n}^2}, \tag{2.135}$$

where

$$\Delta = \tfrac{3}{2} + 2\log 2 - \tfrac{1}{4}\log(2\pi) \approx 2.426825. \tag{2.136}$$

Therefore we see that the squeezed state optimised for minimum intrinsic phase variance has a phase variance scaling as $\log \bar{n}/\bar{n}^2$. This is not quite as good as the result for general optimised states, which scale as $\bar{n}^{-2}$. The factor of $\log \bar{n}$ increases very slowly, however, and it requires very large photon numbers to produce a significant difference between the variances of optimised squeezed and general states.

Note that the expression (2.135) for the variance differs from that given in [45]. The $\frac{1}{4}\log\log(4\bar{n})$ term does not appear in [45], as the approximation $n_0 \sim 1$ is made on the right hand side of the solution for n_0 (2.131). This term increases with photon number, and is therefore more significant than the Δ term.

2.3.2 Numerical Results for Canonical Measurements

This approximate analytic result has been tested numerically by determining the optimum squeezed states for a range of mean photon numbers up to about 8×10^7. Similarly to the case for optimum general states, the state coefficients were considered until the point where they had fallen to about 10^{-16}, where they did not significantly affect the results.

The phase variances of the minimum uncertainty squeezed states are plotted in Fig. 2.15. Also shown is the asymptotic expression (with the $\frac{1}{4}\log\log(4\bar{n})$ term omitted for simplicity). There is good agreement between the calculated values and the analytic expression, with better than 1% agreement for photon numbers above 5×10^4. Also shown is the analytic expression if we omit Δ, and there is much poorer agreement for the smaller photon numbers in that case. This means that although the Δ term is insignificant for large photon numbers, it gives more realistic results for moderate photon numbers.

To demonstrate how close the phase variance is to that for general states, the ratio of the phase variance for optimised squeezed states to that for general states is plotted in Fig. 2.16. For mean photon numbers above about 10^4 the phase uncertainty for general states was estimated using the asymptotic expression, as this is very accurate for large photon numbers. The phase variances are extremely close for photon numbers up to 10 or 100, and it takes a photon number of almost a million for the phase variance for squeezed states to be more than twice that for optimised general states.

In order to see if the term $\frac{1}{4}\log\log(4\bar{n})$ gives a better estimate of the phase variance, $4\bar{n}^2V(\phi) - \log \bar{n}$ is plotted in Fig. 2.17. According to the result (2.135), this should have the asymptotic value $-\frac{1}{4}\log\log(4\bar{n}) + \Delta$, rather than Δ. As can be seen, the exact values do not converge to Δ, but they do not appear to converge to $-\frac{1}{4}\log\log(4\bar{n}) + \Delta$ either. The exact values may converge to $-\frac{1}{4}\log\log(4\bar{n}) + \Delta$ for much larger photon numbers; however, for the range of photon numbers shown Δ gives a better approximation.

This result was checked by calculating the minimum phase variance exactly from (2.128), and these results are also shown in Fig. 2.17. There is close agreement with $-\frac{1}{4}\log\log(4\bar{n})+\Delta$, indicating that the asymptotic expression (2.135) is a good approximation for (2.128). The discrepancy must therefore be due to a slight inaccuracy in Eq. (2.128). In Ref. [45] Collett indicates that a correction term for (2.128) is

$$2(1-4a)\sqrt{\frac{2n_1}{\pi}}2^{-2n_1}, \tag{2.137}$$

where $a \approx 0.1990726$ and n_1 is as defined in Eq. (2.127). Although the results using this correction agree with the exact values for very small photon numbers, and reproduce the peak found in the exact results, the results again do not converge to the exact values for large photon numbers. In fact the asymptotic expression simply using Δ is still the most accuate for photon numbers above

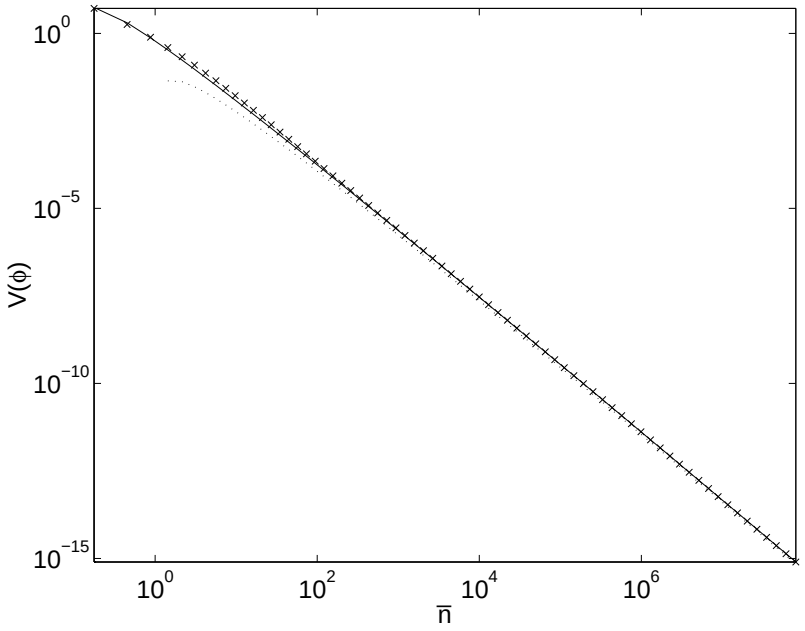

Figure 2.15: The phase variance for optimised squeezed states as a function of mean photon number. The exact phase variance is shown as the crosses, the asymptotic expression $(\log \bar{n} + \Delta)/(4\bar{n}^2)$ is shown as the continuous line, and $\log \bar{n}/(4\bar{n}^2)$ is shown as the dotted line.

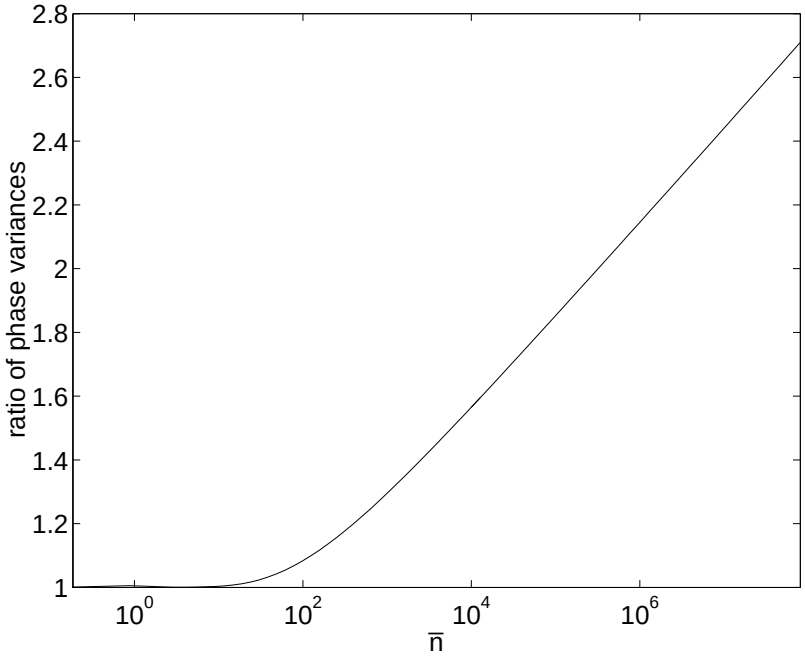

Figure 2.16: The ratio of the phase variance for optimised squeezed states to the phase variance for optimised general states.

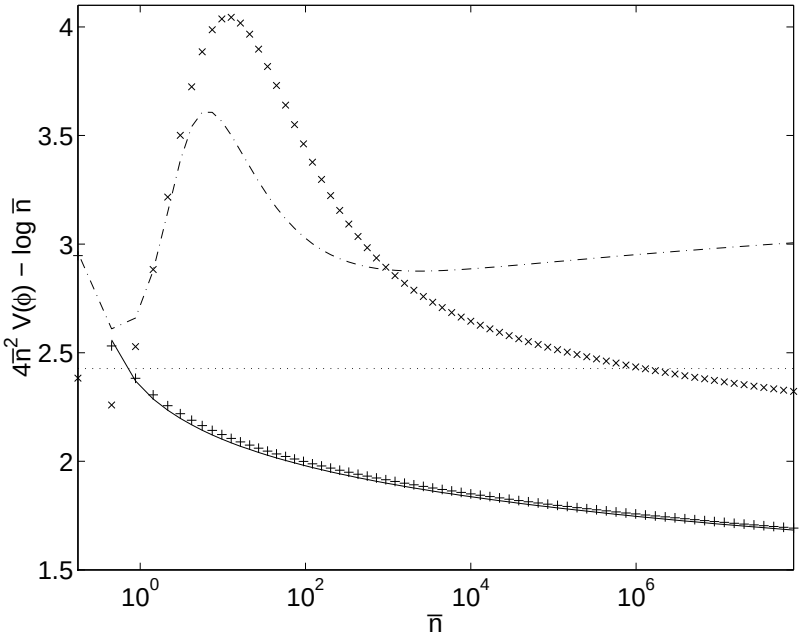

Figure 2.17: The value of $4\bar{n}^2V(\phi) - \log \bar{n}$ as calculated using various methods. The exact values are shown as crosses, Δ is shown as the horizontal dotted line and $-\frac{1}{4}\log\log(4\bar{n}) + \Delta$ is shown as the continuous line. The values determined exactly using Eq. (2.128) are shown as the pluses, and those calculated exactly using the additional correction are shown as the dash-dotted line.

about 10^4. As none of these corrections appear to give better results, I will omit them in further discussion.

Another test of the theory is the optimum value of ζ. From Eq. (2.131) we have

$$n_0 = \bar{n}e^{2\zeta} \approx \log(4\bar{n}) - \tfrac{1}{4}\log(2\pi), \tag{2.138}$$

where the factor of n_0 has been omitted from the second term. Solving for ζ gives

$$\zeta = \tfrac{1}{2}\log\left(\log(4\bar{n}) - \tfrac{1}{4}\log(2\pi)\right) - \tfrac{1}{2}\log\bar{n}. \tag{2.139}$$

The numerically determined optimum values of ζ and the approximate analytic expression (2.139) are plotted in Fig. 2.18. As can be seen, the numerically determined values are extremely close to the analytic result for photon numbers above about 100.

Note also that for very small mean photon numbers, below around 2.6, the optimum value of ζ is actually positive. This is a strange result, as the theory given above indicates that ζ should be negative for reduced phase uncertainty. The reason for this would seem to be that for cases where the photon number is very small, the circular error contour for a coherent state, as shown in Fig. 2.19, would overlap the $X_{\pi/2}$ axis. This means that there would be a significant contribution to the probability distribution for $\phi = \pm\pi$. A squeezed state with negative ζ would have an error contour that extends even further over the $X_{\pi/2}$ axis, resulting in a larger contribution to the probability distribution near $\phi = \pm\pi$.

If ζ is positive, however, the major axis of the error ellipse would be vertical, so the ellipse would avoid overlapping the $X_{\pi/2}$ axis. It is reasonable that this would result in a smaller phase variance. Note that there is no corresponding justification for using a complex value of ζ. Complex values were also tested, but always gave greater phase uncertainties, even for the very low photon number cases.

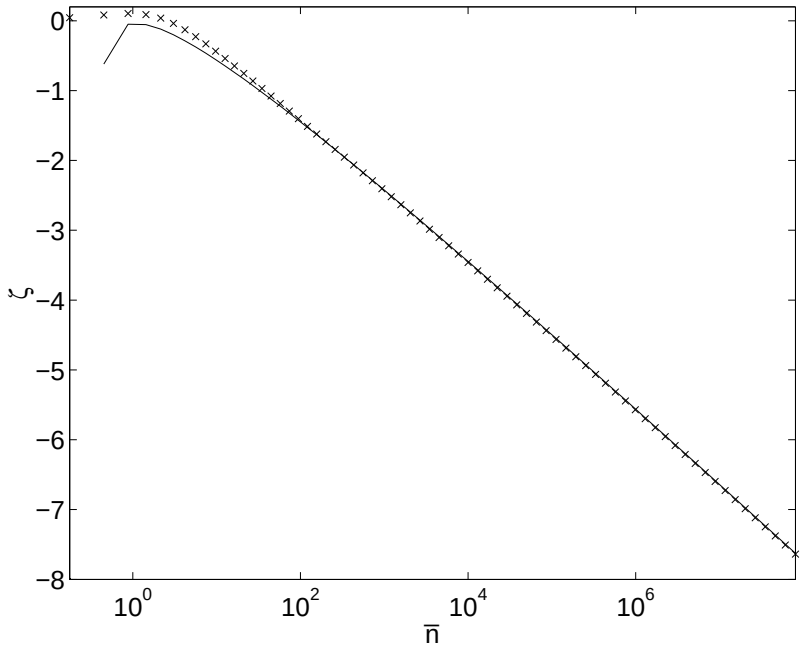

Figure 2.18: The optimum values of ζ for squeezed states as a function of $\bar{n}$. The numerically determined values are shown as crosses, and the values given by Eq. (2.139) are shown as the continuous line.

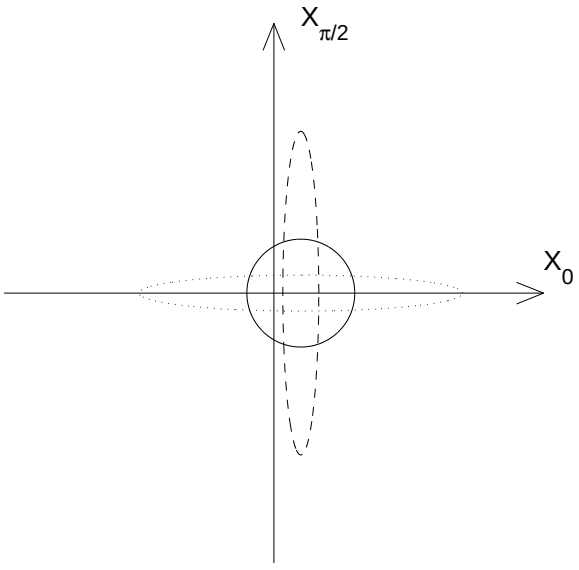

Figure 2.19: Contours of the probability distribution for X_0 and $X_{\pi/2}$. The contour for a coherent state is shown as the continuous line, a contour for a squeezed state with negative ζ is shown as the dotted line, and with positive ζ as the dashed line.

2.3.3 General Dyne Measurements

Now I will consider squeezed states optimised for minimum phase variance under more general dyne measurements. Again we wish to find $\left\langle \widehat{\cos\phi} \right\rangle$, except now we have

$$2\widehat{\cos\phi} = \sum_{n=0}^{\infty}(1-h(n))\left[|n\rangle\langle n+1| + |n+1\rangle\langle n|\right]. \tag{2.140}$$

The measure of the phase variance that I will initially consider is

$$V'(\phi) = 2 - 2\left\langle \widehat{\cos\phi} \right\rangle. \tag{2.141}$$

I will use this rather than the Holevo variance initially, in order to use the results for canonical measurements in a straightforward way. This expression is not a sufficiently accurate estimate of the Holevo variance for the case of mark I measurements. Therefore, at the end of this derivation, the result for this measure of the phase variance will be converted to the Holevo variance. This measure of the phase variance can be used for the intermediate steps, as minimising this measure of the phase variance is equivalent to minimising the Holevo phase variance.

Evaluating this measure of the phase variance gives

$$\begin{aligned} V'(\phi) &= 2 - \sum_{n=0}^{\infty}(1-h(n))\left[\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle + \langle\alpha,\zeta|n+1\rangle\langle n|\alpha,\zeta\rangle\right] \\ &= 2 - \sum_{n=0}^{\infty}\left[\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle + \langle\alpha,\zeta|n+1\rangle\langle n|\alpha,\zeta\rangle\right] \\ &\quad - \sum_{n=0}^{\infty}h(n)\left[\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle + \langle\alpha,\zeta|n+1\rangle\langle n|\alpha,\zeta\rangle\right]. \end{aligned} \tag{2.142}$$

Collett [45] found the first two terms to be approximately

$$\frac{n_0+1}{4\overline{n}^2} + 2\mathrm{erfc}(\sqrt{2n_0}), \tag{2.143}$$

so we therefore only require an expression for the third term. Both α and ζ will be taken to be real in this derivation, so the number state coefficients of the squeezed state are real, and the third term can be expressed as

$$2\sum_{n=0}^{\infty}h(n)\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle. \tag{2.144}$$

Approximating $h(n)$ by cn^{-p}, the sum to be evaluated is

$$\sum_{n=0}^{\infty}cn^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle. \tag{2.145}$$

This expression can be evaluated approximately using

$$\sum_{n=0}^{\infty}n^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle \approx \left\langle n^{-p}\right\rangle. \tag{2.146}$$

Expanding this in a series around $n=\overline{n}$ we find

$$\begin{aligned} \left\langle n^{-p}\right\rangle &= \left\langle(\overline{n}+\Delta n)^{-p}\right\rangle \\ &\approx \overline{n}^{-p}\left(1+\frac{p(p+1)}{2}\frac{\left\langle\Delta n^2\right\rangle}{\overline{n}^2}\right). \end{aligned} \tag{2.147}$$

It is easily shown that for squeezed states

$$\begin{aligned}\left\langle \Delta n^2 \right\rangle &= \alpha^2(\mu-\nu)^2 + 2\mu^2\nu^2, \\ &\approx \frac{\overline{n}^2}{n_0},\end{aligned} \tag{2.148}$$

so we find

$$\sum_{n=0}^{\infty} n^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle \approx \overline{n}^{-p}\left(1+\frac{p(p+1)}{2n_0}\right). \tag{2.149}$$

This should be correct to leading order; however, the second order term is not necessarily correct, as the approximation (2.146) may only be correct to first order. This result is derived in a more rigorous way in Appendix A.2, and it is shown that the second order term is also correct.

Using this result, the phase uncertainty is given by

$$V'(\phi) \approx \frac{n_0+1}{4\overline{n}^2} + 2\mathrm{erfc}(\sqrt{2n_0}) + 2c\overline{n}^{-p}\left[1+\frac{p(p+1)}{2n_0}\right]. \tag{2.150}$$

Taking the derivative with respect to n_0 gives

$$\frac{\partial}{\partial n_0}V'(\phi) \approx \frac{1}{4\overline{n}^2} - \sqrt{\frac{8}{\pi n_0}}e^{-2n_0} + 2c\overline{n}^{-p}\left[\frac{-p(p+1)}{2n_0^2}\right]. \tag{2.151}$$

As the second term falls exponentially with n_0 it can be omitted. Then we find that for minimum phase variance

$$n_0 \approx 2\sqrt{cp(p+1)}\overline{n}^{1-p/2}. \tag{2.152}$$

For $0 < p < 2$, n_0 increases with photon number, but does not increase as rapidly as $\overline{n}$, in agreement with the assumptions used in the appendix.

Substituting this result into Eq. (2.150) gives

$$V'(\phi) \approx 2c\overline{n}^{-p} + \sqrt{cp(p+1)}\overline{n}^{-p/2-1}. \tag{2.153}$$

Converting this to the Holevo phase variance gives an additional correction term:

$$V(\phi) \approx 2c\overline{n}^{-p} + \sqrt{cp(p+1)}\overline{n}^{-p/2-1} + 3c^2\overline{n}^{-2p}. \tag{2.154}$$

This correction term is only significant for mark I measurements, where it is of lower order than the second term. Thus we see that we obtain exactly the same terms for the phase uncertainty when considering squeezed states as we do when considering general states.

2.3.4 Numerical Results

These results have been tested by numerically determining the optimum squeezed states for heterodyne and mark I and II measurements. Because the results are extremely close to those for optimised general states, rather than plotting the phase variance, I have plotted the phase variance as a ratio to the phase variance for optimised general states.

The ratio of the minimum phase variance for squeezed states to that for general states using heterodyne measurements is plotted in Fig. 2.20. The phase variance for optimum squeezed states is never more than 0.3% above the phase variance for optimum general states, and for large photon numbers the phase variances converge. In fact, for some photon numbers the squeezed state variances are closer to the exact general state variances than those calculated using the continuous approximation.

The results for mark I measurements are shown in Fig. 2.21. Here the squeezed state phase variance is never more than about 0.6% above the phase variance for optimum general states, and the phase variances again converge for large photon numbers. The results for mark II measurements

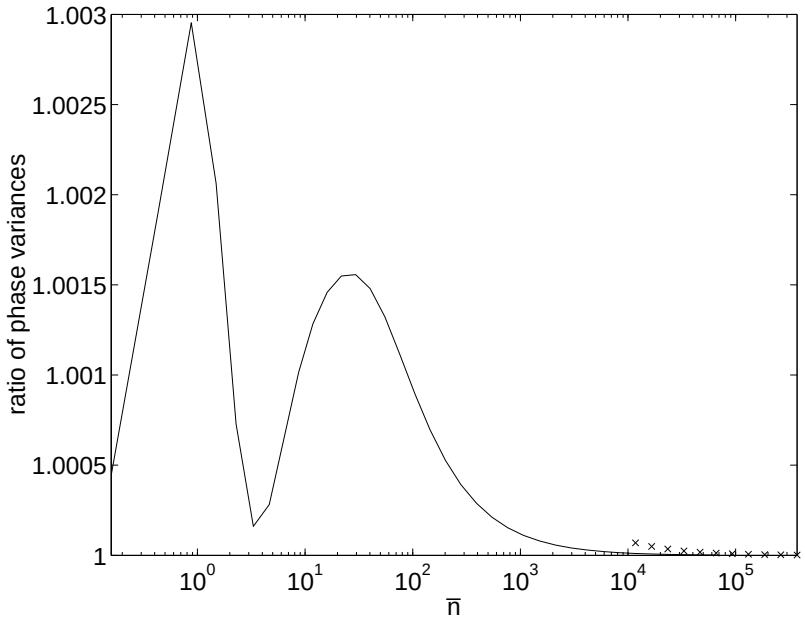

Figure 2.20: The ratio of the phase variance for squeezed states optimised for minimum phase variance under heterodyne measurements to the phase variance for optimum general states. The values using the exactly calculated general states are shown as the continuous line, and the values using the continuous approximation for the general states are shown as the crosses.

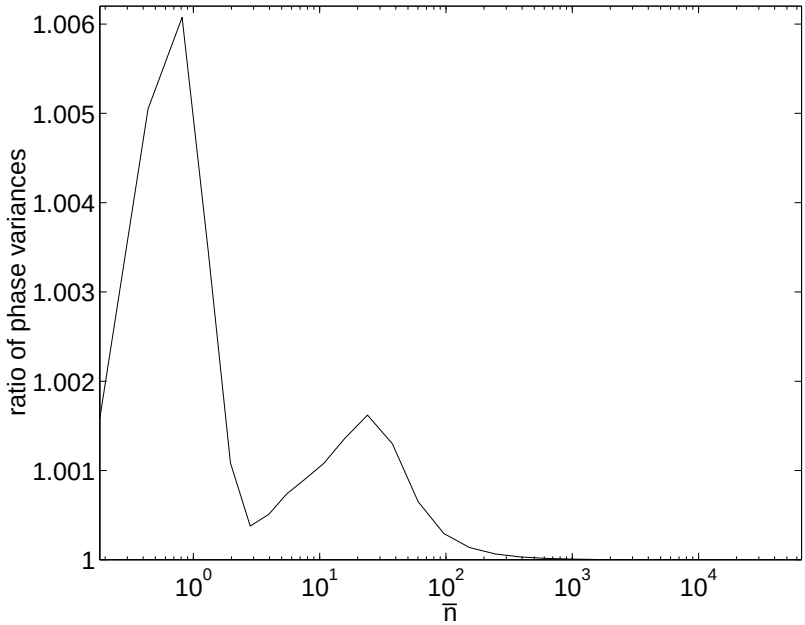

Figure 2.21: The ratio of the phase variance for squeezed states optimised for minimum phase variance under mark I measurements to the phase variance for optimum general states.

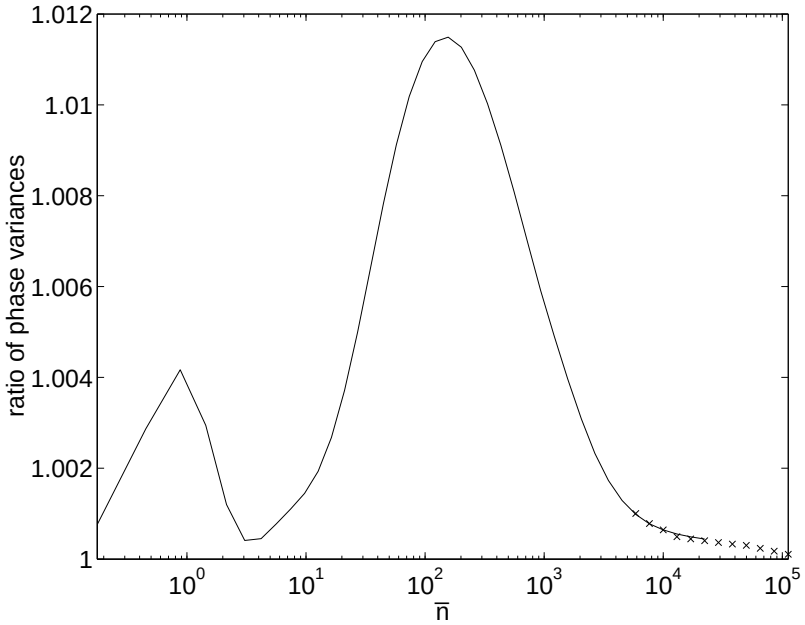

Figure 2.22: The ratio of the phase variance for squeezed states optimised for minimum phase variance under mark II measurements to the phase variance for optimum general states. The values using the exactly calculated general states are shown as the continuous line, and the values using the continuous approximation for the general states are shown as the crosses.

are shown in Fig. 2.22. Here the squeezed state phase variance can be more than 1% above the general state phase variance, but the phase variances still converge for large photon numbers.

Another way of comparing the results for squeezed states and general states is to plot the z parameter defined by (2.107). This is displayed for heterodyne, mark I and mark II measurements in Figs 2.9, 2.11 and 2.13. As can be seen, the results for optimum squeezed states and general states are almost indistinguishable. This means that the phase variances for squeezed and general states agree to better precision than just the first terms that I have derived here.

In order to compare the squeezed states with the states obtained by general optimisation in a more direct way, the number state coefficients for the two cases are plotted in Fig. 2.23(a). This is for the example of heterodyne measurements, and similar results are obtained for the adaptive measurement schemes. The two states are extremely close, indicating that the optimum general states may be converging to squeezed states.

On the other hand, if the state coefficients are plotted logarithmically, there are large differences between the states [see Fig. 2.23(b)]. Although the states are fairly close near the peak, the tails of the two states have different scalings. These differences persist even for very large mean photon numbers.

Another test of the theory is the values of ζ for the optimum squeezed states. Using Eq. (2.152) gives the optimum value of ζ as

$$\zeta = -\frac{p}{4}\log\overline{n} + \tfrac{1}{2}\log\left[2\sqrt{cp(p+1)}\right]. \tag{2.155}$$

Using this the optimal values of ζ for heterodyne, mark I and mark II measurements should be

$$\zeta_{\rm het} = -\frac{1}{4}\log\overline{n} \tag{2.156}$$

$$\zeta_{\rm I} = -\frac{1}{8}\log\overline{n} + \frac{1}{4}\log\frac{3}{8} \tag{2.157}$$

$$\zeta_{\rm II} = -\frac{3}{8}\log\overline{n} + \frac{1}{4}\log\frac{15}{16}. \tag{2.158}$$

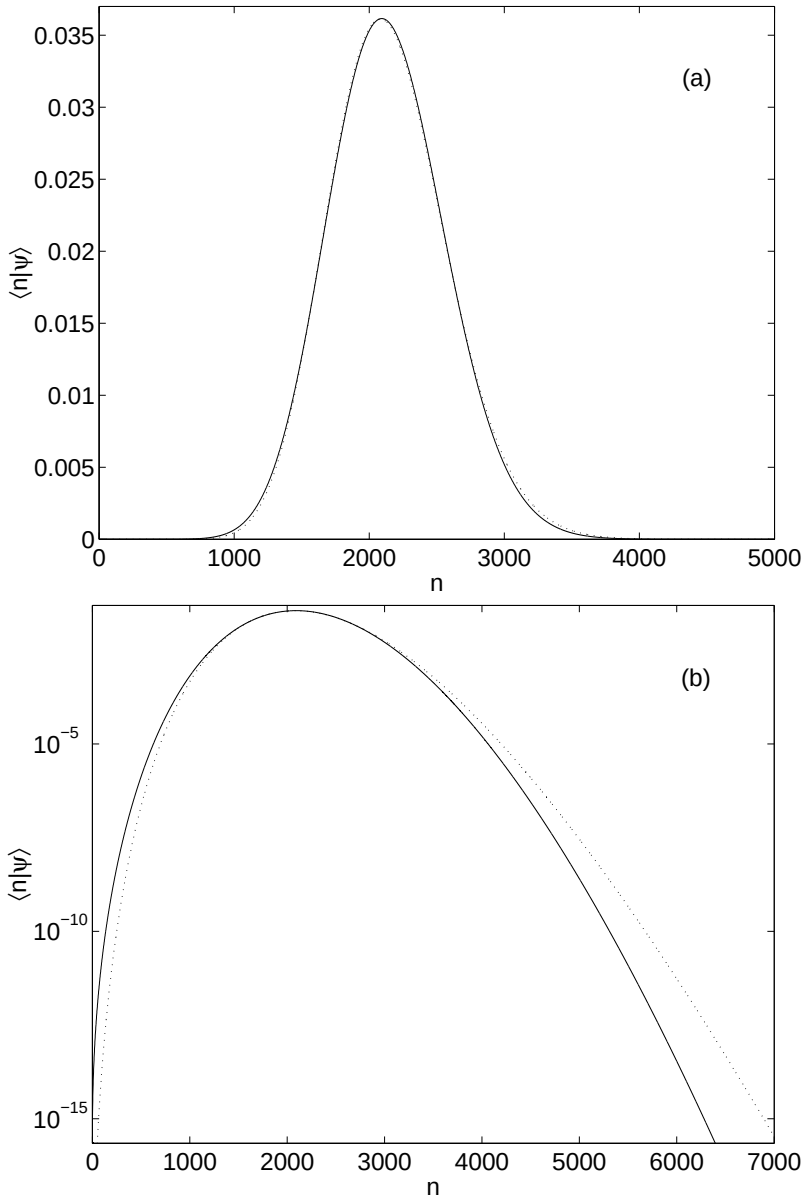

Figure 2.23: The number state coefficients for optimum states for heterodyne measurements and a mean photon number of about 2116. The continuous lines are for the optimised general state with fixed mean photon number, and the dotted lines are for the optimised squeezed state. Plot (a) is a linear plot, and (b) is a semi-log plot.

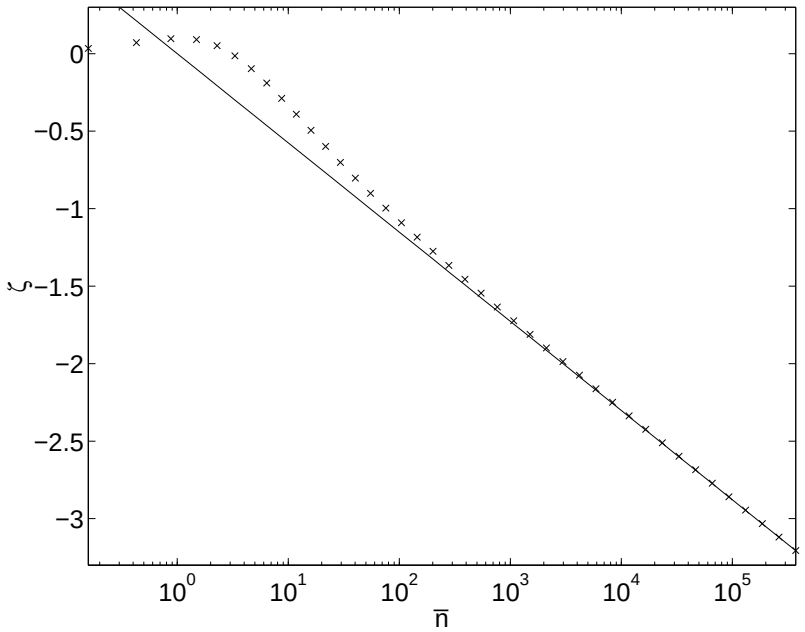

Figure 2.24: The optimum values of ζ for heterodyne phase measurements on squeezed states. The numerically determined values are shown as crosses, and the analytic expression is shown as the continuous line.

The numerically found optimum values of ζ and these asymptotic expressions are plotted in Figs 2.24, 2.25 and 2.26. The values of ζ converge to the asymptotic analytic expressions in all three cases. The value of ζ converges fairly rapidly for heterodyne and mark I measurements, with agreement within 0.05 for photon numbers above around 100. The convergence is much weaker in the mark II case, and a photon number around 10^5 is required for this level of agreement.

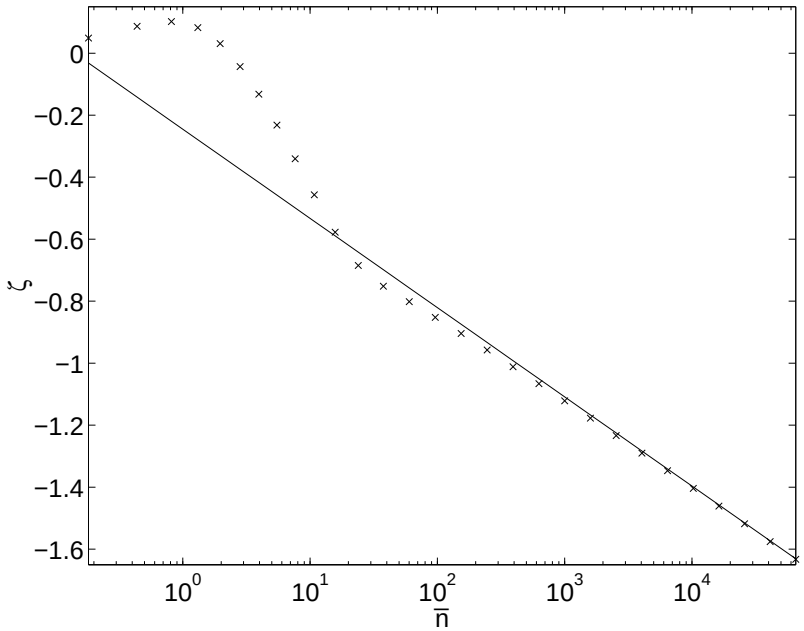

Figure 2.25: The optimum values of ζ for mark I phase measurements on squeezed states. The numerically determined values are shown as crosses, and the analytic expression is shown as the continuous line.

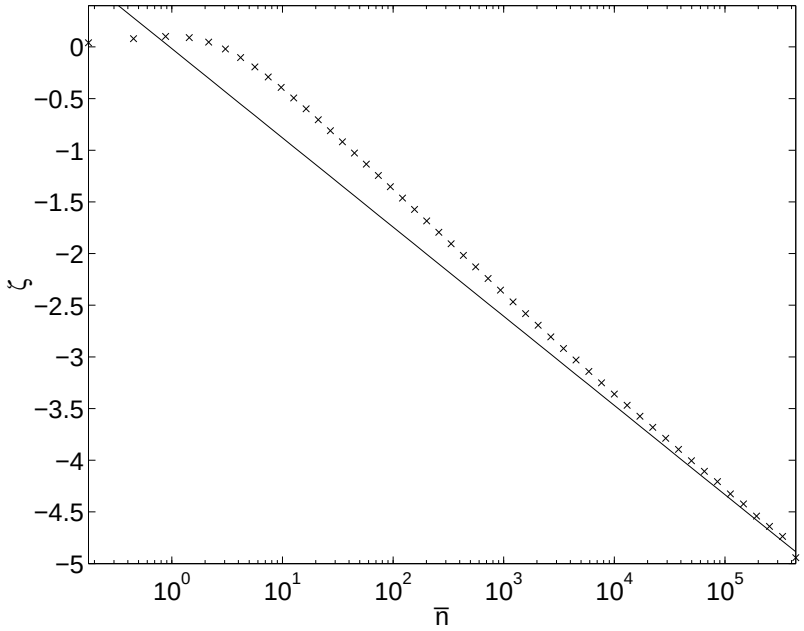

Figure 2.26: The optimum values of ζ for mark II phase measurements on squeezed states. The numerically determined values are shown as crosses, and the analytic expression is shown as the continuous line.

Chapter 3

Optimum Adaptive Dyne Measurements

Now I will consider the second area for improvement of dyne measurements: the measurement technique (as opposed to the input state). This includes both the local oscillator phase used and the final phase estimate. Recall that for a state with a mean photon number of $\overline{n}$, heterodyne measurements introduce a phase variance of about $\frac{1}{4}\overline{n}^{-1}$. This is much larger than the minimum intrinsic phase variance, which scales as $1.89 \times \overline{n}^{-2}$. The mark II adaptive measurements introduced in Ref. [35] give a great improvement on this, introducing a phase variance scaling as $\frac{1}{8}\overline{n}^{-1.5}$.

This scaling is still far worse than the scaling for minimum intrinsic phase variance. In [34] it is shown that the minimum introduced phase variance is the same as for squeezed states optimised for minimum intrinsic phase variance. As discussed in Sec. 2.3.1, this scaling is $\log \overline{n}/(4\overline{n}^2)$. Recall that this is not quite as good as that for general optimised states, but it requires very large photon numbers for the difference to be significant.

In this chapter I describe a feedback scheme that improves on the mark II adaptive scheme, and is in fact very close to this theoretical limit. It is therefore very close to being the best possible phase measurement scheme. It is also possible, under some circumstances, to surpass the theoretical limit. This is discussed in Sec. 3.8.

3.1 The Theoretical Limit

In order to understand how to attain the theoretical limit, we must first understand the reason for the theoretical limit. It can be shown [33] that the probability to obtain the results A, B from an arbitrary (adaptive or non-adaptive) measurement is

$$P(A,B)d^2A\,d^2B = \mathrm{Tr}[\rho F(A,B)]d^2A\,d^2B, \tag{3.1}$$

where ρ is the state of the mode being measured. Here $F(A,B)$ is the POM (probability operator measure) for the measurement, and is given by

$$F(A,B) = Q(A,B)|\tilde{\psi}(A,B)\rangle\langle\tilde{\psi}(A,B)|, \tag{3.2}$$

where $Q(A,B)$ is what the probability distribution $P(A,B)$ would be if ρ were the vacuum state $|0\rangle\langle 0|$, and $|\tilde{\psi}(A,B)\rangle$ is an unnormalised ket defined by

$$|\tilde{\psi}(A,B)\rangle = \exp(\tfrac{1}{2}B(a^\dagger)^2 - Aa^\dagger)|0\rangle. \tag{3.3}$$

This is proportional to a squeezed state [50]:

$$\exp(\tfrac{1}{2}B(a^\dagger)^2 - Aa^\dagger)|0\rangle = \left(1-|B|^2\right)^{-\frac{1}{4}} \exp\left(\frac{A\alpha^{\mathrm{P}*}}{2}\right)|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle, \tag{3.4}$$

where

$$|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle = \exp(\alpha^{\mathrm{P}} a^\dagger - \alpha^{\mathrm{P}*} a) \exp(\tfrac{1}{2}\zeta^{\mathrm{P}*} a^2 - \tfrac{1}{2}\zeta^{\mathrm{P}} (a^\dagger)^2)|0\rangle, \tag{3.5}$$

and the squeezing parameters are

$$\alpha^{\mathrm{P}} = \frac{A + BA^*}{1 - |B|^2}, \tag{3.6}$$

$$\zeta^{\mathrm{P}} = -\frac{B \mathrm{atanh}|B|}{|B|}. \tag{3.7}$$

Here I am using the superscript P on the squeezing parameters to indicate that they are the parameters of the squeezed state in the POM, and not of a physical state. The squeezing parameters for the input state will be given as α and ζ with no superscripts. Note that this differs from the notation in [51], where a superscript S indicates the squeezing parameters for the input state, and no superscript indicates the squeezing parameters for the squeezed state in the POM.

In terms of these the POM is given by

$$F(A, B) = Q'(A, B)|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle\langle\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}|, \tag{3.8}$$

where

$$Q'(A, B) = Q(A, B) \left(1 - |B|^2\right)^{-\frac{1}{2}} \exp\left[\mathrm{Re}\left(A\alpha^{\mathrm{P}*}\right)\right]. \tag{3.9}$$

If the system state is pure, $\rho = |\psi\rangle\langle\psi|$ and the probability distribution is given by

$$P(A, B) = Q'(A, B) \left|\langle\psi|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle\right|^2. \tag{3.10}$$

The greatest overlap between the states, and therefore the highest probability, will be when $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$ has the same phase as the input state. As all the information about the system from the measurement record is contained in the variables A and B, the most probable phase based on the measurement record is

$$\begin{aligned} \phi &= \arg(\alpha^{\mathrm{P}}) \\ &= \arg(A + BA^*). \end{aligned} \tag{3.11}$$

If additional information is known about the system, it is possible to obtain a better phase estimate. This is a far more complicated case, and will be considered in Sec. 3.8.

For an unbiased measurement scheme the probability distribution for this phase estimate resulting from Eq. (3.10) depends entirely on the inner product between the two states, and not on $Q'(A, B)$. To see this, note firstly that if the measurement is unbiased, the vacuum probability distribution $Q(A, B)$ will be independent of the phase. Secondly, recall that for the squeezed state $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$, if we rotate the phase of α^{P} by some angle θ, we can obtain an equivalent phase-shifted state by rotating the phase of ζ^{P} by 2θ. This means that $\zeta^{\mathrm{P}}(\alpha^{\mathrm{P}*})^2$ is independent of the phase. From Eq. (3.7), this means that $B(\alpha^{\mathrm{P}*})^2$ is independent of the phase. Since

$$\begin{aligned} A\alpha^{\mathrm{P}*} &= (\alpha^{\mathrm{P}} - B\alpha^{\mathrm{P}*})\alpha^{\mathrm{P}*} \\ &= \left|\alpha^{\mathrm{P}}\right|^2 - B(\alpha^{\mathrm{P}*})^2, \end{aligned} \tag{3.12}$$

$A\alpha^{\mathrm{P}*}$ and therefore $Q'(A, B)$ are independent of the phase.

Since the probability distribution for the phase estimate depends on the inner product between the two states, the variance in the measured phase will approximately be the sum of the intrinsic phase variance and the phase variance of the squeezed state $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$. The maximum overlap between the states will be when the squeezed state has about the same photon number as the input state. This means that the theoretical limit to the phase variance that is introduced by the measurement is the phase variance of the squeezed state that has the same photon number as the input state and has been optimised for minimum intrinsic phase variance. Since the phase variance of a squeezed

state optimised for minimum intrinsic phase variance is approximately $\log \bar{n}/(4\bar{n}^2)$ in the limit of large $\bar{n}$ [45], this is also the limit to the introduced phase variance.

The photon number of the squeezed state at maximum overlap will be mainly determined by the photon number of the input, but the degree and direction of squeezing (parametrised by ζ^{P}) will be determined by the multiplying factor $Q'(A, B)$. The multiplying factor can be expressed as a function of $\overline{n}^{\mathrm{P}}$ and ξ^{P}, for which the same symbol Q' will be used, even though it is a new function $Q'(\overline{n}^{\mathrm{P}}, \xi^{\mathrm{P}})$. Here $\overline{n}^{\mathrm{P}}$ is the mean photon number of the state $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$ (and will be close to the photon number of the input state), and $\xi^{\mathrm{P}} = \zeta^{\mathrm{P}}\alpha^{\mathrm{P}*}/\alpha^{\mathrm{P}}$ is ζ^{P} with the phase of α^{P} scaled out. In practice the multiplying factor tends to be concentrated along a particular line (for example, see Fig. 3.21), effectively giving ξ^{P} as a function of $\overline{n}^{\mathrm{P}}$. In order to obtain the theoretical limit, the measurement scheme must give a multiplying factor $Q'(\overline{n}^{\mathrm{P}}, \xi^{\mathrm{P}})$ that gives values of ξ^{P} for each $\overline{n}^{\mathrm{P}}$ that are the same as for optimised squeezed states.

We can determine the approximate variation of ξ^{P} with $\overline{n}^{\mathrm{P}}$ in the multiplying factor if we can estimate how it varies for measurements on a coherent state. If the intermediate phase estimates used in the adaptive scheme are unbiased, it is easy to see that the maximum probability will be for B real and therefore also A real. These results imply that

$$\begin{aligned} \alpha^{\mathrm{P}} &\approx \frac{A(1+B)}{1-B^2} \\ &= \frac{A}{1-B}. \end{aligned} \tag{3.13}$$

This means that B should be

$$B \approx 1 - \frac{A}{\alpha^{\mathrm{P}}}, \tag{3.14}$$

so ζ^{P} should be

$$\zeta^{\mathrm{P}} \approx -\mathrm{atanh}(1 - \frac{A}{\alpha^{\mathrm{P}}}). \tag{3.15}$$

Note that when α^{P} is real, ζ^{P} and ξ^{P} are equivalent. Provided the photon number is large, B should be close to 1 so we can use the asymptotic approximation of the atanh function, giving

$$\zeta^{\mathrm{P}} \approx \frac{1}{2}\log\frac{A}{2\alpha^{\mathrm{P}}}. \tag{3.16}$$

The mean photon number for squeezed states is given by

$$\overline{n}^{\mathrm{P}} = \left|\alpha^{\mathrm{P}}\right|^2 + \sinh^2\left|\zeta^{\mathrm{P}}\right|. \tag{3.17}$$

For the states that are considered in this study, $\sinh^2\left|\zeta^{\mathrm{P}}\right|$ is much smaller than $\overline{n}^{\mathrm{P}}$ for large photon numbers, so $\alpha^{\mathrm{P}} \approx \sqrt{\overline{n}^{\mathrm{P}}}$. Using this approximation gives ζ^{P} as

$$\zeta^{\mathrm{P}} \approx \frac{1}{2}\log\frac{A}{2\sqrt{\overline{n}^{\mathrm{P}}}}. \tag{3.18}$$

Since the magnitude of ζ^{P} is governed by the multiplying factor $Q'(\overline{n}^{\mathrm{P}}, \xi^{\mathrm{P}})$, this result for ζ^{P} should hold for more general input states.

From Sec. 2.3.1 and Ref. [45] the intrinsic phase variance of a squeezed state is given approximately by

$$V(\phi) \approx \frac{n_0+1}{4\overline{n}^2} + 2\mathrm{erfc}(\sqrt{2n_0}), \tag{3.19}$$

where $n_0 = \overline{n}e^{2\zeta}$ and ζ is real. This is minimised for

$$n_0 \approx \log(4\overline{n}) - \frac{1}{4}\log(2\pi). \tag{3.20}$$

Using the result obtained for ζ^{P} in Eq. (3.18) gives

$$e^{2\zeta^{\mathrm{P}}} \approx \frac{|A|}{2\sqrt{\overline{n}^{\mathrm{P}}}}, \tag{3.21}$$

so

$$n_0^{\mathrm{P}} \approx \frac{1}{2}|A|\sqrt{\overline{n}^{\mathrm{P}}}. \tag{3.22}$$

As the optimum value of n_0^{P} is given by Eq. (3.20), in order for the measurement to be optimal, $|A|$ should scale with $\overline{n}^{\mathrm{P}}$ as

$$|A| \propto \frac{\log \overline{n}^{\mathrm{P}}}{\sqrt{\overline{n}^{\mathrm{P}}}}. \tag{3.23}$$

Since $\overline{n}^{\mathrm{P}} \approx \overline{n}$, we should get the same scaling with the input photon number. For the case of mark II measurements there is the result that $|A| = 1$ [35], which is why these measurements are not optimal. Note that if we substitute $|A| = 1$ into the expression (3.22) to find n_0^{P}, and substitute that into Eq. (3.19), we obtain the correct result for the mark II introduced phase variance,

$$\Delta V(\phi) \approx \tfrac{1}{8}\overline{n}^{-1.5}. \tag{3.24}$$

3.2 Improved Feedback

Now we have the result that for optimal feedback $|A|$ should decrease with photon number. Therefore, in order to improve the phase measurement scheme, we want one that gives $|A| < 1$. To see in general how this can be achieved, consider a coherent state with amplitude α. The Ito SDE for $|A_v|^2$ will be

$$\begin{aligned} d|A_v|^2 &= A_v^*(dA_v) + (dA_v^*)A_v + (dA_v^*)(dA_v) \\ &= A_v^* e^{i\Phi(v)} I(v)dv + e^{-i\Phi(v)} I(v)dv A_v + dv \\ &= \left[|A_v| I(v)\left(e^{i\Phi(v)}e^{-i\varphi_v^A} + e^{-i\Phi(v)}e^{i\varphi_v^A}\right) + 1\right] dv, \end{aligned} \tag{3.25}$$

where $\varphi_v^A = \arg A_v$. Usually the local oscillator phase will be based on a phase estimate $\hat{\varphi}_v$, so that $\Phi(v) = \hat{\varphi}_v + \pi/2$. In terms of this phase estimate the differential equation becomes

$$\begin{aligned} d|A_v|^2 &= \left[|A_v| I(v)\left(ie^{i\hat{\varphi}_v}e^{-i\varphi_v^A} - ie^{-i\hat{\varphi}_v}e^{i\varphi_v^A}\right) + 1\right] dv \\ &= \left[1 + 2|A_v| I(v)\sin(\varphi_v^A - \hat{\varphi}_v)\right] dv. \end{aligned} \tag{3.26}$$

Taking the expectation value of $I(v)$ and simplifying gives

$$\begin{aligned} \langle I(v)\rangle &= \left\langle 2\mathrm{Re}\left(\alpha e^{-i\Phi(v)}\right) dv + dW(v)\right\rangle \Big/ dv \\ &= -2|\alpha|\sin(\hat{\varphi}_v - \varphi), \end{aligned} \tag{3.27}$$

where $\varphi = \arg\alpha$. Using this result, the expectation value for the increment in $|A_v|^2$ is

$$\left\langle d|A_v|^2\right\rangle = \left[1 - 4|A_v||\alpha|\sin(\hat{\varphi}_v - \varphi)\sin(\varphi_v^A - \hat{\varphi}_v)\right] dv. \tag{3.28}$$

Note that here the average is only over the single increment $dW(v)$, and not over the different trajectories.

The first term on its own will give $|A| = 1$, and in order to get $|A| < 1$ the two sines must have the same sign. This will be the case if the phase estimate is between the actual phase and the phase of A_v. Therefore we would like to use the phase estimate

$$\hat{\varphi}(v) = \arg(\alpha^{1-\varepsilon(v)} A_v^{\varepsilon(v)}), \tag{3.29}$$

where $0 < \varepsilon < 1$. The problem with this phase estimate is that it uses the actual value of the phase. In order to avoid this problem, the best estimate of the phase can be used in place of the actual phase. Therefore the phase estimate that will be considered is

$$\hat{\varphi}(v) = \arg(C_v^{1-\varepsilon(v)} A_v^{\varepsilon(v)}), \tag{3.30}$$

where

$$C_v = A_v v + B_v A_v^*. \tag{3.31}$$

It would at first appear that the best value of ε to use is 1/2, as this will make the phase estimate exactly halfway between the best phase estimate and $\arg A_v$, and this was the initial improvement on mark II feedback that was tried. This is too simplistic, however, as will be shown. In order to estimate what the best values of ε to use are, we can take the actual phase to be zero, and use the phase estimate

$$\hat{\varphi}(v) = \varepsilon \arg A_v. \tag{3.32}$$

For simplicity ε will be taken to be constant. For this value of the phase estimate, $\varphi_v^A = \hat{\varphi}_v/\varepsilon$. Substituting this into Eq. (3.28) gives

$$\left\langle d|A_v|^2 \right\rangle = [1 - 4|A_v|\alpha \sin(\hat{\varphi}_v) \sin[\hat{\varphi}_v(1/\varepsilon - 1)]]\, dv. \tag{3.33}$$

The absolute value symbols on α have been omitted because the phase has been taken to be zero. When the phase estimate is close to zero we can use the linear approximation of the sine function to get

$$\left\langle d|A_v|^2 \right\rangle = \left[1 - 4|A_v|\alpha\hat{\varphi}_v^2\,(1/\varepsilon - 1)\right] dv. \tag{3.34}$$

A differential equation for $\hat{\varphi}_v$ can be obtained in a similar way:

$$\begin{aligned} d\hat{\varphi}_v &= \varepsilon \mathrm{Im}\left[d \ln A_v\right] \\ &= \varepsilon \mathrm{Im}\left[\frac{dA_v}{A_v} - \frac{(dA_v)^2}{2A_v^2}\right] \\ &= \varepsilon \mathrm{Im}\left[\frac{ie^{i\hat{\varphi}_v} I(v) dv}{|A_v| e^{i\hat{\varphi}_v/\varepsilon}} - \frac{\left(ie^{i\hat{\varphi}_v} I(v) dv\right)^2}{2\left(|A_v| e^{i\hat{\varphi}_v/\varepsilon}\right)^2}\right]. \end{aligned} \tag{3.35}$$

Note that here $A_v = |A_v| e^{i\hat{\varphi}_v/\varepsilon}$ has been used. Continuing the derivation we find

$$\begin{aligned} d\hat{\varphi}_v &= \varepsilon \mathrm{Im}\left[\frac{ie^{i\hat{\varphi}_v(1-1/\varepsilon)} I\,(v)\, dv}{|A_v|} + \frac{e^{2i\hat{\varphi}_v(1-1/\varepsilon)} \left(I\,(v)\, dv\right)^2}{2\,|A_v|^2}\right] \\ &= \frac{\varepsilon}{|A_v|}\left[\cos\left[\hat{\varphi}_v\,(1-1/\varepsilon)\right] I\,(v)\, dv + \frac{\sin\left[2\hat{\varphi}_v\,(1-1/\varepsilon)\right] dv}{2\,|A_v|}\right] \\ &= \frac{\varepsilon}{|A_v|}\left[\cos\left[\hat{\varphi}_v\,(1-1/\varepsilon)\right]\left[-2\alpha \sin\hat{\varphi}_v dv + dW\,(v)\right] + \frac{\sin\left[2\hat{\varphi}_v\,(1-1/\varepsilon)\right] dv}{2\,|A_v|}\right] \\ &= \frac{\varepsilon}{|A_v|}\left[-2\alpha \sin\hat{\varphi}_v \cos\left[\hat{\varphi}_v\,(1-1/\varepsilon)\right] dv + \frac{\sin\left[2\hat{\varphi}_v\,(1-1/\varepsilon)\right] dv}{2\,|A_v|} + \cos\left[\hat{\varphi}_v\,(1-1/\varepsilon)\right] dW(v)\right] \\ &= \frac{\varepsilon}{|A_v|}\left[-2\alpha\hat{\varphi}_v dv + \frac{\hat{\varphi}_v\,(1-1/\varepsilon)\, dv}{|A_v|} + dW(v)\right]. \end{aligned} \tag{3.36}$$

In the last line the linear approximation for sine and a constant approximation for cos have been used. Evaluating the increment in the square of the phase gives

$$\begin{aligned} \left\langle d\hat{\varphi}_v^2 \right\rangle &= \left\langle 2\hat{\varphi}_v d\hat{\varphi}_v + (d\hat{\varphi}_v)^2 \right\rangle \\ &= \frac{2\varepsilon}{|A_v|}\left[-2\alpha\hat{\varphi}_v^2 dv + \frac{\hat{\varphi}_v^2\,(1-1/\varepsilon)\, dv}{|A_v|}\right] + \frac{\varepsilon^2}{|A_v|^2} dv \\ &= \frac{\varepsilon}{|A_v|}\left[-4\alpha\hat{\varphi}_v^2 + \frac{2\hat{\varphi}_v^2\,(1-1/\varepsilon)}{|A_v|} + \frac{\varepsilon}{|A_v|}\right] dv. \end{aligned} \tag{3.37}$$

Thus the expectation values for the increments in $|A_v|^2$ and $\hat{\varphi}_v^2$ are:

$$\left\langle d|A_v|^2 \right\rangle = \left[1 - 4\alpha|A_v|\hat{\varphi}_v^2\,(1/\varepsilon - 1)\right] dv, \tag{3.38}$$

$$\left\langle d\hat{\varphi}_v^2 \right\rangle = \frac{\varepsilon}{|A_v|}\left[-4\alpha\hat{\varphi}_v^2 + \frac{2\hat{\varphi}_v^2\,(1-1/\varepsilon)}{|A_v|} + \frac{\varepsilon}{|A_v|}\right] dv. \tag{3.39}$$

Now an approximate solution for $|A_v|^2$ and $\hat{\varphi}_v^2$ can be obtained by solving these as differential equations, ignoring the fact that these are only the expectation values for the increments. The approximate solution for large photon number is

$$|A_v|^2 = \varepsilon v + O\left(\alpha^{-1}\right), \tag{3.40}$$

$$\hat{\varphi}_v^2 = \frac{1}{4\alpha}\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-2}\right). \tag{3.41}$$

To see that these are solutions, we can just substitute them into the above equations. For Eq. (3.38) we have

$$\begin{aligned}
\text{l.h.s.} &= d|A_v|^2 \\
&= \left[\varepsilon + O\left(\alpha^{-1}\right)\right] dv \qquad (3.42)\\
\text{r.h.s.} &= \left[1 - 4\alpha\left(\sqrt{\varepsilon v} + O\left(\alpha^{-1}\right)\right)\left(\frac{1}{4\alpha}\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-2}\right)\right)(1/\varepsilon - 1)\right] dv \\
&= \left[1 - \left(\sqrt{\varepsilon v} + O\left(\alpha^{-1}\right)\right)\left(\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-1}\right)\right)(1/\varepsilon - 1)\right] dv \\
&= \left[1 - \left(\varepsilon + O\left(\alpha^{-1}\right)\right)(1/\varepsilon - 1)\right] dv \\
&= \left[1 - \left(1 - \varepsilon + O\left(\alpha^{-1}\right)\right)\right] dv \\
&= \left[\varepsilon + O\left(\alpha^{-1}\right)\right] dv. \qquad (3.43)
\end{aligned}$$

For Eq. (3.39) we find

$$\begin{aligned}
\text{l.h.s.} &= O\left(\alpha^{-1}\right) dv \qquad (3.44)\\
\text{r.h.s.} &= \frac{\varepsilon}{\left(\sqrt{\varepsilon v} + O\left(\alpha^{-1}\right)\right)}\left[-4\alpha\left(\frac{1}{4\alpha}\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-2}\right)\right)\right. \\
&\qquad \left. + \frac{2\left(\frac{1}{4\alpha}\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-2}\right)\right)(1 - 1/\varepsilon)}{\left(\sqrt{\varepsilon v} + O\left(\alpha^{-1}\right)\right)} + \frac{\varepsilon}{\left(\sqrt{\varepsilon v} + O\left(\alpha^{-1}\right)\right)}\right] dv \\
&= \left(\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-1}\right)\right)\left[-\left(\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-1}\right)\right) + O\left(\alpha^{-1}\right) + \left(\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-1}\right)\right)\right] dv \\
&= \left(\sqrt{\frac{\varepsilon}{v}} + O\left(\alpha^{-1}\right)\right) O\left(\alpha^{-1}\right) dv \\
&= O\left(\alpha^{-1}\right) dv. \qquad (3.45)
\end{aligned}$$

For this equation it is enough that the sides agree to order α^{-1}. To obtain more specific agreement the higher order terms in the solution would be required.

As a simple check on these results, note that the case $\varepsilon = 1$ is just the standard case and the final phase estimate corresponds to the mark I phase estimate. In this case the solutions give to first order

$$|A_1|^2 = 1, \tag{3.46}$$

$$\hat{\varphi}_1^2 = \frac{1}{4\alpha}. \tag{3.47}$$

These are exactly the same results as obtained in Ref. [35].

The important result is that for the final value of A_v, at $v = 1$, we have $|A| \approx \sqrt{\varepsilon}$. This indicates that we can obtain smaller and smaller values of $|A|$ simply by using smaller values of ε, and the minimum is *not* for $\varepsilon = 1/2$. To see the reason for this, note that

$$(\varphi_v^A)^2 \approx \frac{1}{4\alpha\sqrt{\varepsilon^3 v}}. \tag{3.48}$$

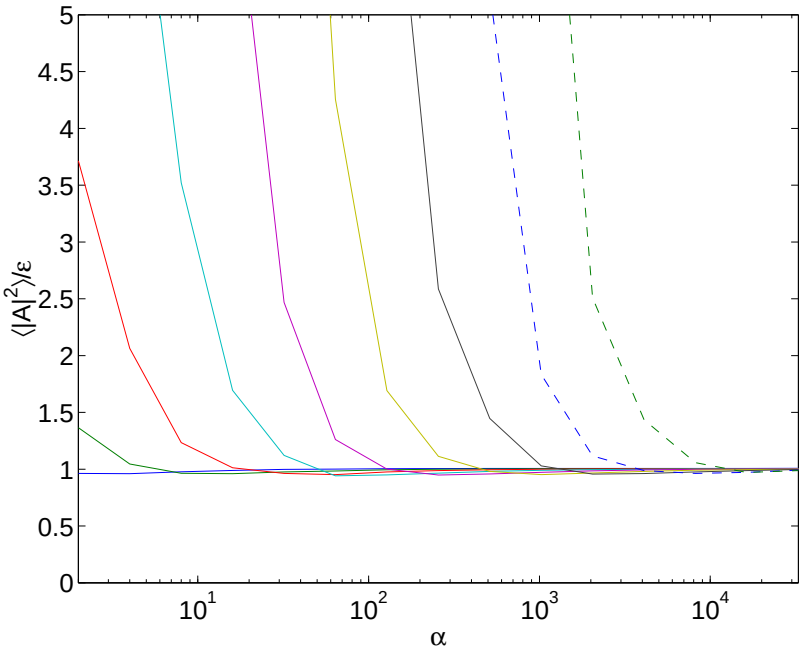

Figure 3.1: The mean values of $|A|^2$ divided by ε for measurements on coherent states with $\varepsilon \arg A_v$ feedback. The results for $\varepsilon = 1/2$ are shown in dark blue, for $\varepsilon = 1/4$ in green, for $\varepsilon = 1/8$ in red, for $\varepsilon = 1/16$ in light blue, for $\varepsilon = 1/32$ in purple, for $\varepsilon = 1/64$ in yellow, for $\varepsilon = 1/128$ in black, for $\varepsilon = 1/256$ as a dashed dark blue line, and for $\varepsilon = 1/512$ as a dashed green line.

This means that for smaller values of ε, the value of φ_v^A will tend to be greater. This means that the second sine in (3.28) will be greater on average, resulting in a smaller value of $|A|$.

For other values of ε these results have been verified by numerically performing the stochastic integrals for coherent states. The numerical technique is very straightforward for coherent states. We simply replace the infinitesimal interval dv with a finite interval δv, and replace the Wiener increment dW with a finite stochastic increment δW that has a Gaussian distribution with zero mean and variance δv. For these calculations 100α time steps and 2^{10} samples were used.

The results for $|A|^2$ are shown in Fig. 3.1. For simplicity, rather than plotting $|A|^2$, I have shown $|A|^2/\varepsilon$. From the theory above this should converge to 1. The results for moderate values of ε agree very well with the theory; however, for smaller values of ε there is poor agreement unless the photon number is very large. For $\varepsilon = 1/64$, a value of α above about 512 is required for good agreement. The photon number required for good agreement with theory appears to scale roughly as ε^{-3}.

In Fig. 3.2 are shown the results for the final variance of the phase estimate. For simplicity, $\hat{\varphi}^2/\sqrt{\varepsilon}$ is plotted. From the theory above this should be approximately $1/(4\alpha)$. Similarly to the case for $|A|^2$, the results for $\varepsilon = 1/2$ agree very well, but those for smaller values of ε require large photon numbers in order to have good agreement.

The next question is whether this good agreement continues if the best phase estimate is used in the feedback, rather than the actual phase, as in Eq. (3.30). The calculations were repeated for this feedback, and the results for $|A|^2$ are shown in Fig. 3.3. The results are very similar to the previous case, except that higher photon numbers than before are required to obtain good agreement in this case. In this case the photon number required for good agreement with theory appears to scale around ε^{-2}.

These results indicate that $|A|$ can be reduced to any value required, provided the photon number is sufficiently large. This implies that the results should be close to optimum for the feedback given by Eq. (3.30), with ε given by

$$\varepsilon = \left(2\alpha_{\rm opt} e^{2\zeta_{\rm opt}}\right)^2, \tag{3.49}$$

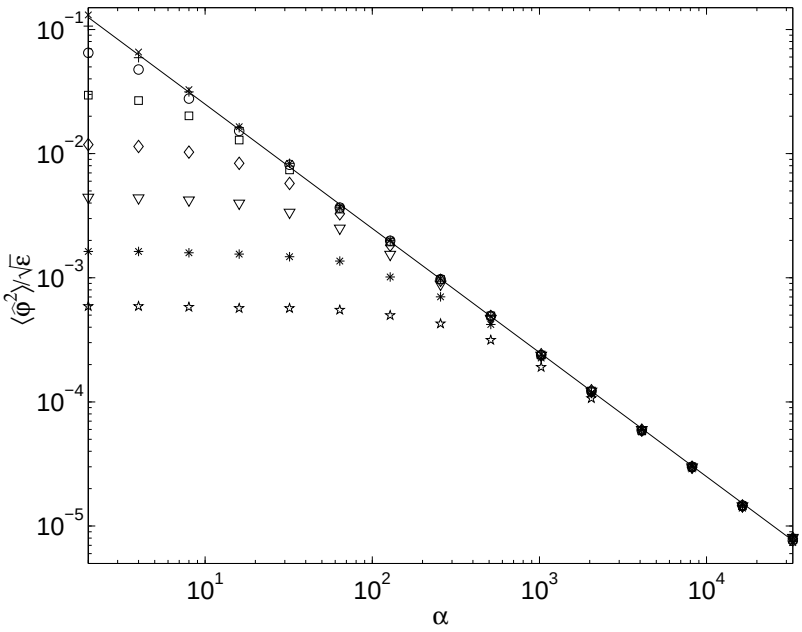

Figure 3.2: The mean values of $\hat{\varphi}^2$ divided by $\sqrt{\varepsilon}$ for measurements on coherent states with $\varepsilon \arg A_v$ feedback. The results for $\varepsilon = 1/2$ are shown as crosses, for $\varepsilon = 1/4$ as pluses, for $\varepsilon = 1/8$ as circles, for $\varepsilon = 1/16$ as squares, for $\varepsilon = 1/32$ as diamonds, for $\varepsilon = 1/64$ as triangles, for $\varepsilon = 1/128$ as asterisks, and for $\varepsilon = 1/256$ as stars. The analytic result of $1/(4\alpha)$ is shown as the continuous line.

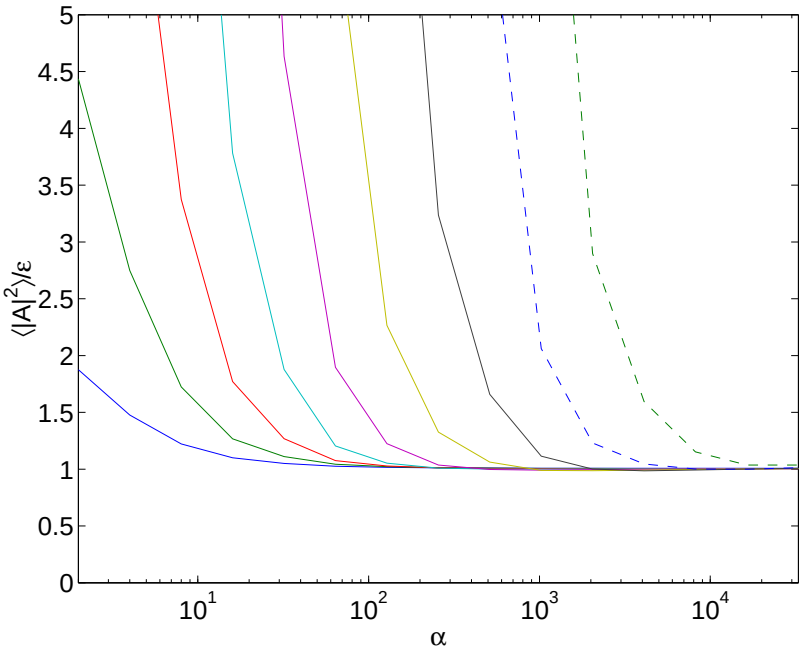

Figure 3.3: The mean values of $|A|^2$ divided by ε for measurements on coherent states with $\arg C_v^{1-\varepsilon} A_v^{\varepsilon}$ feedback. The results for $\varepsilon = 1/2$ are shown in dark blue, for $\varepsilon = 1/4$ in green, for $\varepsilon = 1/8$ in red, for $\varepsilon = 1/16$ in light blue, for $\varepsilon = 1/32$ in purple, for $\varepsilon = 1/64$ in yellow, for $\varepsilon = 1/128$ in black, for $\varepsilon = 1/256$ as a dashed dark blue line, and for $\varepsilon = 1/512$ as a dashed green line.

where $\zeta_{\rm opt}$ is the value of ζ optimised for minimum intrinsic phase variance for the photon number $\overline{n}$. Here α has been used rather than $\sqrt{\overline{n}}$, as it should be closer to $\alpha^{\rm P}$. This is the first feedback technique that I will discuss.

3.3 Simulation Method

The easiest input states to use for numerical simulations are coherent states, as they remain coherent with a deterministically decaying amplitude. We simply have the equations (after the time variation of the amplitude is scaled out)

$$\begin{aligned} I(v)dv &= 2{\rm Re}(\alpha e^{-i\Phi(v)})dv + dW(v) \\ dA_v &= e^{i\Phi(v)}I(v)dv \\ dB_v &= -e^{2i\Phi(v)}dv, \end{aligned} \tag{3.50}$$

where $\Phi(v) = \hat{\varphi}_v + \pi/2$, and $\hat{\varphi}_v$ is obtained by the feedback technique that we are using. These equations can be integrated numerically by using a finite increment δv, and treating $dW(v)$ as a finite random variable with a normal distribution and variance δv.

However, in order to estimate the phase variance that is introduced by the measurement, coherent states would be very inefficient, as the phase variance would be dominated by the intrinsic phase variance. It is almost as easy (and much more efficient) to perform calculations on squeezed states, as squeezed states remain squeezed states under the stochastic evolution, and only the two squeezing parameters need be kept track of. The best squeezed states to use are those optimised for minimum intrinsic phase variance. For these states the total phase variance will be approximately twice the intrinsic phase variance when the measurements are close to optimal.

In order to determine the stochastic evolution of the state, I will use a technique similar to that used in Ref. [33]. The master equation for detection of photons is

$$\dot{\rho} = -i\left[H,\rho\right] + \left[a\rho a^\dagger - \tfrac{1}{2}\left(a^\dagger a\rho + \rho a^\dagger a\right)\right]. \tag{3.51}$$

For a consistent detection scheme the measurement operators must give this master equation. In general the master equation is given by

$$\rho(t+dt) = \sum_{n=0}^{N} \Omega_n \rho(t) \Omega_n^\dagger. \tag{3.52}$$

For simple detection the measurement operator is

$$\Omega_1 = \sqrt{dt}\; a. \tag{3.53}$$

In order for this to be consistent with the master equation, the measurement operator for no detection must be

$$\Omega_0 = 1 - \left(iH + \tfrac{1}{2}a^\dagger a\right) dt. \tag{3.54}$$

To consider measurements where the field is combined with a large amplitude coherent state that is treated classically, we can use a unitary rearrangement of the measurement operators. Using that given in [33] we obtain

$$\begin{aligned} \Omega_1 &= \sqrt{dt}\,(a+\gamma) \\ \Omega_0 &= 1 - \left[iH + \tfrac{1}{2}(a^\dagger+\gamma^*)(a+\gamma) + \tfrac{1}{2}(\gamma^* a - \gamma a^\dagger)\right] dt. \end{aligned} \tag{3.55}$$

The operator for no detection can be simplified to

$$\Omega_0 = 1 - \left(iH + \tfrac{1}{2}a^\dagger a + \gamma^* a + \tfrac{1}{2}|\gamma|^2\right) dt. \tag{3.56}$$

When there is a detection the state changes to

$$|\psi(t+dt)\rangle = \frac{(a+\gamma)|\psi(t)\rangle}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}}. \tag{3.57}$$

Now the state after a detection can be multiplied by an arbitrary complex constant with magnitude 1 without altering the properties of the state. Therefore this state can be altered to

$$|\psi(t+dt)\rangle = \frac{(ae^{-i\Phi}+|\gamma|)|\psi(t)\rangle}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}}. \tag{3.58}$$

where Φ is the phase of the local oscillator. This means that the change in the state is

$$d|\psi(t)\rangle = \left[\frac{(ae^{-i\Phi}+|\gamma|)}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}} - 1\right]|\psi(t)\rangle. \tag{3.59}$$

When there is no detection the unnormalised state changes to

$$|\tilde{\psi}(t+dt)\rangle = \left[1 - \left(iH + \tfrac{1}{2}a^\dagger a + \gamma^* a + \tfrac{1}{2}|\gamma|^2\right)dt\right]|\psi(t)\rangle. \tag{3.60}$$

We find that

$$\begin{aligned}\langle\tilde{\psi}(t+dt)|\tilde{\psi}(t+dt)\rangle &= \left\langle\left[1 - \left(-iH + \tfrac{1}{2}a^\dagger a + \gamma a^\dagger + \tfrac{1}{2}|\gamma|^2\right)dt\right]\right. \\ &\quad \left.\times\left[1 - \left(iH + \tfrac{1}{2}a^\dagger a + \gamma^* a + \tfrac{1}{2}|\gamma|^2\right)dt\right]\right\rangle \\ &= \left\langle 1 - \left(a^\dagger a + \gamma^* a + \gamma a^\dagger + |\gamma|^2\right)dt\right\rangle \\ &= 1 - \left(\langle a^\dagger a\rangle + \gamma^*\langle a\rangle + \gamma\langle a^\dagger\rangle + |\gamma|^2\right)dt. \end{aligned} \tag{3.61}$$

Therefore, to obtain a normalised state, we must multiply by a factor of

$$\langle\tilde{\psi}(t+dt)|\tilde{\psi}(t+dt)\rangle^{-1/2} = 1 + \tfrac{1}{2}\left(\langle a^\dagger a\rangle + \gamma^*\langle a\rangle + \gamma\langle a^\dagger\rangle + |\gamma|^2\right)dt. \tag{3.62}$$

This means that the normalised state changes to

$$|\psi(t+dt)\rangle = \left[1 + \left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle\gamma^* a + \gamma a^\dagger\rangle}{2} - \gamma^* a - iH\right)dt\right]|\psi(t)\rangle, \tag{3.63}$$

so the difference in the state is

$$d|\psi(t)\rangle = dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle\gamma^* a + \gamma a^\dagger\rangle}{2} - \gamma^* a - iH\right)|\psi(t)\rangle. \tag{3.64}$$

The probability of a detection occurring in time dt is

$$P_1 = dt\left\langle(a^\dagger+\gamma^*)(a+\gamma)\right\rangle. \tag{3.65}$$

In order to determine a stochastic differential equation for the state, we define a stochastic increment dN which takes the values 0 (for no detection) and 1 (for detection) and has the expectation value:

$$E(dN) = dt\left\langle(a^\dagger+\gamma^*)(a+\gamma)\right\rangle. \tag{3.66}$$

Using this increment, the SDE for the state taking into account both alternatives is

$$\begin{aligned} d|\psi(t)\rangle = &\left[dN\left(\frac{(ae^{-i\Phi}+|\gamma|)}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}} - 1\right)\right. \\ &\left. + dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle\gamma^* a + \gamma a^\dagger\rangle}{2} - \gamma^* a - iH\right)\right]|\psi(t)\rangle. \end{aligned} \tag{3.67}$$

When dN takes the value 0, we simply have the increment for no detection (3.64). When dN takes the value 1, we can ignore the term for no detection, as it is proportional to dt and is infinitesimal. This means that we obtain the increment for a detection given by Eq. (3.59).

Note that Eq. (3.67) differs slightly from the result given in [33]. This is because the state for a detection was multiplied by $e^{-i\Phi}$ here. This is necessary in order to obtain a simple result in the limit of large $|\gamma|$. To take this limit we can approximate the increment dN in terms of Wiener increments dW:

$$dN = \kappa dt + \sqrt{\kappa} dW, \tag{3.68}$$

where

$$\kappa = \left\langle (a^\dagger + \gamma^*)(a + \gamma) \right\rangle. \tag{3.69}$$

This approximation is used because it gives the correct expectation value (3.66), as well as $\langle dN \rangle = \left\langle dN^2 \right\rangle$. Expanding the SDE for the state in terms of this we find

$$\begin{aligned} d|\psi(t)\rangle = &\left[\left(\kappa dt + \sqrt{\kappa} dW\right) \left(\frac{(ae^{-i\Phi} + |\gamma|)}{\sqrt{\langle (a^\dagger + \gamma^*)(a+\gamma) \rangle}} - 1 \right) \right. \\ &\left. + dt \left(\frac{\langle a^\dagger a \rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle \gamma^* a + \gamma a^\dagger \rangle}{2} - \gamma^* a - iH \right) \right] |\psi(t)\rangle \\ = &\left[(-\kappa dt - \sqrt{\kappa} dW + \sqrt{\kappa}(ae^{-i\Phi} + |\gamma|)dt + (ae^{-i\Phi} + |\gamma|)dW \right) \\ &\left. + dt \left(\frac{\langle a^\dagger a \rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle \gamma^* a + \gamma a^\dagger \rangle}{2} - \gamma^* a - iH \right) \right] |\psi(t)\rangle. \end{aligned} \tag{3.70}$$

Expanding κ gives

$$\begin{aligned} \kappa &= \left\langle \left(a^\dagger + \gamma^*\right)(a + \gamma) \right\rangle \\ &= \left\langle a^\dagger a + a^\dagger \gamma + \gamma^* a + |\gamma|^2 \right\rangle \\ &= |\gamma|^2 \left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2} \right), \end{aligned} \tag{3.71}$$

where

$$\chi = \tfrac{1}{2} \left\langle a e^{-i\Phi} + a^\dagger e^{i\Phi} \right\rangle. \tag{3.72}$$

We can expand $\sqrt{\kappa}$ as

$$\begin{aligned} \sqrt{\kappa} &= |\gamma| \sqrt{\left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2} \right)} \\ &= |\gamma| \left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2} \right). \end{aligned} \tag{3.73}$$

Now using this in the expression for the SDE:

$$\begin{aligned} d|\psi(t)\rangle = &\left[\left(-|\gamma|^2 \left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2} \right) dt - |\gamma| \left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2} \right) dW \right. \\ &+ |\gamma| \left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2} \right) (ae^{-i\Phi} + |\gamma|)dt + (ae^{-i\Phi} + |\gamma|)dW \Big) \\ &\left. + dt \left(\frac{\langle a^\dagger a \rangle}{2} - \frac{a^\dagger a}{2} + \frac{\langle \gamma^* a + \gamma a^\dagger \rangle}{2} - \gamma^* a - iH \right) \right] |\psi(t)\rangle \\ = &\left[\left(-|\gamma|^2 \left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2} \right) dt + (ae^{-i\Phi} - \chi) dW \right. \right. \\ &+ |\gamma|^2 \left(1 + \frac{ae^{-i\Phi}}{|\gamma|} + \frac{\chi}{|\gamma|} + \frac{ae^{-i\Phi}\chi}{|\gamma|^2} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2} \right) dt \Big) \\ &\left. + dt \left(\frac{\langle a^\dagger a \rangle}{2} - \frac{a^\dagger a}{2} + |\gamma|\chi - |\gamma| a e^{-i\Phi} - iH \right) \right] |\psi(t)\rangle \\ = &\left[(ae^{-i\Phi} - \chi) dW - \left(iH + \tfrac{1}{2} a^\dagger a - a e^{-i\Phi}\chi + \tfrac{1}{2}\chi^2 \right) dt \right] |\psi(t)\rangle. \end{aligned} \tag{3.74}$$

These results are for the case where the field is directly combined with a local oscillator field, and there is a single photodetector. This was done for simplicity and for consistency with Ref. [33]. The particular experimental configuration that we would like to consider, however, is that shown in Fig. 1.2. Here the field is combined with the local oscillator field via a 50/50 beam splitter, and there are two photodetectors at the two outputs of the beam splitter. This case turns out to be entirely equivalent, as I will now show.

In order to consider this case we can't perform a simple unitary rearrangement of the measurement operators because we need 3 operators, for detection at one arm, detection at the other arm, and no detection. As a simple extrapolation from the previous case we can take the two measurement operators for detection to be

$$\begin{aligned}\Omega_+ &= \sqrt{\frac{dt}{2}}(a+\gamma)\\ \Omega_- &= \sqrt{\frac{dt}{2}}(a-\gamma).\end{aligned} \tag{3.75}$$

In order for this measurement scheme to give the correct master equation, the measurement operator for no detection must be

$$\Omega_0 = 1 - \left(iH + \tfrac{1}{2}a^\dagger a + \tfrac{1}{2}|\gamma|^2\right)dt. \tag{3.76}$$

The above results for a single photodetector can be simply extended to two photodetectors. The change in state for a detection at photodetector 1 is

$$d|\psi_+(t)\rangle = \left[\frac{(|\gamma| + ae^{-i\Phi})}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}} - 1\right]|\psi(t)\rangle, \tag{3.77}$$

and at photodetector 2 is

$$d|\psi_-(t)\rangle = \left[\frac{(|\gamma| - ae^{-i\Phi})}{\sqrt{\langle(a^\dagger-\gamma^*)(a-\gamma)\rangle}} - 1\right]|\psi(t)\rangle. \tag{3.78}$$

When there is no detection the unnormalised state changes to

$$|\tilde{\psi}(t+dt)\rangle = \left[1 - \left(iH + \tfrac{1}{2}a^\dagger a + \tfrac{1}{2}|\gamma|^2\right)dt\right]|\psi(t)\rangle. \tag{3.79}$$

We find that

$$\begin{aligned}\langle\tilde{\psi}(t+dt)|\tilde{\psi}(t+dt)\rangle &= \left\langle\left[1 - \left(-iH + \tfrac{1}{2}a^\dagger a + \tfrac{1}{2}|\gamma|^2\right)dt\right]\left[1 - \left(iH + \tfrac{1}{2}a^\dagger a + \tfrac{1}{2}|\gamma|^2\right)dt\right]\right\rangle\\ &= \left\langle 1 - (a^\dagger a + |\gamma|^2)dt\right\rangle\\ &= 1 - (\langle a^\dagger a\rangle + |\gamma|^2)dt,\end{aligned} \tag{3.80}$$

so the normalising factor is

$$\langle\tilde{\psi}(t+dt)|\tilde{\psi}(t+dt)\rangle^{-1/2} = 1 + \tfrac{1}{2}(\langle a^\dagger a\rangle + |\gamma|^2)dt. \tag{3.81}$$

This means that the normalised state changes to

$$|\psi(t+dt)\rangle = \left[1 + \left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)dt\right]|\psi(t)\rangle, \tag{3.82}$$

so the difference in the state is

$$d|\psi(t)\rangle = dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)|\psi(t)\rangle. \tag{3.83}$$

Then the increment in the state taking into account all three alternatives is

$$d|\psi(t)\rangle = \left[dN_+\left(\frac{(|\gamma| + ae^{-i\Phi})}{\sqrt{\langle(a^\dagger+\gamma^*)(a+\gamma)\rangle}} - 1\right) + dN_-\left(\frac{(|\gamma| - ae^{-i\Phi})}{\sqrt{\langle(a^\dagger-\gamma^*)(a-\gamma)\rangle}} - 1\right) + dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)\right]|\psi(t)\rangle. \tag{3.84}$$

Now rather than defining a single κ, we can define

$$\kappa_\pm = \left\langle(a^\dagger \pm \gamma^*)(a \pm \gamma)\right\rangle. \tag{3.85}$$

In terms of this the increment is

$$d|\psi(t)\rangle = \left[dN_+\left(\frac{(|\gamma| + ae^{-i\Phi})}{\sqrt{\kappa_+}} - 1\right) + dN_-\left(\frac{(|\gamma| - ae^{-i\Phi})}{\sqrt{\kappa_-}} - 1\right) + dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)\right]|\psi(t)\rangle. \tag{3.86}$$

For large $|\gamma|$ we can use the approximation

$$dN_\pm = \frac{\kappa_\pm}{2}dt + \sqrt{\frac{\kappa_\pm}{2}}dW_\pm. \tag{3.87}$$

Substituting this in gives

$$\begin{aligned} d|\psi(t)\rangle = &\left[\left(\frac{\kappa_+}{2}dt + \sqrt{\frac{\kappa_+}{2}}dW_+\right)\left(\frac{(|\gamma| + ae^{-i\Phi})}{\sqrt{\kappa_+}} - 1\right)\right. \\ &+\left(\frac{\kappa_-}{2}dt + \sqrt{\frac{\kappa_-}{2}}dW_-\right)\left(\frac{(|\gamma| - ae^{-i\Phi})}{\sqrt{\kappa_-}} - 1\right) + dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)\Big]|\psi(t)\rangle \\ = &\left[\frac{\sqrt{\kappa_+}}{2}\left(|\gamma| + ae^{-i\Phi}\right)dt + \frac{1}{\sqrt{2}}\left(|\gamma| + ae^{-i\Phi}\right)dW_+ - \frac{\kappa_+}{2}dt - \sqrt{\frac{\kappa_+}{2}}dW_+\right. \\ &+\frac{\sqrt{\kappa_-}}{2}\left(|\gamma| - ae^{-i\Phi}\right)dt + \frac{1}{\sqrt{2}}\left(|\gamma| - ae^{-i\Phi}\right)dW_- - \frac{\kappa_-}{2}dt - \sqrt{\frac{\kappa_-}{2}}dW_- \\ &\left.+dt\left(\frac{\langle a^\dagger a\rangle}{2} - \frac{a^\dagger a}{2} - iH\right)\right]|\psi(t)\rangle. \end{aligned} \tag{3.88}$$

Similarly to the previous case we have

$$\kappa_\pm = |\gamma|^2\left(1 \pm \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right), \tag{3.89}$$

and

$$\sqrt{\kappa_\pm} = |\gamma|\left(1 \pm \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right). \tag{3.90}$$

Substituting this in gives

$$\begin{aligned} d|\psi(t)\rangle = &\left[\frac{|\gamma|^2}{2}\left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right)\left(1 + \frac{ae^{-i\Phi}}{|\gamma|}\right)dt + \frac{|\gamma|}{\sqrt{2}}\left(1 + \frac{ae^{-i\Phi}}{|\gamma|}\right)dW_+\right. \\ &-\frac{|\gamma|^2}{2}\left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right)dt - \frac{|\gamma|}{\sqrt{2}}\left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right)dW_+ \\ &+\frac{|\gamma|^2}{2}\left(1 - \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a\rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right)\left(1 - \frac{ae^{-i\Phi}}{|\gamma|}\right)dt + \frac{|\gamma|}{\sqrt{2}}\left(1 - \frac{ae^{-i\Phi}}{|\gamma|}\right)dW_- \end{aligned}$$

$$
\begin{aligned}
&-\frac{|\gamma|^2}{2}\left(1-\frac{2\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right)dt-\frac{|\gamma|}{\sqrt{2}}\left(1-\frac{\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dW_-\\
&+dt\left(\frac{\langle a^\dagger a\rangle}{2}-\frac{a^\dagger a}{2}-iH\right)\Bigg]|\psi(t)\rangle\\
=&\Bigg[\frac{|\gamma|^2}{2}\left(1+\frac{ae^{-i\Phi}}{|\gamma|}+\frac{\chi}{|\gamma|}+\frac{a\chi e^{-i\Phi}}{|\gamma|^2}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dt+\frac{|\gamma|}{\sqrt{2}}\left(\frac{ae^{-i\Phi}}{|\gamma|}\right)dW_+\\
&-\frac{|\gamma|^2}{2}\left(1+\frac{2\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right)dt-\frac{|\gamma|}{\sqrt{2}}\left(\frac{\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dW_+\\
&+\frac{|\gamma|^2}{2}\left(1-\frac{ae^{-i\Phi}}{|\gamma|}-\frac{\chi}{|\gamma|}+\frac{a\chi e^{-i\Phi}}{|\gamma|^2}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dt-\frac{|\gamma|}{\sqrt{2}}\left(\frac{ae^{-i\Phi}}{|\gamma|}\right)dW_-\\
&-\frac{|\gamma|^2}{2}\left(1-\frac{2\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right)dt-\frac{|\gamma|}{\sqrt{2}}\left(-\frac{\chi}{|\gamma|}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dW_-\\
&+dt\left(\frac{\langle a^\dagger a\rangle}{2}-\frac{a^\dagger a}{2}-iH\right)\Bigg]|\psi(t)\rangle\\
=&\Bigg[|\gamma|^2\left(1+\frac{a\chi e^{-i\Phi}}{|\gamma|^2}+\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dt+\frac{|\gamma|}{\sqrt{2}}\left(\frac{ae^{-i\Phi}}{|\gamma|}-\frac{\chi}{|\gamma|}\right)dW_+\\
&-|\gamma|^2\left(1+\frac{\langle a^\dagger a\rangle}{|\gamma|^2}\right)dt-\frac{|\gamma|}{\sqrt{2}}\left(\frac{ae^{-i\Phi}}{|\gamma|}-\frac{\chi}{|\gamma|}\right)dW_-\\
&+dt\left(\frac{\langle a^\dagger a\rangle}{2}-\frac{a^\dagger a}{2}-iH\right)\Bigg]|\psi(t)\rangle\\
=&\Bigg[|\gamma|^2\left(\frac{a\chi e^{-i\Phi}}{|\gamma|^2}-\frac{\langle a^\dagger a\rangle}{2|\gamma|^2}-\frac{\chi^2}{2|\gamma|^2}\right)dt+\frac{1}{\sqrt{2}}(ae^{-i\Phi}-\chi)(dW_+-dW_-)\\
&+dt\left(\frac{\langle a^\dagger a\rangle}{2}-\frac{a^\dagger a}{2}-iH\right)\Bigg]|\psi(t)\rangle\\
=&\left[(ae^{-i\Phi}-\chi)\frac{(dW_+-dW_-)}{\sqrt{2}}-\left(iH+\tfrac{1}{2}a^\dagger a-a\chi e^{-i\Phi}+\tfrac{1}{2}\chi^2\right)dt\right]|\psi(t)\rangle. \qquad (3.91)
\end{aligned}
$$

Now we find that the combination of the two stochastic increments is equivalent to a third

$$dW=\frac{(dW_+-dW_-)}{\sqrt{2}}. \qquad (3.92)$$

In terms of this we find

$$d|\psi(t)\rangle=\left[(ae^{-i\Phi}-\chi)dW-\left(iH+\tfrac{1}{2}a^\dagger a-a\chi e^{-i\Phi}+\tfrac{1}{2}\chi^2\right)dt\right]|\psi(t)\rangle. \qquad (3.93)$$

This case is therefore exactly that same as the case where the fields are combined directly and there is only one photodetector.

Note also that the POM derived in Ref. [33] was derived for the case where the signal is combined directly with the local oscillator, and there is one photodetector. It turns out that this POM also applies to the case with a 50/50 beam splitter and two photodetectors. This can be shown by performing the derivation in a similar way to Ref. [33]. In general, we can consider the stochastic evolution of an unnormalised state vector

$$|\tilde{\psi}_r(t+T)\rangle=\frac{\Omega_r(T)|\psi(t)\rangle}{\sqrt{\Lambda_r(T)}}, \qquad (3.94)$$

where $\Lambda_r(T)$ is the ostensible probability for the result r, and the actual probability is given by

$$P_r(T)=\Lambda_r(T)\langle\tilde{\psi}_r(t+T)|\tilde{\psi}_r(t+T)\rangle. \qquad (3.95)$$

We can use this strategy on the above measurements, with the ostensible probabilities given by

$$\Lambda_\pm(dt)=\tfrac{1}{2}|\gamma|^2dt, \qquad (3.96)$$

$$\Lambda_0(dt)=1-|\gamma|^2dt. \qquad (3.97)$$

In this case the increments for the two detection cases are

$$d|\tilde{\psi}_{\pm}(t)\rangle = \pm\frac{a}{\gamma}|\psi(t)\rangle, \tag{3.98}$$

and the increment for the case where there is no detection is

$$d|\tilde{\psi}_0(t)\rangle = -dt\left(\tfrac{1}{2}a^\dagger a + iH\right)|\psi(t)\rangle. \tag{3.99}$$

Writing out the evolution explicitly

$$d|\tilde{\psi}(t)\rangle = \left[dN_+(t)\frac{a}{\gamma} - dN_-(t)\frac{a}{\gamma} - dt\left(iH + \tfrac{1}{2}a^\dagger a\right)\right]|\psi(t)\rangle. \tag{3.100}$$

Now we can approximate the Poisson process by a Gaussian process:

$$dN_\pm = \tfrac{1}{2}|\gamma|^2 dt + \frac{|\gamma|}{\sqrt{2}}dW_\pm(t). \tag{3.101}$$

Using this gives

$$\begin{aligned} d|\tilde{\psi}(t)\rangle &= \left[\frac{1}{\sqrt{2}}dW_+(t)ae^{-i\Phi} - \frac{1}{\sqrt{2}}dW_-(t)ae^{-i\Phi} - dt\left(iH + \tfrac{1}{2}a^\dagger a\right)\right]|\psi(t)\rangle \\ &= \left[\frac{1}{\sqrt{2}}ae^{-i\Phi}\left(dW_+(t) - dW_-(t)\right) - dt\left(iH + \tfrac{1}{2}a^\dagger a\right)\right]|\psi(t)\rangle. \end{aligned} \tag{3.102}$$

Now replacing $(dW_+(t) - dW_-(t))/\sqrt{2}$ with $dW(t)$ as before, we obtain

$$d|\tilde{\psi}(t)\rangle = \left[ae^{-i\Phi}dW(t) - dt\left(iH + \tfrac{1}{2}a^\dagger a\right)\right]|\psi(t)\rangle. \tag{3.103}$$

Then the recorded photocurrent is given by

$$\begin{aligned} I(t) &= \lim_{\delta t\to 0}\lim_{|\gamma|\to\infty}\frac{\delta N_+ - \delta N_-}{|\gamma|\delta t} \\ &= \lim_{|\gamma|\to\infty}\frac{\frac{|\gamma|}{\sqrt{2}}dW_+(t) - \frac{|\gamma|}{\sqrt{2}}dW_-(t)}{|\gamma|dt} \\ &= \frac{\frac{1}{\sqrt{2}}\left(dW_+(t) - dW_-(t)\right)}{dt} \\ &= \frac{dW(t)}{dt}. \end{aligned} \tag{3.104}$$

Note that, although the limits are noncommuting when performed on the increments $\delta N_\pm$, they are commuting for Gaussian increments. It is therefore reasonable to perform the time limit first, as in the second line above.

From this point forward the derivation is identical to that given in Ref. [33], and I will therefore not give it explicitly. Note that the above expression for $I(t)$ is obtained using the same definition as when we are considering the actual probabilities, and not the ostensible probabilities. The increments $dN_\pm$ are chosen with the ostensible probabilities here, resulting in an expression for $I(t)$ that only involves the stochastic part.

Therefore we find that, similarly to the case for a single detector, the POM for the parameters A and B defined by

$$\begin{aligned} A &= \int_0^\infty e^{i\Phi(s)}e^{-s/2}I(s)ds, \\ B &= -\int_0^\infty e^{2i\Phi(s)}e^{-s}ds, \end{aligned} \tag{3.105}$$

is

$$F(A,B) = Q(A,B)|\tilde{\psi}(A,B)\rangle\langle\tilde{\psi}(A,B)|, \tag{3.106}$$

where

$$|\tilde{\psi}(A,B)\rangle = \exp\left(\tfrac{1}{2}B(a^\dagger)^2 + Aa^\dagger\right)|0\rangle, \tag{3.107}$$

and $Q(A,B)$ is the ostensible probability distribution. This derivation is for the case $H = 0$, corresponding to a freely damped cavity. From this point forward I will take $H = 0$, rather than continuing to include this term.

Now we can determine the stochastic evolution of a squeezed state. To do this I use the method of Rigo *et al.* [52]. Squeezed states obey the relation

$$\left(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}\right)|A_t^{\rm S}, B_t^{\rm S}\rangle = 0. \tag{3.108}$$

Here the squeezing parameters $A_t^{\rm S}$ and $B_t^{\rm S}$ are defined so that they are analogous to the parameters A and B. They are related to the usual squeezing parameters by

$$\begin{aligned} B_t^{\rm S} &= -\frac{\zeta_t \tanh|\zeta_t|}{|\zeta_t|}, \\ A_t^{\rm S} &= \alpha_t - B_t^{\rm S}(\alpha_t)^*. \end{aligned} \tag{3.109}$$

This is the same as the relation between A and B and the parameters $\alpha^{\rm P}$ and $\zeta^{\rm P}$ for the squeezed state in the probability distribution.

If the squeezed state remains a squeezed state under the increment $d|\psi\rangle$, then it can be shown from Eq. (3.108) that

$$\left(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}\right)d|\psi\rangle = \left(dB_t^{\rm S}a^\dagger + dA_t^{\rm S}\right)|\psi\rangle, \tag{3.110}$$

where the Stratonovich formalism has been used. If we convert the stochastic increment for the state into the Stratonovich form and substitute it into the above equation, then we can obtain an SDE for the squeezing parameters. To convert from the Ito to the Stratonovich form we make the replacement

$$XdY \to XdY - \tfrac{1}{2}dXdY \tag{3.111}$$

Using this on the above SDE gives the Stratonovich increment

$$\begin{aligned} d|\psi(t)\rangle &= \big[(ae^{-i\Phi} - \chi)dW - (\tfrac{1}{2}a^\dagger a - a\chi e^{-i\Phi} + \tfrac{1}{2}\chi^2)dt \\ &\quad -\tfrac{1}{2}\left(a\,d(e^{-i\Phi}) - d\chi\right)dW\big]|\psi(t)\rangle - \tfrac{1}{2}(ae^{-i\Phi} - \chi)dWd|\psi(t)\rangle \\ &= \big[(ae^{-i\Phi} - \chi)dW - (\tfrac{1}{2}a^\dagger a - a\chi e^{-i\Phi} + \tfrac{1}{2}\chi^2)dt \\ &\quad -\tfrac{1}{2}\left(a\,d(e^{-i\Phi}) - d\chi\right)dW|\psi(t)\rangle - \tfrac{1}{2}(ae^{-i\Phi} - \chi)^2dt|\psi(t)\rangle \\ &= \big[(ae^{-i\Phi} - \chi)dW - (\tfrac{1}{2}a^\dagger a + \tfrac{1}{2}a^2e^{-2i\Phi} - 2a\chi e^{-i\Phi} + \chi^2)dt \\ &\quad -\tfrac{1}{2}\left(a\,d(e^{-i\Phi}) - d\chi\right)dW\big]|\psi(t)\rangle. \end{aligned} \tag{3.112}$$

Here the increments $d(e^{-i\Phi})$ and $d\chi$ have been included because the phase of the local oscillator can vary stochastically.

In order to determine the SDE for the squeezing parameters, it is convenient to firstly determine the effect of $(a - B_t^{\rm S}a^\dagger - A_t^{\rm S})$ on each of the operators in this equation.

$$\begin{aligned} (a - B_t^{\rm S}a^\dagger - A_t^{\rm S})C|B_t^{\rm S}, A_t^{\rm S}\rangle &= C(a - B_t^{\rm S}a^\dagger - A_t^{\rm S})|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= 0 \end{aligned} \tag{3.113}$$

$$\begin{aligned} (a - B_t^{\rm S}a^\dagger - A_t^{\rm S})a|B_t^{\rm S}, A_t^{\rm S}\rangle &= \left[a(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}) + B_t^{\rm S}\right]|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= B_t^{\rm S}|B_t^{\rm S}, A_t^{\rm S}\rangle \end{aligned} \tag{3.114}$$

$$\begin{aligned} (a - B_t^{\rm S}a^\dagger - A_t^{\rm S})a^2|B_t^{\rm S}, A_t^{\rm S}\rangle &= \left[a(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}) + B_t^{\rm S}\right]a|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= \left\{a\left[a(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}) + B_t^{\rm S}\right] + aB_t^{\rm S}\right\}|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= 2aB_t^{\rm S}|B_t^{\rm S}, A_t^{\rm S}\rangle \end{aligned}$$

$$= 2B_t^{\rm S}(B_t^{\rm S}a^\dagger + A_t^{\rm S})|B_t^{\rm S}, A_t^{\rm S}\rangle \tag{3.115}$$

$$\begin{aligned}(a - B_t^{\rm S}a^\dagger - A_t^{\rm S})a^\dagger a|B_t^{\rm S}, A_t^{\rm S}\rangle &= \left[a^\dagger(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}) + 1\right] a|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= \left\{a^\dagger\left[a(a - B_t^{\rm S}a^\dagger - A_t^{\rm S}) + B_t^{\rm S}\right] + a\right\}|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= (a^\dagger B_t^{\rm S} + a)|B_t^{\rm S}, A_t^{\rm S}\rangle \\ &= (2a^\dagger B_t^{\rm S} + A_t^{\rm S})|B_t^{\rm S}, A_t^{\rm S}\rangle \end{aligned} \tag{3.116}$$

Here C is any scalar constant. Using these results it is straightforward to show that

$$\begin{aligned}(a - B_t^{\rm S}a^\dagger - A_t^{\rm S})d|\psi(t)\rangle &= \left\{\left[-\tfrac{1}{2}(2a^\dagger B_t^{\rm S} + A_t^{\rm S}) - B_t^{\rm S}(B_t^{\rm S}a^\dagger + A_t^{\rm S})e^{-2i\Phi} + 2B_t^{\rm S}\chi e^{-i\Phi}\right]dt \right. \\ &\quad \left. + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW\right\}|\psi(t)\rangle \\ &= (dB_t^{\rm S}a^\dagger + dA_t^{\rm S})|\psi(t)\rangle. \end{aligned} \tag{3.117}$$

The Stratonovich SDEs for the squeezing parameters are therefore

$$\begin{aligned} dB_t^{\rm S} &= (-B_t^{\rm S} - (B_t^{\rm S})^2 e^{-2i\Phi})dt \\ &= -B_t^{\rm S}(1 + e^{-2i\Phi}B_t^{\rm S})dt \end{aligned} \tag{3.118}$$

$$\begin{aligned} dA_t^{\rm S} &= (-\tfrac{1}{2}A_t^{\rm S} - B_t^{\rm S}A_t^{\rm S}e^{-2i\Phi} + 2B_t^{\rm S}\chi e^{-i\Phi})dt + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW \\ &= \left[-\tfrac{1}{2}A_t^{\rm S} - B_t^{\rm S}A_t^{\rm S}e^{-2i\Phi} + B_t^{\rm S}(\langle a\rangle e^{-2i\Phi} + \langle a^\dagger\rangle)\right]dt + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW \\ &= \left\{-\tfrac{1}{2}A_t^{\rm S} - B_t^{\rm S}A_t^{\rm S}e^{-2i\Phi} + B_t^{\rm S}\left[(B_t^{\rm S}\langle a^\dagger\rangle + A_t^{\rm S})e^{-2i\Phi} + \langle a^\dagger\rangle\right]\right\}dt \\ &\quad + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW \\ &= -\tfrac{1}{2}A_t^{\rm S}dt + B_t^{\rm S}\langle a^\dagger\rangle(1 + e^{-2i\Phi}B_t^{\rm S})dt + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW. \end{aligned} \tag{3.119}$$

In order to convert back to the SDE for the standard (non-scaled) amplitude, note that

$$\alpha_t = \frac{A_t^{\rm S} + B_t^{\rm S}(A_t^{\rm S})^*}{1 - |B_t^{\rm S}|^2}. \tag{3.120}$$

Using this expression, the Stratonovich increment in α_t is

$$\begin{aligned} d\alpha_t &= \frac{dA_t^{\rm S} + B_t^{\rm S}dA_t^{\rm S*}}{1 - |B_t^{\rm S}|^2} + \frac{A_t^{\rm S*}dB_t^{\rm S}}{1 - |B_t^{\rm S}|^2} + \left(B_t^{\rm S*}dB_t^{\rm S} + B_t^{\rm S}dB_t^{\rm S*}\right)\frac{A_t^{\rm S} + B_t^{\rm S}A_t^{\rm S*}}{\left(1 - |B_t^{\rm S}|^2\right)^2} \\ &= \left[dA_t^{\rm S} + B_t^{\rm S}dA_t^{\rm S*} + A_t^{\rm S*}dB_t^{\rm S} + \alpha_t\left(B_t^{\rm S*}dB_t^{\rm S} + B_t^{\rm S}dB_t^{\rm S*}\right)\right]\frac{1}{1 - |B_t^{\rm S}|^2}. \end{aligned} \tag{3.121}$$

Expanding this gives

$$\begin{aligned} d\alpha_t &= \left\{-\tfrac{1}{2}A_t^{\rm S}dt + B_t^{\rm S}\langle a^\dagger\rangle\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW + B_t^{\rm S}\left[-\tfrac{1}{2}A_t^{\rm S*}dt\right.\right. \\ &\quad \left. + B_t^{\rm S*}\langle a\rangle\left(1 + e^{2i\Phi}B_t^{\rm S*}\right)dt + B_t^{\rm S*}e^{i\Phi}dW - \tfrac{1}{2}B_t^{\rm S*}d(e^{i\Phi})dW\right] - A_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt \\ &\quad \left. -\alpha_t\left(B_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}B_t^{\rm S*}\left(1 + e^{2i\Phi}B_t^{\rm S*}\right)dt\right)\right\}\frac{1}{1 - |B_t^{\rm S}|^2} \\ &= -\tfrac{1}{2}\alpha_t dt + \left\{B_t^{\rm S}\alpha_t^*\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}e^{-i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW\right. \\ &\quad + B_t^{\rm S}\left[B_t^{\rm S*}\alpha_t\left(1 + e^{2i\Phi}B_t^{\rm S*}\right)dt + B_t^{\rm S*}e^{i\Phi}dW - \tfrac{1}{2}B_t^{\rm S*}d(e^{i\Phi})dW\right] - A_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt \\ &\quad \left. -\alpha_t\left(B_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}B_t^{\rm S*}\left(1 + e^{2i\Phi}B_t^{\rm S*}\right)dt\right)\right\}\frac{1}{1 - |B_t^{\rm S}|^2} \\ &= -\tfrac{1}{2}\alpha_t dt + \left\{B_t^{\rm S}\left[\alpha_t^* - B_t^{\rm S*}\alpha_t\right]\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}e^{-i\Phi}dW + B_t^{\rm S}B_t^{\rm S*}e^{i\Phi}dW\right. \\ &\quad \left. -\tfrac{1}{2}B_t^{\rm S}d\left(e^{-i\Phi}\right)dW - \tfrac{1}{2}B_t^{\rm S}B_t^{\rm S*}d(e^{i\Phi})dW - A_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt\right\}\frac{1}{1 - |B_t^{\rm S}|^2} \\ &= -\tfrac{1}{2}\alpha_t dt + \left\{B_t^{\rm S}A_t^{\rm S*}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt + B_t^{\rm S}e^{-i\Phi}dW + B_t^{\rm S}B_t^{\rm S*}e^{i\Phi}dW - \tfrac{1}{2}B_t^{\rm S}d(e^{-i\Phi})dW\right. \\ &\quad \left. -\tfrac{1}{2}B_t^{\rm S}B_t^{\rm S*}d(e^{i\Phi})dW - A_t^{\rm S*}B_t^{\rm S}\left(1 + e^{-2i\Phi}B_t^{\rm S}\right)dt\right\}\frac{1}{1 - |B_t^{\rm S}|^2} \\ &= -\tfrac{1}{2}\alpha_t dt + \frac{B_t^{\rm S}dW}{1 - |B_t^{\rm S}|^2}\left(B_t^{\rm S*}e^{i\Phi} + e^{-i\Phi}\right) - \tfrac{1}{2}\frac{B_t^{\rm S}dW}{1 - |B_t^{\rm S}|^2}\left(B_t^{\rm S*}d(e^{i\Phi}) + d(e^{-i\Phi})\right). \end{aligned} \tag{3.122}$$

Converting back to the Ito form of the SDE gives

$$d\alpha_t = -\tfrac{1}{2}\alpha_t dt + \frac{B_t^{\mathrm{S}} dW}{1 - \left|B_t^{\mathrm{S}}\right|^2}\left(B_t^{\mathrm{S}*} e^{i\Phi} + e^{-i\Phi}\right). \tag{3.123}$$

The SDE for B_t^{S} is unchanged. Note that in converting back to the Ito form, the extra term involving the increment in the phase of the local oscillator cancels out.

Now for consistency with Ref. [33] I will take the signal to be

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_+ - \delta N_-}{|\gamma| \delta t}. \tag{3.124}$$

In order to evaluate this, recall that the expression for the photon number counts at the two photodetectors is

$$dN_\pm = \frac{\kappa_\pm}{2} dt + \sqrt{\frac{\kappa_\pm}{2}} dW_\pm, \tag{3.125}$$

where

$$\kappa_\pm = |\gamma|^2 \left(1 \pm \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2}\right), \tag{3.126}$$

and the expansion of the square root is

$$\sqrt{\kappa_\pm} = |\gamma| \left(1 \pm \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right). \tag{3.127}$$

Using these expressions gives the photocurrent as

$$\begin{aligned}
I(t)dt &= \lim_{|\gamma| \to \infty} \left[\frac{|\gamma|^2}{2} \left(1 + \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2}\right) dt + \frac{|\gamma|}{\sqrt{2}} \left(1 + \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right) dW_+ \right. \\
&\quad \left. - \frac{|\gamma|^2}{2} \left(1 - \frac{2\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{|\gamma|^2}\right) dt - \frac{|\gamma|}{\sqrt{2}} \left(1 - \frac{\chi}{|\gamma|} + \frac{\langle a^\dagger a \rangle}{2|\gamma|^2} - \frac{\chi^2}{2|\gamma|^2}\right) dW_- \right] \frac{1}{|\gamma|} \\
&= \lim_{|\gamma| \to \infty} \left[|\gamma|^2 \left(\frac{2\chi}{|\gamma|}\right) dt + \frac{|\gamma|}{\sqrt{2}} dW_+ - \frac{|\gamma|}{\sqrt{2}} dW_- \right] \frac{1}{|\gamma|} \\
&= 2\chi dt + \frac{1}{\sqrt{2}} \left(dW_+ - dW_-\right) \\
&= 2\chi dt + dW \\
&= 2\mathrm{Re}\left(\langle a \rangle e^{-i\Phi}\right) dt + dW.
\end{aligned} \tag{3.128}$$

This result is entirely general and does not depend on the state being coherent or squeezed. In the case of a squeezed state we have

$$\langle a \rangle = \alpha_t. \tag{3.129}$$

Now note that the deterministic part of the SDE for α_t is

$$\langle d\alpha_t \rangle = -\tfrac{1}{2}\alpha_t dt. \tag{3.130}$$

This means that the deterministic part of the evolution is

$$\alpha_t \propto e^{-t/2}. \tag{3.131}$$

In addition, the parameters A_t and B_t are defined in Ref. [33] by

$$A_t = \int_0^t e^{i\Phi} e^{-s/2} I(s) ds, \tag{3.132}$$

$$B_t = -\int_0^t e^{2i\Phi} e^{-s} ds. \tag{3.133}$$

The exponential factors can be removed by changing the time variable to

$$v = 1 - e^{-t}, \tag{3.134}$$

so that

$$\begin{aligned} dv &= e^{-t}dt \\ &= (1-v)dt. \end{aligned} \tag{3.135}$$

This transformation maps the time to the unit interval $[0, 1)$. It is also convenient to define a new scaled coherent amplitude to remove the systematic variation:

$$\alpha_v = \alpha_t e^{t/2}. \tag{3.136}$$

Here the v subscript indicates the scaled amplitude, and the t subscript indicates the original, unscaled amplitude. Since these are equal to each other at zero time, there is no ambiguity in the initial amplitude α. The definition of $I(t)$ must also be altered to remove the exponential factors:

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_+ - \delta N_-}{|\gamma| e^{-t/2} \delta t}. \tag{3.137}$$

Using this gives the photocurrent as

$$I(v)dv = 2\mathrm{Re}(\alpha_v e^{-i\Phi(v)})dv + dW(v). \tag{3.138}$$

With these changes of variables, the definitions for A_v and B_v become

$$A_v = \int_0^v e^{i\Phi} I(u)du, \tag{3.139}$$

$$B_v = -\int_0^v e^{2i\Phi} du. \tag{3.140}$$

The SDE for the scaled amplitude α_v is

$$\begin{aligned} d\alpha_v &= d\left(\alpha_t e^{t/2}\right) \\ &= \left[-\tfrac{1}{2}\alpha_t dt + \frac{B_t^{\mathrm{S}} dW(t)}{1 - |B_t^{\mathrm{S}}|^2}\left(B_t^{\mathrm{S}*} e^{i\Phi} + e^{-i\Phi}\right)\right] e^{t/2} + \tfrac{1}{2}\alpha_t e^{t/2} dt \\ &= e^{t/2} \frac{B_t^{\mathrm{S}} dW(t)}{1 - |B_t^{\mathrm{S}}|^2}\left(B_t^{\mathrm{S}*} e^{i\Phi} + e^{-i\Phi}\right) \\ &= \frac{1}{1-v} \frac{B_v^{\mathrm{S}} dW(v)}{1 - |B_v^{\mathrm{S}}|^2}\left(B_v^{\mathrm{S}*} e^{i\Phi} + e^{-i\Phi}\right). \end{aligned} \tag{3.141}$$

Similarly the SDE for B_v^{S} in terms of the new time variable is

$$dB_v^{\mathrm{S}} = -\frac{dv}{1-v} B_v^{\mathrm{S}}\left(1 + e^{-2i\Phi} B_v^{\mathrm{S}}\right). \tag{3.142}$$

Initial calculations were performed using these equations, but there is a further simplification that can be made. The solution for B_v^{S} is

$$B_v^{\mathrm{S}} = \frac{1-v}{(B_0^{\mathrm{S}})^{-1} - B_v^*}. \tag{3.143}$$

The only calculations that will be presented in this thesis that were calculated using numerical integration to find B_v^{S} rather than this solution are those for larger photon numbers for the constant

ε case where the optimum value of ε was found numerically. It was found that performing the integration rather than using this solution altered the results by less than 0.5%.

It is also possible to consider input states with more general mode-functions in a similar way [35]. In the general case, the input state has a time-varying mode function $u(t)$ that is real and positive, and is normalised so that

$$\int_0^T u(t)dt = 1, \tag{3.144}$$

where T is the total pulse length. The mode function gives the systematic variation of $\langle a \rangle$, i.e.

$$\langle a \rangle = \alpha_t \propto \sqrt{u(t)}, \tag{3.145}$$

if any stochastic variation is ignored. If the photocurrent is defined as in (3.124), then the definitions for A_t and B_t are

$$A_t = \int_0^t e^{i\Phi(s)} \sqrt{u(s)} I(s) ds, \tag{3.146}$$

$$B_t = -\int_0^t e^{2i\Phi(s)} u(s) ds. \tag{3.147}$$

For the case of squeezed states, the mode-function is

$$u(t) = e^{-t}. \tag{3.148}$$

Using this mode-function these definitions are identical to those used in [33].

In [35] there is no factor of $\sqrt{u(s)}$ in the definition of A_t. This is because $I(t)$ is defined in a slightly different way. In [35] it is assumed that the local oscillator is varied with the same mode-function, so that

$$|\gamma| = \sqrt{u(t)}\beta, \tag{3.149}$$

where β is a constant. The photocurrent is then defined as

$$I(t) = \lim_{\delta t \to 0} \lim_{\beta \to \infty} \frac{\delta N_+ - \delta N_-}{\beta \delta t}. \tag{3.150}$$

This is equivalent to using

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_+ - \delta N_-}{|\gamma| \delta t} \times \sqrt{u(t)}. \tag{3.151}$$

As there is a factor of $\sqrt{u(t)}$ here, there is no need for it in the definition of A_t used in [35].

The mode-function can be removed by redefining $I(t)$ as

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_+ - \delta N_-}{|\gamma| \sqrt{u(t)} \delta t}, \tag{3.152}$$

and changing the time variable to

$$v = \int_0^t u(s) ds. \tag{3.153}$$

With these changes of variables, A_v and B_v are now given by

$$A_v = \int_0^v e^{i\Phi} I(u) du, \tag{3.154}$$

$$B_v = -\int_0^v e^{2i\Phi} du, \tag{3.155}$$

and $I(t)$ simplifies to

$$I(v)dv = 2\mathrm{Re}[\alpha_v e^{-i\Phi(v)}]dv + dW(v), \tag{3.156}$$

where

$$\alpha_v = \alpha_t/\sqrt{u(t)}. \tag{3.157}$$

Here α_v may vary stochastically, but should have no systematic variation. Note that for $u(t) = e^{-t}$, these changes of variables are identical to those used previously to remove the exponential factors.

An additional complication to the numerical technique is that rather than using the feedback in the form

$$\hat{\varphi}_v = \arg C_v^{1-\varepsilon(v)} A_v^{\varepsilon(v)}, \tag{3.158}$$

the form used was

$$\hat{\varphi}_v = \arg C_v \left(\frac{A_v}{C_v}\right)^{\varepsilon(v)}. \tag{3.159}$$

These forms are almost always equivalent; however, for some cases different results are obtained. If the simple form (3.158) is used, this tends to bias the feedback phases towards zero. The reason for this is that fractional powers generally have multiple solutions, and the program does not necessarily give the appropriate one.

To demonstrate this, consider for example the case where $\varepsilon = \frac{1}{2}$. If the phases of A_v and C_v are $2\pi/3$ and $-2\pi/3$, then clearly the phase estimate that we want is halfway between these, at $\hat{\varphi}_v = \pi$. The feedback in the form (3.158) will give

$$\begin{aligned}\hat{\varphi}_v &= \arg(e^{-i2\pi/3})^{0.5}(e^{i2\pi/3})^{0.5} \\ &= \arg e^{-i\pi/3}e^{i\pi/3} \\ &= \arg 1 \\ &= 0.\end{aligned} \tag{3.160}$$

What has happened is that the power of 0.5 has rotated the phases towards zero, so that multiplying $C_v^{0.5}$ and $A_v^{0.5}$ together has yielded a phase of zero. On the other hand, if we use the form given by (3.159), we find

$$\begin{aligned}\hat{\varphi}_v &= \arg e^{-i2\pi/3}(e^{-i2\pi/3})^{0.5} \\ &= \arg e^{-i2\pi/3}e^{-i\pi/3} \\ &= \arg(-1) \\ &= \pi,\end{aligned} \tag{3.161}$$

which is the correct result. I will usually give the feedback in the simple form (3.158), even though it is necessary to use (3.159) in actual calculations.

3.4 Time Steps

Before performing the calculations, it is important to estimate the inaccuracy introduced by the finite time steps. This is necessary to determine how many time steps are required in order to keep the error due to this factor negligible. In Ch. 4 it is shown that the minimum phase variance when there is a time *delay* of τ is $\tau/(8\overline{n})$. It is reasonable that there should be a similar scaling for the error due to the finite time step.

In order to estimate the error due to the time step, the phase variance was determined numerically for mark II measurements on squeezed states optimised for minimum intrinsic phase variance. Calculations were performed for mean photon numbers of approximately 122, 432, 1577 and 5877. For each photon number calculations were performed simultaneously for 2^m intervals, with several different values of m. For photon numbers up to 1577, the value of m was varied from 7 to 16, and for $\overline{n} = 5877$, m was varied from 8 to 17. For each time delay 2^{14} samples were used.

In order to minimise the error between the calculations with different numbers of intervals, the random numbers for the smaller numbers of intervals were determined by combining the random numbers for the larger numbers of intervals. Specifically, for a step size of δv, dW was replaced in the integration with δW, which has a Gaussian distribution with a variance of δv. For the integration with twice the step size, $\delta v' = 2\delta v$, adjacent increments δW were added to give the new increment $\delta W'$, with variance $2\delta v$.

The results for mean photon numbers of 122, 432, 1577 and 5877 are shown in Figs 3.4, 3.5, 3.6 and 3.7, respectively. In these figures the straight lines shown are based on weighted linear regressions. The slopes found for each of these graphs were

$$\begin{aligned} \mathrm{S}_{3.4} &= \frac{0.180 \pm 0.010}{\overline{n}}, \\ \mathrm{S}_{3.5} &= \frac{0.185 \pm 0.009}{\overline{n}}, \\ \mathrm{S}_{3.6} &= \frac{0.205 \pm 0.009}{\overline{n}}, \\ \mathrm{S}_{3.7} &= \frac{0.230 \pm 0.012}{\overline{n}}. \end{aligned} \tag{3.162}$$

These uncertainties are based on the uncertainties of the individual points. As the same set of random numbers was used for each time step size, it would be expected that the uncertainty based on the deviation of the points from the line would be smaller. This is true for the smaller photon numbers, but not for the larger photon numbers where the points tend to fit the line poorly.

Note that there tends to be some nonlinearity in the results for the larger photon numbers. This is most obvious in Fig. 3.7, where the variances for the larger time steps are much higher than would be expected based on a linear interpolation from the results for small time steps. In addition the error bars are larger for these time intervals. The reason for this is that there are a small number of results with large errors in the data sets on which these points are based. This causes both the variance and the uncertainty in the variance to be larger. This is also true to a lesser extent for the smaller photon numbers.

We are only interested in the phase variance for small time steps here, as we wish to use time steps small enough to keep the error due to the finite step size below about 1%. For small step sizes, the effect of this type of nonlinearity will be that the error is less than that expected based on the linear fits for larger time steps. For example, if the last two points in Fig. 3.7 are omitted, the slope obtained is $0.13/\overline{n}$.

There is a fair amount of variation between the slopes found for the different photon numbers, and for different ranges of the linear fit. Nevertheless, the purpose of these calculations is not to find an exact result, but rather to place an upper limit on the error due to the finite step size. In each case, we find that

$$\Delta V(\phi) < \frac{\delta v}{4\overline{n}}. \tag{3.163}$$

This means that if we use time steps of

$$\delta v = \frac{\overline{n} V(\phi)}{25}, \tag{3.164}$$

where $V(\phi)$ is the total phase variance, the error in the phase variance should be less than 1%.

3.5 Naïve Constant ε Feedback

The first method that I will consider is that where the values of ε are the optimum values predicted by the simplified theory:

$$\varepsilon = \left(2\alpha_{\mathrm{opt}} e^{2\zeta_{\mathrm{opt}}}\right)^2. \tag{3.165}$$

For the results presented here, the states used are squeezed states optimised for minimum intrinsic phase variance. For these states, the theoretical limit to the total phase variance is twice the intrinsic

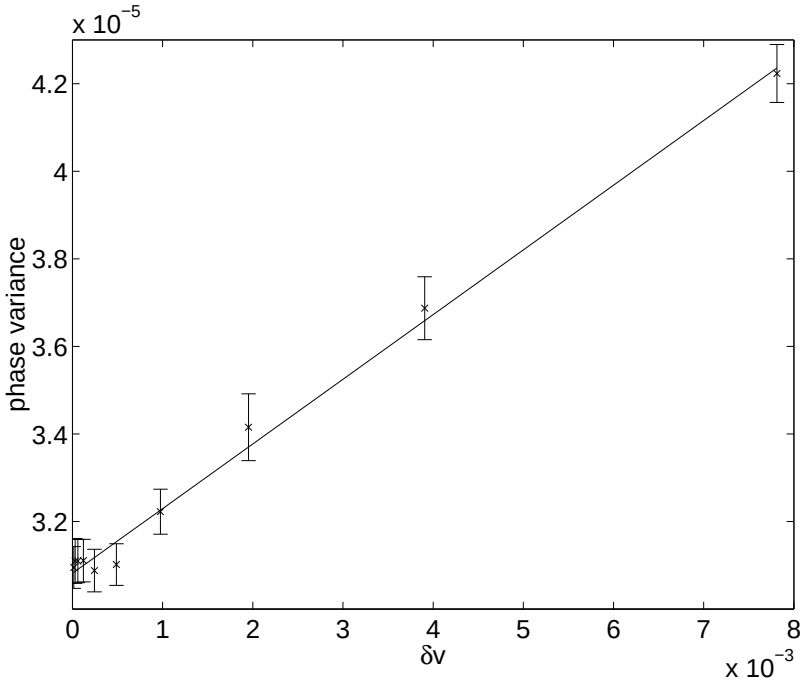

Figure 3.4: The phase variance for mark II measurements on optimised squeezed states with a mean photon number of 122 with various time step sizes. The crosses are the numerical results, and the continuous line is that fitted by a linear regression.

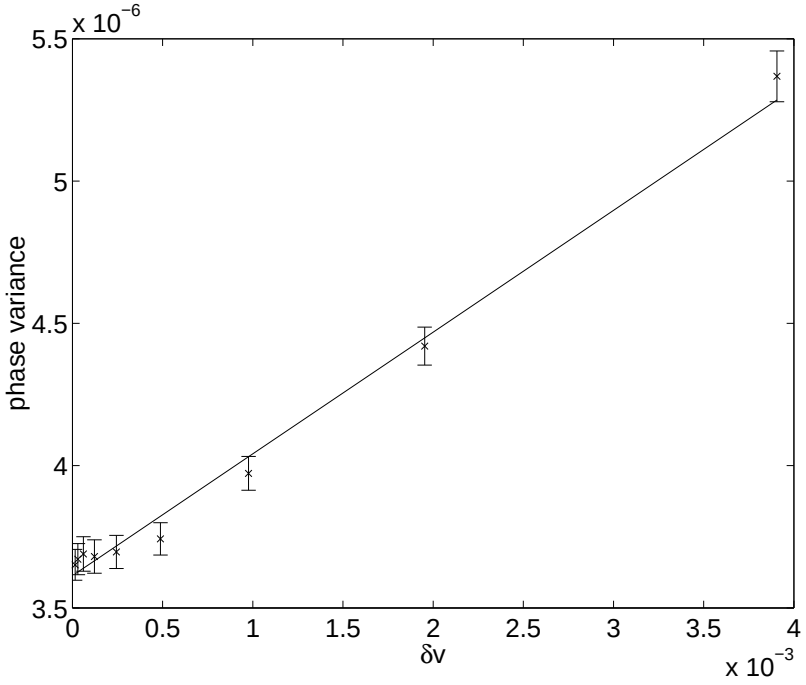

Figure 3.5: The phase variance for mark II measurements on optimised squeezed states with a mean photon number of 432 with various time step sizes. The crosses are the numerical results, and the continuous line is that fitted by a linear regression.

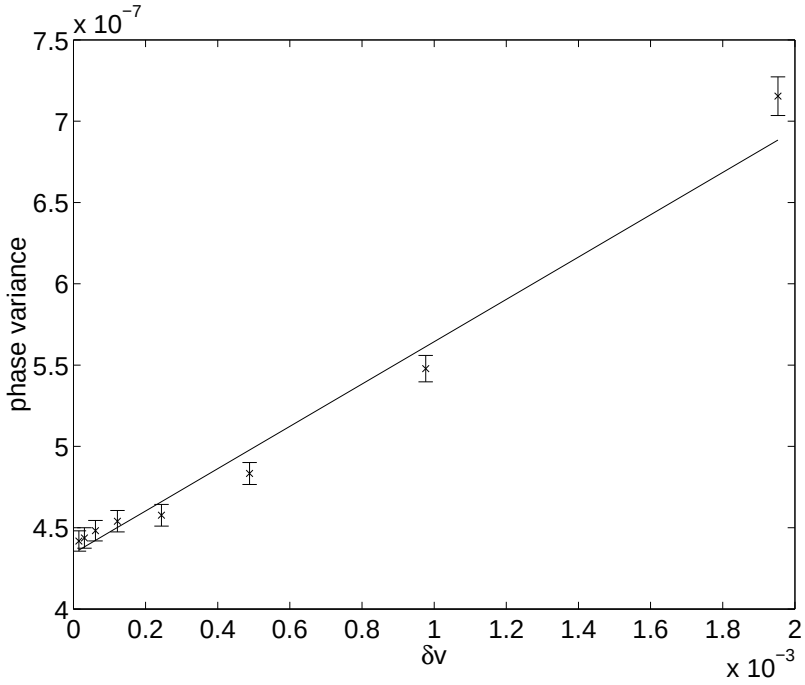

Figure 3.6: The phase variance for mark II measurements on optimised squeezed states with a mean photon number of 1577 with various time step sizes. The crosses are the numerical results, and the continuous line is that fitted by a linear regression.

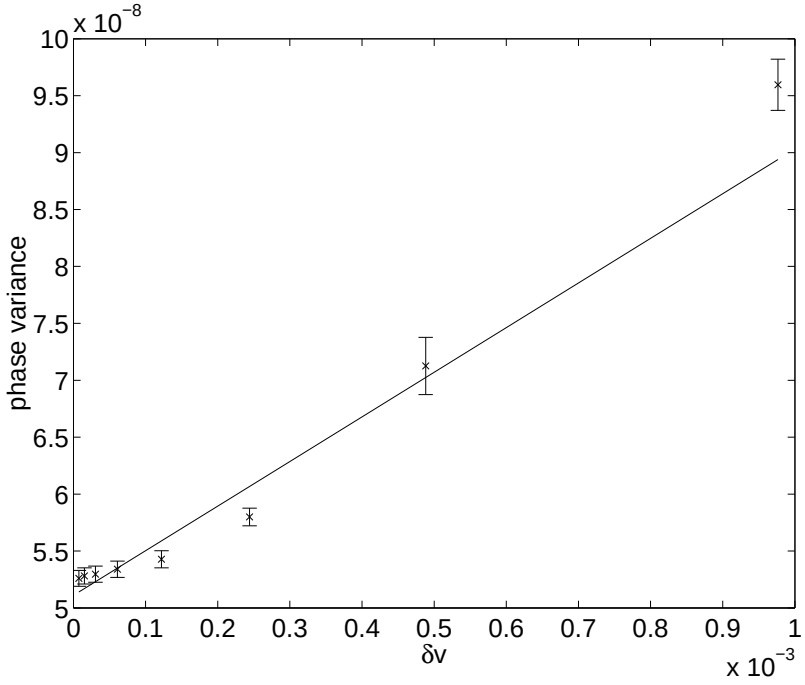

Figure 3.7: The phase variance for mark II measurements on optimised squeezed states with a mean photon number of 5877 with various time step sizes. The crosses are the numerical results, and the continuous line is that fitted by a linear regression.

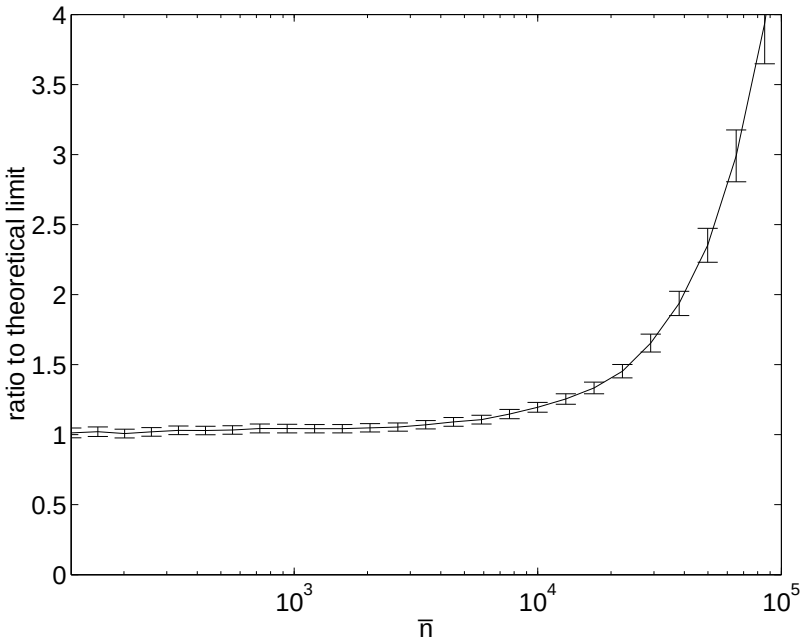

Figure 3.8: The variance for phase measurements using the predicted optimum value of ε as a ratio to the theoretical limit.

phase variance. The variance for mark II measurements (on general states optimised for minimum phase variance under these measurements) does not exceed this limit until a photon number of about 140. This does not contradict the theory, because the theoretical limit is only expected to be accurate for large photon numbers. Because of this, the improved phase measurement schemes will only be considered for photon numbers above about 140.

The results using this predicted value of ε are shown as a ratio to the theoretical limit (of twice the intrinsic phase variance) in Fig. 3.8, and as a ratio to the phase variance for mark II measurements in Fig. 3.9. For the results shown in these figures, 2^{12} samples were used. For moderate photon numbers above 140 this feedback does give an improvement over mark II measurements, and is close to the theoretical limit. This is true only over a small range of photon numbers, and for photon numbers over about 5000 the phase variance is significantly over the theoretical limit. In fact, for photon numbers above about 160000 this feedback technique gives larger phase variances than the mark II technique. The best improvement in the phase variance is at a photon number of about 20000, where the phase variance is about 36% of the mark II phase variance.

The reason for these poor results is that the values of $|A|$ given by this feedback technique are too high. If we plot the mean values of $|A|^2$ (see Fig. 3.10), we see that although they start out close to ε, the agreement gradually gets worse and worse. Now recall from the above results for coherent states that although there is agreement with theory over a large range of values of ε, for smaller values of ε, larger photon numbers are required in order to have good agreement with theory.

In fact, the photon number required scales at least as ε^{-2}. Conversely, we can say that the mimimum value of ε for good agreement with the theory scales roughly as $\overline{n}^{-1/2}$. For the predicted values of ε above, we find that ε scales as $\log^2 \overline{n}/\overline{n}$. Clearly this means that for the larger photon numbers, the value of ε will be too small for good agreement with theory.

3.6 Optimised Constant ε Feedback

In order to obtain smaller values of $|A|$, and therefore phase variances that are closer to the theoretical limit, we can use different values of ε. We would at first expect that reducing ε would give smaller values of $|A|$; however, this is not necessarily the case. As demonstrated in Fig. 3.11, as the value

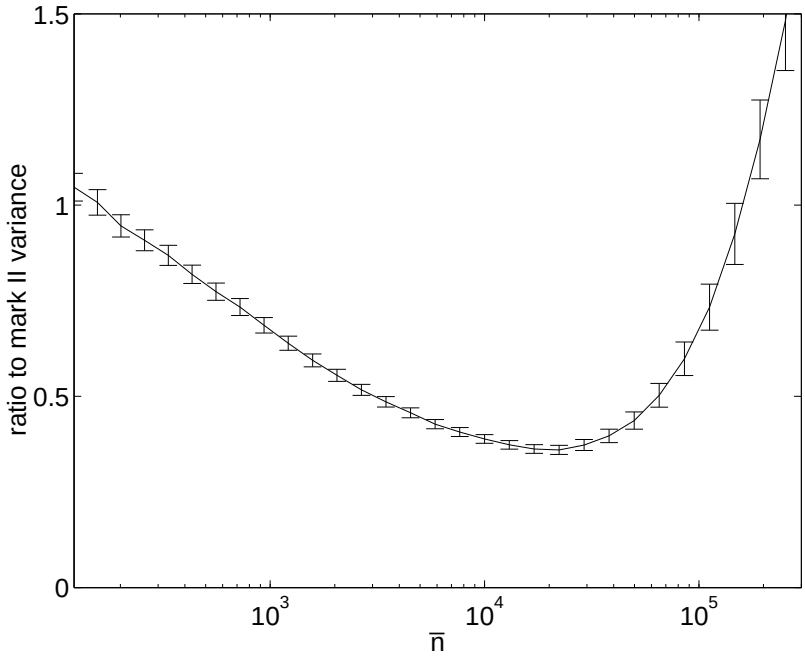

Figure 3.9: The variance for phase measurements using the predicted optimum value of ε as a ratio to the phase variance for mark II measurements on optimum input states.

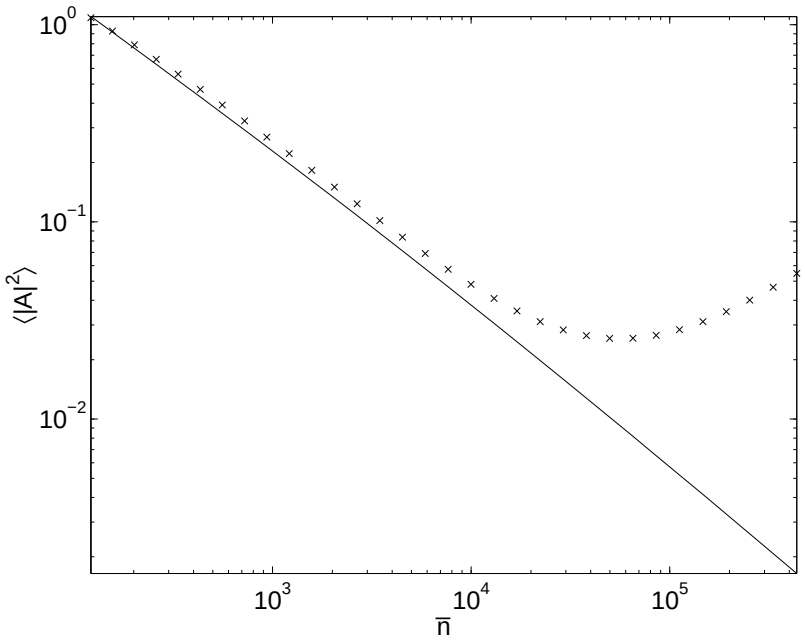

Figure 3.10: The mean values of $|A|^2$ obtained for phase measurements using the predicted optimum value of ε. The analytic result of $\langle |A|^2 \rangle = \varepsilon$ is shown as the continuous line and the numerically obtained values of $\langle |A|^2 \rangle$ are shown as crosses.

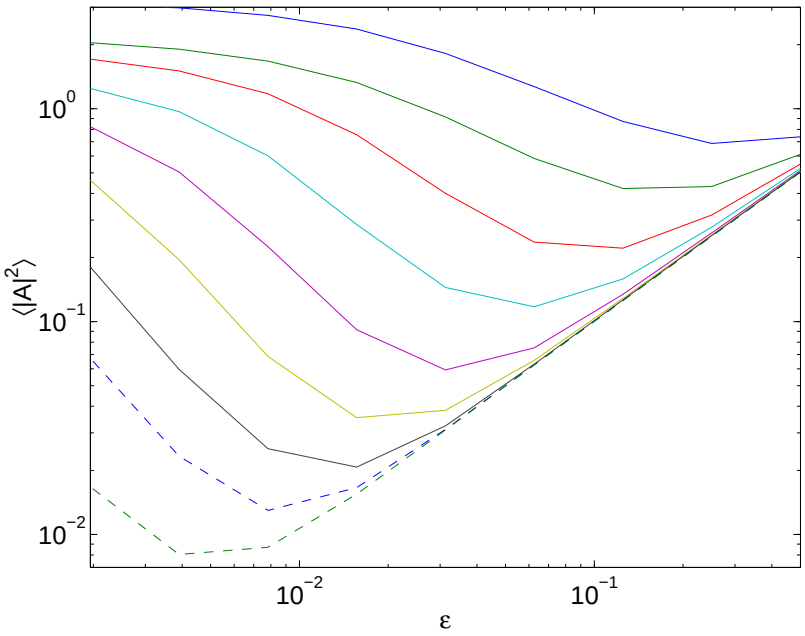

Figure 3.11: The mean values of $|A|^2$ for measurements on coherent states with $\arg(C_v^{1-\varepsilon}A_v^{\varepsilon})$ feedback. The results for $\alpha = 4$ are shown in dark blue, for $\alpha = 8$ in green, for $\alpha = 16$ in red, for $\alpha = 32$ in light blue, for $\alpha = 64$ in purple, for $\alpha = 128$ in yellow, for $\alpha = 256$ in black, for $\alpha = 512$ as a dashed dark blue line, and for $\alpha = 1024$ as a dashed green line.

of ε is decreased for a fixed photon number, the value of $|A|^2$ decreases initially, but it reaches a minimum, and then as ε is decreased further $|A|^2$ increases. (This figure is for the same data as Fig. 3.3.)

As we are in the region where the approximate theory breaks down, rather than using a predicted value of ε, it is better to vary ε to determine which value gives the smallest phase variance for each mean photon number. The phase variances obtained by this method are plotted in Fig. 3.12 as a ratio to the theoretical limit, and in Fig. 3.13 as a ratio to the mark II phase variance. For these results 2^{11} samples were used for moderate photon numbers, and 2^9 samples were used for photon numbers above 2×10^6.

Again the phase variance is close to the theoretical limit for photon numbers up to about 5000, but beyond this the phase variance is greater and greater than the theoretical limit. Unlike the previous case, however, the phase variance continues to get smaller as compared to the mark II phase variance. For the maximum photon number calculations have been performed for, the phase variance is less than 10% of the mark II phase variance.

If we plot the ratio of the numerically determined optimum value of ε to the analytically predicted value (see Fig. 3.14), we find that for smaller photon numbers the optimum value of ε is less than the predicted value. As the photon number is increased, the optimum value of ε is greater and greater as a ratio to the predicted value. The actual value of ε is still decreasing with photon number, though, as demonstrated in Fig. 3.15.

These results indicate that although this phase feedback scheme is not achieving the theoretical limit, it is achieving a better scaling than for mark II measurements. Assuming that $h(n) \approx cn^{-p}$, a numerical fit was performed on the data to determine c and p. Recall from Sec. 2.3.3 that the approximate phase variance for general measurements on squeezed states is

$$V(\phi) \approx V(\phi_{\text{can}}) + 2c\overline{n}^{-p}\left[1 + \frac{p(p+1)}{2n_0}\right]. \tag{3.166}$$

Rather than simply assuming that the introduced phase variance was $2c\overline{n}^{-p}$, this expression was

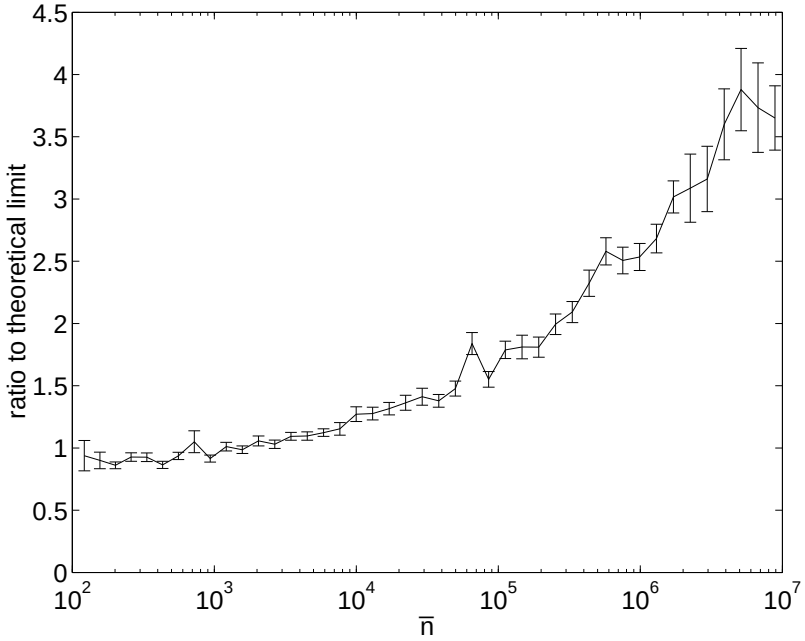

Figure 3.12: The variance for phase measurements using the numerically determined optimum value of ε as a ratio to the theoretical limit.

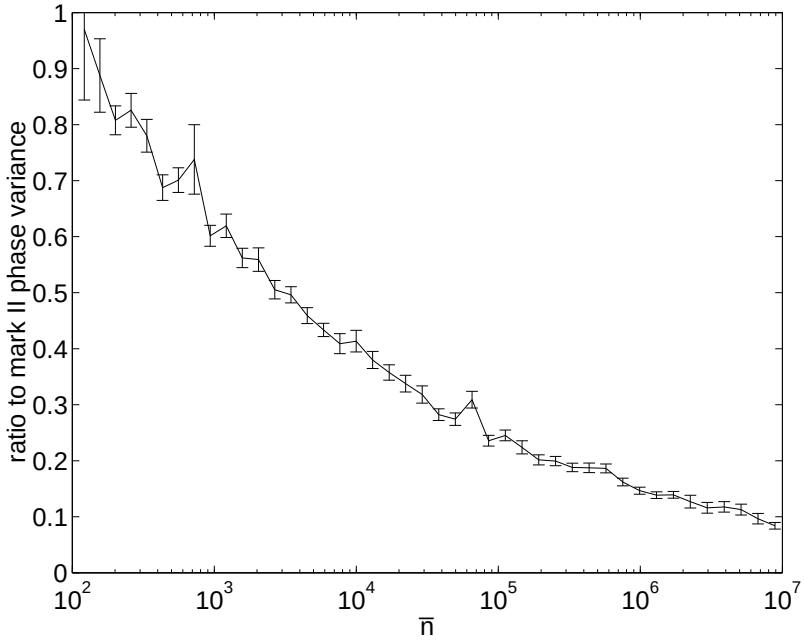

Figure 3.13: The variance for phase measurements using the numerically determined optimum value of ε as a ratio to the mark II phase variance.

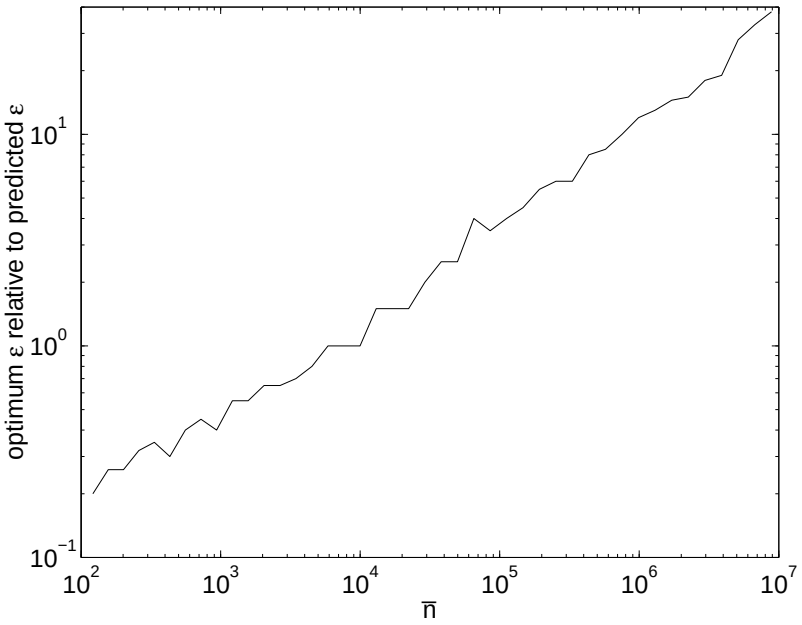

Figure 3.14: The numerically determined optimum values of ε as ratios to the analytically predicted optimum values of ε.

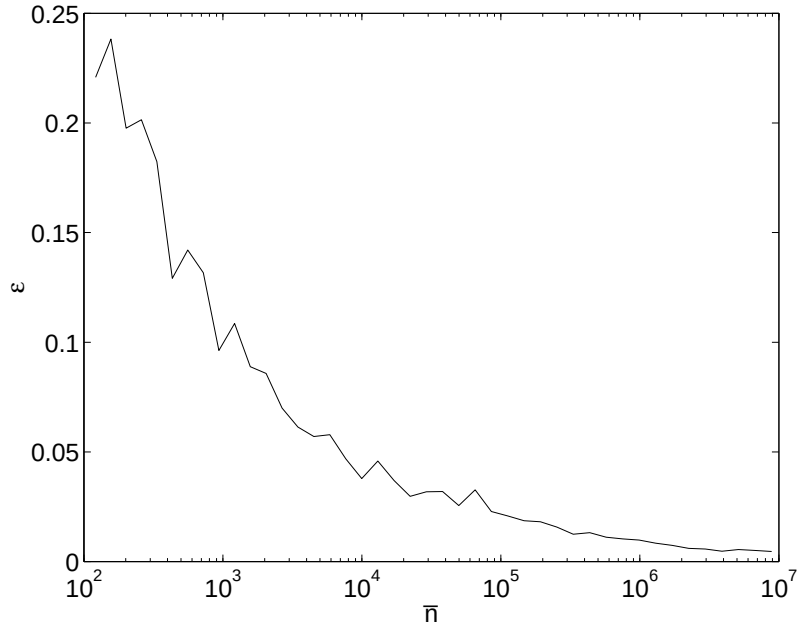

Figure 3.15: The numerically determined optimum values of ε for $\arg(C_v^{1-\varepsilon} A_v^{\varepsilon})$ feedback.

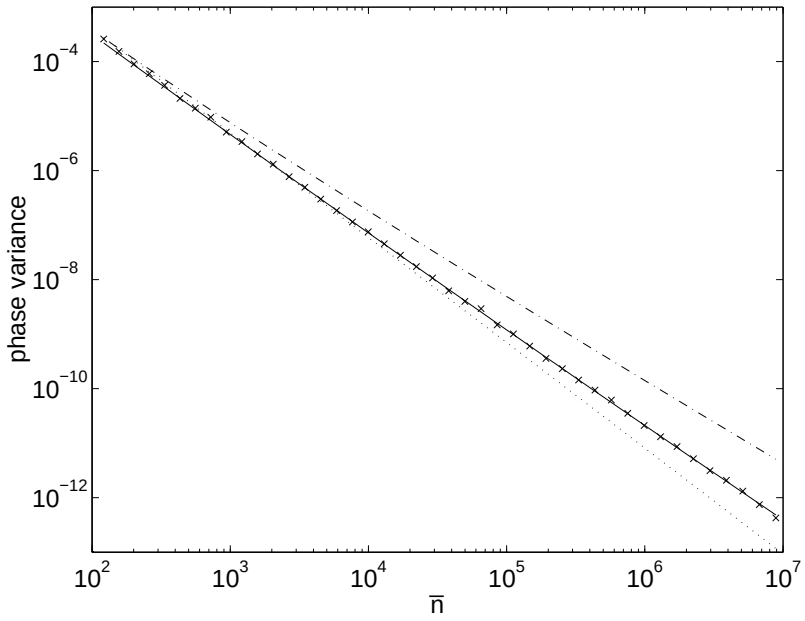

Figure 3.16: The phase variance for $\arg(C_v^{1-\varepsilon} A_v^{\varepsilon})$ feedback using the numerically determined optimum values of ε. The numerical results are shown as crosses and the continuous line is that fitted to the data. The mark II phase variance (dash-dotted line) and theoretical limit (dotted line) are also shown.

used in order to obtain more accurate results. In addition, only data points for photon numbers above about 10000 were used, as the data for small photon numbers gave a poor approximation of the asymptotic result. It was found that the best fit was for $c = 0.095 \pm 0.008$ and $p = 1.685 \pm 0.007$. The data and the fitted line along with the mark II case and the theoretical limit are shown in Fig. 3.16. As can be seen, the power law is a very good fit for the data.

Note that, because the value of ε used depends on the photon number, this phase measurement scheme can not be described by a single POM. Instead, the actual POM and H matrix will depend on the photon number. Nevertheless, it is reasonable to consider these measurements as approximately equivalent to a measurement scheme with a single POM, and the variation of $h(n)$ as given above.

3.7 Time Dependent ε

3.7.1 Simple Method

In order to improve on this result, the remaining alternative is to vary ε during the measurement. As the approximate theory above does not give any information about how ε should be varied during the measurement, trial and error was used initially. It was found that an improvement over the constant ε case could be obtained if ε was increased linearly during the measurement. Even better improvements were obtained when ε was increased proportionally to v^2 or v^3, or some higher power of time.

A possible reason why ε should be increased during the measurement can be deduced from Eq. (3.28). In the case that we are considering a squeezed state, rather than a coherent state, the phase of α_v is now time dependent. The phase estimate $\arg C_v$ gives an estimate of the initial phase, and takes no account of the variation of the phase of the running value α_v. This means that when the variance of $\arg \alpha_v$ is large, the intermediate phase estimate may not be between the phase of α_v and A_v, even though it is between the phases of C_v and A_v.

If the variance of the phase estimate is kept sufficiently above the variance of $\arg \alpha_v$ during the measurement, then this problem should be corrected. Therefore a statistical phase feedback scheme

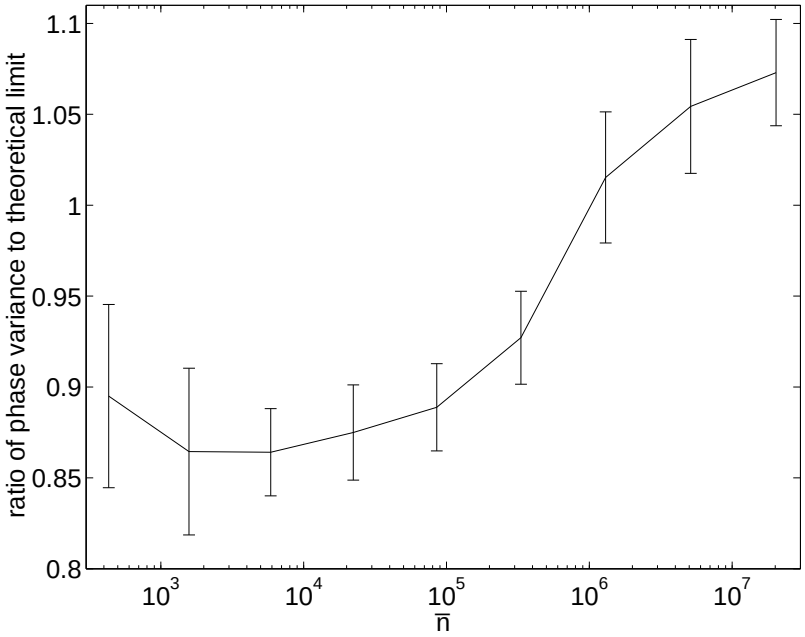

Figure 3.17: The phase variance as a ratio to the theoretical limit for the feedback scheme where ε is chosen such that the variance of the intermediate phase estimate is no less than a fixed multiple of the variance of $\arg \alpha_v$.

was used, where at each time the value of ε was chosen such that the variance in the phase estimate was at or below some fixed multiple of the variance of $\arg \alpha_v$. These variances were determined numerically from a large number (2^{12}) of simultaneous calculations. The best multiple to use cannot be predicted, and was determined numerically from these samples.

The results obtained using this method are shown as a ratio to the theoretical limit of twice the phase variance of an optimised squeezed state in Fig. 3.17. The results are far better than for the case where ε is kept constant, with the phase variance being below or close to the theoretical limit for all the photon numbers tested. Even for the largest photon number it was feasible to perform calculations for the phase variance is less than 8% above the theoretical limit.

The optimum limiting ratio between the phase variances turns out to be different for the different photon numbers. The optimum values obtained are shown in Fig. 3.18. The minimum value is around 2.5, which is about what would be expected in order to prevent the phase estimates being outside the interval between $\arg \alpha_v$ and $\arg A_v$. The optimum limiting ratio can be much higher, however, and generally increases with photon number.

Another factor is that initially $\arg C_v$ will be a poor phase estimate, and therefore the phase estimate may not be between the phase of α_v and A_v if ε is too small. This would seem to indicate that ε should be large initially to take account of this. It was found that this in fact made the measurement poorer. This indicates that the situation is not as simple as indicated in the analysis above.

This phase measurement scheme, although it provides measurements at or very close to the theoretical limit, is unsatisfactory because the values of ε are determined statistically from a large number of measurements, rather than determined from the individual values of A_v, B_v and v. It is possible to determine the variation of ε numerically in this way for a particular photon number, and use this variation to obtain results close to the theoretical limit for independent measurements. This requires knowledge of the photon number, however, and the variation takes a long time to calculate.

In order to determine the values of ε to use in a simple way, we can consider the approximate variation of the variance of $\arg \alpha_v$. Taking the system phase to be zero, and making the approximations

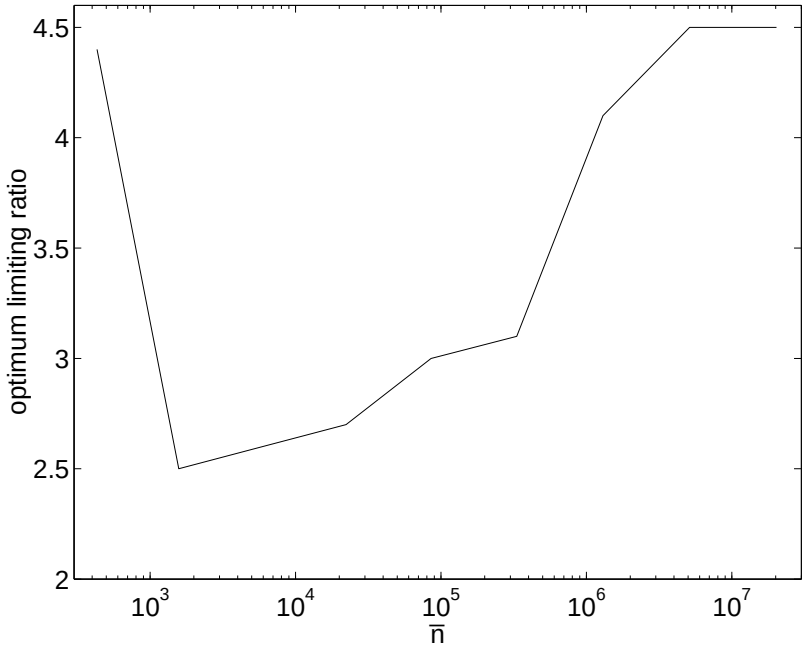

Figure 3.18: The optimum limiting ratio between the variance of the phase estimate and the variance of $\arg \alpha_v$.

that $e^{i\Phi} \approx i$ and B_v^{S} is approximately real, Eq. (3.141) simplifies to

$$d\alpha_v \approx -\frac{i}{1-v}\frac{B_v^{\mathrm{S}} dW}{1+B_v^{\mathrm{S}}}. \tag{3.167}$$

Using the approximation $B_v^{\mathrm{S}} \approx 1$, this becomes

$$d\alpha_v \approx -\frac{i}{1-v}\frac{dW}{2}. \tag{3.168}$$

Using this to determine the variation in the phase of α_v gives

$$\begin{aligned} d\arg\alpha_v &= \mathrm{Im}[d\log\alpha_v] \\ &= \mathrm{Im}\left[\frac{d\alpha_v}{\alpha_v} - \frac{(d\alpha_v)^2}{2\alpha_v^2}\right] \\ &\approx \mathrm{Im}\left[\frac{\frac{-i}{1-v}\frac{dW}{2}}{\alpha_v} + \frac{\frac{1}{(1-v)^2}\frac{dv}{4}}{2\alpha_v^2}\right]. \end{aligned} \tag{3.169}$$

Assuming that α_v remains close to its initial value α, which is real, this simplifies to

$$d\arg\alpha_v \approx \frac{-dW}{2(1-v)\alpha}. \tag{3.170}$$

Using this, the increment in the expectation value of $(\arg\alpha_v)^2$ is

$$d\left\langle(\arg\alpha_v)^2\right\rangle \approx \frac{dv}{4\alpha^2}\frac{1}{(1-v)^2}. \tag{3.171}$$

This gives us

$$\left\langle(\arg\alpha_v)^2\right\rangle \approx \frac{1}{4\alpha^2}\frac{1}{(1-v)}. \tag{3.172}$$

During the course of the measurement we wish to keep the variance of the phase estimate at some constant, k, times the variance of the phase of α_v. Therefore the variance in the phase estimate should be

$$\langle \hat{\varphi}_v^2 \rangle \approx \frac{k}{4\alpha^2} \frac{1}{(1-v)}. \tag{3.173}$$

As the simplest approximation, the variance of $\arg A_v$ can be assumed to be the same as for the mark I case, where it is $1/(4\alpha\sqrt{v})$. Approximating the phase estimate as $\varepsilon(v) \arg A_v$, this gives

$$\varepsilon(v)^2 \frac{1}{4\alpha\sqrt{v}} \approx \frac{k}{4\alpha^2} \frac{1}{(1-v)}. \tag{3.174}$$

This implies that the value of $\varepsilon(v)$ should be

$$\varepsilon(v) \approx \sqrt{\frac{k}{\alpha}} \sqrt{\frac{\sqrt{v}}{1-v}}. \tag{3.175}$$

This approximation is a bit too simplistic, as the variance of $\arg A_v$ increases as ε is decreased. Alternatively, using the result given in Eq. (3.41), we find

$$\varepsilon(v) \approx \left(\frac{k}{\alpha}\right)^2 \frac{v}{(1-v)^2}. \tag{3.176}$$

This result is not necessarily any more accurate than the previous one, however, as the solution (3.41) is only accurate for constant ε.

The full differential equations taking into account the time dependent ε are extremely difficult to solve, and do not appear to have a simple solution of the above form. Therefore we will instead consider the numerical results for the full calculation. The time dependence of $\varepsilon(v)$ for a mean photon number of about 22255 is shown in Fig. 3.19. The analytic approximation shown was found by trial and error, and is

$$\varepsilon(v) = \frac{1}{350} \frac{\sqrt{v}}{(1-v)^{0.8}}. \tag{3.177}$$

The dividing factor of 350 here is around twice the value of α.

The common features of the above expressions for ε are a factor of v to some power in the numerator, and factors of $1-v$ and α to some power in the denominator. Various expressions with these features were tested, but the one that gave the best results was

$$\varepsilon(v) = \frac{1}{\alpha} \sqrt{\frac{v}{1-v}}. \tag{3.178}$$

Note that this is very similar to the above analytic approximation for the exact results, except that $1-v$ is taken to the power of 0.5 rather than 0.8 (as well as a difference of a multiplicative factor).

This expression suffers from the slight drawback that it is dependent on the value of α, rather than the experimentally measured quantities. This is easily corrected by using the slightly modified expression,

$$\varepsilon(v) = \frac{v^2 - |B_v|^2}{|C_v|} \sqrt{\frac{v}{1-v}}. \tag{3.179}$$

Here $|C_v|/(v^2 - |B_v|^2)$ is used as an estimator for α.

The results for this method are shown in Fig. 3.20 as a ratio to the theoretical limit. As this figure shows, the results are very close to the theoretical limit, and even for the largest photon number for which calculations have been performed the phase uncertainty is only about 4% above the theoretical limit. For these results 2^{11} samples were used. This sample size was used for the rest of the results in this chapter, unless otherwise stated.

If the integration time step is reduced, while keeping the time interval at which the phase estimates are updated constant, the phase variance converges. If, however, the phase estimates are

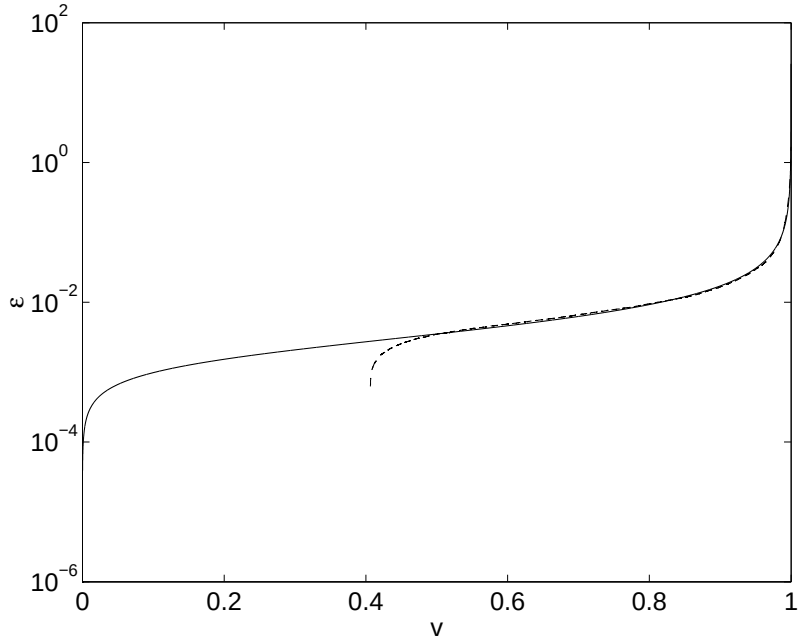

Figure 3.19: The values of ε found for a mean photon number of 22255 and a limiting ratio of 2.7. The dashed line is the numerical results and the continuous line is an analytic approximation.

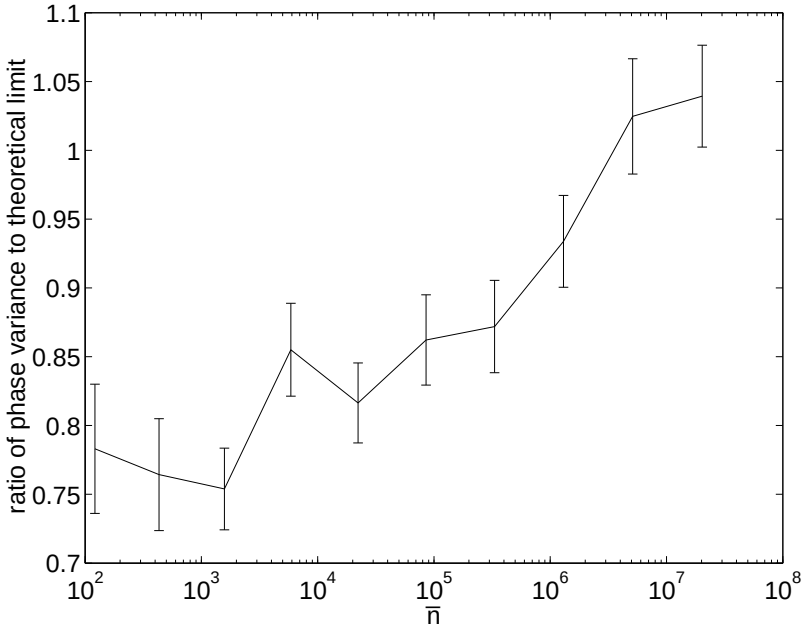

Figure 3.20: Variance for phase measurements with a time dependent ε given by Eq. (3.179) plotted as a function of the photon number of the input state. The phase variance is plotted as a ratio to the theoretical minimum phase variance (i.e. twice the intrinsic phase variance).

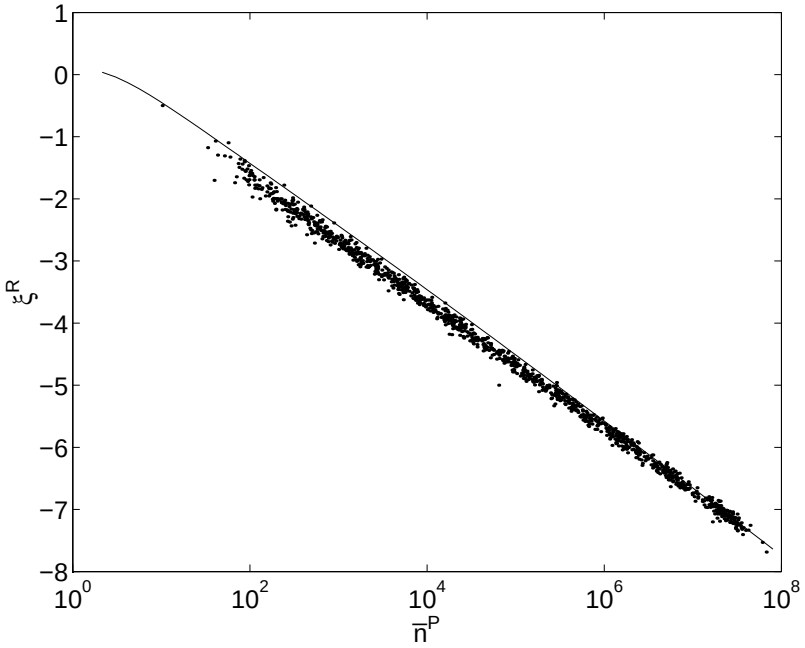

Figure 3.21: Values of ξ^{R} and $\overline{n}^{\mathrm{P}}$ (calculated from A and B) resulting from measurements on squeezed states of various mean photon numbers. The variation of ξ^{P} with $\overline{n}^{\mathrm{P}}$ for optimum squeezed states is also plotted (continuous line).

updated at smaller and smaller time intervals then the phase variance does not converge. For example the phase uncertainty for measurements on an optimised squeezed state with a photon number of 1577 is 1.54×10^{-6} if we use the time steps given above. If we use time steps that are a hundred times smaller, then the phase variance is 1.93×10^{-6}, and if the time steps are a thousand times smaller the phase variance is 2.13×10^{-6}. These results indicate that the phase estimates must be incremented in finite time intervals for this method to give good results, and the size of the time steps that should be used depends on the photon number. The phase variance is not strongly dependent on these time steps, however, and only an order of magnitude estimate of the photon number would be required.

3.7.2 Evaluation of Method

A problem with determining the phase variance by the method above is that for highly squeezed states (that are close to optimised for minimum phase variance), a significant contribution to the phase variance is from low probability results around π. In obtaining numerical results the actual phase variance for the measurement will tend to be underestimated because the results from around π are obtained too rarely for good statistics. It would require an extremely large number of samples to estimate this contribution. However, we can estimate it non-statistically as follows.

Recall that in order to have a measurement that is close to optimum, the multiplying factor $Q'(\overline{n}^{\mathrm{P}}, \xi^{\mathrm{P}})$ should give values of ξ^{P} for each $\overline{n}^{\mathrm{P}}$ that are close to optimised for minimum phase uncertainty. To test this for the phase measurement scheme described above, the $\overline{n}^{\mathrm{P}}$ and ξ^{P} were determined from the values of A and B from the samples. The resulting data along with the line for optimum ξ^{P} are plotted in Fig. 3.21. The imaginary part of ξ^{P} should be zero for optimum measurements, and is small for these results. Therefore in Fig. 3.21 the real part, ξ^{R}, is plotted. As can be seen, the vast majority of the data points are below the line, indicating greater squeezing than optimum. (There are more points above the line for large $\overline{n}$; more will be said about this later.) This means that if the low probability results around π are taken into account the phase variance for these measurements will be above the theoretical limit.

I will firstly consider the effect of variations in the modulus of ξ^{P}, leaving consideration of error in the phase till later. In order to estimate how far above the theoretical limit the actual phase variance is, we can make a quadratic approximation to the expression for the phase variance. From [45] the expression for the intrinsic phase variance of a squeezed state is, for real ζ,

$$V(\phi) \approx \frac{e^{2\zeta}}{4\bar{n}} + \frac{1}{4\bar{n}^2} + 2\mathrm{erfc}(\sqrt{2\bar{n}}e^{\zeta}). \tag{3.180}$$

Taking the derivative with respect to ζ gives

$$\frac{d}{d\zeta}V(\phi) \approx \frac{e^{2\zeta}}{2\overline{n}} - 4e^{\zeta}\sqrt{\frac{2\overline{n}}{\pi}}e^{-2\overline{n}e^{2\zeta}}. \tag{3.181}$$

Taking the second derivative we obtain

$$\frac{d^2}{d\zeta^2}V(\phi) \approx \frac{e^{2\zeta}}{\overline{n}} - 4e^{\zeta}\sqrt{\frac{2\bar{n}}{\pi}}e^{-2\bar{n}e^{2\zeta}} - 4e^{\zeta}\sqrt{\frac{2\bar{n}}{\pi}}(-4\bar{n}e^{2\zeta})e^{-2\bar{n}e^{2\zeta}}. \tag{3.182}$$

Using the fact that the first derivative (3.181) is zero for minimum phase variance gives

$$4e^{\zeta}\sqrt{\frac{2\overline{n}}{\pi}}e^{-2\overline{n}e^{2\zeta}} = \frac{e^{2\zeta}}{2\overline{n}}. \tag{3.183}$$

Using this result, the expression for the second derivative simplifies to

$$\begin{aligned}\frac{d^2}{d\zeta^2}V(\phi) &\approx \frac{e^{2\zeta}}{\overline{n}} - \frac{e^{2\zeta}}{2\overline{n}} + 4\bar{n}e^{2\zeta}\frac{e^{2\zeta}}{2\overline{n}} \\ &= \frac{e^{2\zeta}}{2\overline{n}} + 2e^{4\zeta} \\ &= \frac{n_0}{2\bar{n}^2}\left(1+4n_0\right).\end{aligned} \tag{3.184}$$

This means that for values of ζ close to optimum the increase in the phase variance over the optimum value is

$$\Delta V(\phi) \approx (\Delta|\zeta|)^2\frac{n_0}{4\overline{n}^2}(1+4n_0). \tag{3.185}$$

Here the absolute value of ζ has been used for greater generality.

The main contribution to the phase uncertainty is $n_0/(4\bar{n}^2)$, so the increase in the phase uncertainty as a ratio to the minimum phase uncertainty is approximately

$$\frac{\Delta V(\phi)}{V_{\min}(\phi)} \approx (\Delta|\zeta|)^2(1+4n_0). \tag{3.186}$$

Note that as the photon number is increased, n_0 increases roughly as $\log\overline{n}$. This means that in order for the percentage increase in the excess phase uncertainty to remain limited, the error in $|\zeta|$ must decrease with photon number.

Now I will use the superscript P to indicate specifically the squeezed state in the POM. Since $|\zeta^{\mathrm{P}}| = |\xi^{\mathrm{P}}|$, we obtain

$$\Delta V(\phi) \approx (\Delta|\xi^{\mathrm{P}}|)^2\frac{n_0}{4\overline{n}^2}(1+4n_0). \tag{3.187}$$

Here I have used the values of n_0 and $\overline{n}$ for the input state, rather than the squeezed state in the POM. As the input states considered here are optimum squeezed states, the average values of these quantities for the squeezed state in the POM should be close to those for the input state. For a more general input state we would still use the value of $\overline{n}$ for the input state, but the value of n_0 should be chosen as the corresponding value for an optimum squeezed state with mean photon number $\overline{n}$.

This estimate indicates that the actual phase variance for the measurement scheme described above can be significantly larger than the intrinsic phase variance. For example, for a mean photon

number of about 332000, the rms deviation of $|\xi^{\mathrm{P}}|$ from the optimum value is only about 0.16, but a squeezed state with $|\xi^{\mathrm{P}}|$ differing this much from optimum will have a phase variance more than twice the optimum value. This indicates that if the low probability results around π are taken into account, the introduced phase variance is actually more than twice the theoretical limit.

Now I will estimate the contribution from error in the phase (rather than the modulus) of ξ^{P}. Recall that in attempting to estimate the phase of a state, we measure quadratures close to $\pi/2$. For the $\pi/2$ quadrature

$$\hat{X}_{\pi/2} = -i(a - a^{\dagger}), \tag{3.188}$$

the expectation value is

$$\left\langle \hat{X}_{\pi/2} \right\rangle = 2\,|\alpha| \sin\varphi. \tag{3.189}$$

For small φ, this means that

$$\left\langle \hat{X}_{\pi/2} \right\rangle \approx 2\,|\alpha|\,\varphi. \tag{3.190}$$

Canonical measurements can be considered by assuming that α is known, and that the $\pi/2$ quadrature is proportional to the phase. This would mean that a measurement of $\hat{X}_{\pi/2}$ is equivalent to a direct measurement of the phase.

For simplicity the system phase will now be taken to be zero. The estimated phase is then proportional to the measured value of $\hat{X}_{\pi/2}$, and the variance in the phase estimate will be proportional to the variance in the $\pi/2$ quadrature:

$$\left\langle \Delta \hat{X}^2_{\pi/2} \right\rangle \approx 4\overline{n}\left\langle \Delta\phi^2 \right\rangle. \tag{3.191}$$

Here the approximation $\overline{n} \approx |\alpha|^2$ has been used. This is reasonably accurate for the states that are considered here. Now recall that the uncertainty in this quadrature for a squeezed state is

$$\left\langle \Delta \hat{X}^2_{\pi/2} \right\rangle = e^{2r}\cos^2\tfrac{1}{2}\phi_\zeta + e^{-2r}\sin^2\tfrac{1}{2}\phi_\zeta, \tag{3.192}$$

where r and ϕ_ζ are the magnitude and phase of ζ.

Therefore the uncertainty in the phase is

$$\left\langle \Delta\phi^2 \right\rangle = \frac{e^{2r}\cos^2\tfrac{1}{2}\phi_\zeta + e^{-2r}\sin^2\tfrac{1}{2}\phi_\zeta}{4\overline{n}}, \tag{3.193}$$

If the phase of ζ is close to π (so ζ is close to negative real), we can make the approximation

$$\left\langle \Delta\phi^2 \right\rangle \approx \frac{e^{-2r} + e^{2r\frac{(\Delta\phi_\zeta)^2}{4}}}{4\overline{n}}, \tag{3.194}$$

where $\Delta\phi_\zeta = \phi_\zeta - \pi$. The first term in the numerator is identical to the first term for the intrinsic phase variance of a squeezed state. Clearly the second term is the excess phase variance due to the error in the phase of ζ. Therefore the extra phase variance due to error in the phase of ζ is given by

$$\Delta V(\phi) \approx \frac{(\Delta \arg\zeta)^2}{16 n_0}. \tag{3.195}$$

Note that the error in the phase of ξ is equivalent to the error in the phase of ζ above, as the above case is for a system phase of zero. Now using the superscript P to indicate specifically the squeezed state in the POM, we have

$$\Delta V(\phi) \approx \frac{(\Delta \arg\xi^{\mathrm{P}})^2}{16 n_0}. \tag{3.196}$$

Here we are again using the value of n_0 from the input state. Using this estimate on the example used for the magnitude, with $\overline{n} \approx 332000$, it can be seen that this is not so much of a problem, with the introduced phase uncertainty being increased by less than 3% by this factor.

3.7.3 Corrected Method

The problem of the large contribution from low probability results around π can be effectively eliminated by using corrections near the end of the measurement. In order to describe this we must firstly consider the values of α^{P} and ζ^{P} for intermediate times. From Ref. [33], the POM for intermediate times is

$$F_v(A_v,B_v)=Q_v(A_v,B_v)\exp(\tfrac{1}{2}B_v{a^\dagger}^2+A_va^\dagger)(1-v)^{a^\dagger a}\exp(\tfrac{1}{2}B_v^*a^2+A_v^*a). \tag{3.197}$$

This is mixed, not pure, and it is therefore not possible to simplify it to an expression in terms of squeezed states, as is the case at the end of the measurement. I will therefore define the values of α_v^{P} and ζ_v^{P} by analogy with the definitions at the end of the measurement, but these will not actually correspond to squeezing parameters for a squeezed state in the POM.

Recall from the introduction that for a coherent state we have

$$vA_v+B_vA_v^*=\alpha\left(v^2-|B_v|^2\right)+i\left(v\sigma_v-B_v\sigma_v^*\right), \tag{3.198}$$

so

$$\alpha=\frac{vA_v+B_vA_v^*}{v^2-|B_v|^2}-\frac{i\left(v\sigma_v-B_v\sigma_v^*\right)}{v^2-|B_v|^2}. \tag{3.199}$$

This means that $C_v/(v^2-|B_v|^2)$ can be used as an estimator for α. The situation is more complicated for squeezed states, as there are extra terms due to the stochastic variation of α. It is still possible to use $C_v/(v^2-|B_v|^2)$ as an estimator for α, however. As α^{P} is defined by the value of $C_v/(v^2-|B_v|^2)$ for $v=1$, it is therefore reasonable to define intermediate values α_v^{P} by

$$\alpha_v^{\mathrm{P}}=\frac{A_vv+B_vA_v^*}{v^2-|B_v|^2}. \tag{3.200}$$

For the case of ζ_v^{P}, note that ζ^{P} is very large when all the feedback phases are close to each other, and B is correspondingly close to 1. In order for ζ_v^{P} to have the same property, the argument of the atanh function should be close to 1 when the phase estimates are very close together. As the maximum value B_v can have is v, this will be the case if B_v is divided by v. Using this, the definition of ζ_v^{P} is

$$\zeta_v^{\mathrm{P}}=-\frac{B_v}{|B_v|}\mathrm{atanh}\left(\frac{|B_v|}{v}\right). \tag{3.201}$$

Note that this relation of α_v^{P} and ζ_v^{P} to A_v and B_v is *not* analogous to the relation of α_v and ζ_v to A_v^{S} and B_v^{S}. The quantities A_v^{S} and B_v^{S} are defined by analogy to A_v and B_v at time $v=1$. The time varying values of A_v^{S} and B_v^{S} are found by using the same definition with the time varying squeezing parameters.

Near the end of the measurement, at each time step the photon number is estimated from the values of A_v and B_v. The estimator used is

$$|\alpha_v^{\mathrm{P}}|^2+\sinh^2|\zeta_v^{\mathrm{P}}|. \tag{3.202}$$

I will call this estimator $\overline{n}_v^{\mathrm{P}}$, so that this variable is analogous to $\overline{n}^{\mathrm{P}}$. The optimum value of ξ_v^{P} is then estimated using an asymptotic formula based on the result in [45],

$$\xi_v^{\mathrm{opt}}=-\tfrac{1}{2}\log\left(\frac{\overline{n}_v^{\mathrm{P}}}{\log\overline{n}_v^{\mathrm{P}}+\Delta'}\right), \tag{3.203}$$

where

$$\Delta'=2\log 2-\tfrac{1}{4}\log 2\pi\approx 0.926825. \tag{3.204}$$

Note that this Δ' differs from the Δ defined in Eq. (2.136) of Sec. 2.3.1 by 1.5. The optimum value of ζ_v^{P} will be complex, with a phase dependent on the phase of α_v^{P}.

If ξ_v^{R} (the real part of ξ_v^{P}) is too far below this optimum value, rather than using the feedback phase of Sec. 3.7.1, the feedback phase used is

$$\Phi(v) = \frac{1}{2}\arg\left[B_v - v\frac{C_v}{C_v^*}\tanh\left|\xi_v^{\mathrm{opt}}\right|\right]. \tag{3.205}$$

Using this feedback phase takes B_v directly towards the optimum value. To see this, note firstly that the optimum value of B_v is

$$B_v^{\mathrm{opt}} = v\frac{C_v}{C_v^*}\tanh\left|\xi_v^{\mathrm{opt}}\right|. \tag{3.206}$$

This can be seen by inverting the definition of ζ_v^{P}. The factor of C_v/C_v^* makes the phase of this zero relative to α_v^{P}.

If we then take the exponential of the feedback phase, we find

$$e^{2i\Phi} \propto B_v - B_v^{\mathrm{opt}}, \tag{3.207}$$

so

$$dB_v \propto B_v^{\mathrm{opt}} - B_v. \tag{3.208}$$

Here the proportionality symbol indicates that the phase is equal, but not necessarily the magnitude. This demonstrates that this feedback phase takes the value of B_v directly towards the optimum value.

The details of exactly when ξ_v^{R} is considered too far below optimum can be varied endlessly, but for the results that will be presented here this alternative phase estimate is used after time $v = 0.9$ and when

$$|\xi_v^{\mathrm{R}}| > |\xi_v^{\mathrm{opt}}|e^{\lambda|\alpha_v^{\mathrm{P}}|^2(1-v)}. \tag{3.209}$$

Using the exponential multiplying factor means that this alternative feedback is only used towards the end of the measurement. Only considering the alternative feedback in the last 10% of the measurement is necessary for the smaller photon numbers, where Eq. (3.209) is too weak a restriction. In Ref. [51] the feedback phase

$$\Phi(v) = \frac{1}{2}\arg\left[\frac{B_v}{|B_v|} - \frac{C_v}{C_v^*}\right], \tag{3.210}$$

was also mentioned. It turns out that this is not necessary, and very good results can be obtained by using (3.205) alone.

It was found that good results were obtained for a wide range of photon numbers when the value of λ used was 10^{-3}. It is also possible to adjust the value of λ individually for the different photon numbers; however, this only gives marginal improvements. The results for this value of λ are shown in Fig. 3.22. The contribution due to the error in the phase of ξ^{P} is small for the entire range considered, below 3%. The contribution due to error in the magnitude of ξ^{P} is even smaller for moderate photon numbers, but for the largest photon numbers it rises dramatically.

The reason for this rise is that the above correction only corrects for values of ξ_v^{R} that are below optimum, and for the larger photon numbers many of the uncorrected values of ξ^{R} are above optimum (see Fig. 3.21). In order to make the corrections work well, a dividing factor can be used to bring the uncorrected values below the line. For the second largest photon number tested of around 5×10^6, the best results were obtained when the values of ε as given by (3.179) were divided by 1.1. For the largest mean photon number tested, 2×10^7, the best results were obtained for a dividing factor of 1.2. The value of λ that gave the best results with these dividing factors was 5×10^{-4}.

The results using these dividing factors and altered value of λ for the larger photon numbers are shown in Fig. 3.23. In this case, the contribution due to error in the magnitude of ξ^{P} remains small, around 1.5%, for the largest photon numbers. The contribution due to the error in the phase of ξ^{P} does go up slightly for the largest photon numbers, but it is still well below 3%. The total excess phase variance due to error in ξ_v^{P} does not exceed about 5% for the entire range of photon numbers considered.

The indication is that the optimum dividing factor will continue to increase for photon numbers beyond the maximum for which calculations were performed. Unfortunately, using these dividing

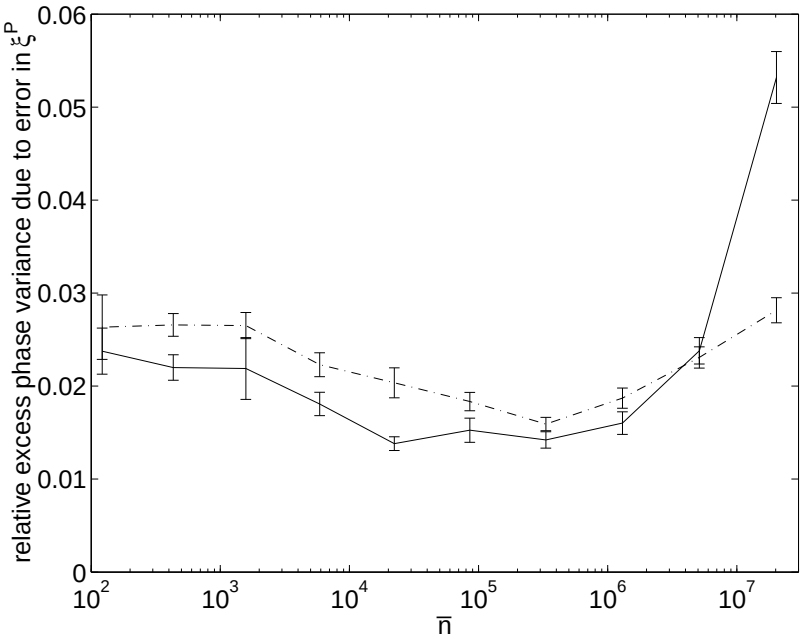

Figure 3.22: Contributions to the phase uncertainty from error in the magnitude of ξ^{P} (continuous line) and the phase of ξ^{P} (dash-dotted line). No dividing factors are used, and $\lambda = 10^{-3}$. The contributions are plotted as a ratio to the theoretical minimum introduced phase uncertainty.

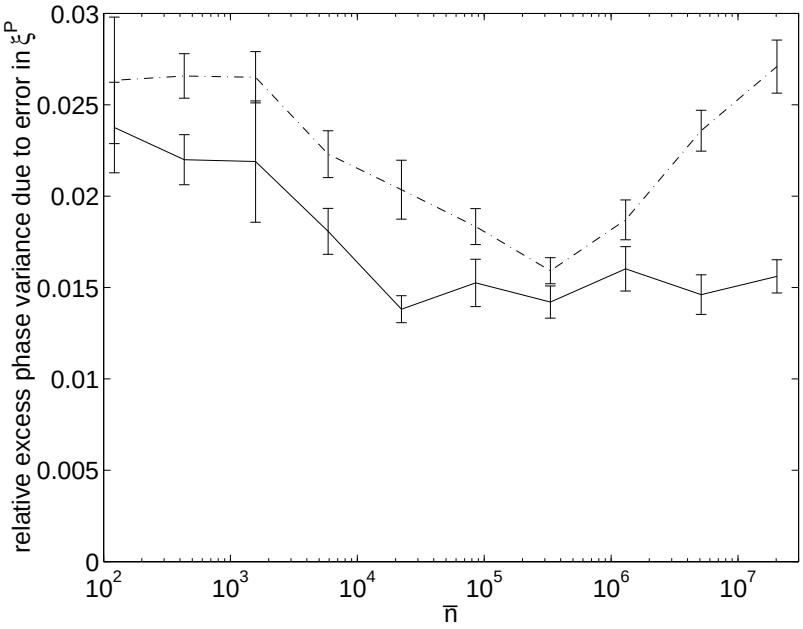

Figure 3.23: Contributions to the phase uncertainty from error in the magnitude of ξ^{P} (continuous line) and the phase of ξ^{P} (dash-dotted line). Dividing factors of 1.1 and 1.2 are used for photon numbers of 5×10^6 and 2×10^7, respectively. For these photon numbers $\lambda = 5 \times 10^{-4}$, otherwise $\lambda = 10^{-3}$. The contributions are plotted as a ratio to the theoretical minimum introduced phase uncertainty.

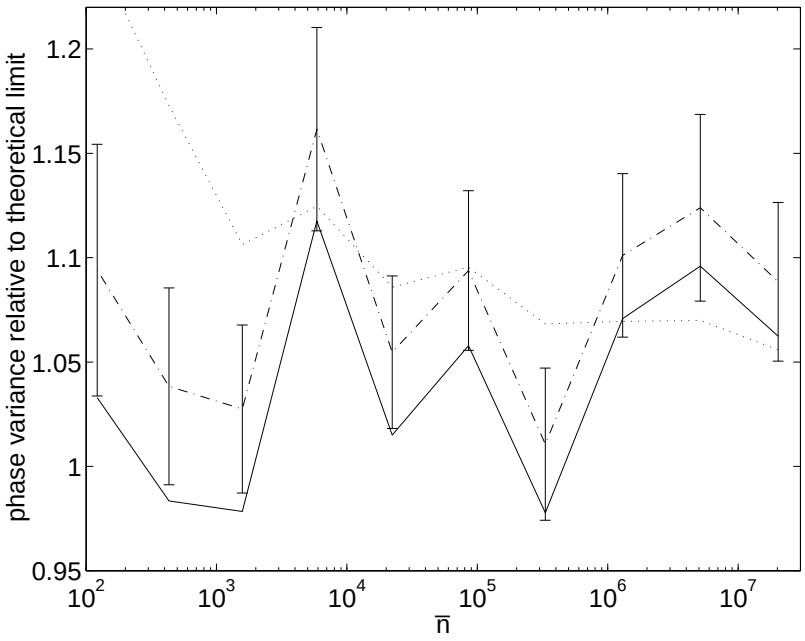

Figure 3.24: The phase variance as estimated using the phase data as a ratio to the theoretical limit. The continuous line is just the variance of the phase data, the dash-dotted line is the variance corrected for the low probability results around $\pm\pi$, and the dotted line is the corrected theoretical limit taking into account the variation in $\overline{n}^{\mathrm{P}}$.

factors means that the photon number must be known beforehand. Nevertheless, as the dividing factor required increases only very slowly with the photon number, only a rough, order of magnitude, estimate of the photon number is required.

With this modified technique the phase variance again does not converge as the feedback phase is updated in smaller and smaller time intervals. The phase variance is less dependent on the time step with this technique, however. For example, for a mean photon number of 1577 the total phase variance for measurements on an optimised squeezed state only increases by about 9% as the time steps are reduced by a factor of 1000. In contrast, the phase variance increases by a factor of 38% for the uncorrected technique.

As an alternative way of evaluating the results we can again consider the variance of the phase estimates obtained, as shown in Fig. 3.24. In order to take account of the low probability phase results with large error, we can add a factor of $1/(4\overline{n})$. The reason for this correction is that, from Ref. [45], the term $2\mathrm{erfc}(\sqrt{2n_0})$ is the contribution due to results with large error. From Eq. (2.134), for squeezed states near the theoretical limit this term becomes approximately $1/(8\overline{n})$. We must add twice this, as phase measurements near the theoretical limit on optimally squeezed states have a variance twice the intrinsic variance of the squeezed state.

As can be seen in Fig. 3.24, when this correction is included the results are noticeably above the theoretical limit. The results are on average around 10% above the theoretical limit, which is slightly more than would be expected from the previous analysis. The problem appears to be that the results where $\overline{n}^{\mathrm{P}}$ is small have disproportionately high errors. Even though $\langle \overline{n}^{\mathrm{P}} \rangle \approx \overline{n}$, we will find that

$$\left\langle \frac{\log \overline{n}^{\mathrm{P}} + \Delta}{4\overline{n}^{\mathrm{P}\,2}} \right\rangle > \frac{\log \overline{n} + \Delta}{4\overline{n}^2}. \tag{3.211}$$

This means that the mean phase variance for the states $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$ will be higher than that for the state $|\alpha, \zeta\rangle$, even if they are minimum phase uncertainty squeezed states.

This seems to be an intrinsic problem with these type of measurements, as any state that has small phase variance will have large uncertainty in the photon number. It is therefore reasonable to

postulate that it is not possible to reduce the introduced phase variance below the mean value of $(\log \overline{n}^{\mathrm{P}} + \Delta)/(4\overline{n}^{\mathrm{P}\,2})$, and that this is therefore the actual limit.

When the theoretical limit to the introduced phase variance is corrected based on the values obtained for $\overline{n}^{\mathrm{P}}$, the theoretical limit to the total phase variance (i.e. the intrinsic phase variance plus the limit to the introduced phase variance), is as shown in Fig. 3.24. The phase variance performs much better when compared with this corrected theoretical limit, and even for the largest photon numbers is no more than about 5% above it.

3.8 Beyond the Theoretical Limit

The final question that will be addressed in this chapter is whether it is possible, in some circumstances, to reduce the introduced phase variance below the theoretical limit. Recall that the theoretical limit is based on the probability distribution for A and B being given by

$$P(A,B) \propto \left|\langle\psi|\alpha^{\mathrm{P}},\zeta^{\mathrm{P}}\rangle\right|^2. \tag{3.212}$$

If the phase estimate that is used is $\arg(A+BA^*)$, then this implies that the introduced variance in this phase estimate is the intrinsic phase variance of the state $|\alpha^{\mathrm{P}},\zeta^{\mathrm{P}}\rangle$. If the phase estimate used is *not* $\arg(A+BA^*)$, however, this limit does not apply.

The main example of this is homodyne measurements. For homodyne measurements, we find that $A+BA^*=0$, so it is not possible to use the phase estimate $\arg(A+BA^*)$. Instead, we must use a phase estimate that relies on prior knowledge of the state. For a homodyne measurement the local oscillator phase is equal to the system phase plus $\pi/2$. We do not need to consider the photocurrent as a function of time $I(t)$, as the measurement is essentially just a measurement of the phase quadrature $\hat{X}_\Phi$, where

$$\hat{X}_\Phi = ae^{-i\Phi} + a^\dagger e^{i\Phi}. \tag{3.213}$$

For arbitrary system phase, the expectation value of the $\pi/2$ quadrature is

$$\left\langle \hat{X}_{\pi/2} \right\rangle = 2\,|\alpha|\sin\varphi. \tag{3.214}$$

This means that, provided $|\alpha|$ is known, we can use

$$\phi = \mathrm{asin}\left(\frac{X_{\pi/2}}{2\,|\alpha|}\right) \tag{3.215}$$

as a phase estimate. As the asin function is very linear near 0, this means that these measurements are extremely close to being direct measurements of the phase.

3.8.1 Fitted Phase Estimates

For the case of adaptive phase measurements, we can consider phase estimates that are based on fitting the phase to the data. The introduced phase uncertainty can then be estimated using the techniques of data analysis. If we consider the photocurrent for a coherent state over a small but finite time interval δv, we have

$$I(v)\delta v = 2\,|\alpha|\cos(\varphi-\Phi)\delta v + \delta W(v), \tag{3.216}$$

where φ is the system phase, and Φ is the local oscillator phase. This is equivalent to a series of data points $y_i = I(v_i)\delta v$, where

$$\langle y_i\rangle = 2\,|\alpha|\cos(\varphi-\Phi)\delta v. \tag{3.217}$$

If the values of $|\alpha|$ and φ are assumed to be $|\alpha^f|$ and ϕ respectively, then the expectation values for the data points are

$$\xi_i = 2\left|\alpha^f\right|\cos(\phi-\Phi)\delta v. \tag{3.218}$$

I will firstly consider the case where both the magnitude and phase of α are fitted for, then the case where $|\alpha|$ is known, and only the phase needs to be fitted. To perform the fit, we wish to minimise

$$M(\alpha^f) = \sum_i (y_i - \xi_i)^2. \tag{3.219}$$

Expanding this gives

$$M(\alpha^f) = \sum_i \left[(I(v_i)\delta v)^2 - 4I(v) \left|\alpha^f\right| \cos(\phi - \Phi)\delta v^2 + 4 \left|\alpha^f\right|^2 \cos^2(\phi - \Phi)\delta v^2 \right]. \tag{3.220}$$

As the $(I(v_i)\delta v)^2$ term does not depend on the fitting values, it can be omitted. We can also remove a constant factor of $2\delta v$. Then taking the limit of small δv gives

$$M'(\alpha^f) = \int_0^1 \left[-2I(v) \left|\alpha^f\right| \cos(\phi - \Phi) + 2 \left|\alpha^f\right|^2 \cos^2(\phi - \Phi) \right] dv. \tag{3.221}$$

This can be simplified to an expression in terms of A and B:

$$M'(\alpha^f) = \mathrm{Re}\left[-2A^*\alpha^f + \left|\alpha^f\right|^2 - B^*(\alpha^f)^2 \right]. \tag{3.222}$$

As this should be minimised for the fitted value of α, the derivatives with respect to the magnitude and phase of α^f should be zero. Thus we have

$$\frac{\partial}{\partial \phi} M'(\alpha^f) = \mathrm{Im}\left[-2A^*\alpha^f - 2B^*(\alpha^f)^2 \right] = 0, \tag{3.223}$$

and

$$\frac{\partial}{\partial \left|\alpha^f\right|} M'(\alpha^f) = \mathrm{Re}\left[-2A^* e^{i\varphi_f} + 2\left|\alpha^f\right| - 2B^* \left|\alpha^f\right| e^{2i\varphi_f} \right] = 0. \tag{3.224}$$

The solution of these equations is given by

$$\alpha^f = \frac{A + BA^*}{1 - |B|^2}. \tag{3.225}$$

Thus we find that, if the coherent amplitude is unknown, fitting gives exactly the same phase estimate $\arg(A + BA^*)$ as has been used previously.

In order to find the uncertainty in the fitted values, we calculate the matrix C, where

$$C_{i1} = \frac{\partial \xi_i}{\partial \phi} \tag{3.226}$$

$$C_{i2} = \frac{\partial \xi_i}{\partial |\alpha^f|}. \tag{3.227}$$

The covariance matrix is then given by

$$\sigma^2 M_C^{-1}, \tag{3.228}$$

where $M_C = C^T C$, and σ^2 is the variance in the individual data points (which is δv in this case). Evaluating M_C gives

$$M_C = \begin{bmatrix} \sum 4|\alpha|^2 \sin^2(\varphi - \Phi)\delta v^2 & -\sum 4|\alpha| \sin(\varphi - \Phi)\cos(\varphi - \Phi)\delta v^2 \\ -\sum 4|\alpha| \sin(\varphi - \Phi)\cos(\varphi - \Phi)\delta v^2 & \sum 4\cos^2(\varphi - \Phi)\delta v^2 \end{bmatrix}. \tag{3.229}$$

In data analysis we would normally evaluate this using the fitted value of α rather than the actual value, as the actual value is unknown, and we wish to estimate the uncertainty in the fitted value. Here, however, we are interested in the variance in the phase estimates for a given value of α, and it is therefore more useful to express the result in terms of the actual value of α.

Taking the inverse to find the variance in the fitted phase, we find

$$\mathrm{var}(\phi) = \frac{\sum 4\cos^2(\varphi-\Phi)\delta v}{\left[\sum 4\left|\alpha\right|^2\sin^2(\varphi-\Phi)\delta v\right]\left[\sum 4\cos^2(\varphi-\Phi)\delta v\right] - \left[\sum 4\left|\alpha\right|\sin(\varphi-\Phi)\cos(\varphi-\Phi)\delta v\right]^2}, \tag{3.230}$$

Taking the limit of small δv, this becomes

$$\mathrm{var}(\phi) = \frac{1}{4\left|\alpha\right|^2}\frac{\int\cos^2(\varphi-\Phi)dv}{\left[\int\sin^2(\varphi-\Phi)dv\right]\left[\int\cos^2(\varphi-\Phi)dv\right] - \left[\int\sin(\varphi-\Phi)\cos(\varphi-\Phi)dv\right]^2}. \tag{3.231}$$

From this expression we can see that smaller phase variances will be obtained if $\sin(\varphi-\Phi)\cos(\varphi-\Phi)$ has a time averaged value close to zero. This is possible because $\sin(\varphi-\Phi)\cos(\varphi-\Phi)$ can take negative values. We will also obtain small phase variances if the time averaged value of $\sin^2(\varphi-\Phi)$ is close to 1. This means that the local oscillator phase should also be close to $\varphi+\pi/2$. Note that if we are using a local oscillator phase of $\hat{\varphi}+\pi/2$, using a better phase estimate $\hat{\varphi}$ does not necessarily result in a smaller variance, as we also want $\sin(\varphi-\Phi)\cos(\varphi-\Phi)$ to average to zero. This gives an alternative explanation of the result found before that using the best intermediate phase estimate does not result in the smallest phase variance.

In order to obtain better phase estimates, we can consider the case where $|\alpha|$ is known, and only the phase is fitted for. Then it is easy to see that we obtain Eq. (3.222), except the actual value of $|\alpha|$ is used:

$$M'(\phi) = \mathrm{Re}\left[-2A^*\left|\alpha\right|e^{i\phi} + \left|\alpha\right|^2 - B^*\left|\alpha\right|^2 e^{2i\phi}\right]. \tag{3.232}$$

As $|\alpha|^2$ does not depend on the fitted phase, it can be omitted, so this simplifies to

$$M'(\phi) = -\mathrm{Re}\left[2A^*\left|\alpha\right|e^{i\phi} + B^*\left|\alpha\right|^2 e^{2i\phi}\right]. \tag{3.233}$$

Note that this is equivalent to the result obtained in Eq. (21) of Ref. [36]. To have a minimum, we require

$$\frac{\partial}{\partial\phi}M'(\phi) = -\mathrm{Im}\left[2A^*\left|\alpha\right|e^{i\phi} + 2B^*\left|\alpha\right|^2 e^{2i\phi}\right] = 0. \tag{3.234}$$

Unlike the previous case there is no simple solution in terms of A and B. The solution to this can be found by finding the roots of the fourth order polynomial

$$\begin{aligned} P(\cos\phi) = {} & 4\left|\alpha\right|^2\left|B\right|^2\cos^4\phi + 4\left|\alpha\right|\mathrm{Re}(AB^*)\cos^3\phi + (\left|A\right|^2 - 4\left|\alpha\right|^2\left|B\right|^2)\cos^2\phi \\ & -2\left|\alpha\right|(\mathrm{Im}A\mathrm{Im}B + 2\mathrm{Re}A\mathrm{Re}B)\cos\phi + \left|\alpha\right|^2(\mathrm{Im}B)^2 - (\mathrm{Re}A)^2. \end{aligned} \tag{3.235}$$

The variance in the phase estimate, on the other hand, can be found far more easily. We simply have

$$M_C = \sum 4\left|\alpha\right|^2\sin^2(\varphi-\Phi)\delta v^2, \tag{3.236}$$

so the variance in the fitted phase is

$$\mathrm{var}(\phi) = \frac{1}{\sum 4\left|\alpha\right|^2\sin^2(\varphi-\Phi)\delta v}. \tag{3.237}$$

In the limit of small δv this becomes

$$\mathrm{var}(\phi) = \frac{1}{4\left|\alpha\right|^2\int\sin^2(\varphi-\Phi)dv}. \tag{3.238}$$

This is much simpler than the case where both the amplitude and the phase were fitted. The smallest phase variance can be obtained by using local oscillator phases as close as possible to $\varphi+\pi/2$, and there are no extra terms to complicate the problem.

It is clear that if an accurate estimate of the phase is known beforehand, it is possible to obtain a phase variance that is extremely close to the intrinsic phase variance of $1/4\left|\alpha\right|^2$. If the local oscillator phase is $\Phi = \hat{\varphi} + \pi/2$, then the expression for the variance becomes

$$\begin{aligned} \mathrm{var}(\phi) &= \frac{1}{4\left|\alpha\right|^2 \int \cos^2(\varphi - \hat{\varphi}) dv} \\ &\approx \frac{1}{4\left|\alpha\right|^2} + \frac{\int (\varphi - \hat{\varphi})^2 dv}{4\left|\alpha\right|^2}. \end{aligned} \tag{3.239}$$

This indicates that the introduced phase variance is proportional to the variance in the intermediate phase estimates, so it should be possible to reduce the introduced phase variance practically indefinitely by using better intermediate phase estimates. In particular, it should be possible to reduce it below the theoretical limit based on $\arg(A + BA^*)$ as the final phase estimate.

If the intermediate phase estimate is based on the measurement results so far, then the variance in the phase estimate cannot be smaller than the canonical phase variance for a coherent state with $\left|\alpha\right|^2 v$ photons, $1/(4\left|\alpha\right|^2 v)$. We cannot use this directly in the above equation, as this variance goes to infinity for zero time. Since the average value of $\cos^2$ for a randomised phase is $1/2$, it is only reasonable to use the approximation

$$\left\langle \cos^2(\varphi - \hat{\varphi}) \right\rangle \approx 1 - \left\langle (\varphi - \hat{\varphi})^2 \right\rangle, \tag{3.240}$$

when it gives a value less than $1/2$. In order to obtain an approximate result, we can use this approximation for $v > 1/(2\left|\alpha\right|^2)$, and use $1/2$ for $v < 1/(2\left|\alpha\right|^2)$. Then we find

$$\begin{aligned} \int \cos^2(\varphi - \hat{\varphi}) dv &\approx \int_0^{\frac{1}{2|\alpha|^2}} \frac{dv}{2} + \int_{\frac{1}{2|\alpha|^2}}^{1} \left[1 - \frac{1}{4\left|\alpha\right|^2 v}\right] dv \\ &= 1 - \frac{1 + \log 2\left|\alpha\right|^2}{4\left|\alpha\right|^2}. \end{aligned} \tag{3.241}$$

Therefore the phase variance should be approximately

$$\mathrm{var}(\phi) = \frac{1}{4\left|\alpha\right|^2} + \frac{1 + \log 2\left|\alpha\right|^2}{16\left|\alpha\right|^4}. \tag{3.242}$$

Thus we find that the introduced phase variance scales roughly as $\log \overline{n}/(16\overline{n}^2)$. This is less than the theoretical limit, which scales as $\log \overline{n}/(4\overline{n}^2)$.

To check the scaling when the calculation is performed exactly, we can use

$$\left\langle \cos^2 \hat{\varphi} \right\rangle = \tfrac{1}{2}(1 + \mathrm{Re}\left\langle e^{2i\hat{\varphi}} \right\rangle). \tag{3.243}$$

Here the system phase has been taken to be zero for simplicity. For a coherent state with amplitude $\alpha\sqrt{v}$, we have

$$\langle n|\alpha\sqrt{v}\rangle = e^{-\alpha^2 v/2} \frac{(\alpha\sqrt{v})^n}{\sqrt{n!}} \tag{3.244}$$

This means that

$$\begin{aligned} \left\langle e^{2i\hat{\varphi}} \right\rangle &= \sum_{n=0}^{\infty} \langle \alpha\sqrt{v}|n\rangle\langle n+2|\alpha\sqrt{v}\rangle \\ &= e^{-\alpha^2 v} \sum_{n=0}^{\infty} \frac{(\alpha^2 v)^{n+1}}{\sqrt{n!(n+2)!}}. \end{aligned} \tag{3.245}$$

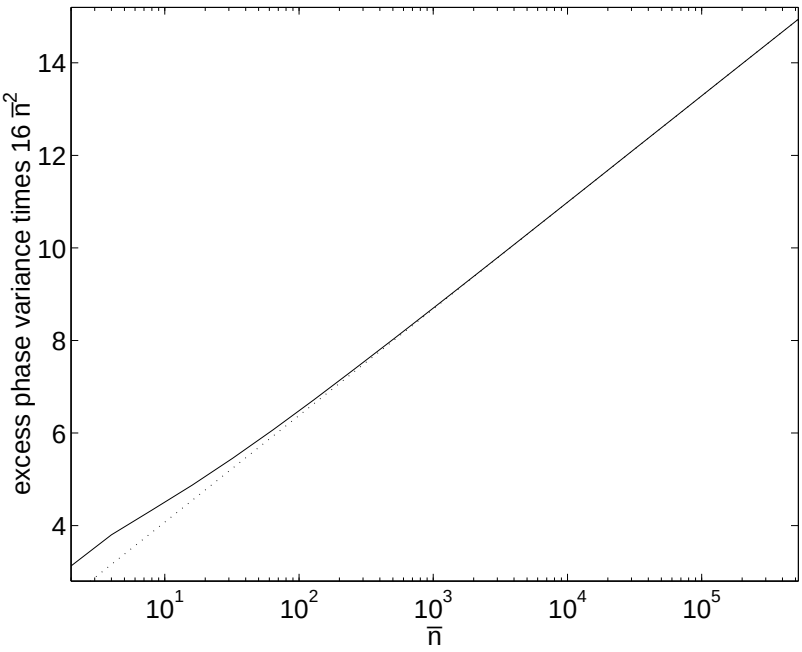

Figure 3.25: The excess phase variance multiplied by $16\overline{n}^2$ for coherent states calculated using Eq. (3.246). The calculated result is shown as the continuous line, and the dotted line is the fitted expression (3.247).

Now taking the integral of this over time gives

$$\begin{aligned}\int_0^1 e^{-\alpha^2 v} \sum_{n=0}^{\infty} \frac{(\alpha^2 v)^{n+1}}{\sqrt{n!(n+2)!}} dv &= \frac{1}{\alpha^2} \int_0^{\alpha^2} e^{-t} \sum_{n=0}^{\infty} \frac{t^{n+1}}{\sqrt{n!(n+2)!}} dt \\ &= \frac{1}{\alpha^2} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha^2, n+2)}{\sqrt{n!(n+2)!}},\end{aligned} \tag{3.246}$$

where $\Gamma(\alpha^2, n+2)$ is the incomplete gamma function. This can then be used to evaluate $\langle \cos^2 \hat{\varphi} \rangle$ exactly. The results calculated in this way are plotted in Fig. 3.25.

It was found that the scaling of the introduced phase variance was again $\log \overline{n}/(16\overline{n}^2)$ when estimated using this more exact method. There was a difference of order $\overline{n}^{-2}$ with the previous result, however. It was found that the introduced phase variance was very close to

$$\frac{\log \overline{n} + 1.7733}{16\overline{n}^2} \tag{3.247}$$

as compared to

$$\frac{\log \overline{n} + 1.693147}{16\overline{n}^2} \tag{3.248}$$

using the very approximate method of Eq. (3.241).

Note that this result is still based on assuming that the intermediate phase estimates are as good as canonical. To determine the result for the real case, where the intermediate phase estimates are based on the preceding data, we must perform stochastic integrals to determine the result numerically. To determine the results in this case, rather than performing the integrals separately for each mean photon number, a continuous calculation with an exponentially increasing time-step was used. Then at various times, the photon number up to that time and the integral of $\langle \cos^2 \hat{\varphi} \rangle$ were determined.

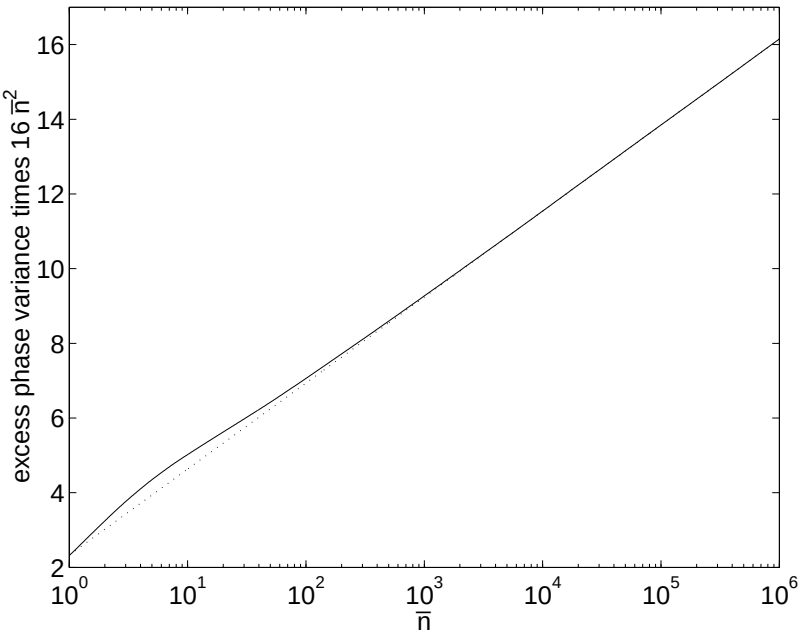

Figure 3.26: The excess phase variance multiplied by $16\overline{n}^2$ for coherent states as determined using stochastic integrals. The calculated result is shown as the continuous line, and the dotted line is the fitted expression.

It is possible to do this in the case of coherent states, because the integral up to time v with coherent amplitude α is equivalent to the integral over the unit interval with amplitude $\alpha\sqrt{v}$. Note also that the integral of $\left\langle \cos^2 \hat{\varphi} \right\rangle$ does not need to be calculated separately, because

$$\mathrm{Re}B_v = \int_0^v \cos^2 \hat{\varphi}_u du. \tag{3.249}$$

For this calculation the intermediate phase estimate was the best phase estimate using the known value of $|\alpha|$, as found by solving Eq. (3.235).

This calculation was performed for 2^{14} samples, and the results are shown in Fig. 3.26. The results again scaled as $\log \overline{n}/(16\overline{n}^2)$, but the term of order $\overline{n}^{-2}$ was different. The scaling obtained was approximately

$$\frac{\log \overline{n} + 2.4}{16\overline{n}^2}. \tag{3.250}$$

As can be expected, this is slightly worse than the result calculated assuming that the intermediate phase estimates are as good as canonical.

This demonstrates that if the coherent amplitude is known, it is possible to reduce the introduced phase variance to a factor of 4 below the theoretical limit (that scales as $\log \overline{n}/(4\overline{n}^2)$). Unfortunately this is not very useful in the case of coherent states, as the introduced phase variance is always (for an adaptive feedback scheme) far less than the intrinsic phase variance.

If the introduced phase variance remained this small for a squeezed state, then this would be a very significant result. Unfortunately, the result that the introduced phase variance should be fairly independent of the input state breaks down for this type of phase estimate, for similar reasons to why the theoretical limit does not apply.

The best phase estimate as found by solving Eq. (3.235) is specific to coherent states, and the best phase estimate for other states will be different. For the case of squeezed states, this is a difficult calculation where the evolution of the state for each input phase must be determined based on the

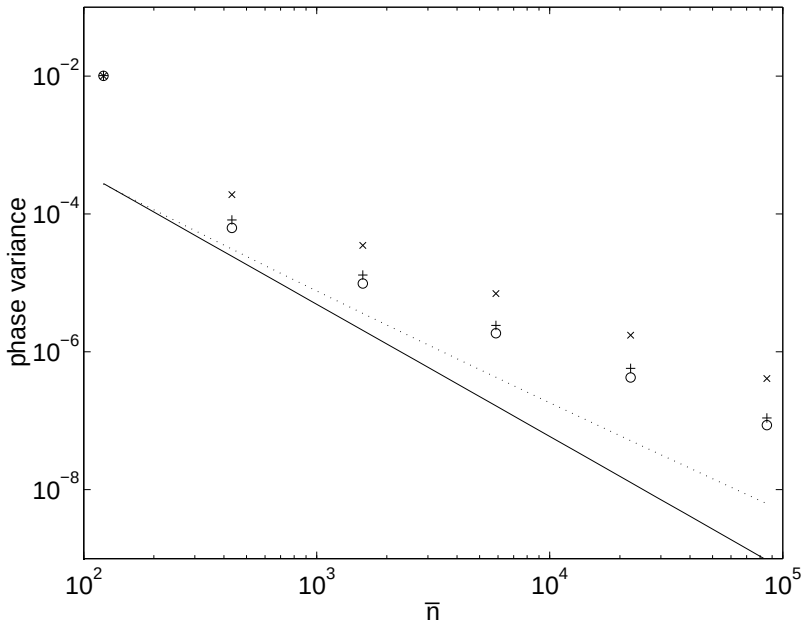

Figure 3.27: The phase variances using $\arg C_v$ feedback on optimally squeezed input states. The crosses are the results for $\arg C$ phase estimates, the pluses are those found by solving Eq. (3.235), and the circles are those for the phase estimates based on numerically minimising Eq. (3.255). The continuous line is the theoretical limit, and the dotted line is the variance for mark II measurements.

measured values of $I(v)$. Specifically, in the discretised calculation we are using

$$I(v)\delta v = 2\mathrm{Re}(\alpha_v e^{-i\Phi(v)})\delta v + \delta W(v). \tag{3.251}$$

For some assumed system phase, we can determine what value $\delta W(v)$ would have using

$$\delta W'(v) = I(v)\delta v - 2\mathrm{Re}(\alpha'_v e^{-i\Phi(v)})\delta v, \tag{3.252}$$

where the primes indicate that these are the values determined based on that assumed system phase (as opposed to the actual phase). For this value of $\delta W'(v)$, the evolution of α'_v is then

$$\delta\alpha'_v = \frac{1}{1-v}\frac{B'^{\mathrm{S}}_v \delta W'(v)}{1-\left|B'^{\mathrm{S}}_v\right|^2}\left(B'^{\mathrm{S}*}_v e^{i\Phi} + e^{-i\Phi}\right), \tag{3.253}$$

where

$$B'^{\mathrm{S}}_v = \frac{1-v}{(B'^{\mathrm{S}}_0)^{-1} - B^*_v}. \tag{3.254}$$

After these values are calculated over the entire time interval $[0,1)$, we then determine

$$\sum(\delta W'(v))^2. \tag{3.255}$$

We wish to minimise this in order to find the phase that gives the best fit to the data.

It is feasible to use this method for the final phase estimates; however, it would make the calculation far too difficult if it was also used for the intermediate phase estimates. Therefore, for the intermediate phase estimates $\arg C_v$ was used initially. The results of using this intermediate phase estimate on optimally squeezed states are plotted in Fig. 3.27.

If the $\arg C$ phase estimate is used at the end of the measurement as well, the phase variance is far greater than the theoretical limit. The phase estimate based on minimising (3.255) gives a much reduced variance, but it is still far above the theoretical limit. In fact, it is still worse than

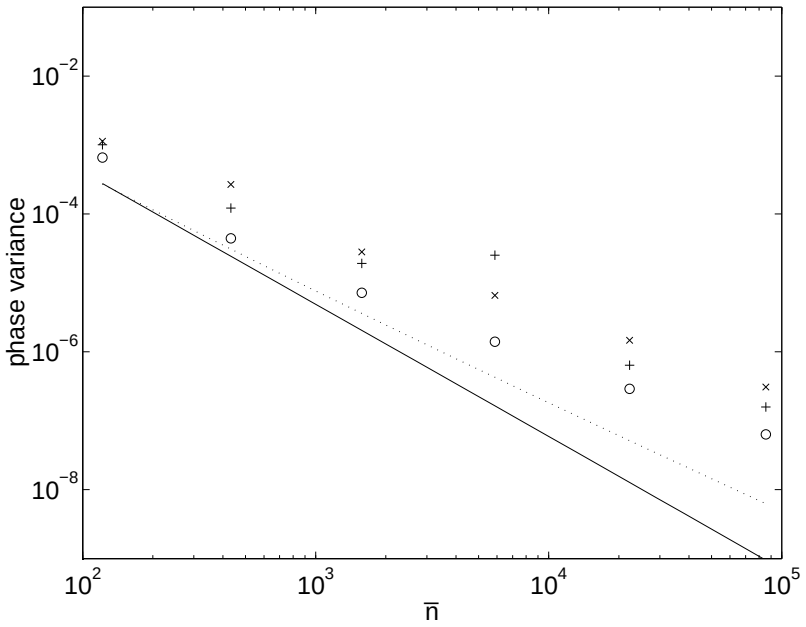

Figure 3.28: The phase variances for measurements on optimally squeezed input states using phase estimates in the feedback found by solving Eq. (3.235). The crosses are the results for $\arg C$ phase estimates, the pluses are those found by solving Eq. (3.235), and the circles are those for the phase estimates based on numerically minimising Eq. (3.255). The continuous line is the theoretical limit, and the dotted line is the variance for mark II measurements.

the phase variance for mark II measurements. Also shown in Fig. 3.27 is the phase estimate based on coherent states, found by solving Eq. (3.235). These phase estimates give variances that are very close to, but slightly above, the variances for the fitted phases.

As an alternative feedback scheme, we can use the phase estimates based on coherent states in the feedback. The results using this feedback are shown in Fig. 3.28. The variances for the fitted phases are slightly less than those using $\arg C_v$ intermediate phase estimates, but are still far above the theoretical limit or even the variances for mark II measurements.

It is possible to obtain phase variances close to the theoretical limit if we use feedback that gives results close to the theoretical limit for the usual $\arg C$ phase estimates. For example, if we use the corrected feedback as in Sec. 3.7.3, we obtain the results shown in Fig. 3.29. As the numerical results do not take account of the low probability results with large errors, a correction factor of $1/(4\overline{n}^2)$ has been added to these results (as was done for Fig. 3.24).

In practice it was found that, for the smaller photon numbers where large error results were obtained, the fitted phases were still close to the $\arg C$ phase estimates, so large errors were obtained for the same samples for both phase estimates. This can also be ensured by only performing the minimisation near $\arg C$. Then the contribution due to the large error results will be the same for both cases. This is why the same correction factor of $1/(4\overline{n}^2)$ has been added to both sets of results in Fig. 3.29.

It is seen in Fig. 3.29 that when the corrections are taken into account, the results for the usual phase estimate are slightly greater than the theoretical limit, but the results for the fitted phase are slightly less than the theoretical limit. This demonstrates that it is possible to surpass the theoretical limit in the case of squeezed states, but only by a very small margin.

3.8.2 Optimal Phase Estimates

The fitting approach considered in the previous section gives the most probable phase, but this is not necessarily the same as the phase estimate that minimises the variance. It is shown in Sec. 6.2

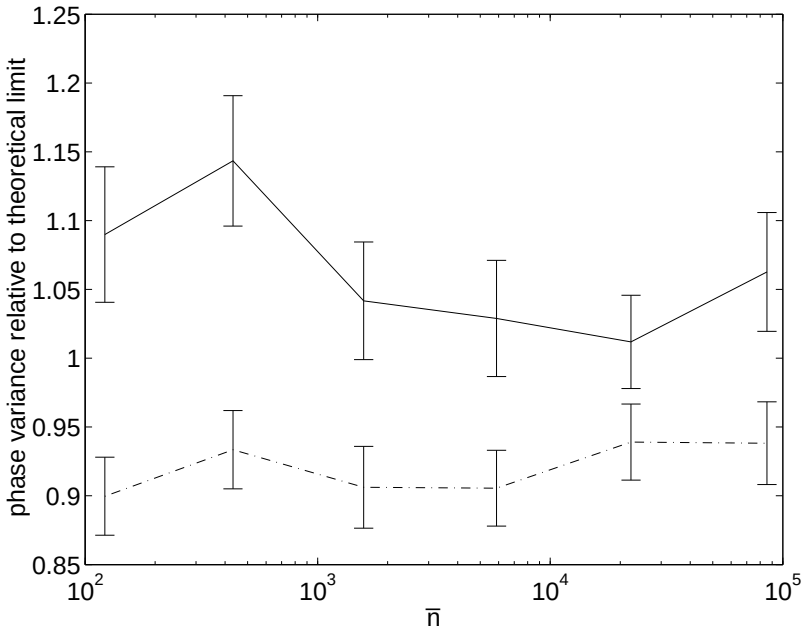

Figure 3.29: The phase variances for measurements on optimally squeezed input states using intermediate phase estimates as in Sec. 3.7.3. The continuous line is the phase variance for $\arg C$ phase estimates and the dash-dotted line is the variance for phase estimates based on fitting. All variances are shown as a ratio to the theoretical limit.

that the phase estimate that minimises the phase variance for interferometric measurements is

$$\hat{\varphi} = \arg \int_{-\pi}^{\pi} e^{i\varphi} P(n_m|\varphi) d\varphi, \tag{3.256}$$

where n_m is the measurement record. This derivation is very general, and should also hold for dyne measurements, where there is a continuous measurement record $\mathbf{I}_{[0,v)}$.

Similarly to the case for interferometry, we have [using Eq. (1.72)]

$$P(\mathbf{I}_{[0,v)}|\varphi) \propto P(\varphi|\mathbf{I}_{[0,v)}), \tag{3.257}$$

so the optimum phase estimate can be expressed as

$$\begin{aligned} \hat{\varphi} &= \arg \int_{-\pi}^{\pi} e^{i\varphi} P(\varphi|\mathbf{I}_{[0,v)}) d\varphi \\ &= \arg \left\langle e^{i\varphi} \right\rangle, \end{aligned} \tag{3.258}$$

where the average is over the probability distribution for the phase based on the measurement record. In order to determine the probability distribution $P(\mathbf{I}_{[0,v)}|\varphi)$, we need to determine the probability of obtaining the increments $dW(v)$. Considering the discretised equation, each increment $\delta W(v)$ has a Gaussian distribution with variance δv. The probability of obtaining $\delta W(v)$ is therefore

$$P(\delta W(v)) \propto e^{-(\delta W(v))^2/(2\delta v)}. \tag{3.259}$$

The probability of obtaining the measurement record will therefore be

$$P(\mathbf{I}_{[0,v)}|\varphi) \propto \exp\left(-\sum (\delta W(v))^2/(2\delta v)\right). \tag{3.260}$$

From Eq. (3.257), this means that we also have

$$P(\varphi|\mathbf{I}_{[0,v)}) \propto \exp\left(-\sum(\delta W'(v))^2/(2\delta v)\right). \tag{3.261}$$

Here the prime on $\delta W'(v)$ indicates that it is calculated from the measurement record based on an assumed system phase, similarly to the previous section.

It is obvious that the fitted phase discussed in the previous section is the most probable phase, as it minimises $\sum(\delta W'(v))^2$ and therefore maximises $P(\varphi|\mathbf{I}_{[0,v)})$. To find the optimal phase that minimises the final phase uncertainty, we need to calculate

$$\begin{aligned}\hat{\varphi} &= \arg\int_{-\pi}^{\pi} e^{i\varphi} P(\varphi|\mathbf{I}_{[0,v)})d\varphi \\ &= \arg\int_{-\pi}^{\pi} e^{i\varphi}\exp\left(-\sum(\delta W'(v))^2/(2\delta v)\right)d\varphi.\end{aligned} \tag{3.262}$$

For a coherent state this simplifies to

$$\hat{\varphi} = \arg\int_{-\pi}^{\pi} e^{i\phi}\exp\left(\mathrm{Re}\left[2A^*\left|\alpha\right|e^{i\phi} + B^*\left|\alpha\right|^2 e^{2i\phi}\right]\right)d\phi. \tag{3.263}$$

It does not appear to be possible to evaluate this integral analytically, and it would need to be evaluated numerically.

The calculation is even more difficult in the case of squeezed states. As was discussed in the previous section, for squeezed states the entire calculation (i.e. determining the time evolution of the state) must be repeated for each value of ϕ. For numerical minimisation the number of values of ϕ for which the calculation must be performed is on the order of 10, which means that the calculations are more lengthy, but not infeasible. For a numerical integral thousands of function evaluations would be required for an accurate result, making this method infeasible. For this reason, these optimal phase estimates were not used for dyne measurements in this study. As will be seen in Ch. 6, however, it is much easier to determine the optimal phase estimates in interferometry.

Chapter 4

The Effect of Time Delays

4.1 Introduction

In practice the adaptive dyne measurements described in the previous chapters cannot be performed exactly. In any experiment there will be imperfections, for example calibration errors. In making phase measurements a major source of error is inefficient photodetectors. This is particularly the case for single photon photodetectors. These are photodetectors that are designed for distinguishing between, for example 1 or 2 photons. These photodetectors currently cannot be made with efficiencies above about 87% [44].

High amplitude photodetectors, on the other hand, can be made with far higher efficiencies, around 98% [43]. This is the sort of photodetector required for dyne measurements, due to the large amplitude local oscillator field. It is fairly straightforward to determine the extra phase uncertainty due to inefficient photodetectors. In the introduction (and Ref. [35]), it is shown that when the photodetector efficiency is η, the extra phase variance is approximately

$$\frac{1-\eta}{4\eta\overline{n}}. \tag{4.1}$$

This extra phase variance means that, for large photon numbers, phase feedback schemes can only reduce the phase variance by a factor of about $1-\eta$. This will be true for mark II measurements, as well as the more advanced phase feedback schemes described in the preceding chapter. For current photodetectors, this extra phase variance is more significant than the introduced phase variance for mark II measurements for photon numbers above about 1000. Below this the extra phase variance for mark II measurements is only marginally above the theoretical limit, and so it is not possible to greatly reduce the phase variance using more advanced feedback schemes.

Another imperfection that is specific to phase measurements with feedback is the time delay in the feedback loop. This contribution is more difficult to estimate. Some highly simplified calculations indicate that the excess phase variance due to time delays is $\tau/2$ for mark I measurements (where τ is the time delay), and $\tau/(2\bar{n})$ for mark II measurements [35].

In Ref. [53] I repeated these derivations more rigorously, and this chapter is based on that paper. While the result for mark I measurements is reasonably accurate, the perturbation approach is inadequate to obtain a consistent result for mark II measurements. In Sec. 4.3 an alternative derivation is considered that gives the minimum phase variance when there is a time delay. The phase variance with time delays is evaluated numerically in Sec. 4.4, and it is shown that for most of the measurement schemes the phase variance approaches this limit for large time delays.

4.2 Perturbation Approach

4.2.1 Mark I

Firstly I will estimate the effect of time delays on simplified mark I measurements in a similar way as in Ref. [35], but using fewer of the simplifications used there. Without a time delay the stochastic differential equation for the phase estimate is

$$\begin{aligned} d\hat{\varphi}_v &= \frac{I(v)dv}{\sqrt{v}} \\ &= v^{-1/2}[-2\alpha \sin\hat{\varphi}_v dv + dW(v)]. \end{aligned} \tag{4.2}$$

Here the input phase has been taken to be zero for simplicity. For some time v_1 the phase will come to lie near 0, so we can linearise around $\hat{\varphi}_v = 0$. The result, which will be valid for $v_1 \leq v \leq 1$ is

$$d\hat{\varphi}_v = v^{-1/2}[-2\alpha\hat{\varphi}_v dv + dW(v)]. \tag{4.3}$$

We wish to consider the limit of large α, so that v_1 is small, and this linearisation is accurate for most of the measurement. Including the time delay the SDE is

$$d\hat{\varphi}_v = v^{-1/2}[-2\alpha\hat{\varphi}_{v-\tau} dv + dW(v)]. \tag{4.4}$$

Now the time delay will be treated perturbatively. The solution to the perturbed equation can be written as

$$\hat{\varphi}_v = \hat{\varphi}_v^{(0)} + \alpha\tau\hat{\varphi}_v^{(1)} + O(\alpha^2\tau^2). \tag{4.5}$$

Note that for this approximation to be accurate, $\alpha\tau$ must be small, in addition to α being large.

The zeroth-order term obeys the SDE for no delay (4.3), so the first-order correction obeys

$$\alpha\tau d\hat{\varphi}_v^{(1)} = 2\alpha v^{-1/2}(\hat{\varphi}_v^{(0)} - \hat{\varphi}_{v-\tau}^{(0)})dv - 2\alpha^2\tau v^{-1/2}\hat{\varphi}_{v-\tau}^{(1)}dv. \tag{4.6}$$

Therefore to first order in τ we have

$$\begin{aligned} d\hat{\varphi}_v^{(1)} &= 2v^{-1/2}d\hat{\varphi}_v^{(0)} - 2\alpha v^{-1/2}\hat{\varphi}_v^{(1)}dv \\ &= 2v^{-1/2}\{v^{-1/2}[-2\alpha\hat{\varphi}_v^{(0)}dv + dW(v)]\} - 2\alpha v^{-1/2}\hat{\varphi}_v^{(1)}dv \\ &= -2\alpha v^{-1/2}\hat{\varphi}_v^{(1)}dv - \frac{4\alpha}{v}\hat{\varphi}_v^{(0)}dv + \frac{2}{v}dW(v). \end{aligned} \tag{4.7}$$

It is straightforward to show that the solution to the zeroth order equation is

$$\hat{\varphi}_v^{(0)} = e^{4\alpha(\sqrt{v_1}-\sqrt{v})}\hat{\varphi}_{v_1}^{(0)} + \int_{v_1}^{v} \frac{e^{4\alpha(\sqrt{u}-\sqrt{v})}}{\sqrt{u}} dW(u). \tag{4.8}$$

Using this in Eq. (4.7) and multiplying on both sides by $e^{4\alpha\sqrt{v}}$ gives

$$d(e^{4\alpha\sqrt{v}}\hat{\varphi}_v^{(1)}) = -\frac{4\alpha}{v}\left[e^{4\alpha\sqrt{v_1}}\hat{\varphi}_{v_1}^{(0)} + \int_{v_1}^{v} \frac{e^{4\alpha\sqrt{u}}}{\sqrt{u}} dW(u)\right] dv + \frac{2}{v}e^{4\alpha\sqrt{v}}dW(v). \tag{4.9}$$

Integrating then gives the solution

$$\begin{aligned} \hat{\varphi}_v^{(1)} = e^{4\alpha(\sqrt{v_1}-\sqrt{v})}\hat{\varphi}_{v_1}^{(1)} &- \int_{v_1}^{v} \frac{4\alpha}{s}\left[e^{4\alpha(\sqrt{v_1}-\sqrt{v})}\hat{\varphi}_{v_1}^{(0)} + \int_{v_1}^{s} \frac{e^{4\alpha(\sqrt{u}-\sqrt{v})}}{\sqrt{u}} dW(u)\right] ds \\ &+ \int_{v_1}^{v} \frac{2}{s} e^{4\alpha(\sqrt{s}-\sqrt{v})} dW(s). \end{aligned} \tag{4.10}$$

In this approximation the mark I phase estimate is given by $\hat{\varphi}_1 = \hat{\varphi}_1^{(0)} + \alpha\tau\hat{\varphi}_1^{(1)}$. To first order in τ and α^{-1} this has a variance of

$$\left\langle \phi_{\mathrm{I}}^2 \right\rangle = \frac{1}{4\alpha} + 2\alpha\tau \left\langle \hat{\varphi}_1^{(0)} \hat{\varphi}_1^{(1)} \right\rangle. \tag{4.11}$$

Here the known variance of $1/(4\alpha)$ of the zeroth order term has been used. Evaluating the second term on the right hand side we find

$$\begin{aligned} \left\langle \hat{\varphi}_1^{(0)} \hat{\varphi}_1^{(1)} \right\rangle &= e^{-8\alpha(1-\sqrt{v_1})} \left\langle \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} \right\rangle + 4\alpha \log(v_1) e^{-8\alpha(1-\sqrt{v_1})} \left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle \\ &\quad - \int_{v_1}^{1} \frac{4\alpha}{s} \int_{v_1}^{s} \frac{e^{8\alpha(\sqrt{u}-1)}}{u} du ds + \int_{v_1}^{1} \frac{2e^{8\alpha(\sqrt{u}-1)}}{u^{1.5}} du. \end{aligned} \tag{4.12}$$

The first two terms decrease exponentially with α and may therefore be omitted. Exchanging the order of the integrals in the third term and integrating gives

$$\left\langle \hat{\varphi}_1^{(0)} \hat{\varphi}_1^{(1)} \right\rangle = 4\alpha \int_{v_1}^{1} \frac{\log u}{u} e^{8\alpha(\sqrt{u}-1)} du + \int_{v_1}^{1} \frac{2e^{8\alpha(\sqrt{u}-1)}}{u^{1.5}} du. \tag{4.13}$$

To perform these integrals it is convenient to change variables to $s = 1 - \sqrt{u}$, so $du = -2(1-s)ds$. Then we obtain

$$\left\langle \hat{\varphi}_1^{(0)} \hat{\varphi}_1^{(1)} \right\rangle = 16\alpha \int_{0}^{1-\sqrt{v_1}} \frac{\log(1-s)}{(1-s)} e^{-8\alpha s} ds + 4 \int_{0}^{1-\sqrt{v_1}} \frac{e^{-8\alpha s}}{(1-s)^2} ds. \tag{4.14}$$

Expanding in a Maclaurin series in s gives

$$\begin{aligned} \left\langle \hat{\varphi}_1^{(0)} \hat{\varphi}_1^{(1)} \right\rangle &= -16\alpha \int_{0}^{1-\sqrt{v_1}} (s + \frac{3}{2}s^2 + O(s^3)) e^{-8\alpha s} ds + 4 \int_{0}^{1-\sqrt{v_1}} (1 + 2s + 3s^2 + O(s^3)) e^{-8\alpha s} ds \\ &= -16\alpha \left[\frac{1}{(8\alpha)^2} + \frac{3}{(8\alpha)^3} + O(\alpha^{-4}) \right] + 4 \left[\frac{1}{8\alpha} + \frac{2}{(8\alpha)^2} + \frac{6}{(8\alpha)^3} + O(\alpha^{-4}) \right] \\ &= \frac{1}{4\alpha} + O(\alpha^{-2}). \end{aligned} \tag{4.15}$$

Note that the upper bound at $1 - \sqrt{v_1}$ has no effect since it gives a term that decays exponentially with α. Using this result, the total phase variance is

$$\left\langle \phi_{\mathrm{I}}^2 \right\rangle = \frac{1}{4\alpha} + \frac{\tau}{2}. \tag{4.16}$$

This provides a good verification of the result obtained by the highly simplified method in [35].

Note that this result is based on continuing to use the intermediate phase estimate at the end of the measurement. If the phase estimate $\arg A$ is used at the end of the measurement, the result will be different, and cannot be predicted using this approach.

4.2.2 Mark II

For the mark II case it does not seem to be possible to obtain a consistent result using this approach. To illustrate this, I will briefly outline the derivation (the details are in Appendix A.3). From [35], the mark II phase estimate is effectively a time average of the mark I phase estimates:

$$\phi_{\mathrm{II}} \approx \int_{0}^{1} \hat{\varphi}_t dt. \tag{4.17}$$

In order to make this consistent with the above theory, we should take the average only from time v_1, then take the limit of small v_1. In perturbation theory the mark II phase estimate is

$$\phi_{\rm II} = \int_{v_1}^{1} [\hat{\varphi}_t^{(0)} + \alpha\tau\hat{\varphi}_t^{(1)}]dt. \tag{4.18}$$

Using this expression the variance is

$$\left\langle \phi_{\rm II}^2 \right\rangle = \int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(0)} \right\rangle + 2\alpha\tau \int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(1)} \right\rangle. \tag{4.19}$$

It is shown in Appendix A.3 that the first term can be simplified to

$$\int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(0)} \right\rangle \approx \frac{1}{4\alpha^2} + \frac{v_1}{4\alpha^2} \left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle, \tag{4.20}$$

which is similar to that obtained in [35]. For the second term, however, we get

$$2\alpha\tau \int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(1)} \right\rangle \approx \frac{\tau}{\alpha} \left(\frac{v_1}{2} \left\langle \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} \right\rangle - \sqrt{v_1} \left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle \right). \tag{4.21}$$

This is radically different to the result obtained in Ref. [35], and seems to cast some doubt on the simplified theory used there.

Unlike the result for mark I measurements, this result depends entirely on the conditions at time v_1, which are unknown. In order to obtain a usable result, we would need the initial conditions to give a neglibible contribution, as is the case for mark I measurements. For this reason I will consider an alternative approach for estimating the increase in the phase variance due to the time delay.

4.3 Theoretical Minimum

The alternative method of obtaining an estimate for the phase variance with a time delay is to consider the squeezed state $|\alpha^{\rm P}, \zeta^{\rm P}\rangle$ in the probability distribution. As was explained in Sec. 3.1, the excess phase variance due to the measurement scheme is approximately the phase variance of this squeezed state.

From Ref. [45], the phase variance of a squeezed state is given by

$$\left\langle \Delta\phi^2 \right\rangle \approx \frac{n_0 + 1}{4\overline{n}^2} + 2{\rm erfc}\left(\sqrt{2n_0}\right), \tag{4.22}$$

where $n_0 = \overline{n}e^{2\zeta}$ for real ζ. To determine the excess phase variance, we would use the values $\overline{n}^{\rm P}$ and $\zeta^{\rm P}$ from the squeezed state in the POM, rather than those for the input state. The average value of $\overline{n}^{\rm P}$ will be close to the photon number of the input state.

For states that are significantly less squeezed than optimum, the second term is negligible and we can omit the term of order $(\overline{n}^{\rm P})^{-2}$. Then this simplifies to

$$\left\langle \Delta\phi^2 \right\rangle \approx \frac{e^{2\zeta^{\rm P}}}{4\overline{n}^{\rm P}}. \tag{4.23}$$

Since $\overline{n}^{\rm P}$ will be close to the photon number of the input state, it is reasonable to replace it with $\overline{n}$.

When there is a delay of τ in the system, before time τ we have no information about the phase of the system to use to adjust the local oscillator phase. Therefore we must use a heterodyne scheme for this time period, rapidly varying the local oscillator phase. This means that B_τ will be equal to

zero, and no matter how good the phase estimate is after time τ, the largest the magnitude of B_v can be made is $v - \tau$. Then at the end of the measurement, the largest $|B|$ can be is $1 - \tau$, and the largest $|\zeta^{\mathrm{P}}|$ can be is $\operatorname{atanh}(1 - \tau)$.

The lower limit to the introduced phase variance when there is a time delay of τ is therefore

$$\begin{aligned} \left\langle \Delta\phi^2 \right\rangle_{\min} &\approx \frac{e^{-2\operatorname{atanh}(1-\tau)}}{4\bar{n}}, \\ &\approx \frac{\tau}{8\bar{n}}. \end{aligned} \tag{4.24}$$

We can expect that the introduced phase variance will be close to this for states of small intrinsic phase variance, as there will quickly be very good phase estimates available for the feedback. In addition, the time delay must be sufficiently large that the phase variance given by this expression is significantly above the introduced phase variance for no time delay.

This result obeys the same scaling law as the result given in [35], but it is a factor of four times smaller. Note, however, that the limit condition for the result given in [35] is that $\tau\alpha$ is small, whereas the above result should only be accurate when both α and τ are reasonably large. The result here also differs in that it is the limit for the total introduced phase variance, rather than just the extra phase variance due to the time delay.

4.4 Numerical Results

These analytic results were also tested numerically. The numerical techniques used were similar to those described in Sec. 3.3. Minimum uncertainty squeezed states were used, with the stochastic differential equations for the squeezing parameter α as given in Eq. (3.141) and the value of B_v^S as given in Eq. (3.143). For all calculations 2^{20} time steps were used, and calculations were performed with time delays of 2^n time steps, where n varies from 0 to 18.

For most of these calculations the same random numbers were used for each time delay in order to see the difference in the variance due to the time delay more accurately. If this is not done, the differences between the variances for the different time delays are primarily due to the variation of the random numbers, rather than the different time delay. Using the same random numbers makes the results correlated, so that the differences are primarily due to the time delay.

For the first 2^n time steps the local oscillator phase was rotated by $\pi/2$ each step. For the following time steps the data up to the time step 2^n before the current time step was used. For a delay of $2^0 = 1$ time steps the data from the previous step was used, corresponding to the technique for no time delay.

Numerical results for four different phase feedback schemes were obtained:
(a) The simplified feedback for mark I and II measurements, where

$$d\hat{\varphi}_v = \frac{dI(v)dv}{\sqrt{v}}. \tag{4.25}$$

(b) The unsimplified feedback, where the phase estimate is

$$\hat{\varphi}(v) = \arg A_v. \tag{4.26}$$

(c) The phase estimate that is intermediate between $\arg A$ and the best phase estimate

$$\hat{\varphi}(v) = \arg\left(A_v^{\varepsilon} C_v^{1-\varepsilon}\right), \tag{4.27}$$

where ε is a constant.
(d) The same as in (c), except that the value of ε varies with time as

$$\varepsilon(v) = \frac{v^2 - |B_v|^2}{|C_v|}\sqrt{\frac{v}{1-v}}. \tag{4.28}$$

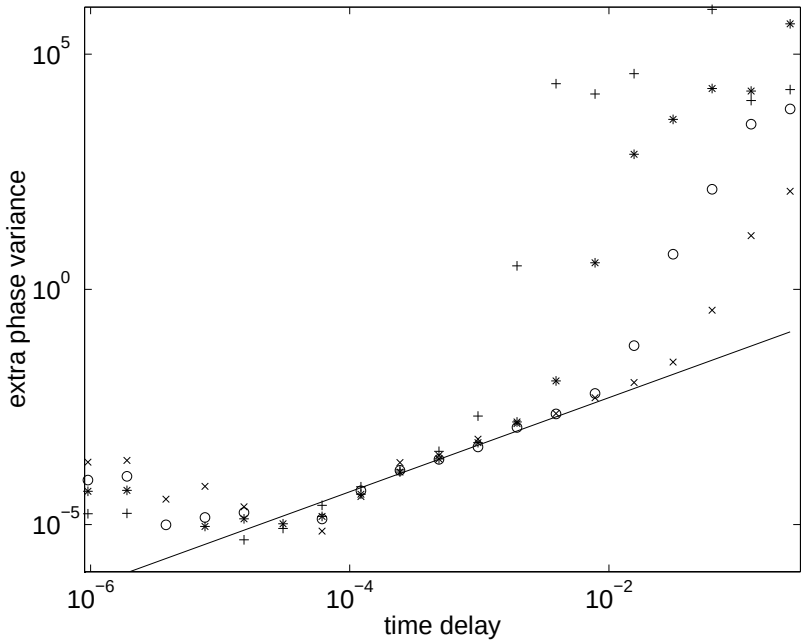

Figure 4.1: The extra phase variance (in the final value of the intermediate phase estimate for the simplified feedback) due to the time delay plotted as a function of time delay for four different mean photon numbers. The data for a mean photon number of 121.590 are shown as crosses, for a photon number of 1576.55 as circles, for a photon number of 22254.8 as asterisks and for a photon number of 332067 as pluses. The approximate analytic result $\tau/2$ is plotted as the continuous line.

4.4.1 Comparison with Perturbative Theory

I will firstly consider the case of simplified feedback, and consider the variance in the final value of the feedback phase, rather than the phase of A or C. This case was examined in Sec. 4.2.1, and the extra phase variance due to the time delay is $\tau/2$ according to that analysis. The extra phase variance is plotted for four different mean photon numbers in Fig. 4.1. For each of the points shown 2^{13} samples were used.

To determine the extra phase variance due to the time delay, an estimate must be made of the phase variance with no time delay. The most convenient estimate to use is the minimum variance obtained, as this prevents negative data points that cannot be plotted on a log-log graph. The minimum variance is not necessarily that for the smallest time delay, due to the stochastic nature of the calculations.

The theoretical asymptotic value of $\tau/2$ is also plotted in Fig. 4.1. As can be seen, many of the results are close to the theoretical line for the intermediate time delays. For small time delays, the extra phase variance due to the time delay is too small a fraction of the total phase variance for the results to be accurate. The reason why the results deviate from the asymptotic result for large time delays is that this result is for the limit of small $\alpha\tau$. Note also that the results for larger photon numbers deviate from the asymptotic result for smaller τ than the results for smaller photon numbers. This is what can be expected from this limit condition.

It is also possible to use fitting techniques to determine how closely the numerical results agree with the theoretical asymptotic result of $\tau/2$. The data and fitted lines for mean photon numbers of 121.590, 1576.55, 22254.8, 332067 and 5122478 are plotted in Figs 4.2, 4.3, 4.4, 4.5 and 4.6 respectively. For each of the points shown 2^{14} samples were used. These results were determined using independent random numbers for each data point. Calculations were also performed using the same random numbers, in order to reduce the relative error between the data points. (These are the data points plotted in Fig. 4.1.) Unfortunately this method tends to produce systematic error in

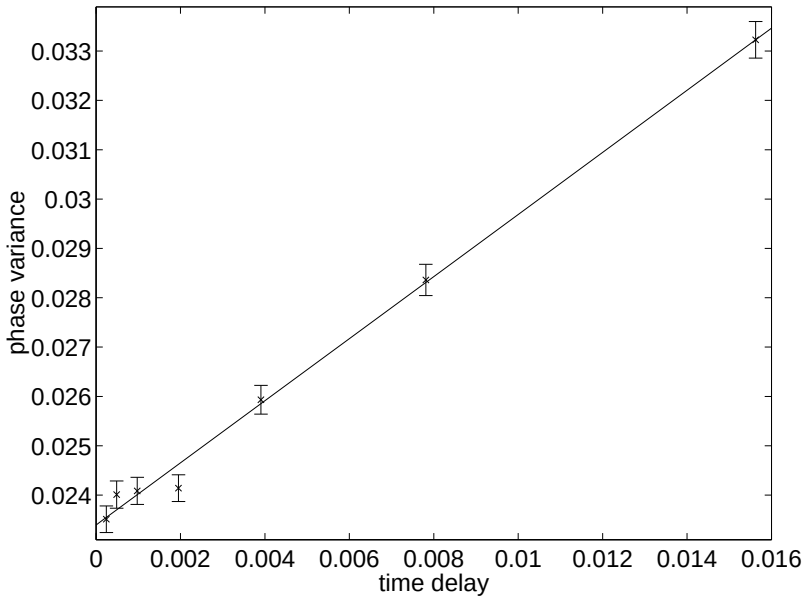

Figure 4.2: The Holevo variance in the final value of the intermediate phase estimate for a mean photon number of 121.590. The results are shown as the crosses and the continuous line is that fitted to the data.

the slope, which is not reflected in the uncertainty. It is therefore better to use independent random numbers but a large number of samples to estimate the slope.

Some of the initial data points have been omitted in each graph, as these were too close to each other to be useful. The data points for larger time delays have also been omitted. These tend to be less accurate, as the approximation is in the limit of small $\alpha\tau$. More specifically, for most of these graphs the data points for $\alpha\tau > 0.08$ were omitted. The slopes of the fitted lines found were

$$\begin{aligned} S_{4.2} &= 0.629 \pm 0.019, \\ S_{4.3} &= 0.490 \pm 0.032, \\ S_{4.4} &= 0.590 \pm 0.098, \\ S_{4.5} &= 0.525 \pm 0.039, \\ S_{4.6} &= 0.498 \pm 0.075. \end{aligned}$$

Except for the first result, these results are all consistent with the theoretical value of 0.5. The third result is higher than 0.5, but also has a large uncertainty. Note that, due to the linearisation in Eq. (4.3), we can only expect the $\tau/2$ result to be accurate in the limit of large α. The reason for the larger slope for the smallest photon number is likely to be that α is not sufficiently large for the linearisation to be accurate.

As was mentioned above, the result for the additional phase variance due to the time delay is only valid for the variance in the final value of the phase estimate, which is not the same as $\arg A$ when there is a time delay. In Fig. 4.7 I have plotted the variation of the phase variance with time delay for three alternative final phase estimates, $\hat{\varphi}_1$, $\arg A$ and $\arg C$. This is for a photon number of approximately 332000, and is fairly representative of the results for other photon numbers. For these results, and the rest of the results in this chapter, 2^{11} samples were used for each data point.

As can be seen, for very small time delays the variances in the $\hat{\varphi}_1$ and $\arg A$ phase estimates are almost identical. As the time delay is increased, however, the variance of $\hat{\varphi}_1$ increases, but the variance of $\arg A$ *decreases*. This is because, as the intermediate phase estimate gets worse, the value of $|B|$ decreases. This means that A is closer to C, so $\arg A$ is closer to the best phase estimate. Note, however, that the variance of $\arg A$ rises again, and does not converge to $\arg C$ for large time delays. This is because $|B|$ does not fall to zero.

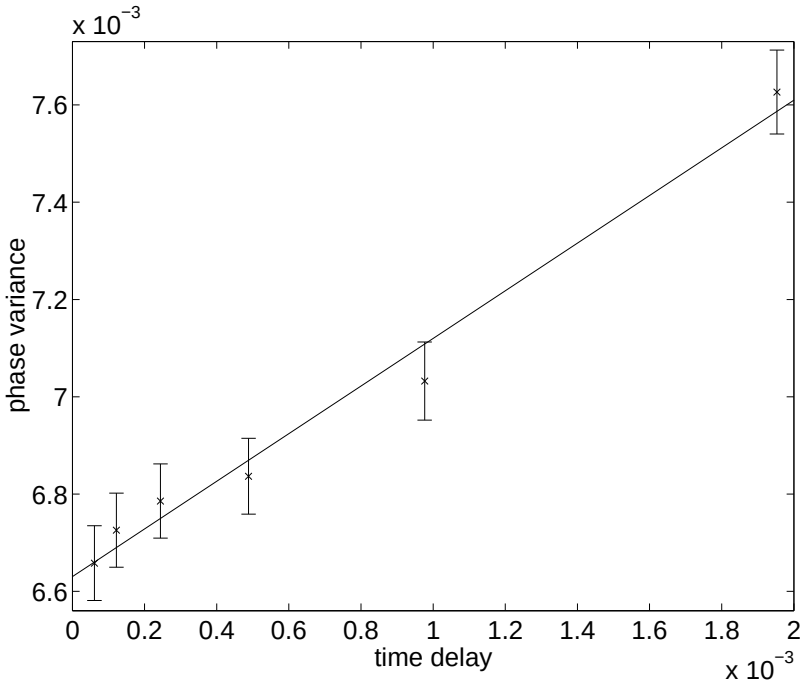

Figure 4.3: The Holevo variance in the final value of the intermediate phase estimate for a mean photon number of 1576.55. The results are shown as the crosses and the continuous line is that fitted to the data.

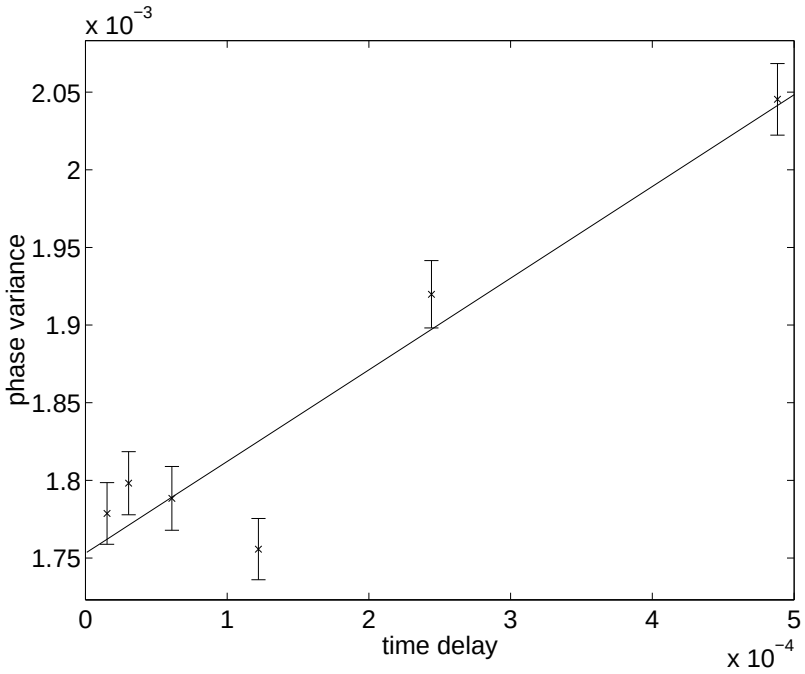

Figure 4.4: The Holevo variance in the final value of the intermediate phase estimate for a mean photon number of 22254.8. The results are shown as the crosses and the continuous line is that fitted to the data.

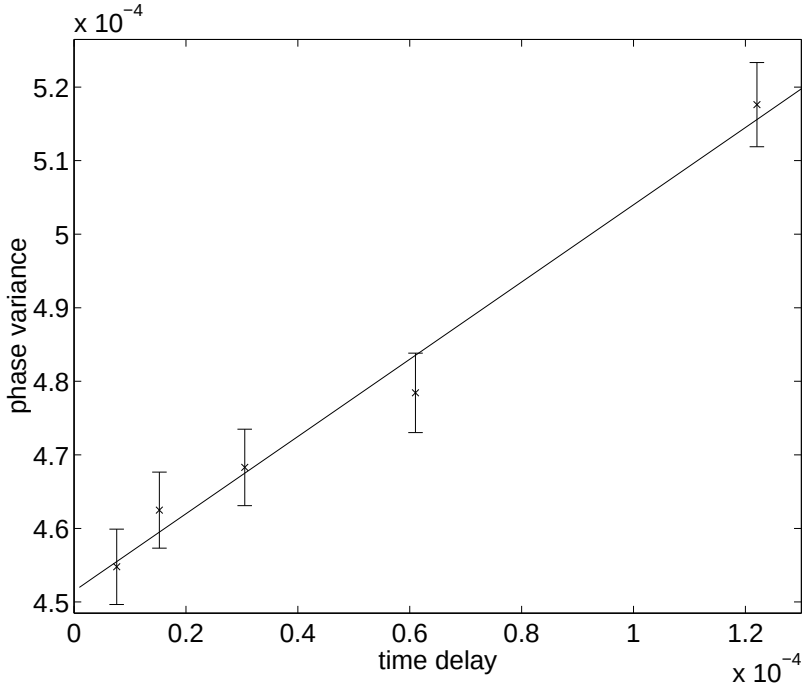

Figure 4.5: The Holevo variance in the final value of the intermediate phase estimate for a mean photon number of 332067. The results are shown as the crosses and the continuous line is that fitted to the data.

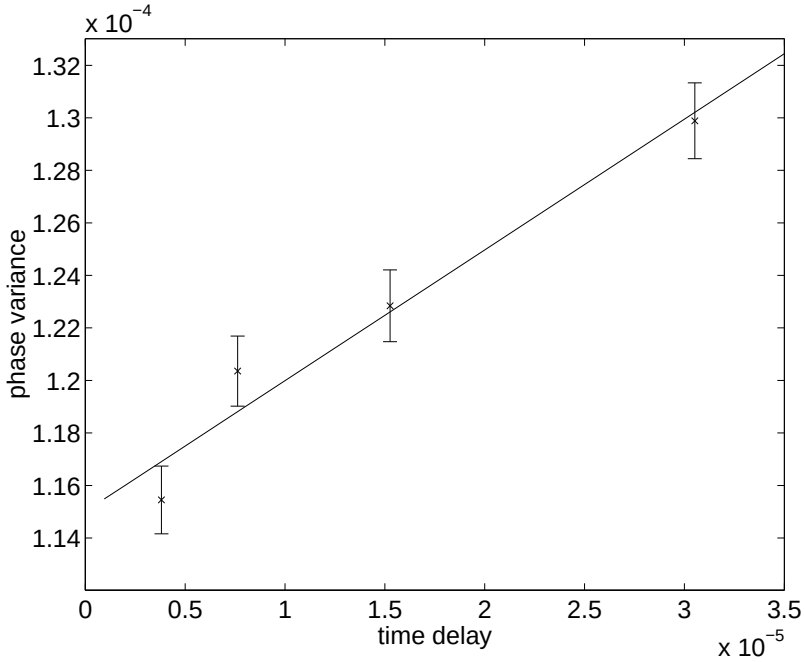

Figure 4.6: The Holevo variance in the final value of the intermediate phase estimate for a mean photon number of 5122478. The results are shown as the crosses and the continuous line is that fitted to the data.

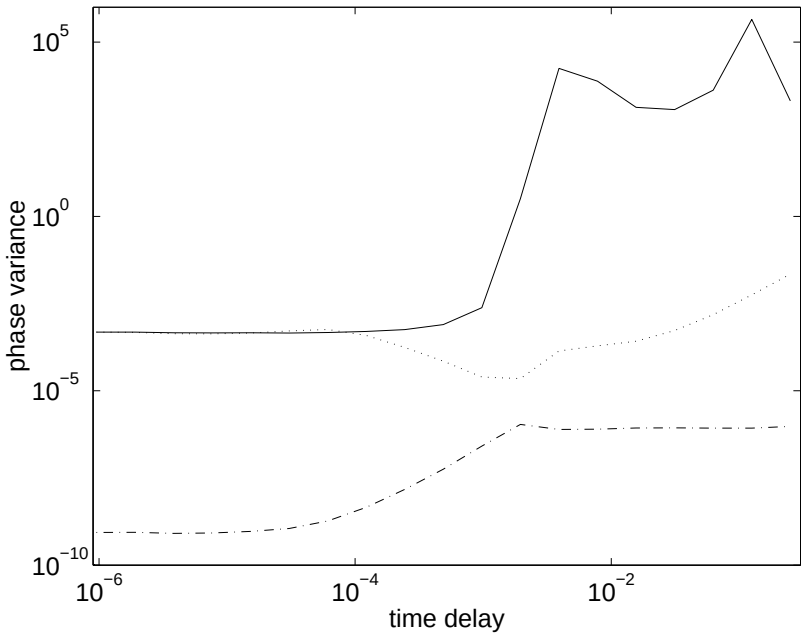

Figure 4.7: The variance of three alternative final phase estimates for simplified feedback with a time delay plotted as a function of time delay. The results for the final value of the intermediate phase estimate are plotted as a continuous line, for $\arg A$ as a dotted line, and for $\arg C$ as a dash-dotted line. All results are for a photon number of 332067.

4.4.2 Comparison with Theoretical Minimum

Lastly the variance in the phase of C will be considered. As was explained above, the theoretical lower limit to the introduced phase variance is $\tau/(8\bar{n})$. I have plotted the introduced variance in the best phase estimate $\arg C$ and the theoretical limit in Fig. 4.8. For additional accuracy I have plotted

$$\frac{e^{-2\mathrm{atanh}(1-\tau)}}{4\bar{n}}, \tag{4.29}$$

as this will continue to be accurate for time delays that are a large fraction of 1. This plot is for a photon number of 332000, and similar results are obtained for other photon numbers. In the case of simplified feedback, the phase variance is well above the theoretical limit. For large time delays the phase variance approximately converges to the heterodyne phase variance, also shown in Fig. 4.8.

The introduced phase variance for mark II measurements with the unsimplified $\arg A_v$ feedback is also shown in Fig. 4.8. The introduced phase variance for this case increases far more slowly with the time delay, and for larger time delays it is very close to the theoretical limit. These results indicate that if there is any significant time delay in the system, the simplified feedback will give a far worse result than using $\arg A_v$.

It is possible to make a correction to the simplified phase feedback scheme that improves this result somewhat. Many different alternatives were tried, and the one that gave the best results was

$$d\hat{\varphi}_v = \frac{I(v)dv}{\sqrt{v+\alpha\tau}}. \tag{4.30}$$

This correction is based on the fact that $|A_v|$ is larger than $\sqrt{v}$ when the phase estimate is worse than $\arg A_v$. (From [35], the factor of $\sqrt{v}$ in the simplified feedback comes from a factor of $|A_v|$.) The results for this correction are also shown in Fig. 4.8. The phase variances obtained in this case are significantly below those for the plain simplified feedback, but are still far above the results for the unsimplified $\arg A_v$ feedback.

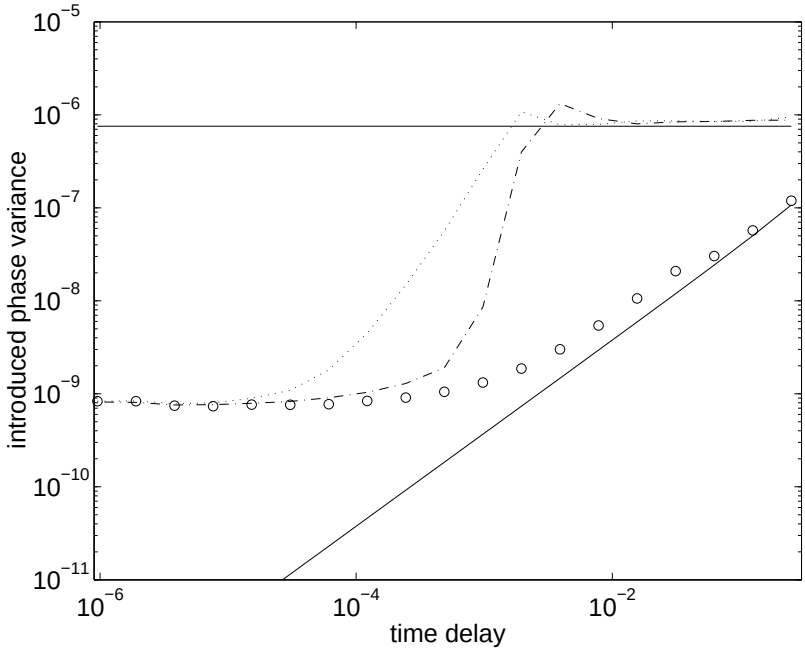

Figure 4.8: The introduced phase variance for three different phase feedback schemes plotted as a function of time delay. The dotted line is for simplified feedback, the dash-dotted line is for the corrected simplified feedback, and the circles are for unsimplified $\arg A_v$ feedback. The best phase estimate $\arg C$ is used in all three cases. The continuous horizontal line is the phase variance for heterodyne measurements, and the continuous diagonal line is the theoretical limit. All results are for a photon number of 332067.

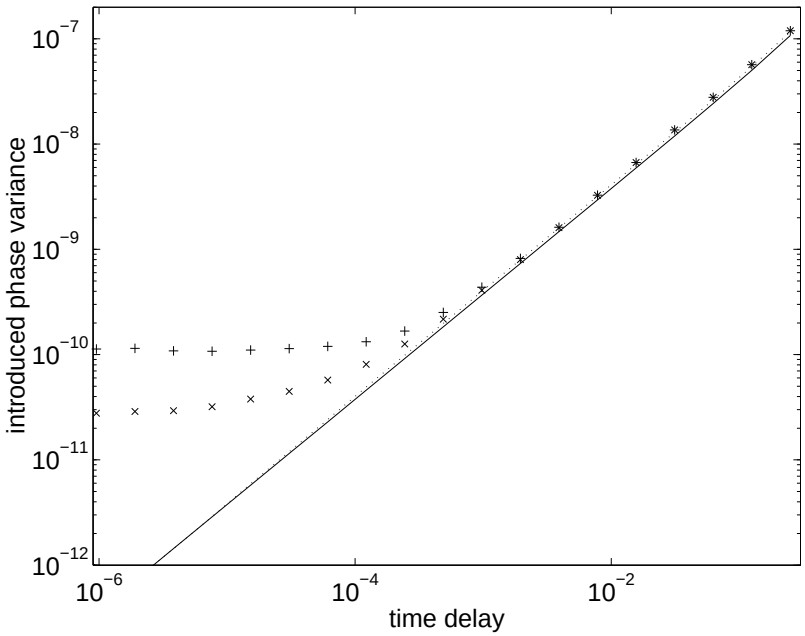

Figure 4.9: The introduced phase variance for better intermediate phase estimates plotted as a function of time delay. The pluses are for the constant ε case and the crosses are for the time dependent ε case. The theoretical limit estimated using the mean inverse photon numbers obtained from the time dependent ε case is plotted as the dotted line, and the theoretical limit using the input photon number is shown as the continuous line. All results are for a photon number of 332067.

Now we will consider the results for better intermediate phase estimates that are between $\arg A_v$ and $\arg C_v$. The introduced phase variance for the constant ε case and the theoretical limit are shown in Fig. 4.9. These results are again for a photon number of about 332000. The results for this case are even closer to the theoretical limit than those for the mark II case.

The introduced phase variance for the feedback with time-dependent ε is also plotted in Fig. 4.9. The results for this case converge to the theoretical limit at smaller time delays than for the constant ε case. For the larger time delays the results for the two cases are about the same, slightly above the theoretical limit.

In both cases the phase variance is still noticeably above the theoretical limit for large time delays, and the values of B obtained are too close to $1-\tau$ to account for this difference. The difference appears to be due to the approximation that the photon number of the state $|\alpha^{\mathrm{P}}, \zeta^{\mathrm{P}}\rangle$ is close to the photon number of the input state. The average value of this photon number is close to the photon number of the input state; however, each individual value is not necessarily close to $\bar{n}$. The expression for the introduced phase variance depends on the inverse of the photon number, and the average of an inverse is not necessarily equal to the inverse of an average. The general expression is

$$\left\langle \frac{1}{n} \right\rangle = \frac{1}{\langle n \rangle} + \frac{\langle \Delta n^2 \rangle}{\langle n \rangle^3} + O(\langle n \rangle^{-4}). \tag{4.31}$$

In Fig. 4.9 I have also plotted the estimated theoretical limit based on the average of $1/\bar{n}^{\mathrm{P}}$ for the data obtained in the time dependent ε case. Specifically, the expression plotted is

$$\frac{1}{4} \left\langle \frac{1}{\bar{n}^{\mathrm{P}}} \right\rangle e^{-2\mathrm{atanh}(1-\tau)}. \tag{4.32}$$

As can be seen, the introduced phase variance converges to this far more closely than to the limit based on the photon number of the input state.

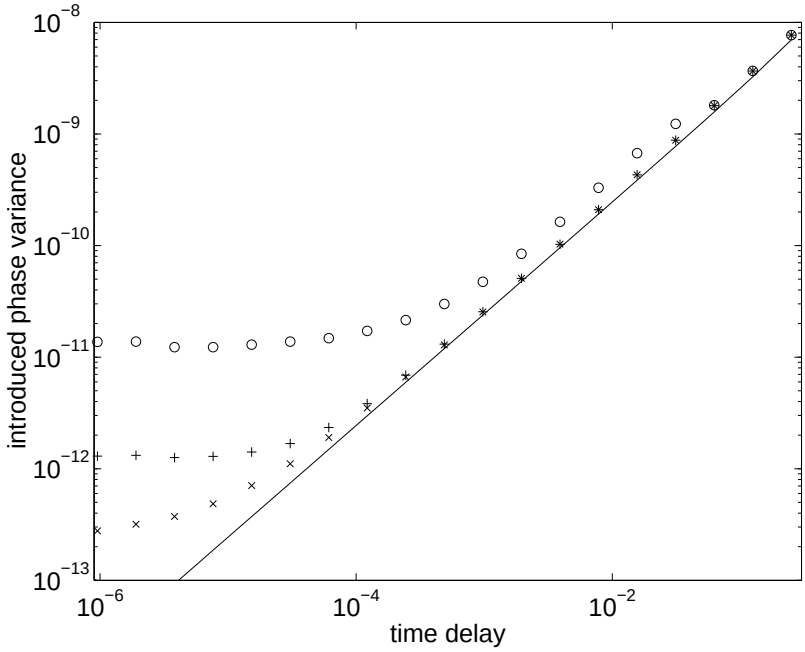

Figure 4.10: The introduced phase variance for three different feedback schemes plotted as a function of time delay. The circles are for mark II measurements, the pluses are for the constant ε case, and the crosses are for the time dependent ε case. The theoretical limit estimated using the input photon number is shown as the continuous line. All results are for a photon number of 5122478.

For larger photon numbers this factor is not so significant, and the difference between the results and the theoretical limit is smaller. For example, the results for the three different feedback schemes (mark II, constant ε and time dependent ε) for a photon number of about 5×10^6 are plotted in Fig. 4.10.

Chapter 5

Optimum Input States for Interferometry

5.1 Introduction

In chapters 2 to 4 just a single mode of the electromagnetic field was considered. The reference phase was provided by a local oscillator field, which was assumed to be sufficiently large amplitude that it could be treated classically. The main alternative to this is to consider two modes, both of which are treated quantum mechanically. Rather than measuring the absolute phase, we now wish to measure the phase difference between the two modes. The simplest way of considering this is via the Mach-Zehnder interferometer, as in Fig. 5.1.

With the Mach-Zehnder interferometer, an initial two-mode input state is fed into a beam splitter. The two beams are then subjected to phase shifts of Φ and φ, then recombined at a second beam splitter. The outputs of this beam splitter are then detected using photodetectors. The counts at the photodetectors are used to obtain an estimate of the phase difference.

These photodetectors will be assumed to be ideal for this analysis, though current photodetectors can only achieve a sensitivity of around 87% [44]. The photodetectors required for this need to be able to distinguish individual photons, in contrast to the large intensity photodetectors required for dyne detection. This is a more difficult task, and the efficiencies currently possible are correspondingly lower. In this study, however, we are concerned with what is possible in principle using foreseeable technologies. I will therefore be considering only the case of unit efficiency.

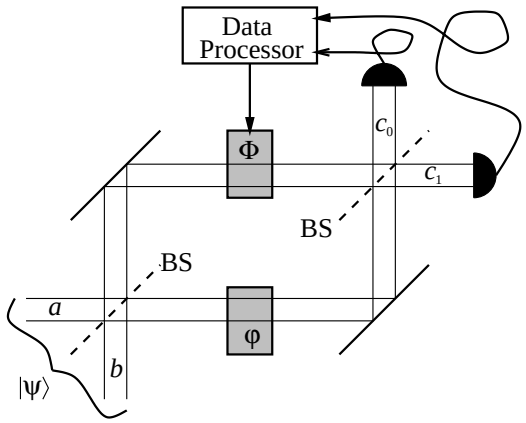

Figure 5.1: The Mach-Zehnder interferometer, with the addition of a controllable phase Φ in one arm. The unknown phase to be estimated is φ. Both beam splitters (BS) are 50/50.

In general the interferometer is used to measure the phase difference between the two arms. Similarly to the single mode case, we wish to add an auxiliary phase shift Φ in order to obtain a more accurate phase measurement. For simplicity I have indicated the phase we wish to measure, φ, as the phase shift in one arm in Fig. 5.1. In the following analysis this phase can equivalently be taken to be the phase difference between the two arms in the absence of the auxiliary phase shift Φ.

It is well known that it is possible to obtain a phase uncertainty scaling as $1/\sqrt{N}$ (the standard quantum limit) when a photon number state with N photons is input to one port of the interferometer. Several authors [1, 39, 40, 41, 42] have proposed ways of reducing the phase uncertainty to the Heisenberg limit of $1/N$. Most of these proposals [1, 39, 40] are essentially for detecting small deviations from some known phase, and therefore only give a $1/N$ scaling for a very small range of phases. Sanders and Milburn [41, 42] considered ideal measurements, that give $1/N$ scaling independent of the phase. Unfortunately they do not discuss how these measurements can be achieved in practice, and as will be shown in Ch. 6 it is not possible to perform these measurements in general, even allowing feedback.

We wish to perform measurements as close as possible to ideal by varying Φ during the measurement, based on the detections. In general there are three different areas available for optimisation:
1. The initial input state.
2. How the feedback phase is changed during the measurement.
3. The final phase estimate.

In this chapter I discuss the optimum initial input states, and in the next chapter the feedback phase and final phase estimate are discussed.

5.2 Optimum Input States

The input states are most conveniently described using the Schwinger representation. The operators in this representation are

$$\begin{aligned}\hat{J}_x &= (a^\dagger b + a b^\dagger)/2,\\ \hat{J}_y &= (a^\dagger b - a b^\dagger)/2i,\\ \hat{J}_z &= (a^\dagger a - b^\dagger b)/2,\\ \hat{J}^2 &= \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2.\end{aligned} \tag{5.1}$$

The operators $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ satisfy the commutation relations for the Lie algebra of SU(2):

$$\begin{aligned}\left[\hat{J}_x, \hat{J}_y\right] &= i\hat{J}_z,\\ \left[\hat{J}_y, \hat{J}_z\right] &= i\hat{J}_x,\\ \left[\hat{J}_z, \hat{J}_x\right] &= i\hat{J}_y.\end{aligned} \tag{5.2}$$

The operator $\hat{J}^2$ is the Casimir invariant for this group (i.e. it commutes with all elements). These commutation relations are the same as those for the operators for components of angular momentum.

I will use the notation $|j\mu\rangle_z$ for the common eigenstate of $\hat{J}_z$ and $\hat{J}^2$, with eigenvalues of μ and $j(j+1)$ respectively. This corresponds to Fock number eigenstates at the interferometer inputs with photon numbers in ports a and b of $j+\mu$ and $j-\mu$ respectively. Similarly the notation $|j\mu\rangle_y$ means the common eigenstate of $\hat{J}_y$ and $\hat{J}^2$ with eigenvalues of μ and $j(j+1)$. This state corresponds to number eigenstates in the interferometer arms with photon numbers of $j+\mu$ and $j-\mu$. To see this, note that the annihilation operators for the modes in the two interferometer arms are

$$(a+ib)/\sqrt{2}, \quad (ia+b)/\sqrt{2}. \tag{5.3}$$

It is simple to show from this that the operator for the difference in the photon numbers beween the arms is

$$\tfrac{1}{2}(-ia^\dagger + b^\dagger)(ia+b) - \tfrac{1}{2}(a^\dagger - ib^\dagger)(a+ib) = (a^\dagger b - a b^\dagger)/i, \tag{5.4}$$

which is twice $\hat{J}_y$.

Note that the scattering matrix used here for the beam splitter is

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \tag{5.5}$$

In contrast, the scattering matrix considered in the case of dyne measurements was

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \tag{5.6}$$

These scattering matrices have been used for consistency with previously published work. The only important consequence of this difference is that for dyne measurements, small errors are obtained when the difference between the signal and local oscillator phases is $\pi/2$, whereas in the interferometer case the phase difference should be close to zero. This is because, in the interferometer case, the i in the scattering matrix gives a $\pi/2$ phase shift.

In order to represent a completely general state we can express it as a sum of input number states:

$$|\psi\rangle = \sum_{2j=0}^{\infty} \sum_{\mu=-j}^{j} \psi_{j\mu} |j\mu\rangle_z. \tag{5.7}$$

Most proposals for reducing the phase uncertainty to the $1/N$ limit consider only input states with a fixed total photon number of $N = 2j$. This restriction is also applied in this study, as it greatly simplifies the analysis. With this restriction, the input state can be represented as

$$|\psi\rangle = \sum_{\mu=-j}^{j} \psi_{\mu} |j\mu\rangle_z. \tag{5.8}$$

Similarly to the single mode case, the probability distribution for the estimate of the phase of a two mode input state, $\hat{\varphi}$, is in general given by

$$P(\hat{\varphi}) = \mathrm{Tr}\left[\rho F(\hat{\varphi})\right], \tag{5.9}$$

where $F(\hat{\varphi})$ is the POM for the measurement. This only depends on the intrinsic phase of the two mode state, and not the phase shift in the interferometer. The interferometer transforms the input state $|\psi\rangle$ to $\hat{I}(\varphi)|\psi\rangle$, where φ is the phase in the interferometer and

$$\hat{I}(\varphi) = \exp\left(-i\varphi \hat{J}_y\right). \tag{5.10}$$

The state matrix therefore transforms to

$$\rho' = \hat{I}(\varphi)\rho\hat{I}^{\dagger}(\varphi). \tag{5.11}$$

When this transformed state is used in Eq. (5.9) the probability distribution will be dependent on the interferometer phase. This transformation acts to shift the phase of the state by φ, so the probability distribution will depend on the sum of the phase of the input state and the interferometer phase shift.

We can alternatively include the interferometer in the POM, and use the input state in Eq. (5.9). Then we would transform the POM to

$$F'(\hat{\varphi}) = \hat{I}^{\dagger}(\varphi) F(\hat{\varphi}) \hat{I}(\varphi). \tag{5.12}$$

If the POM describes a shift invariant phase measurement, then this can be simplified to

$$\begin{aligned} F'(\hat{\varphi}) &= F(\hat{\varphi} - \varphi) \\ &= F(\phi), \end{aligned} \tag{5.13}$$

where ϕ is the error in the phase estimate. Alternatively ϕ can be taken to be the phase estimate when the interferometer phase shift is zero.

Sanders and Milburn [41, 42] considered what they call "optimal" measurements, where the POM is given by

$$F(\phi) = \frac{2j+1}{2\pi}|j\phi\rangle\langle j\phi|, \tag{5.14}$$

where the $|j\phi\rangle$ are normalised phase states given by

$$|j\phi\rangle = (2j+1)^{-1/2}\sum_{\mu=-j}^{j} e^{i\mu\phi}|j\mu\rangle_y. \tag{5.15}$$

As I will show below, this POM is equivalent to the ideal or canonical POM for a single mode. For this reason I will generally use the same terminology for this POM, and reserve the word "optimal" for the best possible measurements that are realisable using photodetection and feedback.

When expressed in terms of the eigenstates of $\hat{J}_y$, the canonical POM is

$$F(\phi) = \frac{1}{2\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\mu-\nu)\phi}|j\mu\rangle_y\langle j\nu|. \tag{5.16}$$

This is very similar to the POM in the single mode case when there is an upper limit of $N = 2j$ on the photon number. This POM is not given explicitly above; however, it can be obtained by taking the single mode canonical POM given by Eq. (1.36), and limiting the sum to N. Explicitly, this POM is

$$F^{\rm can}(\phi) = \frac{1}{2\pi}\sum_{n,m=0}^{N} e^{i(n-m)\phi}|n\rangle\langle m|. \tag{5.17}$$

To show that the interferometer POM of Eq. (5.14) is completely equivalent to this, we can make the change of variables

$$\begin{aligned} \mu' &= \mu + j \\ \nu' &= \nu + j \\ |\mu'\rangle &= |j\mu\rangle_y. \end{aligned} \tag{5.18}$$

With this change in notation, the POM is

$$F(\phi) = \frac{1}{2\pi}\sum_{\mu',\nu'=0}^{N} e^{i(\mu'-\nu')\phi}|\mu'\rangle\langle\nu'|. \tag{5.19}$$

This POM is identical to the POM for the single mode case for an upper limit on the photon number of N. This case was considered in Sec. 2.1.1, and the state that minimises the Holevo phase variance is

$$|\psi_{\rm opt}\rangle = \frac{1}{\sqrt{j+1}}\sum_{\mu'=0}^{2j}\sin\left[\frac{(\mu'+1)\pi}{2j+2}\right]|\mu'\rangle. \tag{5.20}$$

Converting the notation back to the original variables, this state is

$$|\psi_{\rm opt}\rangle = \frac{1}{\sqrt{j+1}}\sum_{\mu=-j}^{j}\sin\left[\frac{(\mu+j+1)\pi}{2j+2}\right]|j\mu\rangle_y. \tag{5.21}$$

The minimum Holevo phase variance corresponding to this state is

$$V(\phi) = \tan^2\left(\frac{\pi}{N+2}\right). \tag{5.22}$$

The states $|j\mu\rangle_y$ correspond to joint number states within the interferometer, and they do not correspond to input number states in a simple way. We therefore wish to re-express the optimum states in terms of the input number states $|j\mu\rangle_z$. To do this, we require the relation from Ref. [41],

$$ {}_y\langle j\mu|j\nu\rangle_z = e^{i(\pi/2)(\nu-\mu)} I^j_{\mu\nu}(\pi/2), \tag{5.23}$$

where $I^j_{\mu\nu}(\pi/2)$ are the interferometer matrix elements given by

$$ I^j_{\mu\nu}(\pi/2) = 2^{-\mu}\left[\frac{(j-\mu)!}{(j-\nu)!}\frac{(j+\mu)!}{(j+\nu)!}\right]^{1/2} P^{(\mu-\nu,\mu+\nu)}_{j-\mu}(0), \tag{5.24}$$

for

$$ \mu-\nu \geq 0, \quad \mu+\nu \geq 0, \tag{5.25}$$

where $P^{(\alpha,\beta)}_n$ are the Jacobi polynomials, given by

$$ P^{(\alpha,\beta)}_n(x) = 2^{-n}\sum_{m=0}^{n}\binom{n+\alpha}{m}\binom{n+\beta}{n-m}(x-1)^{n-m}(x+1)^m. \tag{5.26}$$

Therefore the explicit expression for determining the values of $I^j_{\mu\nu}(\pi/2)$ is

$$\begin{aligned} I^j_{\mu\nu}(\pi/2) &= 2^{-j}\left[\frac{(j-\mu)!}{(j-\nu)!}\frac{(j+\mu)!}{(j+\nu)!}\right]^{1/2}\sum_{m=0}^{j-\mu}\binom{j-\nu}{m}\binom{j+\nu}{j-\mu-m}(-1)^{j-\mu-m} \\ &= 2^{-j}\left[(j-\mu)!(j+\mu)!(j-\nu)!(j+\nu)!\right]^{1/2} \\ &\quad\times\sum_{m=0}^{j-\mu}\left[m!(j-\nu-m)!(j-\mu-m)!(\mu+\nu+m)!\right]^{-1}(-1)^{j-\mu-m}. \end{aligned} \tag{5.27}$$

The values of $I^j_{\mu\nu}(\pi/2)$ for values of μ and ν that do not obey the inequalities (5.25) can be obtained using the symmetry relations

$$ I^j_{\mu\nu}(\theta) = (-1)^{\mu-\nu} I^j_{\nu\mu}(\theta) = I^j_{-\nu,-\mu}(\theta). \tag{5.28}$$

Using the result from Eq. (5.23), the state expressed in terms of the eigenstates of $\hat{J}_z$ is

$$ |\psi_{\rm opt}\rangle = \frac{1}{\sqrt{j+1}}\sum_{\mu,\nu=-j}^{j}\sin\left[\frac{(\mu+j+1)\pi}{2j+2}\right]e^{i(\pi/2)(\mu-\nu)} I^j_{\mu\nu}(\pi/2)|j\nu\rangle_z. \tag{5.29}$$

An example of an optimum state for 40 photons calculated using this formula is shown in Fig. 5.2. As can be seen, the only significant contributions are from 9 or 10 $\hat{J}_z$ eigenstates near $\mu = 0$. In addition, the distribution near the centre is fairly independent of photon number, as can be seen by comparing this state with the state for 1200 photons, also shown in Fig. 5.2. For larger values of μ, the contributions fall approximately exponentially with μ, as can be seen in Fig. 5.3.

In Ref. [39] it is shown that it is possible to obtain states similar to $|j0\rangle_z$ and $|j1\rangle_z$ using a two-mode four-wave mixer. As the optimum input states have their main contributions from these states and other $\hat{J}_z$ eigenstates near $\mu = 0$, this suggests that it may be possible to produce states that are close approximations of the optimum input states using a suitable modification of the apparatus used in Ref. [39]. Unfortunately, rather than producing only the state $|j0\rangle_z$ (for example), the four-wave mixer produces a superposition of these states with a range of values of j.

In [39] the authors state that the value of j can be inferred after the measurement from the number of photons detected. This would not be appropriate for the measurements considered in the next chapter, as these rely on knowing the value of j before the measurement starts. In the limit of large photon number the spread in the values of j will be small compared to the mean, so it may be possible to obtain good measurements using the mean value of j. Unfortunately, the measurement scheme considered in the next chapter is computationally infeasible in the limit of large j.

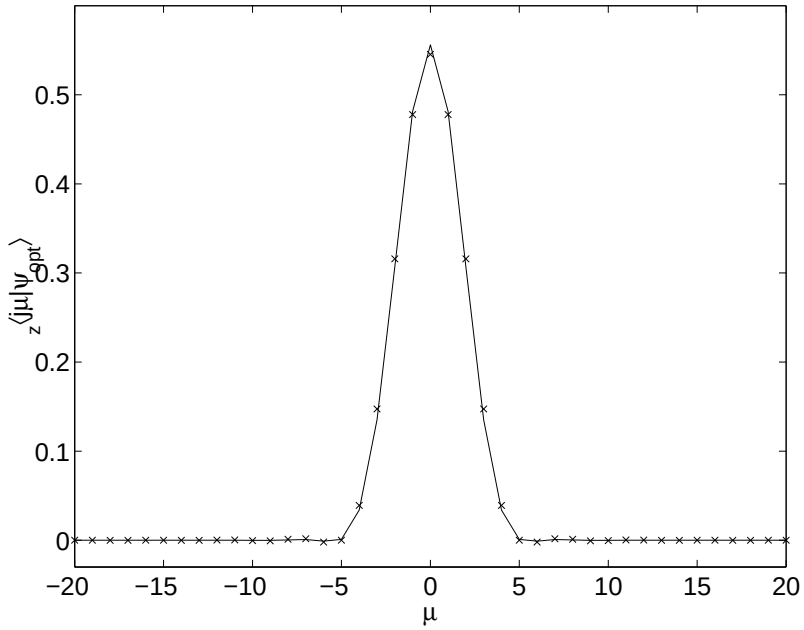

Figure 5.2: The coefficients ${}_z\langle j\mu|\psi_{\rm opt}\rangle$ for the state optimised for minimum phase variance under canonical measurements. All coefficients for a photon number of $2j = 40$ are shown as the continuous line, and those near $\mu = 0$ for a photon number of 1200 as crosses.

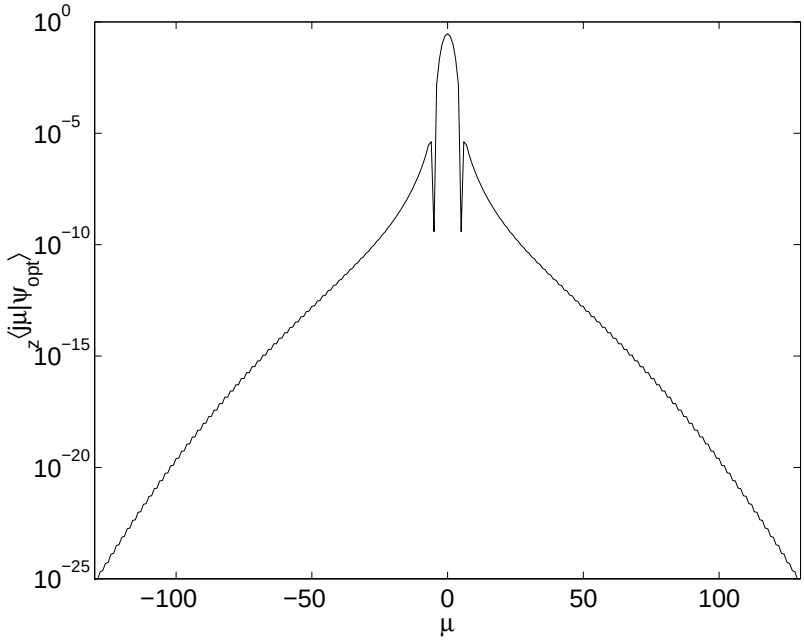

Figure 5.3: The coefficients ${}_z\langle j\mu|\psi_{\rm opt}\rangle$ for larger values of μ for the state optimised for minimum phase variance under canonical measurements for a photon number of 1200.

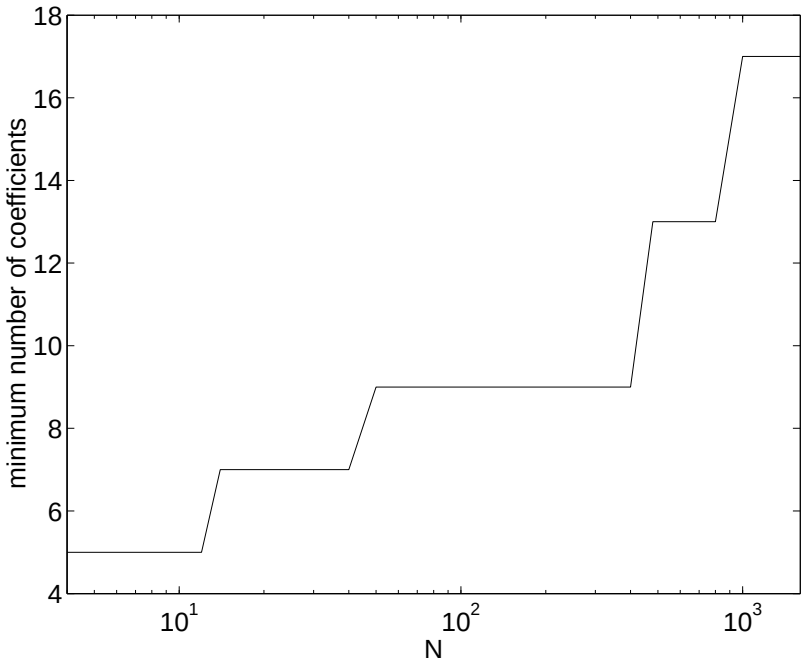

Figure 5.4: The minimum number of $\hat{J}_z$ eigenstates required to approximate the optimum state in order to obtain a phase variance less than twice optimum.

Nevertheless, it should be simpler to produce a small number of $\hat{J}_z$ eigenstates than the entire range. In order to estimate about how many $\hat{J}_z$ eigenstates are required to provide a reasonable approximation of the optimum states for a given photon number, the coefficients for the optimum states were determined, and all except a number of coefficients near $\mu = 0$ were discarded. These remaining coefficients were then normalised, and the phase variance of the resulting state was determined.

In Fig. 5.4 I have plotted the number of $\hat{J}_z$ eigenstates required to approximate the optimum states for a variety of total photon numbers. The criterion used was that the phase variance of the approximate state be less than twice that of the exact state. For photon numbers up to about 400, only 9 $\hat{J}_z$ eigenstates are required, but beyond that the number required increases fairly rapidly, with 17 required for 1600 photons.

5.3 Phase Variances for Other States

I will now consider the phase variances of some other commonly considered states in order to compare them with the optimum states. These states will be expressed in terms of the input photon number eigenstates:

$$|\psi\rangle = \sum_{\mu=-j}^{j} \psi_\mu |j\mu\rangle_z. \tag{5.30}$$

The probability distribution for the phase for these states will be given by

$$\begin{aligned} P(\phi) &= \mathrm{Tr}\left[|\psi\rangle\langle\psi| F(\phi)\right] \\ &= \frac{2j+1}{2\pi} \mathrm{Tr}\left[|\psi\rangle\langle\psi|j\phi\rangle\langle j\phi|\right] \\ &= \frac{2j+1}{2\pi} \left|\langle j\phi|\psi\rangle\right|^2. \end{aligned} \tag{5.31}$$

The inner product, when expressed in terms of the coefficients ψ_μ, is given by

$$\begin{aligned}\langle j\phi|\psi\rangle &= \sum_{\nu=-j}^{j} \psi_\nu \langle j\phi|j\nu\rangle_z \\ &= (2j+1)^{-1/2} \sum_{\nu=-j}^{j} \psi_\nu \sum_{\mu=-j}^{j} e^{-i\mu\phi}{}_y\langle j\mu|j\nu\rangle_z \\ &= (2j+1)^{-1/2} \sum_{\nu=-j}^{j} \psi_\nu \sum_{\mu=-j}^{j} e^{-i\mu\phi} e^{i(\pi/2)(\nu-\mu)} I^j_{\mu\nu}(\pi/2). \end{aligned} \tag{5.32}$$

The state considered in Refs [40, 41, 42] was $|j0\rangle_z$, the state with equal photon numbers in both input ports. This state is the biggest contributor to the optimum states, so it is not unreasonable that this state should have a small phase uncertainty. This state suffers the drawback that it has equal peaks at 0 and π, and therefore must be considered modulo π in order to obtain meaningful results. If we add the state $|j1\rangle_z$ (the next biggest contributor to the optimum state), then we obtain a state for which we can consider the phase modulo 2π. This state, $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$, was considered in Ref. [39].

For the state $|j0\rangle_z$, the only non-zero coefficient is $\psi_0 = 1$, so the inner product is

$$\langle j\phi|j0\rangle_z = (2j+1)^{-1/2} \sum_{\mu=-j}^{j} e^{-i\mu(\phi+\pi/2)} I^j_{\mu 0}\left(\pi/2\right). \tag{5.33}$$

The probability distribution is then given by

$$\begin{aligned} P\left(\phi\right) &= \frac{1}{2\pi} \left| \sum_{\mu=-j}^{j} e^{-i\mu(\phi+\pi/2)} I^j_{\mu 0}\left(\pi/2\right) \right|^2 \\ &= \frac{1}{2\pi} \sum_{\mu,\nu=-j}^{j} e^{-i(\mu-\nu)(\phi+\pi/2)} \left(I^j_{\nu 0}\left(\pi/2\right)\right)^* I^j_{\mu 0}\left(\pi/2\right). \end{aligned} \tag{5.34}$$

Because this state has equal peaks at 0 and π, the usual definition of the Holevo phase variance will give infinite results. Rather than using the usual definition, however, we can use a modified definition that is naturally modulo π. The definition that I will use is

$$V_\pi(\phi) = \left(\left|\left\langle e^{2i\phi}\right\rangle\right|^{-2} - 1\right)\Big/ 4. \tag{5.35}$$

From Eq. (5.34) the value of $\left|\left\langle e^{2i\phi}\right\rangle\right|$ can be determined as

$$\begin{aligned} \left|\left\langle e^{2i\phi}\right\rangle\right| &= \left| \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{\mu,\nu=-j}^{j} e^{i(\nu-\mu+2)\phi} e^{i(\nu-\mu)\pi/2} \left(I^j_{\nu 0}\left(\pi/2\right)\right)^* I^j_{\mu 0}\left(\pi/2\right) d\phi \right| \\ &= \left| \sum_{\mu,\nu=-j}^{j} \delta_{\mu,\nu+2} \left(I^j_{\nu 0}\left(\pi/2\right)\right)^* I^j_{\mu 0}(\pi/2) \right| \\ &= \left| \sum_{\mu=2-j}^{j} \left(I^j_{\mu-2,0}\left(\pi/2\right)\right)^* I^j_{\mu 0}\left(\pi/2\right) \right|. \end{aligned} \tag{5.36}$$

This can be used in Eq. (5.35) to determine the phase variance under this modified definition.

Similarly, for the state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$, we have $\psi_0 = \psi_1 = 1/\sqrt{2}$, so the inner product is

$$\langle j\phi| \left((|j0\rangle_z + |j1\rangle_z)/\sqrt{2}\right) = \frac{1}{\sqrt{2(2j+1)}} \sum_{\mu=-j}^{j} e^{-i\mu(\phi+\pi/2)} \left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}\left(\pi/2\right)\right]. \tag{5.37}$$

The probability distribution is therefore

$$P(\phi) = \frac{1}{4\pi}\left|\sum_{\mu=-j}^{j} e^{-i\mu(\phi+\pi/2)}\left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}(\pi/2)\right]\right|^2$$
$$= \frac{1}{4\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\nu-\mu)(\phi+\pi/2)}\left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}(\pi/2)\right]\left[I^j_{\nu 0}(\pi/2) + iI^j_{\nu 1}(\pi/2)\right]^*. \quad (5.38)$$

Using this expression for the probability distribution we find

$$\left|\left\langle e^{i\phi}\right\rangle\right| = \left|\int_{-\pi}^{\pi}\frac{1}{4\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\nu-\mu+1)(\phi+\pi/2)}e^{-i\pi/2}\left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}(\pi/2)\right]\left[I^j_{\nu 0}(\pi/2) + iI^j_{\nu 1}(\pi/2)\right]^* d\phi\right|$$
$$= \frac{1}{2}\left|\sum_{\mu,\nu=-j}^{j}\delta_{\mu,\nu+1}\left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}(\pi/2)\right]\left[I^j_{\nu 0}(\pi/2) + iI^j_{\nu 1}(\pi/2)\right]^*\right|$$
$$= \frac{1}{2}\left|\sum_{\mu=1-j}^{j}\left[I^j_{\mu 0}(\pi/2) + iI^j_{\mu 1}(\pi/2)\right]\left[I^j_{\mu-1,0}(\pi/2) + iI^j_{\mu-1,1}(\pi/2)\right]^*\right|. \quad (5.39)$$

This result was used to determine the Holevo phase variance for this state. The third state that will be considered that does not have a simple solution for the Holevo phase variance is that with all photons in one port, $|jj\rangle_z$. For this state the only non-zero coefficient is $\psi_j = 1$, so

$$\langle j\phi|jj\rangle_z = (2j+1)^{-1/2}\sum_{\mu=-j}^{j} e^{-i\mu\phi}e^{i(\pi/2)(j-\mu)}I^j_{\mu j}(\pi/2). \quad (5.40)$$

The probability distribution is therefore

$$P(\phi) = \frac{1}{2\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\nu-\mu)(\phi+\pi/2)}\left(I^j_{\nu j}(\pi/2)\right)^* I^j_{\mu j}(\pi/2), \quad (5.41)$$

and the value of $\left|\left\langle e^{i\phi}\right\rangle\right|$ is

$$\left|\left\langle e^{i\phi}\right\rangle\right| = \left|\sum_{\mu=1-j}^{j}\left(I^j_{\mu-1,j}(\pi/2)\right)^* I^j_{\mu j}(\pi/2)\right|. \quad (5.42)$$

The Holevo phase variance was calculated for the above three states using these expressions, for photon numbers up to 25600, and the results are shown in Fig. 5.5. I have also included the analytic expression for the phase variance for optimised states.

The phase variance for $|j0\rangle_z$ scales down with photon number much more slowly than the phase variance for optimal states, and in fact even more slowly than the phase variance for $|jj\rangle_z$, which scales as N^{-1}. In fact, for the range of photon numbers considered the phase variance scales as $N^{-1/2}$. This would seem to imply a phase uncertainty scaling as $N^{-1/4}$, in dramatic contrast to the N^{-1} scaling found in [40, 41, 42]. The state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ is even worse, with a phase variance that does not scale down with photon number at all, and in fact never falls below 1.

Another unusual feature of the graph is that the phase variance for $|j0\rangle_z$ is even smaller than that for optimum states for very small photon numbers. This is not in fact a contradiction, because different measures of the uncertainty are used for the two states. For very small photon numbers, the phase probability distribution for optimum states is significant for phases beyond $\pm\pi/2$. The measure in Eq. (5.35) that is used for the phase variance for $|j0\rangle_z$ effectively ignores the distribution beyond $\pm\pi/2$, so under this measure, this state has a correspondingly lower phase variance.

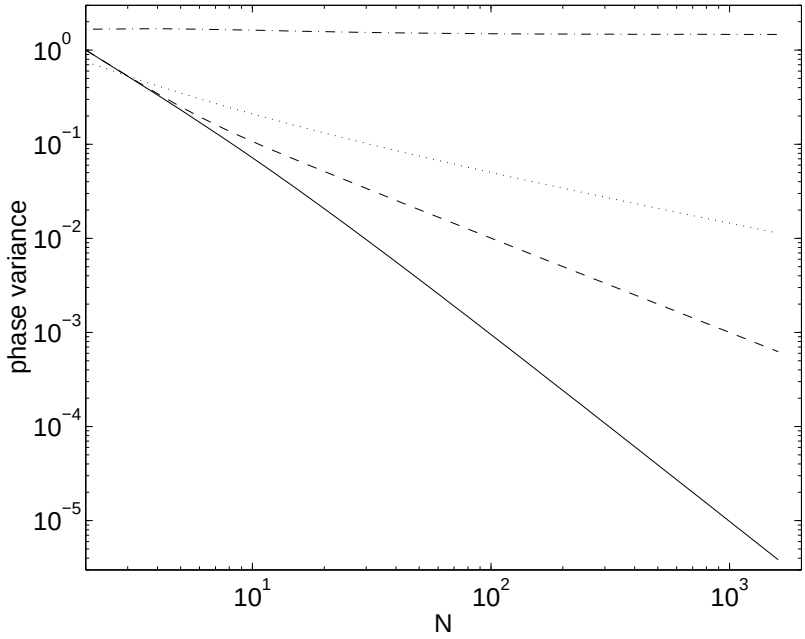

Figure 5.5: The canonical Holevo phase variance versus input photon number $2j$. The continuous line is for optimum states $|\psi_{\rm opt}\rangle$, the dashed line is for all photons in one input port $|jj\rangle_z$, the dotted line is for equal photon numbers in both ports $|j0\rangle_z$, and the dash-dotted line is for the state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$.

The reason for the discrepancies between the results obtained here for the states $|j0\rangle_z$ and $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$, and those obtained in [40, 41, 42], is that the results in [40, 41, 42] are all based on the width of the central peak in the distribution. In contrast, the Holevo phase variance for these states is primarily due to the tails. To demonstrate this for the state $|j0\rangle_z$, in Fig. 5.6 I have plotted the phase distribution multiplied by $\sin^2\phi$.

The reason for multiplying by a factor of $\sin^2\phi$ is that

$$\begin{aligned} V_\pi(\phi) &= \left(\left|\left\langle e^{2i\phi}\right\rangle\right|^{-2} - 1\right)\Big/4 \\ &= \left(\langle\cos(2\phi)\rangle^{-2} - 1\right)\Big/4 \\ &= \left(\left\langle 1 - 2\sin^2\phi\right\rangle^{-2} - 1\right)\Big/4 \\ &= \left[\left(1 - 2\left\langle\sin^2\phi\right\rangle\right)^{-2} - 1\right]\Big/4 \\ &\approx \left[\left(1 + 4\left\langle\sin^2\phi\right\rangle\right) - 1\right]/4 \\ &= \left\langle\sin^2\phi\right\rangle. \end{aligned} \tag{5.43}$$

The above approximation is accurate for small phase variance. This derivation also uses the fact that the phase distribution for this state is unbiased, so $\left\langle e^{2i\phi}\right\rangle$ is real. Note that in this form the variance is very similar to the standard variance, $\left\langle\phi^2\right\rangle$. Since $\phi^2 \geq \sin^2\phi$, the tails of the distribution are even more significant for the standard variance, as illustrated in Fig. 5.7.

There is a similar problem for the state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$. For this state there are peaks at $\pm\pi$, as shown in Fig. 5.8. Although these peaks are smaller than the main peak at 0, they do not get smaller with photon number. This means that the Holevo phase variance is almost entirely due to these peaks, and therefore does not decrease with photon number. As this state is so poor, I will not consider it further, and restrict attention to $|j0\rangle_z$ and the optimum states.

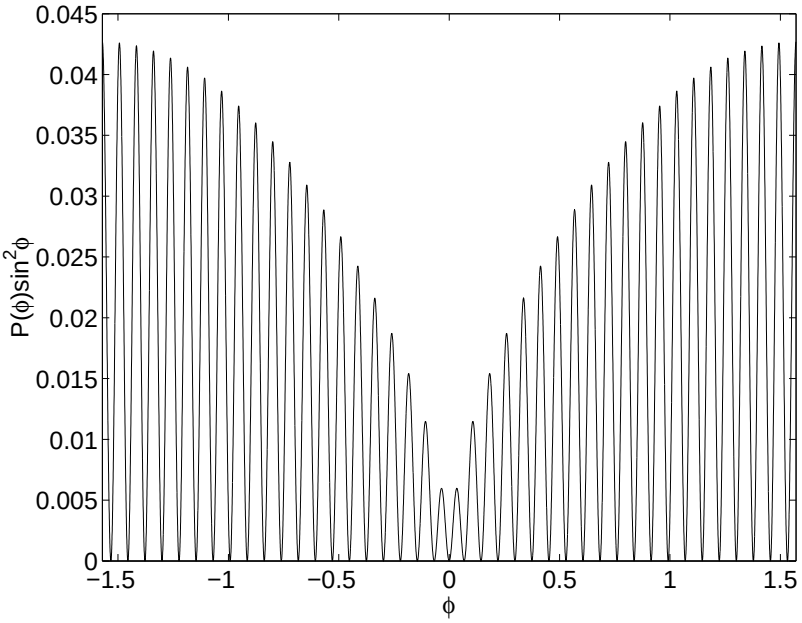

Figure 5.6: The canonical phase probability distribution, multiplied by $\sin^2\phi$, for the state $|j0\rangle_z$ with $2j = 80$ photons.

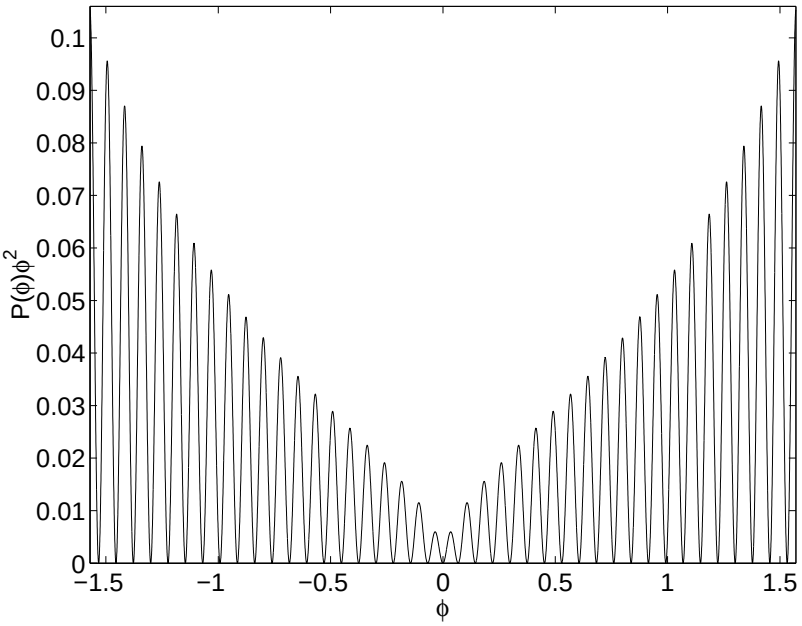

Figure 5.7: The canonical phase probability distribution, multiplied by ϕ^2, for the state $|j0\rangle_z$ with $2j = 80$ photons.

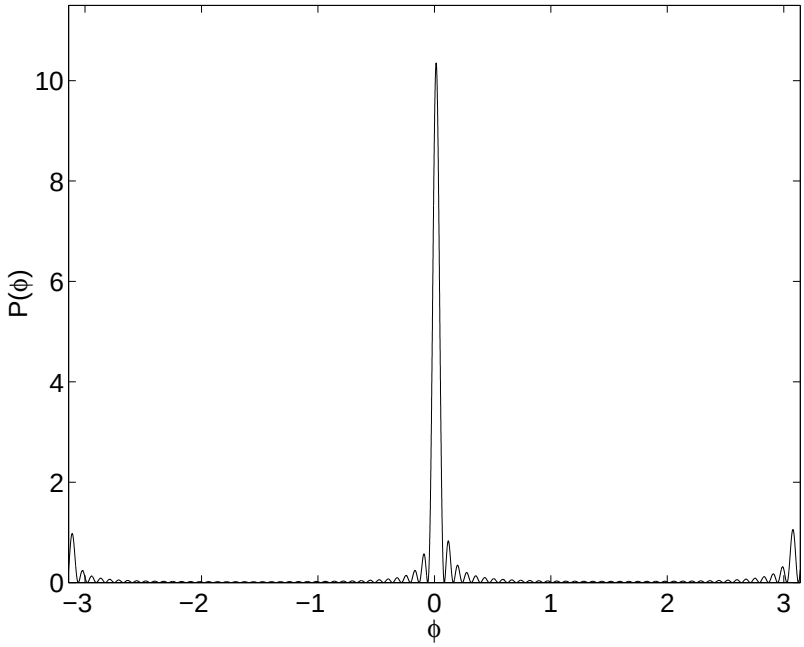

Figure 5.8: The canonical phase probability distribution for $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ for $2j = 80$ photons.

The asymptotic expression for the phase probability distribution for $|j0\rangle_z$ given in [41] is

$$P_j(\phi) = \frac{2j+1}{2\pi} \frac{\left[\Gamma(3/4)\right]^2}{2^{3/2}} \frac{\left[J_{1/4}(j\phi)\right]^2}{\sqrt{j\phi}}. \tag{5.44}$$

This equation is approximately half what it should be in order to be normalised. The reason for this is that it approximates the exact distribution over the interval $[-\pi/2, \pi/2]$, but the exact distribution is normalised over $[-\pi, \pi]$, and repeats modulo π. Since Eq. (5.44) only approximates the distribution over the region $[-\pi/2, \pi/2]$, its integral over this region will be approximately $\frac{1}{2}$. Therefore the expression that I will be using is

$$P_j(\phi) = \frac{2j+1}{2\pi} \frac{\left[\Gamma(3/4)\right]^2}{\sqrt{2}} \frac{\left[J_{1/4}(|j\phi|)\right]^2}{\sqrt{|j\phi|}}. \tag{5.45}$$

This expression is correctly normalised in the limit as j goes to infinity. I have also added absolute value signs so that this expression is correct for negative values of ϕ.

Now for large $j\phi$ there is the approximate proportionality [46]

$$J_{1/4}(|j\phi|) \propto 1/\sqrt{|j\phi|}. \tag{5.46}$$

This implies that for large ϕ

$$P_j(\phi) \propto \frac{1}{|\phi|^{3/2}}. \tag{5.47}$$

This means that we should have

$$P_j(\phi)\phi^2 \propto |\phi|^{1/2}. \tag{5.48}$$

It can be seen in Fig. 5.7 that the scaling is nothing like this, and is closer to

$$P_j(\phi)\phi^2 \propto |\phi|. \tag{5.49}$$

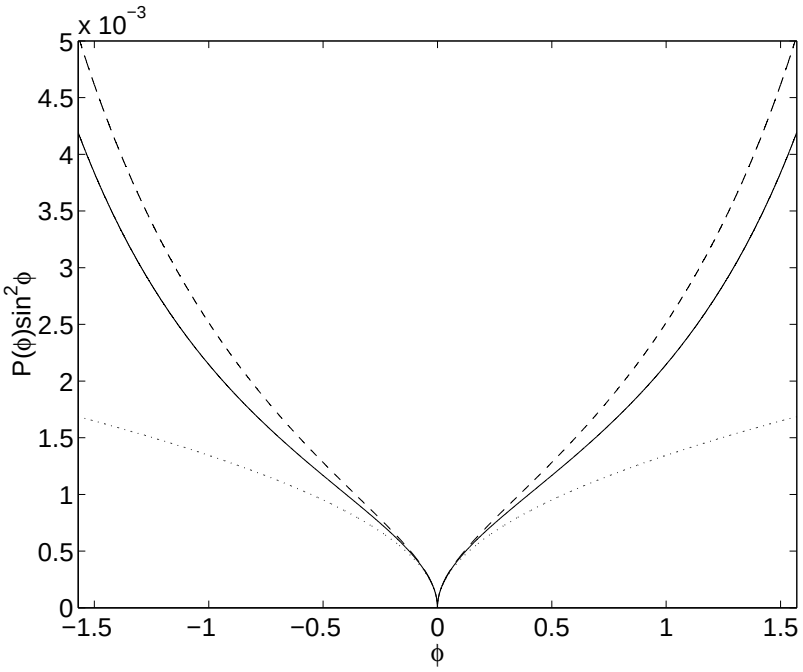

Figure 5.9: The canonical phase probability distribution envelope, multiplied by $\sin^2\phi$, for the state $|j0\rangle_z$ with $2j = 51200$ photons, calculated using three different methods. The exact calculation is indicated by the continuous line, the first approximation (5.51) is shown as the dashed line, and the Bessel function approximation (5.45) is shown as the dotted line.

To see where this discrepancy originates, consider the intermediate approximation made in Ref. [41]:

$$\langle j\phi|j0\rangle_z \approx \frac{(-1)^{j/2}}{\sqrt{2j+1}}\sqrt{\frac{2}{\pi}}\sum_{\mu=-j/2}^{j/2}\frac{e^{-2i\mu\phi}}{\left[j(j+1)-(2\mu)^2\right]^{1/4}}. \tag{5.50}$$

Using this approximation the probability distribution is

$$P(\phi) \approx \frac{2}{\pi^2}\left|\sum_{\mu=-j/2}^{j/2}\frac{e^{-2i\mu\phi}}{\left[j(j+1)-(2\mu)^2\right]^{1/4}}\right|^2. \tag{5.51}$$

Here I have multiplied by a factor of 2 so the probability distribution is normalised over the interval $[-\pi/2, \pi/2]$, for consistency with Eq. (5.45). In Ref. [41] the sum is then approximated by an integral in order to obtain the Bessel function approximation.

In order to see where the approximation is deviating from the exact expression, in Fig. 5.9 I have plotted the probability distribution multiplied by ϕ^2 for the exact expression, the approximation (5.51), and the Bessel function approximation (5.45). For the exact expression I have used twice Eq. (5.34), so that the distribution is normalised over the interval $[-\pi/2, \pi/2]$. The functions are very rapidly oscillating, so to make the three curves legible only the peaks are plotted.

Near $\phi = 0$ the three expressions for the distribution give very similar results, but there are large differences in the tails. As can be seen in the figure, the first approximation (5.51) has tails that are fairly close to the exact expression, but still noticeably higher. In contrast, the tails for the Bessel function approximation are much different, and *lower* than the exact expression. These results indicate that it is not primarily the initial approximation for $I^j_{\mu 0}(\pi/2)$ that is giving the incorrect scaling for the tails, but the approximation where the sum is approximated by an integral.

On the other hand, if we look at the results close to the centre of the distribution, we find that there is very good agreement between the curves. To illustrate this, the distribution near the centre

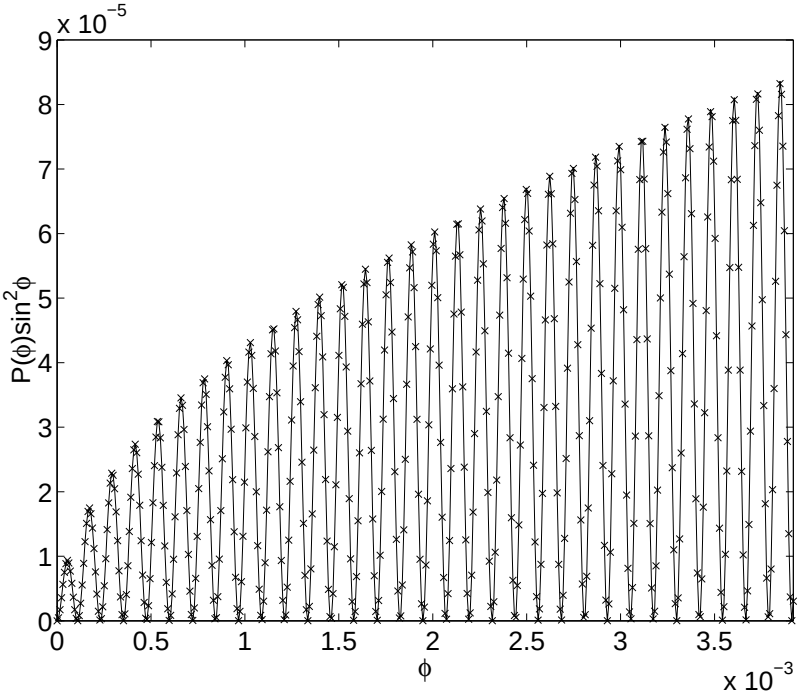

Figure 5.10: The canonical phase probability distribution, multiplied by $\sin^2\phi$, for the state $|j0\rangle_z$ with $2j = 51200$ photons, calculated using two different methods and restricted to a range of small ϕ. The exact calculation is indicated by the continuous line, and the Bessel function approximation (5.45) is shown as the crosses.

multiplied by $\sin^2\phi$ is plotted in Fig. 5.10. For large photon numbers there is good agreement over a region that is large compared to the central peak, but the agreement is always poor for phases that are significant compared to $\pi/2$.

In the light of these results, any results based on the Bessel function approximation should be treated very carefully if they depend on the distribution for large phases. For example, the Bessel function approximation can be used to show that the Holevo phase variance should scale as $N^{-1/2}$ (as stated in [54]), but it does not give the correct scaling constant. From Eq. (5.43) the Holevo phase variance is given approximately by

$$\left\langle \sin^2\phi \right\rangle = \int_{-\pi/2}^{\pi/2} P(\phi)\sin^2\phi d\phi. \tag{5.52}$$

Note that this approximation is accurate provided that $\left\langle \sin^2\phi \right\rangle$ is small, but does not depend on the main part of the contribution being from small ϕ. Using the Bessel function approximation (5.45), we find that

$$\begin{aligned}\left\langle \sin^2\phi \right\rangle &= \frac{2j+1}{2\pi}\frac{\left[\Gamma(3/4)\right]^2}{\sqrt{2}}\int_{-\pi/2}^{\pi/2}\frac{\left[J_{1/4}\left(|j\phi|\right)\right]^2}{\sqrt{|j\phi|}}\sin^2\phi d\phi \\ &= \frac{2j+1}{\pi}\frac{\left[\Gamma(3/4)\right]^2}{\sqrt{2}}\int_{0}^{\pi/2}\frac{\left[J_{1/4}(j\phi)\right]^2}{\sqrt{j\phi}}\sin^2\phi d\phi. \end{aligned}\tag{5.53}$$

From Ref. [46], for large ϕ the Bessel function can be approximated by

$$J_{1/4}(j\phi) \approx \sqrt{\frac{2}{\pi}} \frac{\sin\left(j\phi + \pi/8\right)}{\sqrt{j\phi}}. \tag{5.54}$$

Therefore, since the majority of the contribution to the phase variance is from large ϕ, we have

$$\left\langle \sin^2\phi \right\rangle \approx \frac{2j+1}{\pi} \frac{\left[\Gamma\left(3/4\right)\right]^2}{\sqrt{2}} \int_0^{\pi/2} \frac{2}{\pi} \frac{\sin^2\left(j\phi + \pi/8\right)}{\left(j\phi\right)^{3/2}} \sin^2\phi d\phi. \tag{5.55}$$

Using the average value of $\sin^2(j\phi + \pi/8)$ and taking the limit of large j, this becomes

$$\begin{aligned} \left\langle \sin^2\phi \right\rangle &\approx \frac{2j}{\pi^2 j^{3/2}} \frac{\left[\Gamma(3/4)\right]^2}{\sqrt{2}} \int_0^{\pi/2} \frac{\sin^2\phi}{\phi^{3/2}} d\phi \\ &= \frac{2}{\pi^2} \frac{\left[\Gamma(3/4)\right]^2}{\sqrt{N}} \int_0^{\pi/2} \frac{\sin^2\phi}{\phi^{3/2}} d\phi. \end{aligned} \tag{5.56}$$

This demonstrates that the Bessel function approximation predicts that the Holevo phase variance should scale as $N^{-1/2}$. The scaling constant is given by

$$\begin{aligned} \frac{2}{\pi^2}\left[\Gamma(3/4)\right]^2 \int_0^{\pi/2} \frac{\sin^2\phi}{\phi^{3/2}} d\phi &= \frac{2}{\pi^2}\left[\Gamma(3/4)\right]^2 \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\pi^{2n}\sqrt{2/\pi}}{(2n)!(4n-1)} \\ &= 0.2845775946062444\ldots \end{aligned} \tag{5.57}$$

The above expression is easily evaluated by numerical evaluation of the integral, or by summing the first dozen or so terms of the sum. The scaling constant obtained is about 0.28; however, from the numerical results shown in Fig. 5.5, the actual scaling constant is 0.44.

Note that for the standard variance, $\left\langle \phi^2 \right\rangle$, we are simply replacing $\sin^2\phi$ with ϕ^2, so the scaling is again $N^{-1/2}$, except with a scaling constant of

$$\begin{aligned} \frac{2}{\pi^2}\left[\Gamma(3/4)\right]^2 \int_0^{\pi/2} \phi^{1/2} d\phi &= \frac{2}{\pi^2}\left[\Gamma(3/4)\right]^2 \left[\frac{\phi^{3/2}}{3/2}\right]_0^{\pi/2} \\ &= \sqrt{\frac{2}{\pi}} \frac{\left[\Gamma(3/4)\right]^2}{3} \\ &= 0.3993800782451976\ldots \end{aligned} \tag{5.58}$$

This scaling constant of about 0.40 is again different than the actual scaling constant of about 0.66 (the results demonstrating this scaling constant will be discussed below). It is significant that the scaling constants for the Holevo phase variance and the standard phase variance are different, because it demonstrates that the Holevo phase variance is not necessarily the same as the standard phase variance, even in the limit of very sharply peaked distributions. For the Holevo phase variance to be the same as the standard variance, the phase distribution must not only be narrowly peaked, but the tails must scale down rapidly enough that there is no significant contribution to the phase variance from large ϕ.

5.4 Other Measures of the Phase Uncertainty

In practice the high tails of the $|j0\rangle_z$ state mean that although most of the results of phase measurements will have small errors, scaling as N^{-1}, there will always be a significant number of results

with large phase errors. This means that we would need to be very careful analysing results obtained from this state. For example, if we take the mean of a large number of results obtained using this state, the error in the mean will scale as $N^{-1/4}$ rather than N^{-1}. In order to obtain results with error scaling as N^{-1}, we would need to use some more sophisticated data analysis technique. In contrast, because the optimum states derived here have a Holevo phase variance that scales as N^{-2}, we can use all the standard data analysis techniques and still get an error scaling as N^{-1}.

Another issue is that although the phase uncertainty for $|j0\rangle_z$ as indicated by the Holevo or standard phase variance does not scale as N^{-1}, it does scale as N^{-1} under other measures of uncertainty. The phase variance is a very stringent measure of uncertainty, and generally gives an upper limit on other measures of uncertainty. For example, there are the inequalities [55, 56]

$$\begin{aligned} \sqrt{2\pi e}\Delta\phi \geq L_{\mathrm{H}} \geq \sqrt{2\pi e}L_{\mathrm{F}} \\ \frac{\Delta\phi}{\sqrt{1-C}} \geq L_C, \end{aligned} \tag{5.59}$$

where $\Delta\phi$ is the square root of the variance, L_{H} is the entropic length, L_{F} is the Fisher length, and L_C is a $C \times 100\%$ confidence interval. These inequalities mean that if $\Delta\phi$ scales as N^{-1}, as is the case for optimum states, then the entropic length, Fisher length, and confidence intervals must also scale at least as N^{-1}.

Two other measures of the phase uncertainty are the reciprocal-of-peak-value, L_{rp}, and the Süssman measure, L_{S}. Specifically, the definitions for each of these measures are

$$\begin{aligned} \Delta\phi^2 &= \int_{-\pi}^{\pi} P(\phi)\left(\phi - \bar{\phi}\right)^2 d\phi \\ \log L_{\mathrm{H}} &= -\int_{-\pi}^{\pi} P(\phi)\log P(\phi)\, d\phi \\ L_{\mathrm{F}}^{-2} &= \int_{-\pi}^{\pi}\left[\frac{dP(\phi)}{d\phi}\right]^2 \frac{1}{P(\phi)} d\phi \\ 1 - C &= \int_{\bar{\phi}-L_C}^{\bar{\phi}+L_C} P(\phi)\, d\phi \\ L_{\mathrm{rp}}^{-1} &= \max\left[P(\phi)\right] \\ L_{\mathrm{S}}^{-1} &= \int_{-\pi}^{\pi}\left[P(\phi)\right]^2 d\phi \end{aligned} \tag{5.60}$$

where $\bar{\phi}$ is the mean phase, defined as

$$\bar{\phi} = \int_{-\pi}^{\pi} P(\phi)\phi d\phi. \tag{5.61}$$

In addition, there is the usual Holevo phase variance

$$V(\phi) = \left|\int_{-\pi}^{\pi} P(\phi)e^{i\phi}d\phi\right|^{-2} - 1. \tag{5.62}$$

Most of these measures are discussed in more detail in Ref. [57].

In order to calculate these measures, the full probability distribution for the measurement scheme is required. In the case of the optimum input state, the inner product is given by

$$\begin{aligned}\langle j\phi|\psi_{\rm opt}\rangle &= \frac{1}{\sqrt{j+1}}\sum_{\nu=-j}^{j}\sin\left[\frac{(\nu+j+1)\,\pi}{2j+2}\right]\langle j\phi|j\nu\rangle_y \\ &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\sum_{\mu,\nu=-j}^{j}\sin\left[\frac{(\nu+j+1)\pi}{2j+2}\right]e^{-i\mu\phi}{}_y\langle j\mu|j\nu\rangle_y \\ &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\sum_{\mu=-j}^{j}\sin\left[\frac{(\mu+j+1)\pi}{2j+2}\right]e^{-i\mu\phi}. \end{aligned} \tag{5.63}$$

Expressing the sine in terms of exponentials gives a sum that can be evaluated:

$$\begin{aligned}\langle j\phi|\psi_{\rm opt}\rangle &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\sum_{\mu=-j}^{j}\frac{1}{2i}\left\{\exp\left[i\frac{(\mu+j+1)\,\pi}{2j+2}\right]-\exp\left[-i\frac{(\mu+j+1)\,\pi}{2j+2}\right]\right\}e^{-i\mu\phi} \\ &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\frac{e^{i(j\phi)}}{2i}\left\{\sum_{\mu'=0}^{2j}\exp\left[i\frac{(\mu'+1)\pi}{2j+2}-i\mu'\phi\right]\right. \\ &\quad\left.-\sum_{\mu'=0}^{2j}\exp\left[-i\frac{(\mu'+1)\pi}{2j+2}-i\mu'\phi\right]\right\} \\ &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\frac{e^{i(j\phi)}}{2i}\left\{\sum_{\mu'=0}^{N}\exp\left[i\mu'\left(\frac{\pi}{N+2}-\phi\right)+i\frac{\pi}{N+2}\right]\right. \\ &\quad\left.-\sum_{\mu'=0}^{N}\exp\left[-i\mu'\left(\frac{\pi}{N+2}+\phi\right)-i\frac{\pi}{N+2}\right]\right\}. \end{aligned} \tag{5.64}$$

Using the summation formula

$$\sum_{m=0}^{N}a^m=\frac{1-a^{N+1}}{1-a}, \tag{5.65}$$

this simplifies to

$$\begin{aligned}\langle j\phi|\psi_{\rm opt}\rangle &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\frac{e^{i(j\phi)}}{2i}\left\{e^{i\frac{\pi}{N+2}}\frac{1-e^{i\left(\frac{\pi}{N+2}-\phi\right)(N+1)}}{1-e^{i\left(\frac{\pi}{N+2}-\phi\right)}}-e^{-i\frac{\pi}{N+2}}\frac{1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)(N+1)}}{1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)}}\right\} \\ &= \frac{1}{\sqrt{2j+1}}\frac{1}{\sqrt{j+1}}\frac{e^{i(j\phi)}}{2i}\left\{\frac{e^{i\frac{\pi}{N+2}}+e^{-i\phi(N+1)}}{1-e^{i\left(\frac{\pi}{N+2}-\phi\right)}}-\frac{e^{-i\frac{\pi}{N+2}}+e^{-i\phi(N+1)}}{1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)}}\right\}. \end{aligned} \tag{5.66}$$

The probability distribution is therefore

$$P(\phi)=\frac{1}{4\pi}\frac{1}{N+2}\left|\frac{e^{i\frac{\pi}{N+2}}+e^{-i\phi(N+1)}}{1-e^{i\left(\frac{\pi}{N+2}-\phi\right)}}-\frac{e^{-i\frac{\pi}{N+2}}+e^{-i\phi(N+1)}}{1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)}}\right|^2. \tag{5.67}$$

Simplifying this we find

$$P(\phi)=\frac{1}{4\pi}\frac{1}{N+2}\left|\frac{e^{i\frac{\pi}{N+2}}-e^{-i\left(\frac{\pi}{N+2}+\phi(N+2)\right)}-e^{-i\frac{\pi}{N+2}}+e^{i\left(\frac{\pi}{N+2}-\phi(N+2)\right)}}{\left(1-e^{i\left(\frac{\pi}{N+2}-\phi\right)}\right)\left(1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)}\right)}\right|^2$$

$$
\begin{aligned}
&= \frac{1}{4\pi}\frac{1}{N+2}\left|\frac{\left(e^{i\frac{\pi}{N+2}}-e^{-i\frac{\pi}{N+2}}\right)\left(1+e^{-i\phi(N+2)}\right)}{\left(1-e^{i\left(\frac{\pi}{N+2}-\phi\right)}\right)\left(1-e^{-i\left(\frac{\pi}{N+2}+\phi\right)}\right)}\right|^2 \\
&= \frac{1}{2\pi}\frac{1}{N+2}\frac{\sin^2\left(\frac{\pi}{N+2}\right)\left(1+\cos\phi(N+2)\right)}{\left(1-\cos\left(\frac{\pi}{N+2}-\phi\right)\right)\left(1-\cos\left(\frac{\pi}{N+2}+\phi\right)\right)} \\
&= \frac{1}{2\pi}\frac{1}{N+2}\frac{\sin^2\left(\frac{\pi}{N+2}\right)\left(1+\cos\phi(N+2)\right)}{\left(\cos\phi-\cos\left(\frac{\pi}{N+2}\right)\right)^2}.
\end{aligned}
\tag{5.68}
$$

This simple result for the probability distribution can be used to determine the phase uncertainty under each of the measures in (5.60) using numerical integrals.

For the state with equal numbers of photons in each port, the probability distribution is given by Eq. (5.34). This expression cannot be evaluated to a simple expression like Eq. (5.68), and the full sum must be used in order to treat the distribution exactly. This becomes prohibitively time consuming if the integrals must be performed numerically; however, all of the integrals can be evaluated using sums, except for the entropic length. To see this, note that the probability distribution can be expressed as

$$
P(\phi) = \frac{1}{2\pi}\sum_{k=-2j}^{2j} P_k e^{ik\phi}, \tag{5.69}
$$

where

$$
P_k = \sum_{\mu,\nu=-j}^{j} e^{-i(\mu-\nu)(\pi/2)}\left(I^j_{\nu 0}(\pi/2)\right)^* I^j_{\mu 0}(\pi/2)\delta_{k,\nu-\mu}. \tag{5.70}
$$

Recall that the modified Holevo phase variance for this state can be evaluated using (5.35) and (5.36). For the other measures of the phase uncertainty the integrals will be performed over the interval $[-\pi/2, \pi/2]$, so the above distribution must be multiplied by a factor of 2 to be correctly normalised. The modified probability distribution is therefore

$$
P(\phi) = \frac{1}{\pi}\sum_{k=-2j}^{2j} P_k e^{ik\phi}. \tag{5.71}
$$

In addition, it can be shown that all the P_k are real, and $P_k = 0$ for odd k. For the standard phase variance, it is simpler to express the probability distribution as

$$
P(\phi) = \frac{1}{\pi}P_0 + \frac{2}{\pi}\sum_{k=2}^{2j} P_k \cos(k\phi). \tag{5.72}
$$

The standard phase variance can then be determined using

$$
\Delta\phi^2 = \frac{1}{\pi}P_0\int_{-\pi/2}^{\pi/2}\phi^2 d\phi + \frac{2}{\pi}\sum_{k=2}^{2j} P_k \int_{-\pi/2}^{\pi/2}\cos(k\phi)\phi^2 d\phi. \tag{5.73}
$$

Evaluating this we find that

$$
\begin{aligned}
\Delta\phi^2 &= \frac{1}{\pi}P_0\frac{2}{3}\left(\frac{\pi}{2}\right)^3 + \frac{2}{\pi}\sum_{k=2}^{2j} P_k \frac{\pi}{2}\frac{(-1)^{k/2}}{(k/2)^2} \\
&= \frac{\pi^2}{12}P_0 + \sum_{k=2}^{2j} P_k\left(\frac{2}{k}\right)^2(-1)^{k/2}.
\end{aligned}
\tag{5.74}
$$

This provides a simple formula to determine the standard phase variance, once the coefficients P_k have been determined.

Next, to determine the Fisher information, note that for the state with equal photon numbers in both ports, the probability distribution can be expressed as

$$P(\phi) = \psi^2(\phi), \tag{5.75}$$

with

$$\psi(\phi) = \frac{i^j}{\sqrt{\pi}} \sum_{\mu=-j}^{j} e^{-i\mu(\phi+\pi/2)} I^j_{\mu 0}(\pi/2). \tag{5.76}$$

Here the factor of i^j ensures that this is always real, so the absolute value sign is not required. Note also that a factor of $1/\sqrt{\pi}$ (rather than $1/\sqrt{2\pi}$) is required for the probability distribution to be normalised over the interval $[-\pi/2, \pi/2]$. In terms of this function the Fisher information is

$$\begin{aligned} L_{\rm F}^{-2} &= \int_{-\pi/2}^{\pi/2} \left[\frac{d\psi^2(\phi)}{d\phi}\right]^2 \frac{1}{\psi^2(\phi)} d\phi \\ &= \int_{-\pi/2}^{\pi/2} \left[2\psi(\phi)\frac{d\psi(\phi)}{d\phi}\right]^2 \frac{1}{\psi^2(\phi)} d\phi \\ &= 4\int_{-\pi/2}^{\pi/2} \left[\frac{d\psi(\phi)}{d\phi}\right]^2 d\phi. \end{aligned} \tag{5.77}$$

Evaluating this gives

$$\begin{aligned} L_{\rm F}^{-2} &= 4(-1)^j \int_{-\pi/2}^{\pi/2} \left[\frac{1}{\sqrt{\pi}} \sum_{\mu=-j}^{j} (-i\mu) e^{-i\mu(\phi+\pi/2)} I^j_{\mu 0}(\pi/2)\right]^2 d\phi \\ &= \frac{4}{\pi}(-1)^j \int_{-\pi/2}^{\pi/2} \sum_{\mu,\nu=-j}^{j} \mu(-\nu) e^{-i(\mu+\nu)(\phi+\pi/2)} I^j_{\mu 0}(\pi/2) I^j_{\nu 0}(\pi/2) d\phi \\ &= 4(-1)^j \sum_{\mu=-j}^{j} \mu^2 I^j_{\mu 0}(\pi/2) I^j_{-\mu 0}(\pi/2) d\phi \\ &= 4\sum_{\mu=-j}^{j} \mu^2 \left[I^j_{\mu 0}(\pi/2)\right]^2 d\phi. \end{aligned} \tag{5.78}$$

In the last line the result that $I^j_{-\mu 0}(\pi/2) = (-1)^j I^j_{\mu 0}(\pi/2)$ has been used. This expression gives a simple method for calculating the Fisher length.

Next, in order to evaluate the confidence interval we wish to perform the integral

$$1 - C = \int_{-L_C}^{L_C} P(\phi) d\phi. \tag{5.79}$$

This is easily evaluated as

$$\begin{aligned} 1 - C &= \int_{-L_C}^{L_C} \left[\frac{P_0}{\pi} + \frac{2}{\pi}\sum_{k=2}^{2j} P_k \cos(k\phi)\right] d\phi \\ &= \frac{2L_C P_0}{\pi} + \frac{4}{\pi}\sum_{k=2}^{2j} \frac{P_k}{k} \sin(kL_c). \end{aligned} \tag{5.80}$$

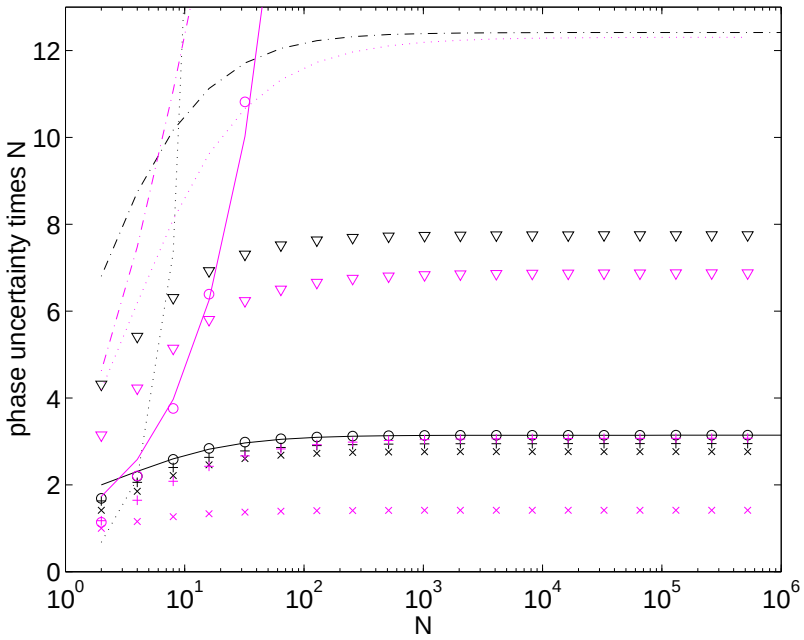

Figure 5.11: The phase uncertainty of the $|j0\rangle_z$ state (purple) and states optimised for minimum Holevo phase variance (black) under several measures multiplied by N. The square root of the Holevo phase variance is shown as the continuous lines, the square root of the standard variance as the circles, the inverse-of-maximal-value as the triangles, the Süssman measure as the dotted lines, the entropic length as the dash-dotted lines, the Fisher length as the crosses, and the 67% confidence interval is shown as the pluses.

This expression can then be used to find L_C for a given C numerically.

The evaluation of the reciprocal-of-peak-value is simple, as it merely requires the evaluation of the probability distribution at $\phi = 0$. Lastly, the Süssman measure can be evaluated using

$$\begin{aligned} L_{\mathrm{S}}^{-1} &= \int_{-\pi/2}^{\pi/2} \frac{1}{\pi^2} \sum_{k,k'=-2j}^{2j} P_k P_{k'} e^{i(k+k')\phi} d\phi \\ &= \frac{1}{\pi} \sum_{k,k'=-2j}^{2j} P_k^2. \end{aligned} \tag{5.81}$$

These methods were used to calculate each of the above measures of uncertainty for the optimum input state and $|j0\rangle_z$ for a large range of photon numbers, and the results multiplied by N are plotted in Fig. 5.11. It is clear from this plot that the asymptotic values for most of these measures are good approximations for photon numbers of order 100 or greater. The results for the Holevo and standard variance for $|j0\rangle_z$ do not converge, as these measures of the phase uncertainty scale as $N^{-1/4}$. The results for the entropic length for $|j0\rangle_z$ do not appear to converge at this scale; however, if we increase the scale of the plot, as in Fig. 5.12, we find that the results converge to a large value for very large photon number.

It can be seen from Fig. 5.11 that the optimal state is actually worse than the $|j0\rangle_z$ state as evaluated using the reciprocal peak L_{rp} or the Fisher length L_F, but it is better under all the other measures of phase uncertainty. The Süssman measure and the entropic length are smaller for $|j0\rangle_z$ for smaller photon numbers, but the asymptotic values are larger.

It can be argued that the reciprocal peak and Fisher length do not give very meaningful estimates of the phase uncertainty in the case of $|j0\rangle_z$. The reciprocal peak gives good scaling because the

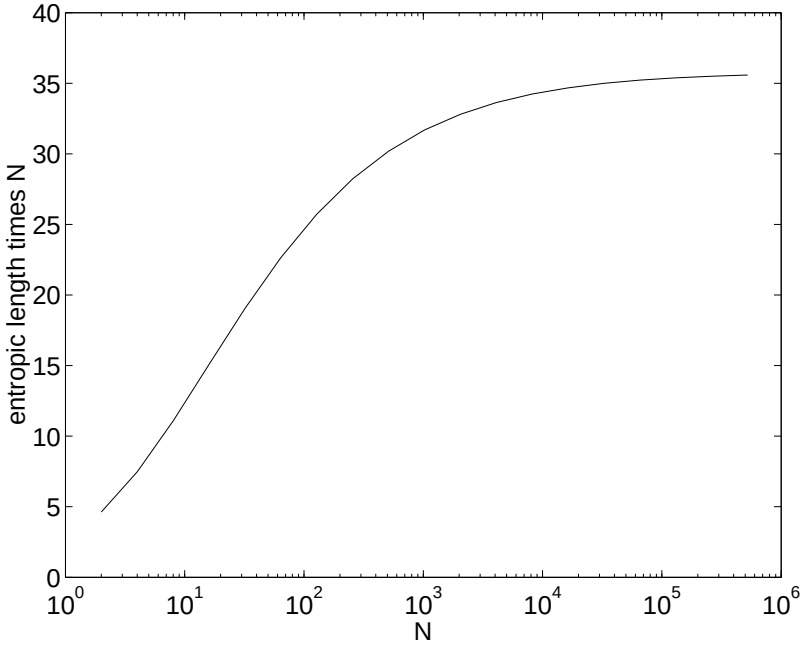

Figure 5.12: The entropic length for the $|j0\rangle_z$ state multiplied by N on a larger scale.

probability distribution has a high peak, but it does not take the tails into account at all. As an extreme example of where this happens, consider the probability distribution given by

$$P(\phi) = \frac{1-\lambda}{2\pi} + \lambda\delta(\phi), \tag{5.82}$$

where $\lambda \ll 1$. Under the reciprocal peak measure, the phase uncertainty is zero; however, it is clear that the actual phase uncertainty is very large.

The Fisher length is small when the probability distribution has a large first derivative. This means that the Fisher length gives a meaningful estimate of the uncertainty when the probability distribution has a smooth peak, as the larger the derivative is, the sharper the peak. The Fisher length is not so meaningful when the probability distribution oscillates, as is the case for $|j0\rangle_z$. The small Fisher length is then just due to the rapid oscillations, rather than a single narrow peak. As an extreme example of this, consider

$$P(\phi) = \frac{\cos^2(n\phi)}{\pi}, \tag{5.83}$$

where $n \gg 1$. The Fisher length will be very small for this distribution; however, it is clear that the actual phase uncertainty is very large. In view of these considerations, it appears that the reciprocal peak and Fisher length give misleadingly small estimates of the phase uncertainty for the $|j0\rangle_z$ state. The better scaling using these measures should not be taken to imply that $|j0\rangle_z$ states are better than the optimal states.

Using the results plotted in Fig. 5.11 we can determine the scaling constants for each of these measures for the two states, and the results are listed in Table 5.1. In this table $\Delta\phi_{\rm H}$ denotes the square root of the Holevo phase variance. Note that for $|\psi_{\rm opt}\rangle$ the square roots of both the standard variance and the Holevo variance scale as π/N, in agreement with the scaling predicted analytically. For $|j0\rangle_z$, the standard and Holevo variance scale as

$$\Delta\phi^2 \approx \frac{0.6573863}{\sqrt{N}} \tag{5.84}$$

$$V(\phi) \approx \frac{0.4395}{\sqrt{N}}. \tag{5.85}$$

Table 5.1: The scaling constants for each of the measures of phase uncertainty for the state optimised for minimum Holevo phase variance and $|j0\rangle_z$.

Measure	$\|\psi_{\rm opt}\rangle$	$\|j0\rangle_z$
$N\Delta\phi$	3.14159265359	$0.8107937 \times N^{3/4}$
$N\Delta\phi_{\rm H}$	3.1415927	$0.66292 \times N^{3/4}$
$NL_{\rm rp}$	7.7515691701	6.87519
$NL_{\rm S}$	10.710529485	12.30505
$NL_{\rm H}$	12.414819836	35.79
$NL_{2/3}$	2.9481552495	3.07129
$NL_{\rm F}$	2.76615948	1.4142136

These are different from each other, and also from the values obtained previously using the Bessel function approximation.

It is interesting to note that the coefficients found here for the confidence interval (3.071) and the Fisher length (1.414) for the state $|j0\rangle_z$ differ from those found in [42], of 3.36 and 2 respectively. In order to see the reason for this difference we can use the asymptotic approximation (5.45). In the large j limit the $+1$ can be ignored, so this becomes

$$P_j(\phi) = \frac{j}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}\frac{\left[J_{1/4}(|j\phi|)\right]^2}{\sqrt{|j\phi|}}. \tag{5.86}$$

Therefore the 2/3 confidence interval can be determined as

$$\begin{aligned}\frac{1}{3} &= \int_0^{L_{2/3}} \frac{j}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}\frac{J_{1/4}^2(j\phi)}{\sqrt{j\phi}}d\phi \\ &= \frac{1}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}\int_0^{jL_{2/3}} \frac{J_{1/4}^2(x)}{\sqrt{x}}dx\end{aligned} \tag{5.87}$$

where $x = j\phi$. Rather than using the asymptotic approximation of the Bessel function, as used in [42], it is fairly straightforward to numerically evaluate the integral and use Newton's method to find $jL_{2/3}$. This gives $jL_{2/3} \approx 1.53564794820384$, so $NL_{2/3} \approx 3.07129589640767$. This agrees with the value found from the exact calculations to six significant digits. Thus we see that the difference in the scaling constant is due to the asymptotic approximation for the Bessel function used in [42].

For the Fisher length, if we perform the derivation in the same way as in [42], we find

$$\begin{aligned}\frac{1}{P_j(\phi)}\left[\frac{d}{d\phi}P_j(\phi)\right]^2 &= \frac{j}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}\frac{\sqrt{j\phi}}{J_{1/4}^2(j\phi)}\left[\frac{d}{d\phi}\frac{J_{1/4}^2(j\phi)}{\sqrt{j\phi}}\right]^2 \\ &= \frac{j}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}4j^2\frac{J_{5/4}^2(j\phi)}{\sqrt{j\phi}}.\end{aligned} \tag{5.88}$$

So the Fisher information is

$$\begin{aligned}F_j &= 2\int_0^{\pi/2} \frac{j}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}4j^2\frac{J_{5/4}^2(j\phi)}{\sqrt{j\phi}}d\phi \\ &\approx \frac{j^2}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}8\int_0^{\infty}\frac{J_{5/4}^2(x)}{\sqrt{x}}dx \\ &= \frac{j^2}{\pi}\frac{\Gamma^2(3/4)}{\sqrt{2}}8\frac{\pi}{\sqrt{8}\Gamma^2(3/4)}\end{aligned}$$

Table 5.2: The scaling constants for each of the measures of phase uncertainty for the state optimised for minimum Holevo phase variance and $|j0\rangle_z$ as determined using the asymptotic approximations.

Measure	$\|\psi_{\rm opt}\rangle$	$\|j0\rangle_z$
$N\Delta\phi$	$\pi = 3.14159265359265\ldots$	$0.631965250820959 \times N^{3/4}$
$N\Delta\phi_{\rm H}$	$\pi = 3.14159265359265\ldots$	$0.533458147005221 \times N^{3/4}$
$NL_{\rm rp}$	$\pi^3/4 = 7.75156917007495\ldots$	$4\pi\left[\frac{\Gamma(5/4)}{\Gamma(3/4)}\right]^2 = 6.87518581802037\ldots$
$NL_{\rm S}$	10.7105294850660	12.305050002393
$NL_{\rm H}$	12.4148198362985	35.78817
$NL_{2/3}$	2.94815524951393	3.07129589640767
$NL_{\rm F}$	2.7661594839	$\sqrt{2} = 1.41421356237309\ldots$

$$= 2j^2. \tag{5.89}$$

This is twice what was found in [42], and using this the Fisher length is approximately $L_{\rm F} \approx \sqrt{2}/N$. The reason for this factor of two is that the approximation used for the probability distribution in [42] is not correctly normalised. Here the normalised version has been used, which is twice that given in [42].

Note that this approximate analytic result agrees very precisely with the result found using the exact calculations. We can also determine the scaling constants for the other measures of phase variance using the analytic approximation (5.86). The scaling constants for the standard and Holevo phase variances were found above. The reciprocal peak can be evaluated as

$$\begin{aligned} L_{\rm rp} &= \frac{\pi}{j}\frac{\sqrt{2}}{\Gamma^2(3/4)}\lim_{\phi\to 0}\frac{\sqrt{j\phi}}{J^2_{1/4}(j\phi)} \\ &= \frac{2\pi}{j}\left[\frac{\Gamma(5/4)}{\Gamma(3/4)}\right]^2. \end{aligned} \tag{5.90}$$

The scaling is therefore

$$L_{\rm rp} \approx \frac{1}{N} 4\pi \left[\frac{\Gamma(5/4)}{\Gamma(3/4)}\right]^2. \tag{5.91}$$

This scaling constant is approximately 6.87518581802037, agreeing very accurately with that found from the exact calculations.

The Süssman measure and entropic length can also be determined by numerical integrals. The scaling constants found, along with the scaling constants found above for the other measures, are given in Table 5.2. Note that there is very good agreement with the results based on the exactly calculated data in these two cases as well. The numerical integral for the entropic length is particularly difficult to calculate, as the probability distribution falls very slowly with the phase, and the integral must be calculated to very large values of $j\phi$. For this reason only 7 significant digits could be found.

It is also possible to determine the scaling constants for the state optimised for minimum intrinsic phase variance using an asymptotic approximation. It is easy to see that in the limit of large N and small ϕ, Eq. (5.68) becomes

$$\begin{aligned} P(\phi) &\approx \frac{1}{2\pi}\frac{1}{N+2}\frac{\left(\frac{\pi}{N+2}\right)^2\left(1+\cos\phi(N+2)\right)}{\left[\frac{1}{2}\phi^2 - \frac{1}{2}\left(\frac{\pi}{N+2}\right)^2\right]^2} \\ &\approx 2\pi N \frac{1+\cos N\phi}{\left[(N\phi)^2-\pi^2\right]^2}. \end{aligned} \tag{5.92}$$

In the second line it has been assumed that $N+2 \approx N$ in the limit of large photon number.

Note that the asymptotic approximations for both of these states have the general form

$$P(\phi) = Nf(N\phi). \tag{5.93}$$

For asymptotic functions of this form, we obtain $1/N$ scaling for each of the measures of phase uncertainty, provided that the integrals are bounded. To see this, note that each of the measures is, in terms of the asymptotic function,

$$\begin{aligned}
(N\Delta\phi)^2 &= \int_{-N\pi}^{N\pi} f(x)(x-\bar{x})^2 dx \\
\log NL_{\mathrm{H}} &= -\int_{-N\pi}^{N\pi} f(x)\log f(x) dx \\
(NL_{\mathrm{F}})^{-2} &= \int_{-N\pi}^{N\pi} \frac{[f'(x)]^2}{f(x)} dx \\
1-C &= \int_{\bar{x}-NL_C}^{\bar{x}+NL_C} f(x) dx \\
(NL_{\mathrm{rp}})^{-1} &= \max f(x) \\
(NL_{\mathrm{S}})^{-1} &= \int_{-N\pi}^{N\pi} f^2(x) dx.
\end{aligned} \tag{5.94}$$

In the asymptotic limit the Holevo variance is given by the same expression as the standard variance. For the standard or Holevo variance in the case of the $|j0\rangle_z$ state, the integral is not bounded, and must be considered up to the limit $\phi = \pi/2$. That is why $1/N$ scaling is not obtained in that case.

For the asymptotic expression for the optimal input states, the scaling constant for the variance is already known to be π, using the analytic result for the Holevo phase variance. The scaling constant for the reciprocal-peak measure of the phase uncertainty is easily determined analytically as $\pi^3/4$. For the other measures, the scaling constants can be determined using numerical integrals, and the results are as shown in Table 5.2. Comparing the results in Tables 5.1 and 5.2, it can be seen that there is excellent agreement between the scaling constants for all of the measures.

We therefore find that the asymptotic expressions accurately give the scaling constants for all the cases considered, except the Holevo or standard variance for the $|j0\rangle_z$ state. For these cases the main contribution is from significant values of ϕ, where the asymptotic expression is not accurate. The scaling constants are higher for $|j0\rangle_z$ than the optimal state for all of the measures except the reciprocal-of-peak and Fisher length. As explained above, these measures appear to give unrealistically low estimates of the phase uncertainty for the $|j0\rangle_z$ state.

Similarly the Holevo and standard variances give unrealistically high estimates of the phase uncertainty for $|j0\rangle_z$. A more accurate way of comparing the two states is through confidence intervals. For the 2/3 confidence interval, the scaling constant is only slightly higher for $|j0\rangle_z$ than the optimal state. The difference is far more pronounced if we consider a higher probability confidence interval. For a 90% confidence interval, the scaling constant is only a little higher for the optimal state, at about 4.88, but it is 37.69 for $|j0\rangle_z$.

Probably the most accurate way of comparing the two distributions is through the entropic length, as this corresponds exactly with the phase information contained in the distribution. The scaling constant for the entropic length is almost three times higher for the $|j0\rangle_z$ state than for the optimal state. These results demonstrate that although the phase uncertainty scales as N^{-1} for $|j0\rangle_z$, it is not as good as the optimal state.

Chapter 6

Optimum Adaptive Interferometry

Now that the input states have been fully considered, the next factor to consider is the optimum way of measuring these states. There are basically two components to this: the feedback phase that is used during the measurement, and the phase estimate that is used at the end of the measurement. It has been shown [35, 51, 34] that in the single mode case it is possible to make very good phase measurements by using feedback to an auxiliary phase shift. In order to apply the same principle here, I consider adaptive measurements where the phase to be measured, φ, is in one arm of the interferometer, and a known phase shift, Φ, is introduced into the other arm of the interferometer, as in Fig. 5.1.

The initial feedback scheme that will be considered is that where the introduced phase shift is adjusted in order to minimise the variance in the phase estimate after the next photodetection. In order to evaluate this feedback scheme, the optimal phase estimates are required; these are derived in Sec. 6.2. It is also possible to select the feedback phases in order to minimise the final variance, using numerical minimisation techniques.

For these adaptive schemes to work, the feedback that adjusts Φ must act much faster than the average time between photon arrivals. For simplicity I make the assumption that the feedback is arbitrarily fast, which simply means that the phase Φ is always assumed to have been changed before the next detection occurs. It is the ability to change Φ during the passage of a single (two-mode) pulse that makes photon counting measurements more general than a measurement of the output $\hat{J}_z$ considered in [39, 40].

As mentioned in the previous chapter, it is assumed that the photodetectors have unit efficiency. In addition, only single detection events are considered, rather than multiple detections. Physically, simultaneous detections correspond to individual detections that are too close together to be resolved by the apparatus. Here it will be assumed that the apparatus is arbitrarily fast, so that all the detections can be resolved, and the feedback phase updated between detections. This is not necessarily realistic for current technology, particularly for larger photon numbers. However, the purpose of this study is to examine how far phase measurements can, in principle, be improved, rather than to examine the limits of current technology.

6.1 Preliminary Theory

To describe the measurement, it is convenient to denote the result u from the mth detection as u_m (which is 0 or 1 depending on which output the photon is detected in), and the measurement record up to and including the mth detection as the binary string $n_m \equiv u_m \ldots u_2 u_1$. The input state after m detections will be a function of the measurement record and φ, and is denoted as $|\psi(n_m, \varphi)\rangle$. In the calculations these states are not normalised (except for the initial state), in order to express the state as a power series in $e^{i\varphi}$.

The time between detections does not give any information, and is not included in the measurement record. To see this, note that the measurement operator for no detection is

$$\Omega_0 = 1 - \tfrac{1}{2}(a^\dagger a + b^\dagger b)dt, \tag{6.1}$$

where a and b are the annihilation operators for the two input modes. As this does not depend on the phase, the probability distribution for the phase is unchanged between detections. In addition, we consider only states with a fixed total photon number N. This means that, after m detections, the system will be in an eigenstate of $a^\dagger a + b^\dagger b$ with eigenvalue $N - m$. Therefore the measurement operator Ω_0 does not produce any system evolution. As neither the probability distribution for the phase nor the system state changes between detections, the time between detections may be ignored.

Now the evolution produced by detections will be considered. After the first beam splitter the operators for the two beams are (in the Heisenberg picture)

$$(a + ib)/\sqrt{2}, \quad (ia + b)/\sqrt{2}. \tag{6.2}$$

The two beams are then subjected to phase shifts of φ and Φ, so the operators become

$$e^{i\varphi}(a + ib)/\sqrt{2}, \quad e^{i\Phi}(ia + b)/\sqrt{2}. \tag{6.3}$$

Lastly, the effect of the second beam splitter is the same as the first, giving the two operators

$$\begin{aligned} &ie^{i(\varphi+\Phi)/2}(a\sin\theta + b\cos\theta), \\ &ie^{i(\varphi+\Phi)/2}(a\cos\theta - b\sin\theta), \end{aligned} \tag{6.4}$$

where $\theta = (\varphi - \Phi)/2$. Ignoring the unimportant initial phase factors, these can be represented as the operators $\hat{c}_0, \hat{c}_1$, where

$$\hat{c}_u = a\sin(\theta + u\pi/2) + b\cos(\theta + u\pi/2). \tag{6.5}$$

The input state is then determined by the initial condition

$$|\psi(n_0, \varphi)\rangle = |\psi\rangle, \tag{6.6}$$

where n_0 is the empty string, and the recurrence relation

$$|\psi(u_m n_{m-1}, \varphi)\rangle = \hat{c}_{u_m}(\varphi)\frac{|\psi(n_{m-1}, \varphi)\rangle}{\sqrt{2j + 1 - m}}. \tag{6.7}$$

These states are not normalised, and their norm represents the probability for the measurement record n_m given φ. That is,

$$P(n_m|\varphi) = \langle\psi(n_m, \varphi)|\psi(n_m, \varphi)\rangle. \tag{6.8}$$

An arbitrary input state with $N = 2j$ photons can be expressed as a sum of $\hat{J}_z$ eigenstates

$$|\psi\rangle = \sum_{\mu=-j}^{j} \psi_\mu |j\mu\rangle_z. \tag{6.9}$$

The state after m detections will be (for $m \geq 1$) a function of φ. It will be denoted as follows:

$$|\psi(n_m, \varphi)\rangle = \sum_{\mu=-j+m/2}^{j-m/2} \psi_{\mu,m}(n_m, \varphi)|j - m/2, \mu\rangle_z. \tag{6.10}$$

Using the recurrence relation (6.7), it can be shown that the functional form of $\psi_{\mu,m}(n_m, \varphi)$ is always

$$\psi_{\mu,m}(n_m, \varphi) = \sum_{k=-m/2}^{m/2} \psi_{\mu,m,k}(n_m)e^{ik\varphi}. \tag{6.11}$$

The recurrence relation for the coefficients $\psi_{\mu,m,k}$ is

$$\psi_{\mu,m,k}(n_m) = \frac{e^{-i(\Phi_m - u_m\pi)/2}}{2\sqrt{2j-m+1}}\left[s_-\psi_{\mu-\frac{1}{2},m-1,k-\frac{1}{2}}(n_{m-1}) - is_+\psi_{\mu+\frac{1}{2},m-1,k-\frac{1}{2}}(n_{m-1})\right] + \frac{e^{i(\Phi_m - u_m\pi)/2}}{2\sqrt{2j-m+1}}\left[s_-\psi_{\mu-\frac{1}{2},m-1,k+\frac{1}{2}}(n_{m-1}) + is_+\psi_{\mu+\frac{1}{2},m-1,k+\frac{1}{2}}(n_{m-1})\right], \quad (6.12)$$

where

$$s_\pm = \sqrt{j - \frac{m}{2} \pm \mu + 1}. \quad (6.13)$$

We need to express the system state in this form for two main reasons. Firstly the feedback scheme should not depend on the actual value of the phase, only on the measurement record. To determine the feedback phase, the probability distribution for the system phase is required. This can be determined from the variation of the state with the unknown system phase. This requires the above coefficients, as repeating the calculation for each individual value of the phase would be too inefficient. Secondly, we can perform the entire calculation independently of the system phase, and average over the system phase at the end of the measurement. This allows us to take account exactly of the full range of input phases.

The probability distribution for the unknown phase can be determined using Bayes' theorem

$$P(\varphi|n_m) = \frac{P(\varphi)P(n_m|\varphi)}{P(n_m)}. \quad (6.14)$$

The probability distribution for the phase at the start of the measurement, $P(\varphi)$, will be flat, as it is assumed that there is no prior knowledge about the phase. The divisor $P(n_m)$ is independent of the phase, and therefore only provides a normalising factor to the phase distribution. Ignoring these terms gives

$$P(\varphi|n_m) \propto P(n_m|\varphi), \quad (6.15)$$

and therefore

$$P(\varphi|n_m) \propto \langle\psi(n_m,\varphi)|\psi(n_m,\varphi)\rangle. \quad (6.16)$$

Note that this result is equivalent to Eq. (1.72) for the case of continuous measurements. A similar Bayesian approach to interferometry has been considered before, and used in analysing experimental results [58]. However, this was done only with non-adaptive measurements and with all particles incident on one port.

6.2 Optimum Phase Estimates

Before describing the phase feedback technique, I will firstly describe how to select the final phase estimate. This is necessary because it is not possible to determine the phase variance produced by a phase feedback technique without specific final phase estimates. The best estimate to use is that which minimises the Holevo variance in the final phase estimate. This can be determined by summing over the 2^{2j} combinations of results, then averaging over the input phase.

For greater generality, the optimum phase estimate after m detections will be considered, rather than specifically the phase estimate at the end of the measurement. The result for the final phase estimate can then be obtained by substituting $m = 2j$. First, summing over the combinations of results gives the probability distribution for the error in the phase estimate as

$$P(\phi|\varphi) = \sum_{[n_m]=0}^{2^m-1} P(n_m|\varphi)\delta(\phi - (\hat{\varphi}(n_m) - \varphi)), \quad (6.17)$$

where the square brackets in $[n_m]$ denote the numerical value of this binary string interpreted as a binary number, and $\hat{\varphi}(n_m)$ is the final phase estimate.

Next we wish to average over the system phase and the initial feedback phase. For a feedback scheme to be unbiased (in that it treats all input phases equivalently), the initial feedback phase should be chosen at random, as there is no information to base this phase on. This phase should therefore be averaged over in order to determine the average probability distribution. In determining the coefficients $\psi_{\mu,m,k}(n_m)$, a specific initial feedback phase must be chosen, so at first it might appear that this phase cannot be averaged over. Note, however, that all the successive feedback phases are relative to the initial feedback phase. If the initial feedback phase is altered by some amount $\Delta\Phi$, then all the successive feedback phases will be altered by the same amount (for the same detection results). This is because, if the feedback scheme is unbiased, the initial feedback phase provides a reference phase.

Therefore it is only the difference between the system phase and the initial feedback phase that is significant, and we need only average over one of them. I will take the initial feedback phase to be zero, which is equivalent to measuring all phases relative to the initial feedback phase, and average over the system phase. Performing this average gives

$$\begin{aligned} P(\phi) &= \int_{-\pi}^{\pi} d\varphi \frac{1}{2\pi} \sum_{[n_m]=0}^{2^m-1} P(n_m|\varphi)\delta(\phi - (\hat{\varphi}(n_m) - \varphi)), \\ &= \frac{1}{2\pi} \sum_{[n_m]=0}^{2^m-1} P(n_m|\hat{\varphi}(n_m) - \phi). \end{aligned} \tag{6.18}$$

The exact phase variance for the measurement scheme can be determined from this probability distribution. Evaluating $\langle e^{i\phi} \rangle$ gives

$$\begin{aligned} \langle e^{i\phi} \rangle &= \int_{-\pi}^{\pi} d\phi e^{i\phi} \frac{1}{2\pi} \sum_{[n_m]=0}^{2^m-1} P(n_m|\hat{\varphi}(n_m) - \phi) \\ &= \frac{1}{2\pi} \sum_{[n_m]=0}^{2^m-1} e^{i\hat{\varphi}(n_m)} \int_{-\pi}^{\pi} d\phi e^{-i(\hat{\varphi}(n_m)-\phi)} P(n_m|\hat{\varphi}(n_m) - \phi) \\ &= \frac{1}{2\pi} \sum_{[n_m]=0}^{2^m-1} e^{i\hat{\varphi}(n_m)} \int_{-\pi}^{\pi} dx e^{-ix} P(n_m|x). \end{aligned} \tag{6.19}$$

In order to minimise the Holevo phase variance, we wish to maximise $\left|\langle e^{i\phi} \rangle\right|$. A phase estimate that maximises this is

$$\begin{aligned} \hat{\varphi}(n_m) &= \arg \int_{-\pi}^{\pi} e^{i\varphi} P(n_m|\varphi) d\varphi \\ &= \arg \int_{-\pi}^{\pi} e^{i\varphi} \langle \psi(n_m, \varphi) | \psi(n_m, \varphi) \rangle d\varphi \\ &= \arg \int_{-\pi}^{\pi} e^{i\varphi} P(\varphi|n_m) d\varphi \\ &= \arg \langle e^{i\varphi} \rangle, \end{aligned} \tag{6.20}$$

where the expectation value is determined from the probability distribution for the phase based on the measurement record. For the specific case of the phase estimate at the end of the measurement, this can be calculated as

$$\hat{\varphi}(n_{2j}) = \arg \sum_{k=-j}^{j-1} \psi_{0,2j,k} \psi^*_{0,2j,k+1}. \tag{6.21}$$

Unfortunately there is a slight ambiguity here, as the same Holevo phase variance will be obtained if a constant is added to these phase estimates. This would make the probability distribution biased, and as discussed in the introduction, a better way of evaluating the variance when the measurements may be biased is

$$V_0(\phi) = \left[\mathrm{Re}\langle e^{i\phi}\rangle\right]^{-2} - 1. \tag{6.22}$$

Note that $e^{i\phi}$ is used here, rather than $e^{i(\phi-\varphi)}$, because ϕ is the deviation from the system phase. Minimising this estimate of the phase variance is equivalent to maximising $\mathrm{Re}\langle e^{i\phi}\rangle$. If any constant is added to the phase estimates given by Eq. (6.20), the value of $\left|\langle e^{i\phi}\rangle\right|$ will be the same, but the value of $\mathrm{Re}\langle e^{i\phi}\rangle$ will be smaller, as $\langle e^{i\phi}\rangle$ is complex. Therefore the phase estimates given by Eq. (6.20) are the unique solution that maximises $\mathrm{Re}\langle e^{i\phi}\rangle$.

With these phase estimates, the exact variance after m detections can be determined using the simple expression

$$\left|\langle e^{i\phi}\rangle\right| = \frac{1}{2\pi}\sum_{[n_m]=0}^{2^m-1}\left|\int_{-\pi}^{\pi} e^{i\varphi}\langle\psi(n_m,\varphi)|\psi(n_m,\varphi)\rangle d\varphi\right|. \tag{6.23}$$

As these phase estimates are unbiased, the absolute value will be used, rather than the real part. The final variance can be determined using

$$\left|\langle e^{i\phi}\rangle\right| = \sum_{[n_{2j}]=0}^{2^{2j}-1}\left|\sum_{k=-j}^{j-1}\psi_{0,2j,k}\psi^*_{0,2j,k+1}\right|. \tag{6.24}$$

This means that during calculations, after each sequence of measurements, the phase estimate and contribution to $\left|\langle e^{i\phi}\rangle\right|$ can be determined as the phase and magnitude respectively of

$$\sum_{k=-j}^{j-1}\psi_{0,2j,k}\psi^*_{0,2j,k+1}. \tag{6.25}$$

6.3 The Feedback Technique

In order to find the optimum phase feedback technique, we need to find the feedback phase Φ_m for each measurement record n_{m-1} that maximises the value of $\left|\langle e^{i\phi}\rangle\right|$ as determined using Eq. (6.24) (and therefore minimises the Holevo phase variance). This is, in general, a very difficult problem, and I will initially consider a much simpler feedback technique that can be determined analytically.

Rather than choosing each feedback phase to minimise the final phase variance at the end of the measurement, we can choose the feedback phase that minimises the phase variance after the next detection. This will mean that the last feedback phases are optimal, but not necessarily the intermediate feedback phases. This is because they minimise the intermediate phase variances, not the final phase variance.

In order to see how to choose the phase estimate, recall that the value of $\left|\langle e^{i\phi}\rangle\right|$ after the mth detection is given by Eq. (6.23). For the feedback phase before the mth detection, Φ_m, we need only consider the part of the sum for the measurement record n_{m-1}, as this is the only part of the sum that is affected by the feedback phase. We must still sum over the mth detection result, as this is still unknown. The expression to be maximised is therefore

$$M(\Phi_m) = \sum_{u_m=0,1}\left|\int_{-\pi}^{\pi}\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle e^{i\varphi}d\varphi\right|. \tag{6.26}$$

In order to use this expression we require $\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle$ explicitly in terms of u_m. It is straightforward to show from Eq. (6.12) that

$$\begin{aligned}\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle = \frac{1}{2}[&\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle \\ &+\lambda_{m-1}(\varphi)e^{i\varphi}e^{-i(\Phi_m-u_m\pi)} + \lambda^*_{m-1}(\varphi)e^{-i\varphi}e^{i(\Phi_m-u_m\pi)}],\end{aligned} \tag{6.27}$$

where $\lambda_{m-1}(\varphi)$ is defined by

$$\lambda_{m-1}(\varphi) = \sum_{n=-m+1}^{m-1} \lambda_{m-1,n} e^{in\varphi}, \tag{6.28}$$

where

$$\lambda_{m-1,n} = -\frac{\xi_{m-1,n} + i\zeta_{m-1,n}}{2j-m+1}, \tag{6.29}$$

and where

$$\xi_{m-1,n} = \sum_{k,k'=-\frac{m-1}{2}}^{\frac{m-1}{2}} \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \mu \psi_{\mu,m-1,k} \psi^*_{\mu,m-1,k'} \delta_{n,k-k'}, \tag{6.30}$$

$$\zeta_{m-1,n} = \sum_{k,k'=-\frac{m-1}{2}}^{\frac{m-1}{2}} \sum_{\mu=-j+\frac{m}{2}}^{j-\frac{m}{2}} \frac{s_+ s_-}{2} (\psi_{\mu-\frac{1}{2},m-1,k} \psi^*_{\mu+\frac{1}{2},m-1,k'} + \psi_{\mu+\frac{1}{2},m-1,k} \psi^*_{\mu-\frac{1}{2},m-1,k'}) \delta_{n,k-k'}. \tag{6.31}$$

Using this result it is possible to show that

$$\begin{aligned} M(\Phi_m) = & \sum_{u_m=0,1} \left| \int_{-\pi}^{\pi} \frac{1}{2} [\langle \psi(n_{m-1},\varphi) | \psi(n_{m-1},\varphi) \rangle + \lambda_{m-1}(\varphi) e^{i\varphi} e^{-i(\Phi_m - u_m\pi)} \right. \\ & \left. + \lambda^*_{m-1}(\varphi) e^{-i\varphi} e^{i(\Phi_m - u_m\pi)}] e^{i\varphi} d\varphi \right| \\ = & \frac{1}{2} \sum_{u_m=0,1} \left| \int_{-\pi}^{\pi} \left[\sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \sum_{k,k'=-\frac{m-1}{2}}^{\frac{m-1}{2}} \psi_{\mu,m-1,k} \psi^*_{\mu,m-1,k'} e^{i(k-k'+1)\varphi} \right. \right. \\ & \left. \left. + e^{-i(\Phi_m - u_m\pi)} \sum_{n=1-m}^{m-1} \lambda_{m-1,n} e^{i(n+2)\varphi} + e^{i(\Phi_m - u_m\pi)} \sum_{n=1-m}^{m-1} \lambda^*_{m-1,n} e^{-in\varphi} \right] d\varphi \right| \\ = & \pi \sum_{u_m=0,1} \left| \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \sum_{k=-\frac{m-1}{2}}^{\frac{m-3}{2}} \psi_{\mu,m-1,k} \psi^*_{\mu,m-1,k+1} \right. \\ & \left. + \lambda_{m-1,-2} e^{-i(\Phi_m - u_m\pi)} + \lambda^*_{m-1,0} e^{i(\Phi_m - u_m\pi)} \right|. \end{aligned} \tag{6.32}$$

This can be expressed in the form

$$M(\Phi_m) = \left| a + b e^{-i\Phi_m} + c e^{i\Phi_m} \right| + \left| a - b e^{-i\Phi_m} - c e^{i\Phi_m} \right|, \tag{6.33}$$

where

$$a = \pi \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \sum_{k=-\frac{m-1}{2}}^{\frac{m-3}{2}} \psi_{\mu,m-1,k} \psi^*_{\mu,m-1,k+1}, \tag{6.34}$$

$$b = \pi \lambda_{m-1,-2}, \tag{6.35}$$

$$c = \pi \lambda^*_{m-1,0}. \tag{6.36}$$

There is an analytic solution for the Φ_m that maximises $M(\Phi_m)$. This solution gives three phases, Φ_0 and $\Phi_\pm$, and the phase that is optimal must be found by substituting into Eq. (6.33). These phases are given by

$$\Phi_0 = \arg(ba^* - c^* a), \tag{6.37}$$

and

$$\Phi_{\pm} = \arg \sqrt{\frac{c_2 \pm \sqrt{c_2^2 + |c_1|^2}}{c_1}}, \tag{6.38}$$

where

$$c_1 = (a^* c)^2 - (ab^*)^2 + 4(|b|^2 - |c|^2) b^* c, \tag{6.39}$$

$$c_2 = -2i\mathrm{Im}(a^2 b^* c^*). \tag{6.40}$$

Note that $M(\Phi) = M(\Phi + \pi)$ so that in addition to the solution found by this method there will be another differing by π. It does not matter which of these is chosen; it simply reverses the significance of the two alternative detection results $u_m = 0$ and 1.

As mentioned above, the initial feedback phase Φ_1 should be chosen at random, as there is no information to base it on. At each following step, we determine the optimal feedback phase by the method described above, then determine the evolution of the state for that feedback phase. This process continues until all photons have been counted. The measurement record is then the binary string $n_{2j} = u_{2j} \ldots u_2 u_1$ and the result is a posterior distribution $P(\varphi|n_{2j})$ that is proportional to $\langle \psi(n_{2j}, \varphi) | \psi(n_{2j}, \varphi) \rangle$ and is characterised by the $2j + 1$ numbers $\psi_{0,2j,k}(n_{2j})$.

6.4 Stochastic Method

For a moderate number of photons it is possible to determine the exact phase variance by systematically determining the evolution of the state for each combination of measurement results and using Eq. (6.24). As the number of possible measurement records increases as 2^{2j}, the exact variance can be determined only for moderate photon numbers (up to 20 or 30). For larger photon numbers it is necessary to determine the phase distribution stochastically.

The initial feedback phase is selected at random, and the results of the detections are selected by their probability of occurring. In order to determine the probabilities, a specific system phase must be selected. For simplicity, the system phase was taken to be zero in the results that will be presented here. This leads to no loss of generality, as the initial feedback phases were selected at random.

In order to determine the probabilities, it is simplest to determine the phase dependent state coefficients $\psi_{\mu,m}(n_m, \varphi)$. The state coefficients will change as

$$\psi_{\mu,m}(n_m, \varphi) = s_+ \sin\Theta_m \psi_{\mu+\frac{1}{2},m-1}(n_{m-1}, \varphi) + s_- \cos\Theta_m \psi_{\mu-\frac{1}{2},m-1}(n_{m-1}, \varphi), \tag{6.41}$$

where

$$\Theta_m = \frac{\varphi - \Phi_m + u_m \pi}{2}. \tag{6.42}$$

The probability of obtaining the result u_m is given by

$$P(u_m|\varphi, n_{m-1}) = \frac{1}{2} + \frac{e^{iu_m\pi}}{2j - m + 1} \left[-\cos(\varphi - \Phi_m) \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \mu \left| \psi_{\mu,m-1}(n_{m-1}, \varphi) \right|^2 \right.$$
$$\left. + \sin(\varphi - \Phi_m) \sum_{\mu=-j+\frac{m}{2}}^{j-\frac{m}{2}} s_+ s_- \mathrm{Re}\left(\psi_{\mu+\frac{1}{2},m-1}(n_{m-1}, \varphi) \psi^*_{\mu-\frac{1}{2},m-1}(n_{m-1}, \varphi) \right) \right]. \tag{6.43}$$

Here it has been assumed that the coefficients $\psi_{\mu,m-1}(n_{m-1}, \varphi)$ are normalised. The recurrence relation (6.41) does not give the normalised coefficients. The coefficients determined from Eq. (6.41) should be normalised in a separate calculation.

The detection results were chosen with probabilities determined using $\varphi = 0$ and these formulae, and the final phase estimates were determined using Eq. (6.21). For the ensemble $\{\phi_\mu\}_{\mu=0}^{M}$ of M

final phase estimates the Holevo phase variance was estimated by

$$\left[\mathrm{Re}\left(M^{-1}\sum_{n=1}^{M}e^{i\phi_n}\right)\right]^{-2}-1. \tag{6.44}$$

6.5 Modulo π States

In the case of the state with equal photon numbers in both ports, the above techniques will not apply exactly, as the phase distribution repeats modulo π. This means, for example, that $\langle e^{i\phi}\rangle = 0$. In the previous chapter the measure of the phase variance used for this state was

$$V_\pi(\phi) = (|\langle e^{2i\phi}\rangle|^{-2}-1)/4. \tag{6.45}$$

Most of the above analysis must also be modified to consider the phase modulo π.

Firstly the distribution (6.17) should be averaged over the interval $[-\pi/2,\pi/2]$ rather than $[-\pi,\pi]$. This gives

$$P(\phi) = \frac{1}{\pi}\sum_{[n_m]=0}^{2^m-1}P(n_m|\hat{\varphi}(n_m)-\phi). \tag{6.46}$$

Now, rather than evaluating $\langle e^{i\phi}\rangle$, we wish to find $\langle e^{2i\phi}\rangle$. This can be evaluated as

$$\begin{aligned}\langle e^{2i\phi}\rangle &= \left|\int_{-\pi/2}^{\pi/2}d\phi e^{2i\phi}\frac{1}{\pi}\sum_{[n_m]=0}^{2^m-1}P(n_m|\hat{\varphi}(n_m)-\phi)\right| \\ &= \left|\frac{1}{\pi}\sum_{[n_m]=0}^{2^m-1}e^{2i\hat{\varphi}(n_m)}\int_{-\pi/2}^{\pi/2}dx e^{-2ix}\langle\psi(n_m,x)|\psi(n_m,x)\rangle\right|.\end{aligned} \tag{6.47}$$

This means that the optimal phase estimate is

$$\hat{\varphi}(n_m) = \tfrac{1}{2}\arg\int_{-\pi/2}^{\pi/2}e^{2i\varphi}\langle\psi(n_m,\varphi)|\psi(n_m,\varphi)\rangle d\varphi. \tag{6.48}$$

This can be evaluated at the end of the measurement using

$$\hat{\varphi}(n_{2j}) = \tfrac{1}{2}\arg\sum_{k=-j}^{j-2}\psi_{0,2j,k}\psi^*_{0,2j,k+2}, \tag{6.49}$$

and the final variance is calculated using

$$|\langle e^{2i\phi}\rangle| = \sum_{[n_{2j}]=0}^{2^{2j}-1}\left|\sum_{k=-j}^{j-2}\psi_{0,2j,k}\psi^*_{0,2j,k+2}\right|. \tag{6.50}$$

Now instead of Eq. (6.26), the expression to be maximised is

$$M(\Phi_m) = \sum_{u_m=0,1}\left|\int_{-\pi/2}^{\pi/2}\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle e^{2i\varphi}d\varphi\right|. \tag{6.51}$$

Using Eq. (6.27), this becomes

$$\begin{aligned}
M(\Phi_m) = \sum_{u_m=0,1} \Bigg| \int_{-\pi/2}^{\pi/2} & \frac{1}{2}[\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle + \lambda_{m-1}(\varphi)e^{i\varphi}e^{-i(\Phi_m-u_m\pi)} \\
& + \lambda^*_{m-1}(\varphi)e^{-i\varphi}e^{i(\Phi_m-u_m\pi)}]e^{2i\varphi}d\varphi\Bigg| \\
= \frac{1}{2}\sum_{u_m=0,1}\Bigg|\int_{-\pi/2}^{\pi/2}\Bigg[& \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}}\sum_{k,k'=-\frac{m-1}{2}}^{\frac{m-1}{2}} \psi_{\mu,m-1,k}\psi^*_{\mu,m-1,k'}e^{i(k-k'+2)\varphi} \\
& + e^{-i(\Phi_m-u_m\pi)}\sum_{n=1-m}^{m-1}\lambda_{m-1,n}e^{i(n+3)\varphi} + e^{i(\Phi_m-u_m\pi)}\sum_{n=1-m}^{m-1}\lambda^*_{m-1,n}e^{-i(n-1)\varphi}\Bigg]d\varphi\Bigg| \\
= \frac{\pi}{2}\sum_{u_m=0,1}\Bigg| & \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}}\sum_{k=-\frac{m-1}{2}}^{\frac{m-5}{2}} \psi_{\mu,m-1,k}\psi^*_{\mu,m-1,k+2} \\
& + \lambda_{m-1,-3}e^{-i(\Phi_m-u_m\pi)} + \lambda^*_{m-1,1}e^{i(\Phi_m-u_m\pi)}\Bigg|. \qquad (6.52)
\end{aligned}$$

Thus a, b and c are given by

$$a = \frac{\pi}{2}\sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}}\sum_{k=-\frac{m-1}{2}}^{\frac{m-5}{2}} \psi_{\mu,m-1,k}\psi^*_{\mu,m-1,k+2}, \qquad (6.53)$$

$$b = \frac{\pi}{2}\lambda_{m-1,-3}, \qquad (6.54)$$

$$c = \frac{\pi}{2}\lambda^*_{m-1,1}. \qquad (6.55)$$

The last equation that must be modified for this case is the formula for estimating the phase variance from the data, which becomes

$$\frac{1}{4}\left\{\left[\mathrm{Re}\left(M^{-1}\sum_{n=1}^{M}e^{2i\phi_n}\right)\right]^{-2} - 1\right\}. \qquad (6.56)$$

6.6 Results

The results of using this adaptive phase measurement scheme on the four alternative input states are shown in Fig. 6.1. The phase variances for states up to $N = 20$ (or $N = 30$ for $|jj\rangle_z$) were determined exactly using Eq. (6.24), whereas those for larger photon numbers were determined using the stochastic method described in Sec. 6.4. The sample sizes used were about 2^{15} for the smallest photon numbers, down to 2^{10} for the larger photon numbers. The phase variances are very close to the phase variances for canonical measurements for all of these input states.

For the optimal input states described in Ch. 5, the scaling is close to $1/N^2$, but the phase variances do differ relatively more from the canonical values for larger photon numbers. If we plot the phase variances as a ratio to the canonical phase variance (see Fig. 6.2), we find that the ratio of the phase variance to the canonical phase variance increases fairly regularly with photon number. This ratio possibly increases proportional to $\log N$. In that case the introduced phase variance would increase as $\log N/N^2$, as is the case for the optimal single mode phase measurements considered in Ch. 3.

For $|j0\rangle_z$ the variances are very close to those for canonical measurements, scaling as $1/N^{1/2}$. If we look at the distribution of the phases resulting from these measurements, we find that there is a sharp peak, but a significant number of results with large error that produce the large variance (see

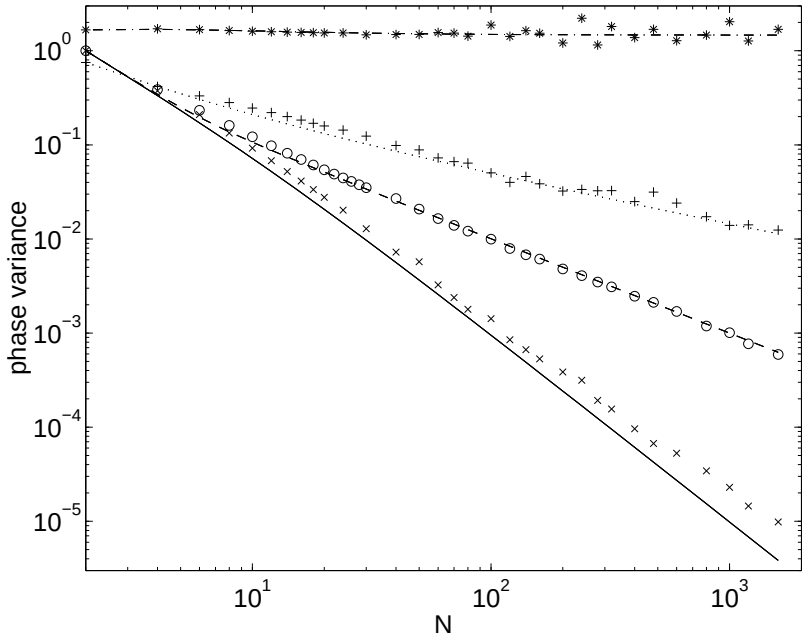

Figure 6.1: Variances in the phase estimate versus input photon number $2j$. The lines are exact results for canonical measurements on optimal states $|\psi_{\rm opt}\rangle$ (continuous line), on states with all photons incident on one input port $|jj\rangle_z$ (dashed line), on states with equal photon numbers incident on both input ports $|j0\rangle_z$ (dotted line), and the state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ (dash-dotted line). The crosses are the numerical results for the adaptive phase measurement scheme on $|\psi_{\rm opt}\rangle$, the circles are those on $|jj\rangle_z$, the pluses are those on $|j0\rangle_z$, and the asterisks are those on a $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ input state. All variances for the $|j0\rangle_z$ state are for phase modulo π.

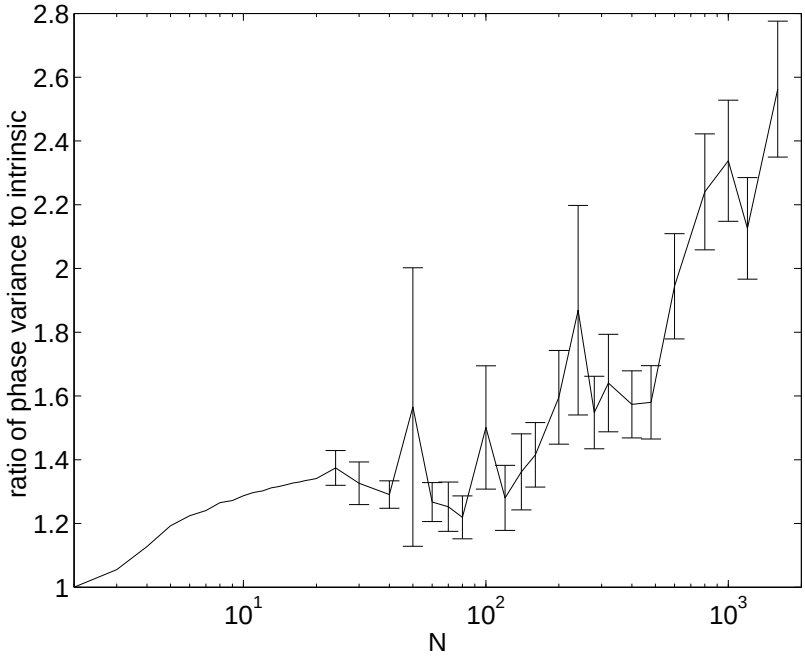

Figure 6.2: The ratio of the phase variance using the feedback scheme of Sec. 6.3 to the canonical phase variance for optimal input states.

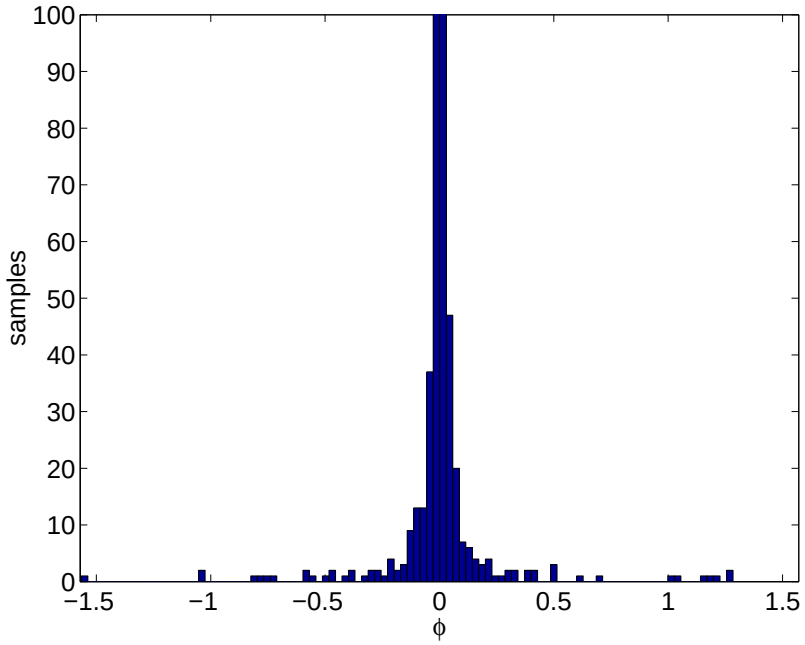

Figure 6.3: The phase distribution resulting from using adaptive phase measurements on an input state of $|j0\rangle_z$ for 800 photons. The vertical axis has been cut off at 100 (the peak count is almost 500) to show the tails more clearly.

Fig. 6.3). Similarly, for the case of the state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$, there are a large number of results at $\pm\pi$, as seen in Fig. 6.4. As for the canonical distribution, this is why the phase variance does not decrease significantly with photon number.

Now recall that although the phase uncertainty of the $|j0\rangle_z$ state as measured by the square root of the variance does not scale as N^{-1}, other measures of the phase uncertainty do scale as N^{-1}. This indicates that if the data is analysed in a way corresponding to one of these other measures, the uncertainty should scale down more rapidly with photon number. I will consider 2/3 confidence intervals, as the other measures do not make sense for discrete data.

The phase uncertainty for adaptive measurements on the $|j0\rangle_z$ state as measured by the 2/3 confidence interval is plotted in Fig. 6.5. A power law of the form cN^{-p} was fitted to this data set (for photon numbers of 20 and over), and it was found that the best fit was for $c = 1.39 \pm 0.08$ and $p = 0.69 \pm 0.01$. This means that the scaling is not as good as the N^{-1} scaling for the canonical phase distribution, but it is still far better than the $N^{-1/4}$ scaling indicated by the variance.

In contrast, the confidence interval for measurements on the optimal state (also shown in Fig. 6.5) scales very close to N^{-1}. In fact for the range of photon number considered the confidence interval for the measurements is never more than about 30% above the canonical confidence interval. This means that although the $|j0\rangle_z$ state has 2/3 confidence intervals for the canonical distribution that are close to those for the optimal state, the confidence intervals for the measurements are far worse.

As mentioned in the previous chapter, the uncertainty in the mean phase will scale as $N^{-1/4}$ rather than N^{-1} for $|j0\rangle_z$. According to the Central Limit Theorem, the probability distribution for the mean of a large number of measurements will be close to a normal distribution, even if the probability distribution for each individual sample is far from a normal distribution. In addition, the variance in the mean is the variance for an individual sample divided by the number of samples. As the probability distribution for the mean is approximately Gaussian, this means that the uncertainty for any other measure will have the same scaling as the square root of the variance, $N^{-1/4}$.

For example, for a photon number of 1600 and a sample size of 2^{10}, the uncertainty in the mean phase as indicated by the standard error (the standard deviation divided by the square root of the number of samples) is 4.0×10^{-3}. The mean phase is also around this value, at 3.5×10^{-3}. On the

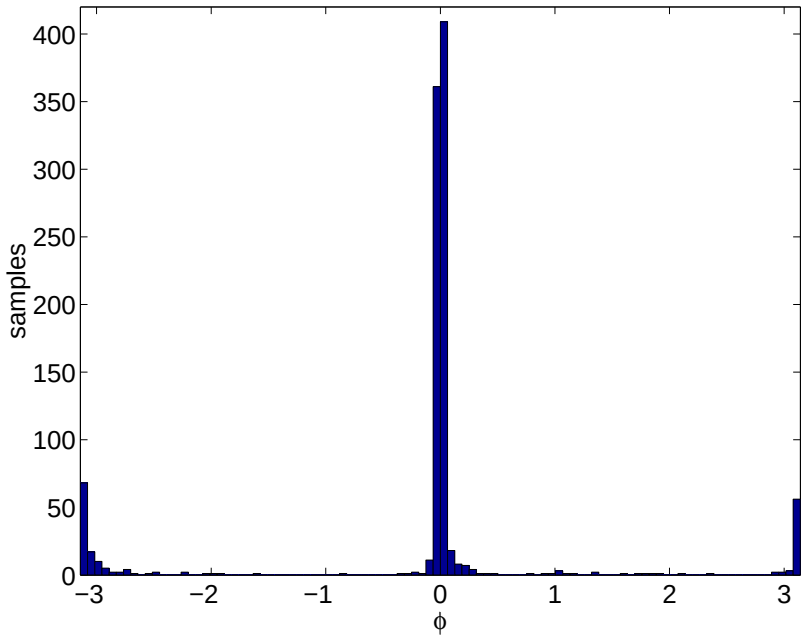

Figure 6.4: The phase distribution resulting from using adaptive phase measurements on an input state of $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ for 800 photons.

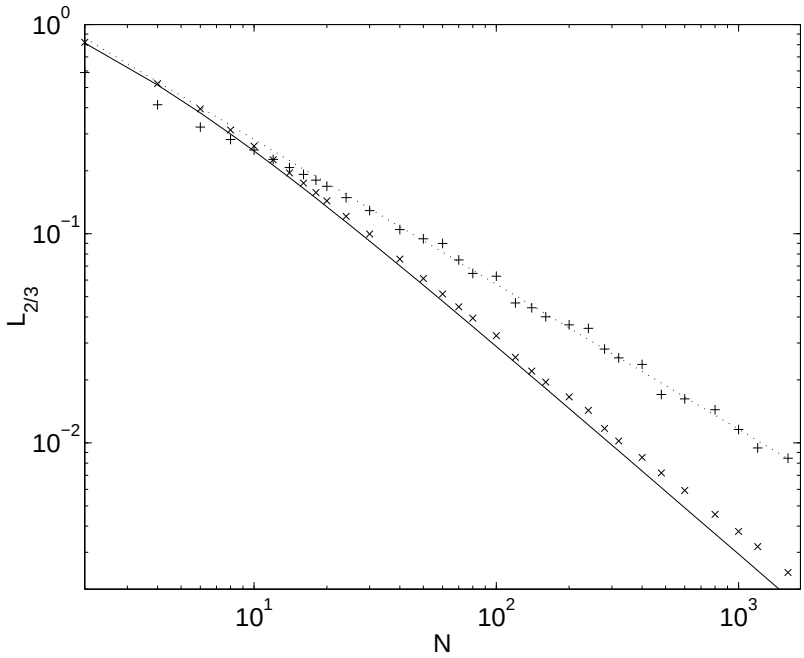

Figure 6.5: The phase uncertainty resulting from using adaptive phase measurements as measured using 2/3 confidence intervals. The pluses are the numerical results for an input state of $|j0\rangle_z$ and the crosses are those for optimal states. The continuous line is the confidence interval for the canonical distribution for $|\psi_{\rm opt}\rangle$, and the dotted line is the function fitted to the $|j0\rangle_z$ data.

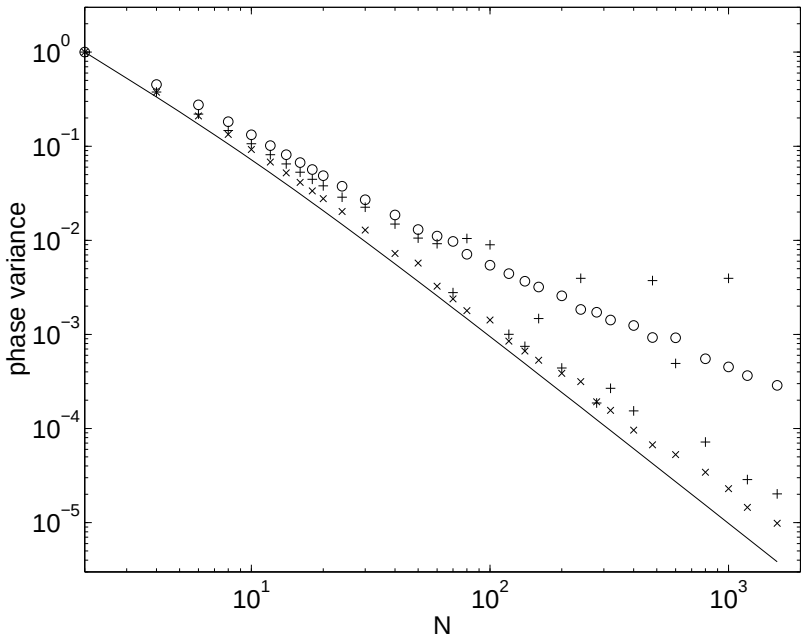

Figure 6.6: The Holevo phase variance for optimal input states under various measurement schemes. The canonical phase variance is shown as the continuous line, the results for the adaptive measurement scheme of Sec. 6.3 as crosses, the non-adaptive measurement scheme of Eq. (6.57) as circles, and the feedback scheme of Eq. (6.58) as pluses.

other hand the median is much closer to zero, at about 5.9×10^{-5}.

In order to find the best phase estimate based on the data, we would need to multiply together the probability distributions for the phase from each of the samples. Similarly to the case for the individual samples, the optimal phase estimate would then be given by $\frac{1}{2} \arg \langle e^{2i\phi} \rangle$, where the average is based on that distribution. The problem with this method is that it is extremely computationally intensive. For the example given, we would need to multiply together 1024 sums, each with 3200 terms, resulting in a total of more than 3.2 million terms.

I have also considered phase measurements using two other measurement schemes. The first is a non-adaptive phase measurement introduced in Ref. [54] and defined by

$$\Phi_m = \Phi_0 + \frac{m\pi}{N}. \tag{6.57}$$

This is analogous to heterodyne detection [34] on a single mode, in that the phase Φ equally weights all relevant values over the course of the measurement. The second scheme is a simple adaptive feedback scheme using a running estimate of the phase:

$$\begin{aligned} \Phi_m &= \arg \langle e^{i\phi} \rangle \\ &= \arg \sum_{\mu=-j+\frac{m-1}{2}}^{j-\frac{m-1}{2}} \sum_{k=-\frac{m-1}{2}}^{\frac{m-3}{2}} \psi_{\mu,m-1,k} \psi^*_{\mu,m-1,k+1}. \end{aligned} \tag{6.58}$$

This is motivated by the relative success of the analogous simple feedback scheme [34] for phase measurement of a single mode. The difference here is that there is no $\pi/2$ term. For interferometry, small phase uncertainties are obtained when the phase difference between the arms is close to zero [38, 39, 40], rather than $\pi/2$. As mentioned in the previous chapter, the reason for this difference is that a different scattering matrix for the beam splitters has been assumed for interferometry than for dyne measurements.

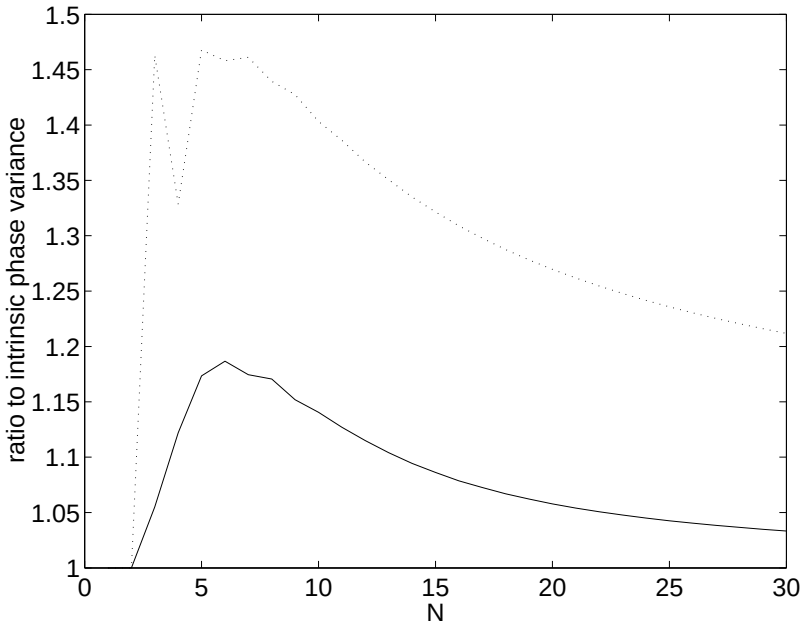

Figure 6.7: The exact phase variance for $|jj\rangle_z$ input states under two different measurement schemes as a ratio to the canonical phase variance. The results for the adaptive measurement scheme of Sec. 6.3 are shown as the continuous line, and the non-adaptive measurement scheme of Eq. (6.57) as the dotted line.

The results of using these two measurement schemes, as well as the adaptive measurement scheme of Sec. 6.3, on the optimal input states, are shown in Fig. 6.6. The non-adaptive measurement scheme is far inferior to the adaptive measurement scheme of Sec. 6.3, and the variance scales as N^{-1}. The simple adaptive feedback scheme also gives poor results. Although most of the phase results for this feedback scheme have small error, there are a small number of results with very large error. This also means that the results shown in Fig. 6.6 are fairly erratic, as the results for which a large error sample was obtained have much larger phase variance.

I also considered the non-adaptive measurement scheme on the state with all photons in one port. The exact results for that case for N up to 30 are shown in Fig. 6.7. The phase variance is not much more than the canonical phase variance, about 20% more for $N = 30$ and still decreasing. This demonstrates that for this state there is relatively little improvement in using a more advanced feedback scheme for larger photon numbers. The biggest improvement is a reduction in the variance of about 24% for $N = 3$.

6.7 Optimal Feedback

The next question is whether the adaptive measurement technique described above is optimal. Note firstly that the initial feedback phase has no effect, because it is effectively averaged over by averaging over the system phase. Secondly the last feedback phase is always optimal, as was noted above. This means that for states with 1 or 2 photons the measurement technique must be optimal. In fact, for the states considered here the phase variance was equal to the canonical phase variance for 1 or 2 photons.

To see if the phase variance was equal to canonical for arbitrary states, the complete range of possible states was considered. For 1 photon the state can be expressed as

$$|\psi\rangle = \sum_{\mu=-1/2}^{1/2} \psi_\mu |1/2, \mu\rangle_y. \tag{6.59}$$

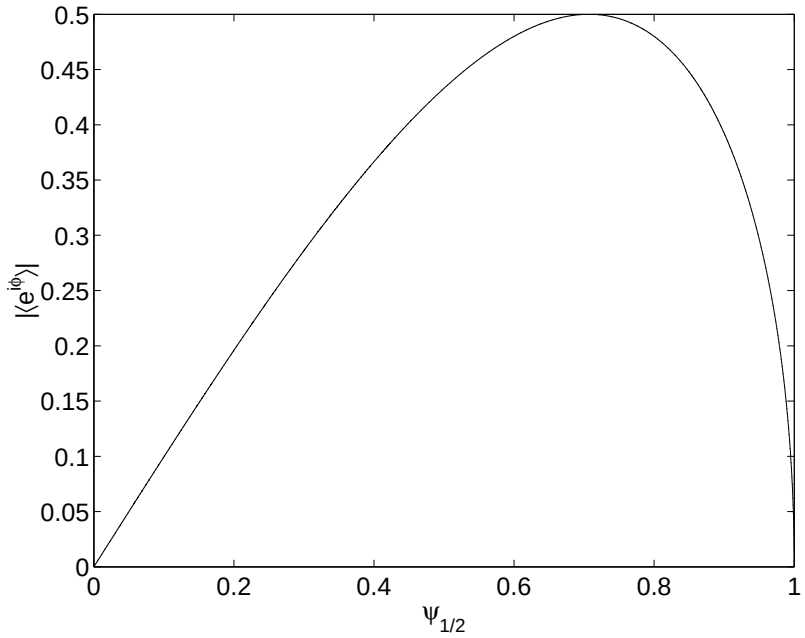

Figure 6.8: The value of $|\langle e^{i\phi}\rangle|$ for a 1 photon input state as a function of $\psi_{1/2}$. The canonical variance is shown as the continuous line and the value for the measurements is shown as the dotted line (these lines overlap).

There are therefore two coefficients, $\psi_{\pm 1/2}$, which can in general take complex values. The magnitude and phase of these coefficients give four real numbers that can be varied for these states. There is the restriction, however, that the states must be normalised, which removes one degree of freedom.

Also the absolute phase of the coefficients is irrelevant. That is, changing the phase of both coefficients $\psi_{\pm 1/2}$ by the same amount gives an equivalent state. In addition, the relative phases of $\psi_{\pm 1/2}$ are irrelevant. This is because adding a phase difference between the coefficients simply gives an equivalent input state with a phase shift. These considerations mean that only real $\psi_{\pm 1/2}$ need be considered, and there is therefore only one degree of freedom.

The value of $\psi_{1/2}$ was varied in 10000 steps from 0 to 1 (see Fig. 6.8), and it was found that the phase variance obtained was identical to the canonical phase variance for the entire range. Note that this case is independent of the feedback scheme, as there is only the initial feedback phase, which has no effect. This means that the variance is the same as the canonical variance, regardless of the input state or feedback scheme.

For the case of 2 photons the state has three coefficients. Taking account of the magnitude and phase this gives 6 real numbers that can be varied. The same considerations as for the 1 photon case apply, leaving only 3 degrees of freedom. We can vary the magnitude of two of the coefficients independently, and the phase of one.

The magnitudes of ψ_{-1} and ψ_1 and the phase of ψ_0 were varied over the complete range with 100 steps in each of the variables, and it was found that the phase variance obtained was identical to the canonical phase variance. This demonstrates that the phase variance is as good as canonical for 2 photons, independent of the input state. This case is not independent of the feedback scheme, and if any other feedback phase is used the feedback is not quite as good as canonical. These results were also obtained when the states were selected completely at random.

For the case of optimal input states with 3 or 4 photons, it was found that it is not possible to decrease the variance by altering the intermediate feedback slightly, so showing that the feedback technique is locally optimal for the phase variance. For more than 4 photons it is possible to reduce the phase variance by varying the intermediate feedback phases, and so the feedback is not optimal.

In order to show that the feedback is globally optimal for 3 or 4 photons, it is necessary to test

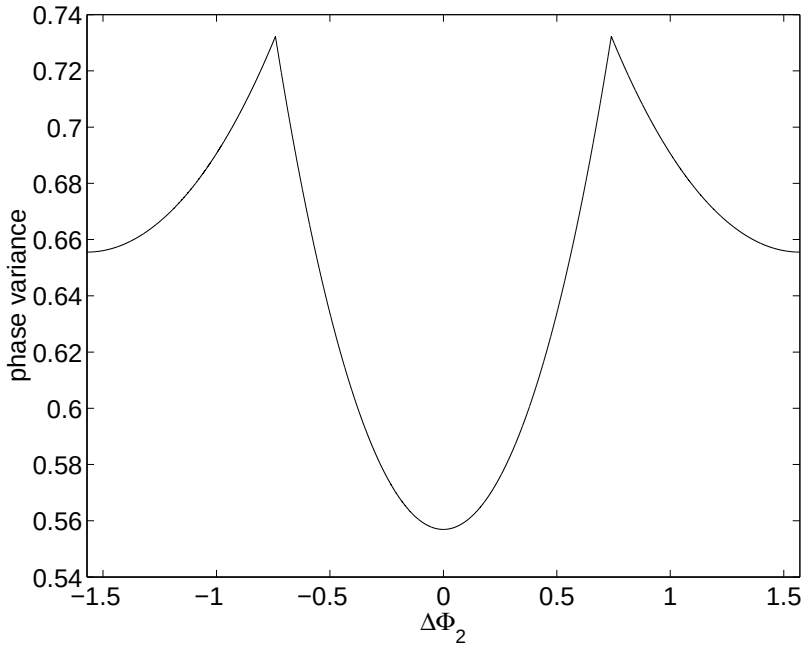

Figure 6.9: The phase variance for the 3 photon optimum input state as a function of the second feedback phase. The phase given here is relative to the second feedback phase given by the feedback scheme of Sec. 6.3. The other feedback phases are as given by this feedback scheme.

the entire phase range. Only optimal input states will be considered here, and the more general case will be considered later. There are three factors that reduce the number of phases that need be varied. The first two are as noted above: the first feedback phase has no effect (and so may be ignored), and the last feedback phase is always optimal (and so need not be varied). The third is that the contribution to the phase variance for a sequence of detections is independent of the first detection result. This is because changing the first feedback phase by π reverses the significance of the first detection results, and the first feedback phase is arbitrary.

The consequence of these three factors is that for 3 photons the variation of only one feedback phase needs to be considered, and for 4 photons the variation of three feedback phases needs to be considered. The phase variance for the 3 photon case as the second feedback phase is varied from its value for the feedback technique of Sec. 6.3 is shown in Fig. 6.9. This figure shows that this feedback technique is globally optimal for 3 photons. Since the phase variance is above the canonical phase variance in this case, this demonstrates that it is not possible to perform canonical measurements using photodetection and feedback alone. For 4 photons the second feedback phase and two third feedback phases must be varied. This case was tested with 100 steps in each of the three variables, and it was found that the feedback technique is globally optimal in this case also.

In order to see how far the phase variance could be improved for photon numbers above 4, the feedback phases were optimised using function minimisation techniques, and the results are shown in Fig. 6.10. Unfortunately the number of feedback phases increases exponentially with the photon number, making this technique infeasible very rapidly, and therefore only results up to $N = 12$ are shown. As can be seen, this optimisation only gives minor improvements in the phase variance, with the maximum reduction in the phase variance being about 3.5%.

Note that the input states used here give minimum canonical phase variance, but as the phase variances obtained for these feedback schemes differ from the canonical phase variance (above 2 photons), these input states are not necessarily optimum for these measurements. Therefore numerical optimisations were performed where both the input state and the feedback were optimised for, and the results are also shown in Fig. 6.10.

It was found that the phase variance was reduced below the results where only the feedback was

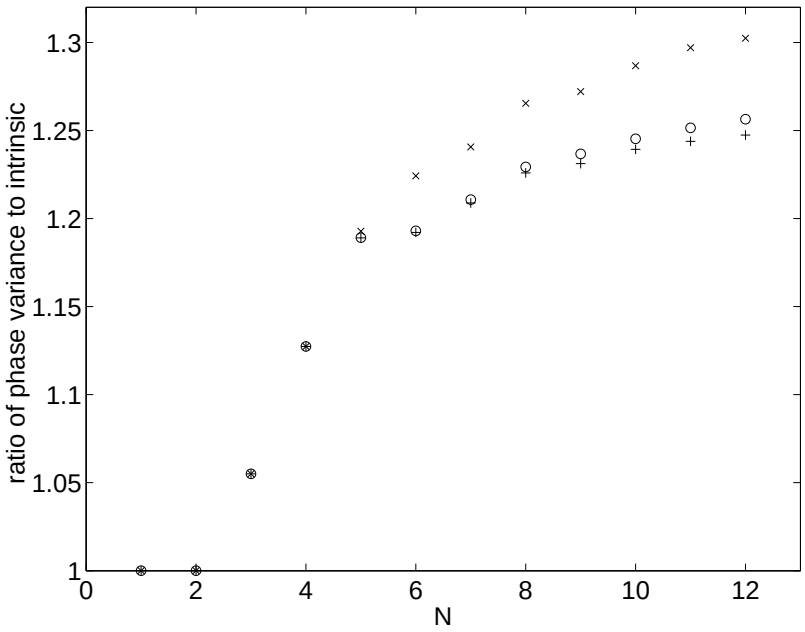

Figure 6.10: The phase variance for the feedback scheme of Sec. 6.3 and two numerically optimised feedback schemes as ratios to the minimum intrinsic phase variance. The variance for the feedback scheme of Sec. 6.3 is shown as the crosses, the case where the feedback alone is numerically optimised is shown as circles, and the case where both the state and the feedback are numerically optimised is shown as pluses.

optimised for, even for the case with 3 photons. In the case with 3 photons, the improvement is only about 0.0005%, which is not visible on the graph. The improvements for the larger photon numbers are slightly larger, almost 1% for the largest photon number calculations have been performed for. These improvements are still very minor, and much less than the improvements obtained by solving for the feedback.

For 3 or 4 photons it was found that the feedback scheme of Sec. 6.3 is still optimum for the numerically optimised states. For more general states, however, this is not the case. For example, for the 3 photon state given by

$$|\psi\rangle = \frac{1}{\sqrt{3}}\left(|3/2,-3/2\rangle_z + |3/2,1/2\rangle_z + |3/2,3/2\rangle_z\right), \tag{6.60}$$

the variation of the phase variance with the second feedback phase is as given in Fig. 6.11. The phase variance here is much smaller when the feedback phase is altered by π. In fact, the variance is a maximum when this feedback phase is at the value given by the feedback scheme of Sec. 6.3. Not only this, but the variance is less than the canonical variance for the entire range.

This is a strange result, as we would normally expect that it is impossible for the phase variance for the measurements to be smaller than the canonical phase variance. To understand the reason for this, consider the expression for $\langle e^{i\phi}\rangle$. From Eq. (1.52) of the introduction, for measurements of a single mode we have

$$\langle e^{i\phi}\rangle = \sum_{n=0}^{\infty}\langle\psi|n\rangle\langle n+1|\psi\rangle H_{n,n+1}. \tag{6.61}$$

The corresponding expression for interferometric measurements is

$$\langle e^{i\phi}\rangle = \sum_{\mu=-j}^{j-1}\langle\psi|j\mu\rangle_y\langle j,\mu+1|\psi\rangle H_{\mu,\mu+1}. \tag{6.62}$$

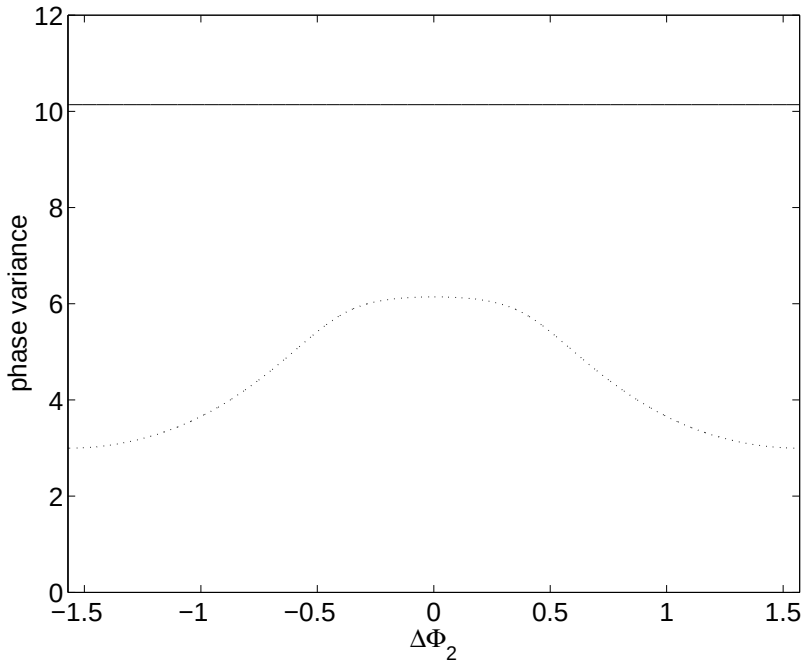

Figure 6.11: The phase variance for the 3 photon input state of Eq. (6.60) as a function of the second feedback phase. The phase given here is relative to the second feedback phase given by the feedback scheme of Sec. 6.3. The other feedback phases are as given by this feedback scheme. The continuous line is the canonical phase variance, and the dotted line is the variance for the measurements.

For canonical measurements all of the $H_{\mu,\mu+1}$ are equal to one. We would normally expect that any smaller value of $H_{\mu,\mu+1}$ would lead to a smaller value of $\left|\left\langle e^{i\phi}\right\rangle\right|$ (and a larger Holevo phase variance). This is not necessarily the case, however. For example, if all the values of ${}_y\langle j\mu|\psi\rangle$ are positive except for one negative value for $\mu = \mu_1$, then $\langle\psi|j\mu_1\rangle_y\langle j_1, \mu_1 + 1|\psi\rangle$ will be negative. In that case, it is obvious that a smaller value of H_{μ_1,μ_1+1} will lead to a smaller variance.

In general, the POM with $H_{nm} = 1$ gives the smallest variance only for states where the phases of ${}_y\langle j\mu|\psi\rangle$ vary as

$$\arg {}_y\langle j\mu|\psi\rangle = \mu\varphi + \phi_0. \tag{6.63}$$

As is discussed in Ref. [19], if the phases of ${}_y\langle j\mu|\psi\rangle$ have some more arbitrary variation

$$\arg {}_y\langle j\mu|\psi\rangle = \phi_\mu, \tag{6.64}$$

then the optimum POM (for the error in the phase) is given by

$$F(\phi) = \frac{1}{2\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\mu-\nu)\phi}|j\mu\rangle_y\langle j\nu|e^{i(\phi_\mu-\phi_\nu)}. \tag{6.65}$$

I will call this the corrected canonical POM, to avoid confusion with the optimum measurements based on feedback. With this POM we find

$$\left\langle e^{i\phi}\right\rangle = \sum_{\mu=-j}^{j-1}\left|\langle\psi|j\mu\rangle_y\langle j, \mu+1|\psi\rangle\right|. \tag{6.66}$$

Using this expression, the variance for the state of Eq. (6.60) is approximately 0.6863, which is much smaller than the variance for the measurements. This is possible because, although the ${}_z\langle j\mu|\psi\rangle$ are positive, the ${}_y\langle j\mu|\psi\rangle$ are complex.

For most of the states considered in this study, the phases of ${}_y\langle j\mu|\psi\rangle$ vary linearly as in Eq. (6.63), so using the corrected canonical POM does not change the results. The only exceptions [apart from the state of Eq. (6.60)] are:
1. The state $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$. This has a slightly smaller variance for corrected canonical measurements than for the usual canonical measurements. The results are not qualitatively changed, however, as the phase variance still remains on the order of 1 for very large photon numbers, due to the large peak at $\pm\pi$. This state is therefore still too poor to be useful.
2. Arbitrary 2 photon states. For any 2 photon state where the phases of ${}_y\langle j\mu|\psi\rangle$ do not vary linearly, although the feedback scheme of Sec. 6.3 gives a variance equal to the canonical variance, it does not give a variance as small as that for corrected canonical measurements. For other 2 photon states and for 1 photon states there is no change.

Chapter 7

Continuous Phase Measurements

Now I will consider the case of continuous adaptive measurements, where we are continuously measuring a phase that is varying. This case is closer to what is usually done in practice, where rather than measuring a single, fixed phase of a pulse, a signal is transmitted in the varying phase of a continuous beam, and we wish to measure that varying phase.

7.1 Continuous Dyne Measurements on Coherent States

Firstly I will consider the case of continuous measurements on a single mode field. For this case it is simplest to consider coherent states. For continuous coherent states, the coherent amplitude α has a constant magnitude, but varying phase. As in the single-shot case, a quadrature of the field is measured by combining the mode to be measured with a large amplitude local oscillator field that is treated classically. The two fields at the two output ports of the beam splitter are given by

$$b_{\pm}(t) = \frac{1}{\sqrt{2}}\left(a \pm \gamma\right), \tag{7.1}$$

where a is the operator for the mode to be measured and γ is the amplitude of the local oscillator field. The instantaneous rate of photodetection at each photodetector is

$$\begin{aligned}\left\langle b_{\pm}^{\dagger}(t)b_{\pm}(t)\right\rangle &= \frac{1}{2}\left\langle \left(a^{\dagger} \pm \gamma^{*}\right)\left(a \pm \gamma\right)\right\rangle \\ &= \frac{1}{2}\left(|\alpha|^{2} + |\gamma|^{2} \pm 2\mathrm{Re}\left(\alpha\gamma^{*}\right)\right).\end{aligned} \tag{7.2}$$

The signal of interest is the difference between the two photocurrents at the two detectors. The number of photocounts at each of the detectors in the time interval $[t, t+\delta t)$ will be denoted by $\delta N_{\pm}(t)$. I will use the usual definition of the signal photocurrent

$$I(t) = \lim_{\delta t \to 0} \lim_{|\gamma| \to \infty} \frac{\delta N_{+}(t) - \delta N_{-}(t)}{|\gamma|\delta t}. \tag{7.3}$$

The expectation values of the increments $\delta N_{\pm}(t)$ in the infinitesimal limit are

$$\begin{aligned}\left\langle dN_{\pm}(t)\right\rangle &= \left\langle b_{\pm}^{\dagger}(t)b_{\pm}(t)\right\rangle dt \\ \left\langle dN_{\pm}^{2}(t)\right\rangle &= \left\langle dN_{\pm}(t)\right\rangle.\end{aligned} \tag{7.4}$$

For large $|\gamma|$, $\delta N_{\pm}(t)$ can be approximated by the Gaussian increments $\delta W_{\pm}(t)$:

$$\delta N_{\pm}(t) \approx \kappa_{\pm}\delta t + \sqrt{\kappa_{\pm}}\delta W_{\pm}(t), \tag{7.5}$$

where

$$\kappa_{\pm} = \frac{1}{2}\left(|\alpha|^{2} + |\gamma|^{2} \pm 2\mathrm{Re}\left(\alpha\gamma^{*}\right)\right). \tag{7.6}$$

Using this in the definition of the photocurrent (7.3) gives

$$\begin{aligned} I(t) &= \lim_{\delta t\to 0}\lim_{|\gamma|\to\infty}\frac{\delta N_+(t)-\delta N_-(t)}{|\gamma|\delta t} \\ &= \lim_{\delta t\to 0}\frac{2\mathrm{Re}\left(\alpha e^{-i\Phi(t)}\right)\delta t+\frac{1}{\sqrt{2}}\left(\delta W_+(t)-\delta W_-(t)\right)}{\delta t} \\ &= 2\mathrm{Re}\left(\alpha e^{-i\Phi(t)}\right)+\frac{dW(t)}{dt}, \end{aligned} \tag{7.7}$$

where $\Phi(t)$ is the phase of γ. The result is therefore

$$I(t)dt = 2\mathrm{Re}\left(\alpha e^{-i\Phi(t)}\right)dt + dW(t). \tag{7.8}$$

This is identical to the result in the case of single-shot measurements, except in this case the time is not scaled to the unit interval.

In making adaptive phase measurements the phase of the local oscillator is usually taken to be

$$\Phi(t) = \hat{\varphi}(t) + \pi/2, \tag{7.9}$$

where $\hat{\varphi}(t)$ is some estimate of the system phase $\varphi(t)$. The only case where the feedback phase is not based on this (for dyne measurements) is for the corrections to the close-to-optimal measurements considered in Ch. 3. With this, the signal becomes

$$I(t)dt = 2\left|\alpha\right|\sin\left(\varphi(t)-\hat{\varphi}(t)\right)dt + dW(t). \tag{7.10}$$

7.1.1 Linear Approximation

Provided that the estimated system phase is sufficiently close to the actual system phase, we can make the linear approximation

$$I(t)dt = 2\left|\alpha\right|\left(\varphi(t)-\hat{\varphi}(t)\right)dt + dW(t). \tag{7.11}$$

Rearranging this equation gives

$$\begin{aligned} 2\left|\alpha\right|\hat{\varphi}(t)dt + I(t)dt &= 2\left|\alpha\right|\varphi(t)dt + dW(t) \\ \hat{\varphi}(t) + \frac{I(t)}{2\left|\alpha\right|} &= \varphi(t) + \frac{dW(t)}{2\left|\alpha\right|dt}, \end{aligned} \tag{7.12}$$

which gives

$$\left\langle \hat{\varphi}(t) + \frac{I(t)}{2|\alpha|} \right\rangle = \varphi(t). \tag{7.13}$$

Therefore, under the linear approximation, each data point $I(t)$ can be used to obtain an independent estimate of the phase.

Now I will denote the best phase estimate based on all the data up to time t by $\Theta(t)$. I will also denote the best phase estimate based on the data in the infinitesimal time interval $[t, t+dt)$ by $\theta(t)$. Note that these are the *best* phase estimates, in contrast to the phase estimate used in the feedback $\hat{\varphi}(t)$. The phase estimates $\theta(t)$ are given by

$$\theta(t) = \hat{\varphi}(t) + \frac{I(t)}{2|\alpha|}, \tag{7.14}$$

which has the expectation value $\varphi(t)$ and the variance

$$\begin{aligned} \left\langle \left(\theta(t)-\varphi(t)\right)^2 \right\rangle &= \left\langle \left(\frac{dW(t)}{2\left|\alpha\right|dt}\right)^2 \right\rangle \\ &= \frac{1}{4\left|\alpha\right|^2 dt}. \end{aligned} \tag{7.15}$$

Here the simple definition of the variance has been used, rather than the Holevo phase variance, because we are using the linear approximation.

If we were considering a measurement over the time interval $[0, 1]$, during which there are $\overline{n} = |\alpha|^2$ photons from the signal mode, then we would be effectively averaging over $1/dt$ estimates of the phase, each with a variance given by Eq. (7.15). As usual with averages, the variance in the mean is the variance of each data point divided by the number of data points. This means that the variance in the final phase estimate is

$$\begin{aligned} \frac{1\Big/\left(4\left|\alpha\right|^2 dt\right)}{1/dt} &= \frac{1}{4\left|\alpha\right|^2} \\ &= \frac{1}{4\overline{n}}, \end{aligned} \tag{7.16}$$

which is the standard result for a coherent state under adaptive phase measurements.

In this case we want to consider a measurement that is continued indefinitely. If the system phase is not varied, then the variance in the phase estimate will just go down indefinitely. In that case, however, no information is transmitted beyond a single real number. What we wish to do is vary the system phase in order to transmit information, as is the case for FM radio.

Next is the question of how the system phase $\varphi(t)$ will be varied. The simplest way of varying the system phase is to vary it stochastically via Wiener increments. I will therefore take the variation in the system phase to be

$$\varphi(t + dt) = \varphi(t) + \kappa dW'(t). \tag{7.17}$$

This Wiener increment is independent from that used previously for the photocurrent, as indicated by the prime.

It is clear that the most recent data will be least affected by the variation in the system phase, and the data from further and further back will be less and less accurate. To quantify this, note that when the current system phase $\varphi(t)$ is used as a reference, the variance in the system phase at some previous time t' is $\kappa^2(t - t')$. As the phase estimate $\theta(t')$ based on the time interval $[t', t' + dt')$ is an estimate of the system phase at time t', when it is considered relative to the current system phase its variance will be increased by $\kappa^2(t - t')$. Therefore the total variance in each individual phase estimate $\theta(t')$ is

$$\frac{1}{4\left|\alpha\right|^2 dt} + \kappa^2(t - t'). \tag{7.18}$$

In order to determine the current best phase estimate $\Theta(t)$, we would like to form a weighted average of each of the individual phase estimates $\theta(t')$. If this is done in the usual way for weighted averages, using the variance given by (7.18), the results obtained are ridiculous, as the contribution to the variance from the variation in the system phase is infinitesimal as compared to the quantum noise. The problem is that the variation in the system phase makes the error in the phase estimates $\theta(t')$ correlated, so the contribution from the variation in the system phase is infinitesimal over infinitesimal time intervals, but becomes significant over finite time intervals. This means that the usual method of performing weighted averages will not work.

As an alternative approach, we can consider just the weighting of the latest phase estimate $\theta(t)$ as compared to the phase estimate from all the previous data $\Theta(t)$. In this case $\theta(t)$ has no variance from the variation in the system phase, and so it is uncorrelated with the previous phase estimate and we can use a weighted average in the usual way.

The equilibrium value of the variance of $\Theta(t)$, with all the individual phase estimates correctly weighted, will be denoted by $\Delta\Theta^2$. After a time dt the phase variance of $\Theta(t)$ with respect to the new system phase $\varphi(t + dt)$, i.e.

$$\left\langle \left(\Theta(t) - \varphi(t + dt)\right)^2 \right\rangle, \tag{7.19}$$

will be $\Delta\Theta^2 + \kappa^2 dt$. The variance in the phase estimate from the latest time interval, $\theta(t)$, will be given by Eq. (7.15). If we take a weighted average of $\Theta(t)$ and $\theta(t)$, then the contributions from

each of the phase estimates from the individual time intervals should be correctly weighted, and the variance in the weighted average should be the equilibrium value, $\Delta\Theta^2$. This implies that

$$\frac{1}{\Delta\Theta^2 + \kappa^2 dt} + 4\left|\alpha\right|^2 dt = \frac{1}{\Delta\Theta^2}. \tag{7.20}$$

This expression can be used to determine the equilibrium value of the phase variance. Solving for $\Delta\Theta^2$ gives

$$\Delta\Theta^2 = \frac{\kappa}{2\left|\alpha\right|}. \tag{7.21}$$

Thus we find that for continuous measurements the phase variance scales as $1/|\alpha|$ rather than $1/|\alpha|^2$.

Showing explicitly how the weighted average is performed,

$$\Theta(t+dt) = \frac{\left(4\left|\alpha\right|^2 dt\right)\theta(t) + \frac{1}{\Delta\Theta^2+\kappa^2 dt}\Theta(t)}{1/\Delta\Theta^2}. \tag{7.22}$$

Simplifying this gives

$$\Theta\left(t+dt\right) = \left(2\left|\alpha\right|\kappa\, dt\right)\theta(t) + \left(1 - 2\left|\alpha\right|\kappa\, dt\right)\Theta\left(t\right). \tag{7.23}$$

In terms of the increment in the phase estimate this is

$$d\Theta(t) + \left(2\left|\alpha\right|\kappa dt\right)\Theta\left(t\right) = \left(2\left|\alpha\right|\kappa dt\right)\theta(t). \tag{7.24}$$

Solving this we find that

$$\begin{aligned} d\Theta(t)e^{2|\alpha|\kappa t} + \left(2\left|\alpha\right|\kappa\, dt\right)\Theta(t)e^{2|\alpha|\kappa t} &= \left(2\left|\alpha\right|\kappa\, dt\right)\theta(t)e^{2|\alpha|\kappa t} \\ d\left(\Theta(t)e^{2|\alpha|\kappa t}\right) &= \left(2\left|\alpha\right|\kappa\right)\theta(t)e^{2|\alpha|\kappa t}dt \\ \Theta(t) &= 2\left|\alpha\right|\kappa \int_{-\infty}^{t} \theta(s)e^{2|\alpha|\kappa(s-t)}ds. \end{aligned} \tag{7.25}$$

Therefore this method corresponds to a simple negative exponential scaling of the weighting.

We can also consider a more general negative exponential scaling given by

$$\Theta(t) = \chi \int_{-\infty}^{t} \theta(s)e^{\chi(s-t)}ds. \tag{7.26}$$

Note that with this more general scaling, $\Theta(t)$ is no longer necessarily the best phase estimate. For most of the remainder of this chapter, $\Theta(t)$ will be used in this more general sense, rather than as specifically the best phase estimate. The best phase estimate will be found by finding the optimum value of χ. Taking the derivative of this expression with respect to time gives

$$\begin{aligned} d\left(\Theta(t)e^{\chi t}\right) &= \chi\theta(t)e^{\chi t}dt \\ d\Theta(t) &= \chi dt\left(\theta(t) - \Theta(t)\right) \\ \Theta(t+dt) &= \chi dt\theta(t) + \left(1 - \chi dt\right)\Theta(t). \end{aligned} \tag{7.27}$$

This means that this method is again a weighted average, except with a weighting that is not optimum. If we find the variance of both sides of this equation we obtain

$$\begin{aligned} \mathrm{var}\left(\Theta(t+dt)\right) &= \left(\chi dt\right)^2 \mathrm{var}\left(\theta(t)\right) + \left(1-\chi dt\right)^2 \mathrm{var}\left(\Theta(t)\right) \\ \Delta\Theta^2 &= \frac{\left(\chi dt\right)^2}{4\left|\alpha\right|^2 dt} + \left(1 - 2\chi dt\right)\left(\Delta\Theta^2 + \kappa^2 dt\right) \\ \Delta\Theta^2 &= \frac{\chi^2 dt}{4\left|\alpha\right|^2} + \Delta\Theta^2 + \kappa^2 dt - 2\chi dt\Delta\Theta^2. \end{aligned} \tag{7.28}$$

Solving for $\Delta\Theta^2$ gives

$$\Delta\Theta^2 = \frac{\chi}{8\left|\alpha\right|^2} + \frac{\kappa^2}{2\chi}. \tag{7.29}$$

The optimum value of χ can be verified from this equation. Taking the derivative with respect to χ gives

$$\frac{\partial}{\partial\chi}\left(\Delta\Theta^2\right) = \frac{1}{8\left|\alpha\right|^2} - \frac{\kappa^2}{2\chi^2}. \tag{7.30}$$

For the variance to be minimised this must be zero, so

$$\begin{aligned} \frac{1}{8\left|\alpha\right|^2} &= \frac{\kappa^2}{2\chi^2} \\ \chi &= 2\left|\alpha\right|\kappa. \end{aligned} \tag{7.31}$$

This is the exponential constant which was found directly. Substituting this value of χ into the expression for the variance gives

$$\begin{aligned} \Delta\Theta^2 &= \frac{2\left|\alpha\right|\kappa}{8\left|\alpha\right|^2} + \frac{\kappa^2}{4\left|\alpha\right|\kappa} \\ &= \frac{\kappa}{2\left|\alpha\right|}, \end{aligned} \tag{7.32}$$

which is the result found for the phase variance in Eq. (7.21).

7.1.2 Exact Case

The results of the previous section are all using the linear approximation (7.11). Although this approximation is very useful for obtaining the asymptotic value of the variance, it does not directly tell us what to do in the exact case. In the exact case for single-shot measurements (see Sec. 1.5), rather than forming independent phase estimates from each time interval dt and then averaging them, we determine A_v and B_v, and the phase estimate is given by

$$\Theta(v) = \arg\left(vA_v + B_v A_v^*\right). \tag{7.33}$$

Therefore the average phase estimate with exponential weighting does not make much sense in this case. In addition the intermediate phase estimate must be considered. An alternative approach is to use an exponential weighting in determining A_v and B_v, and then use these to determine the phase estimate. Specifically, I will replace the definitions of A_v and B_v,

$$\begin{aligned} A_v &= \int_0^v e^{i\Phi} I(u) du \\ B_v &= -\int_0^v e^{2i\Phi} du, \end{aligned} \tag{7.34}$$

by

$$\begin{aligned} A_t &= \int_{-\infty}^t e^{\chi(u-t)} e^{i\Phi} I\left(u\right) du \\ B_t &= -\int_{-\infty}^t e^{\chi(u-t)} e^{2i\Phi} du. \end{aligned} \tag{7.35}$$

Then $\arg A_t$ can still be used as the intermediate phase estimate. I will not consider any better intermediate phase estimates here, as these only give very small improvements over the mark II case for coherent states. To find a phase estimate to use for $\Theta(t)$, we can use a similar approach to that used in Ref. [35]. If the system phase is constant, then we find

$$\begin{aligned} A_t &= \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi} I(u) du \\ &= \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi} \left[\left(\alpha e^{-i\Phi} + \alpha^* e^{i\Phi} \right) du + dW(u) \right] \\ &= \alpha \int_{-\infty}^{t} e^{\chi(u-t)} du + \alpha^* \int_{-\infty}^{t} e^{\chi(u-t)} e^{2i\Phi} du + \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi} dW(u) \\ &= \frac{\alpha}{\chi} - \alpha^* B_t + i\sigma_t. \end{aligned} \tag{7.36}$$

where

$$\sigma_t = \int_{-\infty}^{t} e^{\chi(u-t)} e^{i(\Phi-\pi/2)} dW(u). \tag{7.37}$$

This result is analogous to the result (1.77) for the case of single-shot measurements, except with v replaced with $1/\chi$. Note that from this derivation it naturally emerges that we should use the same exponential scaling for B_t as for A_t. From Eq. (7.36) it can be shown that

$$A_t + \chi B_t A_t^* = \alpha \left(\frac{1}{\chi} - \chi \left| B_t \right|^2 \right) + i\sigma_t - i\chi B_t \sigma_t^*. \tag{7.38}$$

This means that

$$\langle A_t + \chi B_t A_t^* \rangle \approx \alpha \left(\frac{1}{\chi} - \chi \left| B_t \right|^2 \right). \tag{7.39}$$

Similarly to the single-shot case this is not necessarily exact if the local oscillator phase is dependent on the measurement record, but it should still be approximately true. Therefore the phase estimate that will be used here is

$$\Theta(t) = \arg(A_t + \chi B_t A_t^*). \tag{7.40}$$

Note that the factor of χ makes sense, as the time over which the data is used is effectively $1/\chi$. Similarly to the single-shot case, I will define the variable $C_t = A_t + \chi B_t A_t^*$, so $\Theta(t) = \arg C_t$. The above derivation is not exact if the system phase is not constant; however, $\arg C_t$ should still be a good estimator for the phase.

A differential equation for the feedback phase can be determined in a similar way as in Ref. [35]. Using Eq. (7.35), we can determine the increment in A_t:

$$\begin{aligned} e^{\chi t} A_t &= \int_{-\infty}^{t} e^{\chi u} e^{i\Phi} I\left(u\right) du \\ d\left(e^{\chi t} A_t \right) &= e^{\chi t} e^{i\Phi} I(t) dt \\ dA_t &= e^{i\Phi} I(t) dt - \chi A_t dt. \end{aligned} \tag{7.41}$$

Taking the local oscillator phase to be

$$\Phi(t) = \arg A_t + \frac{\pi}{2}, \tag{7.42}$$

as in the case of mark II measurements, we find that

$$dA_t = i\frac{A_t}{|A_t|}I(t)dt - \chi A_t dt. \tag{7.43}$$

So the magnitude of A_t varies as

$$\begin{aligned} d\left|A_t\right|^2 &= A_t^*\left(dA_t\right) + \left(dA_t^*\right)A_t + \left(dA_t^*\right)\left(dA_t\right) \\ &= A_t^*\left(i\frac{A_t}{|A_t|}I(t)dt - \chi A_t dt\right) + \left(-i\frac{A_t^*}{|A_t|}I(t)dt - \chi A_t^* dt\right)A_t + dt \\ &= \left(1 - 2\chi\left|A_t\right|^2\right)dt. \end{aligned} \tag{7.44}$$

This demonstrates that, rather than increasing linearly as in the standard case, $|A_t|$ increases up to an equilibrium value given by

$$\left|A_t\right|^2 = \frac{1}{2\chi}. \tag{7.45}$$

Using this result, the increment in the feedback phase is

$$\begin{aligned} d\Phi(t) &= \mathrm{Im}\left[d\ln A_t\right] \\ &= \mathrm{Im}\left[\frac{dA_t}{A_t} - \frac{\left(dA_t\right)^2}{2A_t^2}\right] \\ &= \mathrm{Im}\left[\frac{i\frac{A_t}{|A_t|}I(t)dt - \chi A_t dt}{A_t} + \frac{\frac{A_t^2}{|A_t|^2}dt}{2A_t^2}\right] \\ &= \frac{I(t)dt}{|A_t|} \\ &= \sqrt{2\chi}I(t)dt. \end{aligned} \tag{7.46}$$

Therefore the feedback for this case is actually much simpler than for the single-shot case. The feedback phase just changes linearly with the signal, and there is no $\sqrt{t}$ scaling as there is in the standard case.

Using this result gives the stochastic differential equation for the phase estimate $\hat{\varphi}(t)$ as

$$\begin{aligned} d\hat{\varphi}(t) &= \sqrt{2\chi}\left[I(t)dt\right] \\ &= \sqrt{2\chi}\left[2\left|\alpha\right|\sin(\varphi(t) - \hat{\varphi}(t))dt + dW(t)\right]. \end{aligned} \tag{7.47}$$

Unlike the standard case, no change of variables is required, as there is no factor of $\sqrt{t}$. Making a linear approximation gives

$$d\hat{\varphi}(t) = \sqrt{2\chi}\left[2\left|\alpha\right|\left(\varphi(t) - \hat{\varphi}(t)\right)dt + dW(t)\right]. \tag{7.48}$$

Rearranging this gives

$$\begin{aligned} d\hat{\varphi}(t) + 2\left|\alpha\right|\sqrt{2\chi}\hat{\varphi}(t)dt &= 2\left|\alpha\right|\sqrt{2\chi}\varphi(t)dt + \sqrt{2\chi}dW(t) \\ d\left[e^{2|\alpha|\sqrt{2\chi}t}\hat{\varphi}(t)\right] &= e^{2|\alpha|\sqrt{2\chi}t}\left[2\left|\alpha\right|\sqrt{2\chi}\varphi(t)dt + \sqrt{2\chi}dW(t)\right]. \end{aligned}$$

Integrating then gives the solution as

$$\hat{\varphi}(t) = \sqrt{2\chi}\int_{-\infty}^{t} e^{2|\alpha|\sqrt{2\chi}(u-t)}\left[2\left|\alpha\right|\varphi(u)du + dW(u)\right]. \tag{7.49}$$

If the phase is measured relative to the current system phase, then

$$\varphi(u) = -\kappa \int_u^t dW'(s). \tag{7.50}$$

Using this, the solution for the phase estimate is

$$\hat{\varphi}(t) = \sqrt{2\chi} \int_{-\infty}^t e^{2|\alpha|\sqrt{2\chi}(u-t)} \left[dW(u) - 2|\alpha| \kappa \int_u^t dW'(s) du \right]. \tag{7.51}$$

The variance in this phase estimate will be

$$\begin{aligned}
\left\langle \hat{\varphi}^2(t) \right\rangle &= 2\chi \left\langle \int_{-\infty}^t du_1 \int_{-\infty}^t du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} 4|\alpha|^2 \kappa^2 \int_{u_1}^t \int_{u_2}^t dW'(s_2) dW'(s_1) \right\rangle \\
&\quad + 2\chi \left\langle \int_{-\infty}^t \int_{-\infty}^t e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} dW(u_2) dW(u_1) \right\rangle \\
&= 8\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{-\infty}^t du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} \int_{\max(u_1,u_2)}^t ds + 2\chi \int_{-\infty}^t e^{4|\alpha|\sqrt{2\chi}(u-t)} du \\
&= 8\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{-\infty}^{u_1} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} \int_{u_1}^t ds \\
&\quad + 8\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{u_1}^t du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} \int_{u_2}^t ds + 2\chi \left[\frac{e^{4|\alpha|\sqrt{2\chi}(u-t)}}{4|\alpha|\sqrt{2\chi}} \right]_{-\infty}^t \\
&= 8\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{-\infty}^{u_1} e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} (t-u_1) du_2 \\
&\quad + 8\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{u_1}^t e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} (t-u_2) du_2 + \frac{\sqrt{2\chi}}{4|\alpha|} \\
&= 16\chi |\alpha|^2 \kappa^2 \int_{-\infty}^t du_1 \int_{-\infty}^{u_1} e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-2t)} (t-u_1) du_2 + \frac{\sqrt{2\chi}}{4|\alpha|} \\
&= \frac{16\chi |\alpha|^2 \kappa^2}{2|\alpha|\sqrt{2\chi}} \int_{-\infty}^t e^{4|\alpha|\sqrt{2\chi}(u_1-t)} (t-u_1) du_1 + \frac{\sqrt{2\chi}}{4|\alpha|} \\
&= \frac{16\chi |\alpha|^2 \kappa^2}{2|\alpha|\sqrt{2\chi}\left(4|\alpha|\sqrt{2\chi}\right)^2} + \frac{\sqrt{2\chi}}{4|\alpha|} \\
&= \frac{\kappa^2}{4|\alpha|\sqrt{2\chi}} + \frac{\sqrt{2\chi}}{4|\alpha|}.
\end{aligned} \tag{7.52}$$

This result for the variance of the intermediate phase estimate is quite different from that for the variance of $\Theta(t)$ given in Eq. (7.29). Of particular interest is the fact that the contribution due to the variation in the system phase is not simply $\kappa^2/(2\chi)$, as is usually the case.

At first it would seem that the method used to find the mark II phase variance in [35] would be applicable here also. It turns out that this method is not useable, because when the system phase

varies it is not possible to separate the different terms as in Eq. (7.36). A more promising way is to use a method similar to that in Eq. (7.52). The phase estimate can be simplified to

$$\begin{aligned}\Theta(t) &= \arg\left(A_t + \chi B_t A_t^*\right) \\ &= \arg A_t + \arg\left(1 + \chi B_t A_t^*/A_t\right) \\ &= \hat{\varphi}(t) + \arg\left(1 + \chi e^{-2i\hat{\varphi}(t)} B_t\right).\end{aligned} \tag{7.53}$$

Expressing B_v as an integral gives

$$\Theta(t) = \hat{\varphi}(t) + \arg\left(1 + \chi e^{-2i\hat{\varphi}(t)} \int_{-\infty}^{t} e^{\chi(u-t)} e^{2i\hat{\varphi}(u)} du\right). \tag{7.54}$$

Expanding the exponentials to first order we get

$$\begin{aligned}\Theta(t) &\approx \hat{\varphi}(t) + \arg\left(1 + \chi\left(1 - 2i\hat{\varphi}(t)\right) \int_{-\infty}^{t} e^{\chi(u-t)} \left(1 + 2i\hat{\varphi}(u)\right) du\right) \\ &= \hat{\varphi}(t) + \arg\left[1 + \left(1 - 2i\hat{\varphi}(t)\right)\left(1 + 2i\chi \int_{-\infty}^{t} e^{\chi(u-t)} \hat{\varphi}(u) du\right)\right] \\ &\approx \hat{\varphi}(t) + \arg\left(2 - 2i\hat{\varphi}(t) + 2i\chi \int_{-\infty}^{t} e^{\chi(u-t)} \hat{\varphi}(u) du\right) \\ &= \hat{\varphi}(t) + \arg\left(1 - i\hat{\varphi}(t) + i\chi \int_{-\infty}^{t} e^{\chi(u-t)} \hat{\varphi}(u) du\right) \\ &\approx \hat{\varphi}(t) - \hat{\varphi}(t) + \chi \int_{-\infty}^{t} e^{\chi(u-t)} \hat{\varphi}(u) du \\ &= \chi \int_{-\infty}^{t} \hat{\varphi}(u) e^{\chi(u-t)} du.\end{aligned} \tag{7.55}$$

This demonstrates that the mark II phase estimate is approximately a weighted average of the intermediate phase estimates, just as in the standard case it is approximately a normal average. Note also the similarity of this result to the result for the linear case (7.26). Unfortunately the simple technique used in the linear case cannot be applied here. The problem here is that the phase estimates $\hat{\varphi}(t)$ are based on the previous data, not just on the data from the infinitesimal time interval $[t, t+dt)$. This means that $\hat{\varphi}(t)$ is not independent of $\Theta(t)$, and we therefore cannot use the simple techniques based on weighted averages.

Using the result (7.51) for the intermediate phase estimate gives

$$\Theta(t) \approx \sqrt{2\chi^3} \int_{-\infty}^{t} e^{\chi(v-t)} \left[\int_{-\infty}^{v} e^{2|\alpha|\sqrt{2\chi}(u-v)} \left(dW(u) - 2|\alpha|\kappa \int_{u}^{t} dW'(s) du\right)\right] dv. \tag{7.56}$$

Note that the integral for the system phase variation is taken up to time t, rather than the time of the intermediate phase estimate. This is because the phase is measured relative to the current system phase, rather than the system phase at the time of that intermediate phase estimate.

In order to determine the variance, we need to evaluate

$$\left\langle \Theta^2(t) \right\rangle = 2\chi^3 \left\langle \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} dW(u_1) \int_{-\infty}^{v_2} dW(u_2) e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \right\rangle$$

$$+8\chi^3 |\alpha|^2 \kappa^2 \left\langle \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{u_1}^{t} dW'(s_1) \int_{u_2}^{t} dW'(s_2) \right\rangle . \tag{7.57}$$

This is a lengthy calculation, and is performed in Appendix A.4. The result found is

$$\left\langle \Theta^2(t) \right\rangle \approx \frac{\chi}{8|\alpha|^2} + \frac{\kappa^2}{2\chi}. \tag{7.58}$$

This is exactly the same result as for the simplified linear case (7.29), and the minimum phase variance is

$$\left\langle \Theta^2(t) \right\rangle_{\min} \approx \frac{\kappa}{2|\alpha|} \tag{7.59}$$

for

$$\chi = 2|\alpha|\kappa. \tag{7.60}$$

7.2 Continuous Heterodyne Measurements

In order to determine how much of an improvement feedback gives for continuous measurements, I will compare it with the case of continuous heterodyne measurements. For heterodyne measurements on a pulsed coherent state, the introduced phase variance is equal to the intrinsic phase variance. This indicates that the first term in Eq. (7.58) should be double for the heterodyne case, so the phase variance is

$$\left\langle \Theta^2(t) \right\rangle \approx \frac{\chi}{4|\alpha|^2} + \frac{\kappa^2}{2\chi}. \tag{7.61}$$

This can be shown more rigorously using a similar technique to that used in [35]. Expanding A_v gives

$$\begin{aligned} A_t &= \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi(u)} I(u) du \\ &= \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi(u)} \left[\left(\alpha e^{-i\Phi} + \alpha^* e^{i\Phi} \right) du + dW(u) \right] \\ &= |\alpha| \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\varphi(u)} du + |\alpha| \int_{-\infty}^{t} e^{\chi(u-t)} e^{2i\Phi(u)-i\varphi(u)} du + \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\Phi(u)} dW(u). \end{aligned} \tag{7.62}$$

For the heterodyne case, the local oscillator phase $\Phi(t)$ varies very rapidly, so the second term above will be very small. This means that A_t simplifies to

$$A_t = |\alpha| \int_{-\infty}^{t} e^{\chi(u-t)} e^{i\varphi(u)} du + i\sigma_t. \tag{7.63}$$

Since B_v is negligible, the phase estimate $\Theta(t)$ simplifies to

$$\Theta(t) = \arg A_v. \tag{7.64}$$

As above, the phase will be measured relative to the current system phase. In the limit of small phase variance, the system phase does not vary significantly during the time $1/\chi$, so we can take

the linear approximation, giving

$$\begin{aligned} A_t &\approx |\alpha| \int_{-\infty}^{t} e^{\chi(u-t)}(1+i\varphi(u))du + i\sigma_t \\ &= \frac{|\alpha|}{\chi} + i\,|\alpha| \int_{-\infty}^{t} e^{\chi(u-t)}\varphi(u)du + i\sigma_t. \end{aligned} \tag{7.65}$$

Using this the phase estimate is

$$\begin{aligned} \Theta(t) &= \mathrm{Im}\left[\log\left(\frac{|\alpha|}{\chi} + i\,|\alpha| \int_{-\infty}^{t} e^{\chi(u-t)}\varphi(u)du + i\sigma_t\right)\right] \\ &= \mathrm{Im}\left[\log\left(1 + i\chi \int_{-\infty}^{t} e^{\chi(u-t)}\varphi(u)du + i\chi\sigma_t/\,|\alpha|\right)\right] \\ &\approx \mathrm{Im}\left[i\chi \int_{-\infty}^{t} e^{\chi(u-t)}\varphi(u)du + i\chi\sigma_t/\,|\alpha|\right]. \end{aligned} \tag{7.66}$$

In the last line the linear approximation has again been used. Further evaluating this gives

$$\begin{aligned} \Theta(t) &= \chi \int_{-\infty}^{t} e^{\chi(u-t)}\varphi(u)du + \frac{\chi}{2\,|\alpha|}(\sigma_t + \sigma_t^*) \\ &= -\kappa\chi \int_{-\infty}^{t} du\, e^{\chi(u-t)} \int_{u}^{t} dW'(s) + \frac{\chi}{2\,|\alpha|}(\sigma_t + \sigma_t^*). \end{aligned} \tag{7.67}$$

The variance is therefore

$$\left\langle \Theta^2(t) \right\rangle = \kappa^2\chi^2 \left\langle \int_{-\infty}^{t} du_1 \int_{-\infty}^{t} du_2 e^{\chi(u_1+u_2-2t)} \int_{u_1}^{t} dW'(s_1) \int_{u_2}^{t} dW'(s_2) \right\rangle + \frac{\chi^2}{4\,|\alpha|^2} \left\langle (\sigma_t + \sigma_t^*)^2 \right\rangle. \tag{7.68}$$

The first term here can be evaluated to give

$$\begin{aligned} &\left\langle \int_{-\infty}^{t} du_1 \int_{-\infty}^{t} du_2 e^{\chi(u_1+u_2-2t)} \int_{u_1}^{t} dW'(s_1) \int_{u_2}^{t} dW'(s_2) \right\rangle \\ &= \int_{-\infty}^{t} ds_1 \int_{-\infty}^{t} ds_2 e^{\chi(s_1+s_2-2t)} \int_{\max(s_1,s_2)}^{t} du \\ &= \int_{-\infty}^{t} ds_1 \int_{-\infty}^{s_1} ds_2 e^{\chi(s_1+s_2-2t)} \int_{s_1}^{t} du + \int_{-\infty}^{t} ds_1 \int_{s_1}^{t} ds_2 e^{\chi(s_1+s_2-2t)} \int_{s_2}^{t} du \\ &= \int_{-\infty}^{t} ds_1 \int_{-\infty}^{s_1} ds_2 e^{\chi(s_1+s_2-2t)}(t-s_1) + \int_{-\infty}^{t} ds_1 \int_{s_1}^{t} ds_2 e^{\chi(s_1+s_2-2t)}(t-s_2) \\ &= 2\int_{-\infty}^{t} ds_1 \int_{-\infty}^{s_1} ds_2 e^{\chi(s_1+s_2-2t)}(t-s_1) \end{aligned}$$

$$= \frac{2}{\chi}\int_{-\infty}^{t} ds_1 e^{2\chi(s_1-t)}(t-s_1)$$
$$= \frac{1}{2\chi^3}. \tag{7.69}$$

In addition, it is easy to show that

$$\left\langle \sigma_t^2 \right\rangle = -\int_{-\infty}^{t} e^{2\chi(u-t)} e^{2i\Phi(u)} du. \tag{7.70}$$

Note that $\left\langle \sigma_t^2 \right\rangle \neq B_t$, so these variables are not completely analogous to those defined in the single-shot case. Nevertheless, as Φ is rotating rapidly in the heterodyne case, we should still find that $\left\langle \sigma_t^2 \right\rangle \approx 0$. Similarly, evaluating $\left\langle |\sigma_t|^2 \right\rangle$ gives

$$\left\langle |\sigma_t|^2 \right\rangle = \int_{-\infty}^{t} e^{\chi(u_1+u_2-2t)} e^{i(\Phi(u_1)-\Phi(u_2))} dW(u_1) dW(u_2)$$
$$= \int_{-\infty}^{t} e^{2\chi(u-t)} du$$
$$= \frac{1}{2\chi}. \tag{7.71}$$

Using these results, as well as Eq. (7.69), the variance is

$$\left\langle \Theta^2(t) \right\rangle = \frac{\kappa^2}{2\chi} + \frac{\chi^2}{4|\alpha|^2}\left\langle \sigma_t^2 + 2|\sigma_t|^2 + \sigma_t^{*2} \right\rangle$$
$$= \frac{\kappa^2}{2\chi} + \frac{\chi}{4|\alpha|^2}. \tag{7.72}$$

This shows that Eq. (7.61) is correct. Taking the derivative of Eq. (7.61) gives

$$\frac{\partial}{\partial\chi}\left\langle \Theta^2(t) \right\rangle = \frac{1}{4|\alpha|^2} - \frac{\kappa^2}{2\chi^2}. \tag{7.73}$$

Therefore the variance is minimised by

$$\chi = \sqrt{2}\kappa|\alpha|, \tag{7.74}$$

and the minimum variance is

$$\left\langle \Theta^2(t) \right\rangle_{\min} = \frac{\kappa}{\sqrt{2}|\alpha|}. \tag{7.75}$$

The minimum phase variance is therefore $\sqrt{2}$ times the minimum phase variance for the adaptive case.

7.3 Results for Continuous Dyne Measurements

In order to verify these approximate analytic results, the equilibrium phase variance was determined numerically for a variety of parameters. Although it at first appears that there are three parameters that should be varied, $|\alpha|$, κ and χ, these parameters are not completely independent, and we need

only consider variation in two parameters. To see this, consider the equations for this system:

$$\begin{aligned} I(t)dt &= 2\left|\alpha\right|\sin\left(\varphi(t)-\hat{\varphi}(t)\right)dt + dW(t) \\ \varphi(t+dt) &= \varphi(t)+\kappa dW'(t) \\ A_t &= \int_{-\infty}^{t} e^{\chi(u-t)}e^{i\Phi}I(u)du \\ B_t &= -\int_{-\infty}^{t} e^{\chi(u-t)}e^{2i\Phi}du \\ \Theta(t) &= \arg(A_t+\chi B_t A_t^*). \end{aligned} \tag{7.76}$$

Consider a change in the time variable

$$t' = t/\lambda^2. \tag{7.77}$$

For this change in the time variable, the variables $I(t)$, A_t and B_t should be scaled to

$$\begin{aligned} I'(t') &= \lambda I(t') \\ A'_{t'} &= A_{t'}/\lambda \\ B'_{t'} &= B_{t'}/\lambda^2. \end{aligned} \tag{7.78}$$

This can be done because these are merely intermediate variables. With these substitutions the equations become

$$\begin{aligned} I'(t')dt' &= 2\lambda\left|\alpha\right|\sin\left(\varphi(t')-\hat{\varphi}(t')\right)dt' + dW(t') \\ \varphi(t'+dt') &= \varphi(t')+\lambda\kappa dW'(t') \\ A'_{t'} &= \int_{-\infty}^{t'} e^{\lambda^2\chi(u-t')}e^{i\Phi}I'(u)du \\ B'_{t'} &= -\int_{-\infty}^{t'} e^{\lambda^2\chi(u-t')}e^{2i\Phi}du \\ \Theta(t') &= \arg(A'_{t'}+\lambda^2\chi B'_{t'}A'^{*}_{t'}). \end{aligned} \tag{7.79}$$

This time scaling is therefore equivalent to making the substitutions

$$\begin{aligned} \left|\alpha\right| &\rightarrow \lambda\left|\alpha\right| \\ \kappa &\rightarrow \lambda\kappa \\ \chi &\rightarrow \lambda^2\chi. \end{aligned} \tag{7.80}$$

If $\lambda = 1/\left|\alpha\right|$, then $\left|\alpha\right|$ is replaced in the equations with 1, κ is replaced with $\kappa/\left|\alpha\right|$, and χ is replaced with $\chi/\left|\alpha\right|^2$. This means that the results depend only on the ratios $\kappa/\left|\alpha\right|$ and $\chi/\left|\alpha\right|^2$. If the individual values of κ, $\left|\alpha\right|$ and χ are varied while keeping these ratios the same, then the results will not change. It is therefore convenient to define the variables

$$\mathrm{K} = \frac{\kappa}{\left|\alpha\right|} \tag{7.81}$$

$$\mathrm{X} = \frac{\chi}{\left|\alpha\right|^2}. \tag{7.82}$$

Note that these variables are equal to the values of κ and χ if $\left|\alpha\right|$ is equal to 1. In terms of these variables, the minimum phase variance for adaptive measurements is K/2, and the optimum value of X is 2K.

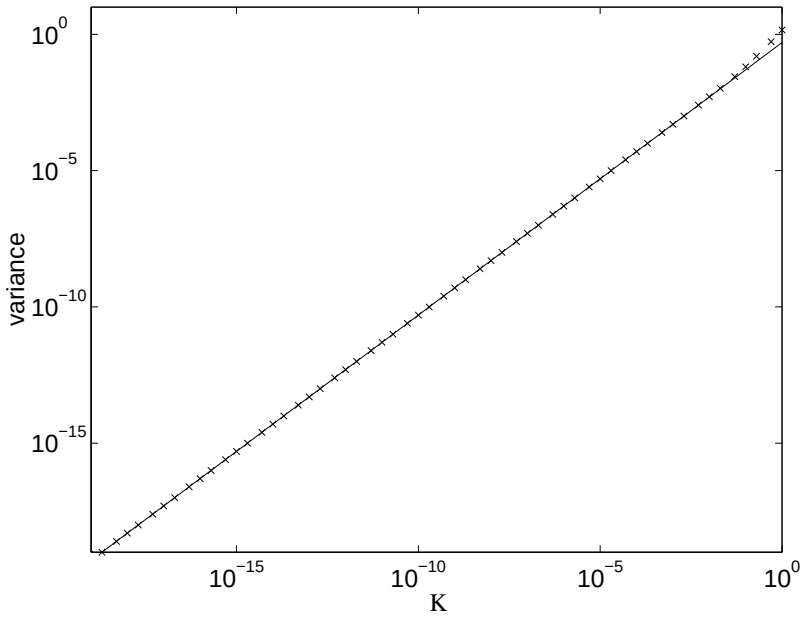

Figure 7.1: The phase variance for continuous adaptive measurements for X = 2K. The numerical results are shown as crosses and the approximate analytic relation K/2 is shown as the continuous line.

The value of K was varied from 1 down to 2×10^{-19}. For each value of K, X was varied from a quarter to four times its optimum value of 2K. The time steps used were

$$\Delta t = \frac{1}{10^3 \mathrm{X}}. \tag{7.83}$$

For these calculations 1024 simultaneous integrations were performed and the variance was sampled repeatedly. The integrations were taken up to time 10/X, in order for the variance to reach its equilibrium value, then the variance was sampled at time intervals of 1/X up until time 100/X.

The results for X = 2K are plotted in Fig. 7.1. The variances for K = 1 to 5×10^{-7} are the Holevo variances, and for below 5×10^{-7} are the standard variances. As can be seen, the results are very close to the analytic expression. To show the improvement over heterodyne measurements, the ratio of the minimum phase variance for adaptive measurements to the minimum phase variance for heterodyne measurements (with $\mathrm{X} = \sqrt{2}\mathrm{K}$) is plotted in Fig. 7.2. The ratio is close to 1 for large K, but for smaller K the ratio gets closer and closer to $1/\sqrt{2}$.

In order to see the differences from the analytic expression more clearly, the ratios of the variances to the analytic values for adaptive and heterodyne measurements are plotted in Fig. 7.3. As can be seen the phase variances do differ significantly from the analytic values for large K, but the agreement is extremely good for small K. This can be expected, because the approximations made are for the limit of large $|\alpha|$, which is equivalent to small K. Also the agreement at large K is slightly better for heterodyne measurements than for adaptive measurements.

Alternatively we can plot the phase variance as a function of X for fixed K. In Fig. 7.4 I have shown the phase variance as a function of X for K = 0.001 for adaptive and heterodyne measurements. The numerical results agree reasonably closely with the analytic values, although there is a noticeable difference for adaptive measurements for the larger values of X. Note that the minimum phase variance for adaptive measurements is at X = 2K, and the minimum phase variance for heterodyne measurements is larger and at a smaller value of X. When the value of K is reduced further, as in Fig. 7.5, the numerical results agree even more closely with the analytic values.

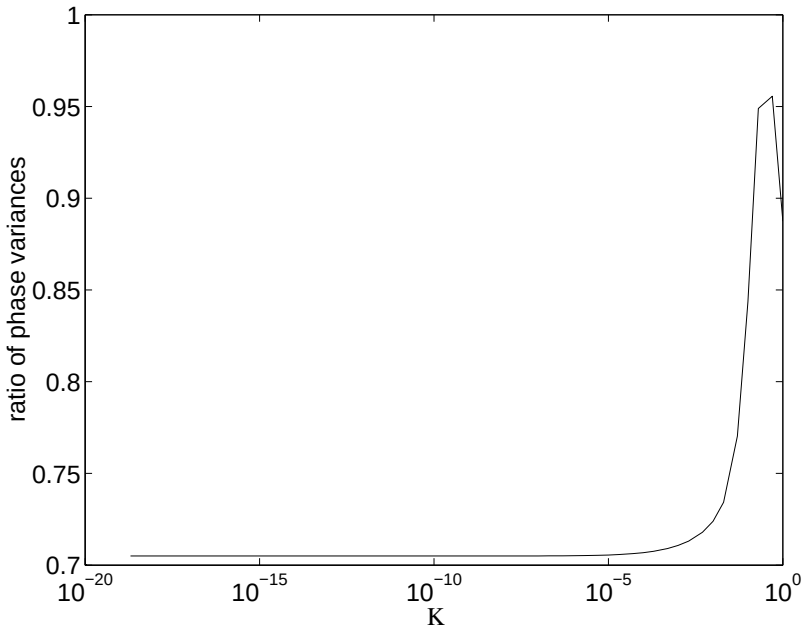

Figure 7.2: The ratio of the minimum phase variance for continuous adaptive measurements to the minimum phase variance for continuous heterodyne phase measurements.

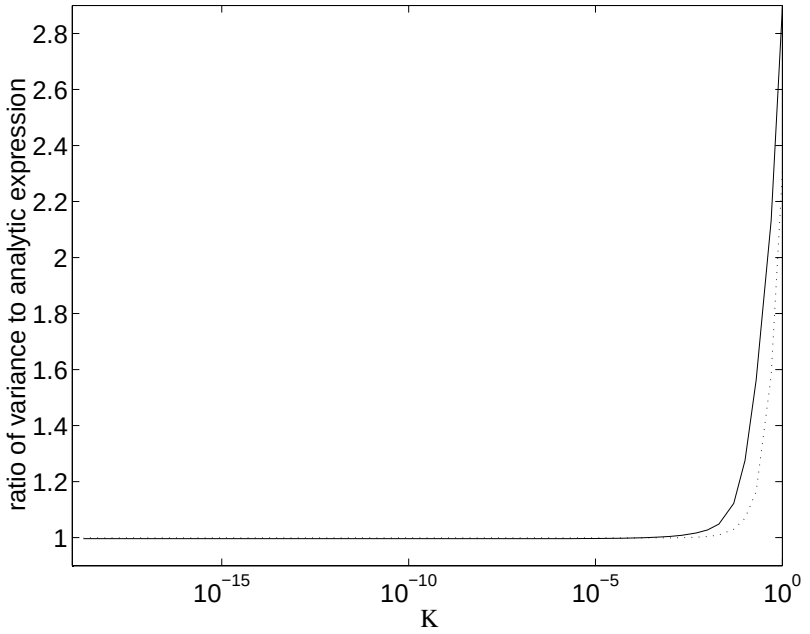

Figure 7.3: The ratio between the numerically obtained phase variance and the analytic expression K/2 for adaptive measurements (continuous line), and K/$\sqrt{2}$ for heterodyne measurements (dotted line).

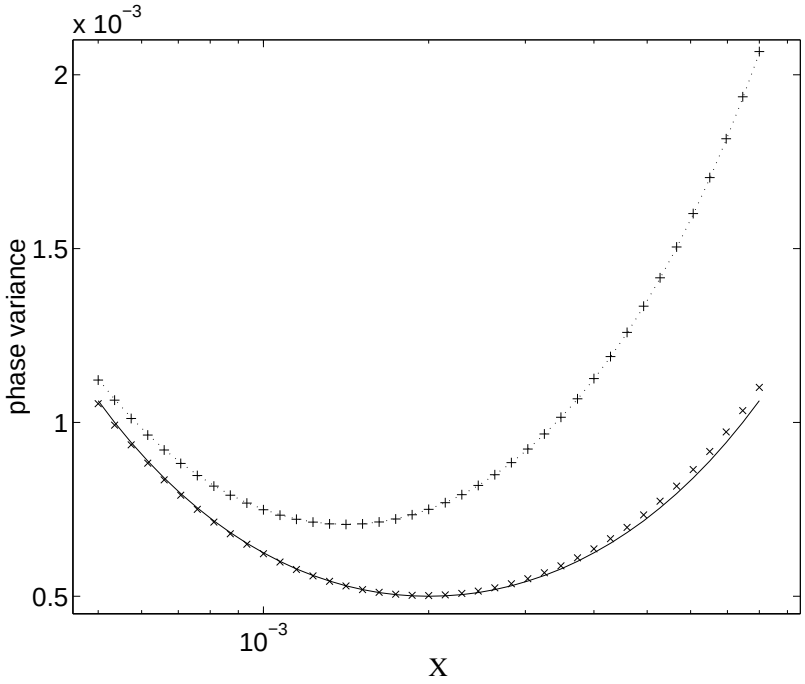

Figure 7.4: The phase variance as a function of X for K = 0.001. The numerical results for adaptive and heterodyne measurements are shown as the crosses and pluses respectively and the approximate analytic results for adaptive and heterodyne measurements are shown as the continuous line and dotted line respectively.

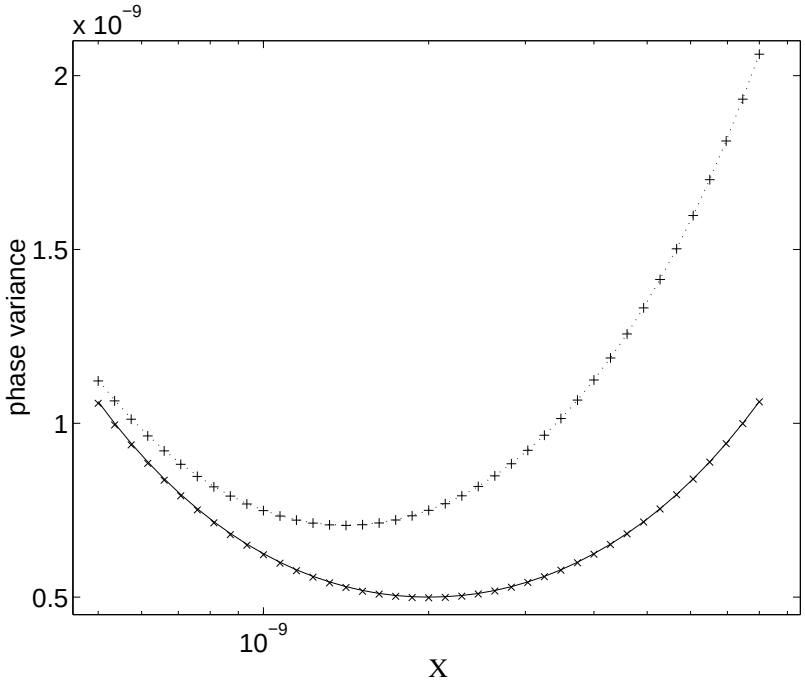

Figure 7.5: The phase variance as a function of X for K = 10^{-9}. The numerical results for adaptive and heterodyne measurements are shown as the crosses and pluses respectively and the approximate analytic results for adaptive and heterodyne measurements are shown as the continuous line and dotted line respectively.

7.4 Continuous Squeezed State Measurements

It is also possible to consider dyne measurements on continuous squeezed states. At first it appears that it does not make sense to consider squeezed states in the continuous case. This is because, in the single-shot case, the reduced variance is due to the back-action of the measurement on the state. The photocurrent is given by

$$I(t)dt = 2\mathrm{Re}(\langle a\rangle\, e^{-i\Phi(t)})dt + dW(t). \tag{7.84}$$

Here there is no factor multiplying dW, so the variances in the individual values of $I(t)$ are not reduced for a squeezed state. Instead, the reduced phase variance comes from the variation of $\langle a\rangle$ during the measurement. If we are considering continuous measurements, then the state should remain constant during the measurement. This means that $\langle a\rangle$ is constant, and the phase variance cannot be reduced by this factor.

For the state to remain constant, we must consider a squeezed state produced by a driven parametric oscillator in the limit that the decay time of the cavity is extremely short. This limit must be taken before the limit $\delta t \to 0$ in the definition for the photocurrent. In this limit, the changes in the state over the time scale of the decay constant result in reduced noise, but these changes do not persist over the time interval δt, which is much larger than the decay time.

The net result of this is that when we take the limit $\delta t \to 0$ the photocurrent is given by

$$I(t)dt = 2\mathrm{Re}(\alpha e^{-i\Phi(t)})dt + \sqrt{e^{-2r}\cos^2(\Phi-\phi_\zeta/2) + e^{2r}\sin^2(\Phi-\phi_\zeta/2)}dW(t), \tag{7.85}$$

where α is the amplitude of the squeezed state, and r and ϕ_ζ are the magnitude and direction of the squeezing. Therefore in this case, rather than the uncertainty being reduced by variation in $\langle a\rangle$, it is reduced by a multiplying factor for dW.

For reduced phase uncertainty, the phase of the squeezing should be $\phi_\zeta = 2\varphi + \pi$, where φ is the system phase. If we are using feedback given by

$$\Phi = \hat{\varphi} + \pi/2, \tag{7.86}$$

where $\hat{\varphi}$ is an estimate of the phase, then the photocurrent can be expressed as

$$I(t)dt = 2|\alpha|\sin(\varphi-\hat{\varphi})dt + \sqrt{e^{-2r}\cos^2(\hat{\varphi}-\varphi) + e^{2r}\sin^2(\hat{\varphi}-\varphi)}dW(t). \tag{7.87}$$

It is clear that if the intermediate phase estimate used is very close to the system phase, then the factor multiplying dW will be close to e^{-r} and will be at a minimum. The better the intermediate phase estimate is, the smaller this multiplying factor will be. If the intermediate phase estimate is not perfect, it is clear that increasing the squeezing past a certain level will not reduce the multiplying factor. This is because the e^{2r} term will start to dominate.

It is possible to estimate the optimum squeezing and the minimum phase variance under these measurements using the linear approximation. In this approximation, the variance in the individual phase estimates $\theta(t)$ is equal to

$$\frac{e^{-2r}\cos^2(\hat{\varphi}-\varphi) + e^{2r}\sin^2(\hat{\varphi}-\varphi)}{4\,|\alpha|^2\,dt}. \tag{7.88}$$

It is clear that the minimum phase variance (in this approximation) will be obtained when the best phase estimates are used for $\hat{\varphi}$. It is therefore reasonable to use the phase estimates $\Theta(t)$ for $\hat{\varphi}$. These will be the best phase estimates when the correct value of χ is used. As the variance of these estimates is $\Delta\Theta^2$, we obtain

$$\left\langle e^{-2r}\cos^2(\hat{\varphi}-\varphi) + e^{2r}\sin^2(\hat{\varphi}-\varphi)\right\rangle \approx e^{-2r} + e^{2r}\Delta\Theta^2. \tag{7.89}$$

This approximation will be true for small phase variances and large squeezing. Following the same derivation as for the coherent state case, the only difference is the multiplying factor, so we obtain

$$\Delta\Theta^2 = \frac{\chi}{8\,|\alpha|^2}\left(e^{-2r} + e^{2r}\Delta\Theta^2\right) + \frac{\kappa^2}{2\chi}. \tag{7.90}$$

Solving this for $\Delta\Theta^2$ gives

$$\Delta\Theta^2 = \frac{\frac{\chi e^{-2r}}{8|\alpha|^2} + \frac{\kappa^2}{2\chi}}{1 - \frac{\chi e^{2r}}{8|\alpha|^2}}. \tag{7.91}$$

This expression has two independent variables, χ and r, that can be varied in order to find the minimum phase variance. Taking the derivative of Eq. (7.90) with respect to χ gives

$$\frac{\partial\Delta\Theta^2}{\partial\chi} = \frac{1}{8|\alpha|^2}\left(e^{-2r} + e^{2r}\Delta\Theta^2\right) + \frac{\chi e^{2r}}{8|\alpha|^2}\frac{\partial\Delta\Theta^2}{\partial\chi} - \frac{\kappa^2}{2\chi^2}. \tag{7.92}$$

Since the derivative is zero for the minimum this gives

$$0 = \frac{1}{8|\alpha|^2}\left(e^{-2r} + e^{2r}\Delta\Theta^2\right) - \frac{\kappa^2}{2\chi^2}. \tag{7.93}$$

Together with Eq. (7.90), this gives

$$\chi = \frac{\kappa^2}{\Delta\Theta^2}. \tag{7.94}$$

Substituting this into Eq. (7.90) gives

$$\Delta\Theta^2 = \frac{\kappa^2}{4|\alpha|^2}\left(e^{2r} + \frac{e^{-2r}}{\Delta\Theta^2}\right). \tag{7.95}$$

Taking the derivative of this with respect to r gives

$$\frac{\partial\Delta\Theta^2}{\partial r} = \frac{\kappa^2}{2|\alpha|^2}\left(e^{2r} - \frac{e^{-2r}}{\Delta\Theta^2} - \frac{e^{-2r}}{(\Delta\Theta^2)^2}\frac{\partial\Delta\Theta^2}{\partial r}\right). \tag{7.96}$$

Since the derivative is zero at the minimum, this becomes

$$0 = \frac{\kappa^2}{2|\alpha|^2}\left(e^{2r} - \frac{e^{-2r}}{\Delta\Theta^2}\right). \tag{7.97}$$

This can be solved to give

$$e^{-4r} = \Delta\Theta^2. \tag{7.98}$$

Substituting this back into Eq. (7.95) gives the phase variance as

$$\Delta\Theta^2 = \left(\frac{\kappa}{\sqrt{2}|\alpha|}\right)^{4/3}. \tag{7.99}$$

Thus we see that even for an arbitrarily squeezed state, the best scaling we can obtain for the phase variance is $|\alpha|^{-4/3}$, as compared to $|\alpha|^{-1}$ for a coherent state. This difference is less than for single-shot measurements, where optimum squeezed states scale as almost $\overline{n}^{-2}$, as compared to $\overline{n}^{-1}$ for coherent states.

7.5 Results for Squeezed States

The results for the continuous squeezed state case were obtained by a similar method as for the coherent state case. Similarly to the case for coherent states, only variation in the variables K and X was considered, rather than in all three variables $|\alpha|$, κ and χ. The step sizes used were

$$\Delta t = \frac{1}{10^3\mathrm{X}}. \tag{7.100}$$

The integrations were taken up to time 30/X, then the variance was sampled every time step until time 130/X. The integration was performed using the photocurrent given in Eq. (7.85), except with

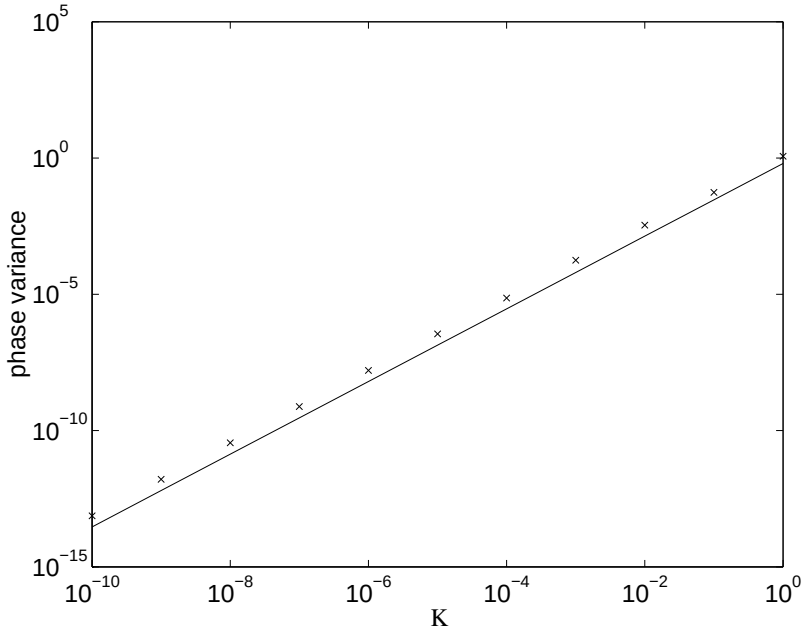

Figure 7.6: The phase variance as a function of K for continuous squeezed states. The continuous line is the theoretical analytic relation, and the crosses are the numerical results.

the time scaled such that $|\alpha|$ was not required (similarly to the coherent state case). The phase of the squeezing was assumed at all times to be in the correct direction for reduced phase variance, i.e. negative with respect to the coherent amplitude.

It was found that when the phase estimate $\arg C_t$ was used in the feedback, very poor results were obtained. This is a similar result to the case for single-shot measurements, where using $\arg C_v$ feedback results in large phase variances. This is because, when the intermediate phase estimates are extremely good, the results do not distinguish easily between the real system phase and the system phase plus π. The means that many of the results are out by π, resulting in a large overall phase variance.

In order to avoid this problem, rather than using $\arg C_t$ in the feedback, an intermediate phase estimate given by

$$\hat{\varphi}(t) = \arg(C_t^{1-\varepsilon} A_t^{\varepsilon}), \tag{7.101}$$

was used. Note that this is similar to the phase estimate used to obtain phase measurements close to optimum in the single-shot case. Here a constant ε was used, as the system state is not changing like in the single-shot case.

For each value of K there are three variables that can be altered to minimise the phase variance: X, r and ε. It is not calculationally feasible to consider a range of values for all three variables. Rather than considering a range of values, the values of the variables were varied in order to find the values that gave the minimum phase variance.

The minimum phase variances obtained by this method are plotted as a function of K in Fig. 7.6. The values given by the approximate analytic expression (7.99) are also shown in this figure. The numerical results are higher than the analytic expression, but for small K they have the same scaling. If we plot the ratio of the numerical results to the analytic values as in Fig. 7.7, we find that for the smallest values of K the ratio levels off at about 2.6. Thus we see that the stochastic results also give a scaling of $|\alpha|^{-4/3}$.

Now note that, from Eqs (7.98) and (7.99), the optimum value of e^{-2r} should be

$$e^{-2r} = \left(\frac{\kappa}{\sqrt{2}\,|\alpha|}\right)^{2/3}$$

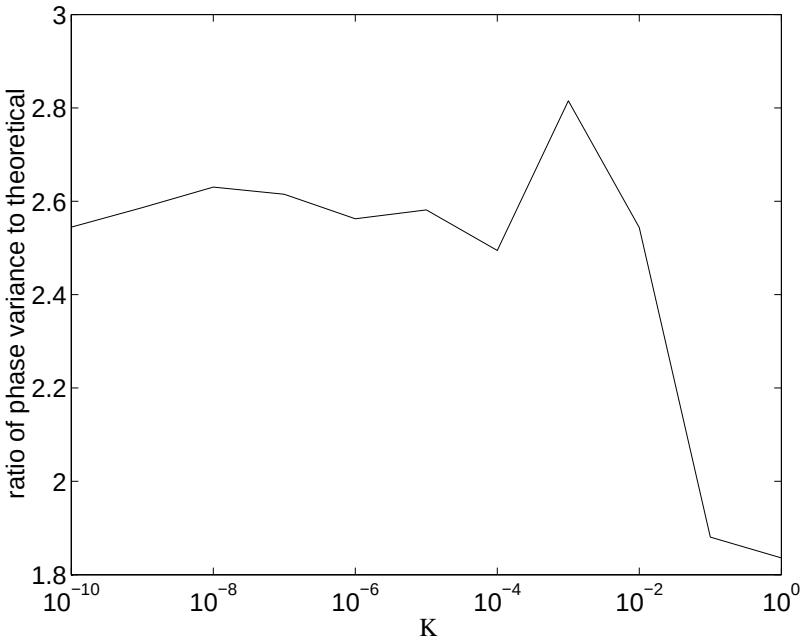

Figure 7.7: The ratio of the numerically obtained phase variance to the analytic relation as a function of K for continuous squeezed states.

$$= \left(\frac{\mathrm{K}}{\sqrt{2}}\right)^{2/3}. \tag{7.102}$$

Similarly, from Eqs (7.94) and (7.99), the optimum value of χ should be

$$\chi = \left(2\left|\alpha\right|^2\kappa\right)^{2/3}. \tag{7.103}$$

The corresponding optimum value of X is

$$\mathrm{X} = (2\mathrm{K})^{2/3}. \tag{7.104}$$

This indicates that the optimum values for both e^{-2r} and X should scale as $\mathrm{K}^{2/3}$.

The numerically obtained optimum values of e^{-2r} and X, as well as these analytic expressions, are plotted in Fig. 7.8. Similarly to the case for the phase variance, the scaling is the same as that predicted analytically, but the scaling constants are different. For the case of e^{-2r}, the optimum values are about 8 times those analytically predicted, whereas the values of X are around a third of those analytically predicted.

For the case of ε there is no analytic prediction for the optimum value. The numerically obtained values are shown in Fig. 7.9, and as can be seen ε decreases in a regular way with κ. A power law was fitted to these values (for $\mathrm{K} < 1$), and the power found was 0.70 ± 0.01. This is very similar to the $\mathrm{K}^{2/3}$ scaling found for e^{-2r} and X.

A problem with these results is that they do not take account of the low probability results with large error. Unfortunately it is not possible to take account of these results in an analytic way as in the case of single shot measurements on squeezed states. The best that can be done is to use a very large number of samples. For these calculations the phase variance was sampled every time step, resulting in about 10^8 samples. Unfortunately the phase estimates from adjacent time steps are strongly correlated, so there is only around 10^5 independent samples.

Despite the large number of samples used, there were generally no large error samples for the results shown. The optimum parameters found usually gave the minimum variance because varying the parameters beyond these values, such that the variance would be smaller according to the linear theory, resulted in samples with large error.

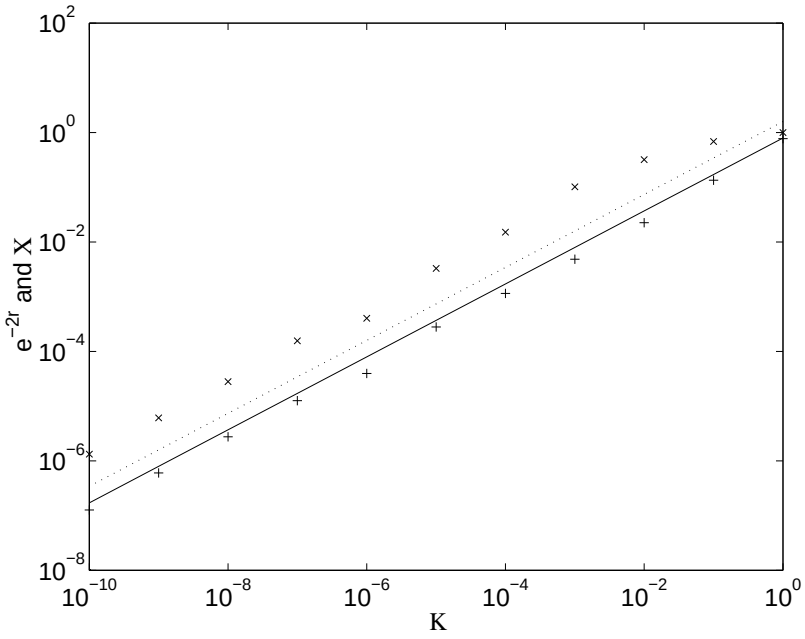

Figure 7.8: The optimum values of e^{-2r} and X for measurements on continuous squeezed states. The numerically found values of e^{-2r} are plotted as crosses, and the approximate analytic expression as a continuous line. The numerically found values of X are plotted as pluses, and the approximate analytic expression as a dotted line.

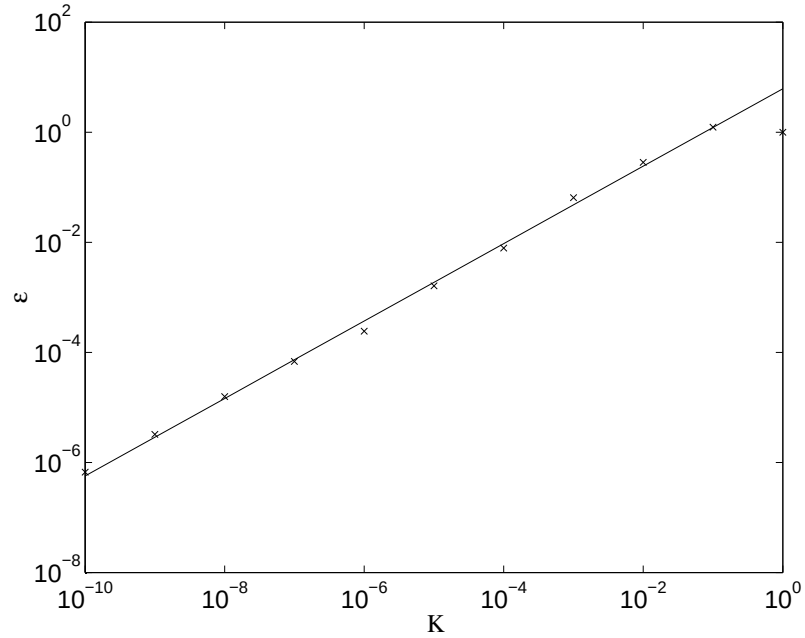

Figure 7.9: The optimum values of ε for measurements on continuous squeezed states. The crosses are the numerically found values, and the continuous line is the expression fitted to the data.

7.6 Continuous Interferometric Measurements

Now I will consider the case of continuous measurements for two-mode interferometry. In this case we have a Mach-Zehnder interferometer as in Ch. 5, and are attempting to continuously track the phase in one arm, and control the phase in the other arm in order to obtain the best possible estimate of the phase. The optimum states discussed in Ch. 5 make little sense in this context, so instead I will consider the state with all the photons in one arm. This state can be generalised to the continuous case by considering a state with N photons per unit time.

This case is essentially semiclassical, and the detections can be considered independently. Therefore, consider a single photon incident on port a, so the state is $|\frac{1}{2},\frac{1}{2}\rangle_z$. Upon detection the unnormalised state changes as

$$|\psi(u,\varphi)\rangle = \hat{c}_u(\varphi)|\psi\rangle. \tag{7.105}$$

Using Eq. (6.5) for the detection operator, the probability of detecting the photon in detector u is given by

$$\langle\psi(u,\varphi)|\psi(u,\varphi)\rangle = \sin^2\left(\frac{\varphi-\Phi+u\pi}{2}\right). \tag{7.106}$$

Using Bayes' theorem, the probability distribution for the system phase after the detection is proportional to this probability times the initial phase distribution.

The probability distribution for the phase based on m of these independent detections, $P(\varphi|n_m)$, can be expressed as

$$P(\varphi|n_m) = \sum_{k=-m}^{m} P_{mk}(n_m)e^{ik\varphi}. \tag{7.107}$$

It is straightforward to show from Eq. (7.106) that the coefficients $P_{mk}(n_m)$ can be determined by

$$P_{mk}(n_m) \propto P_{m-1,k}(n_{m-1}) - \tfrac{1}{2}e^{-i(\Phi_m-u_m\pi)}P_{m-1,k-1}(n_{m-1}) - \tfrac{1}{2}e^{i(\Phi_m-u_m\pi)}P_{m-1,k+1}(n_{m-1}). \tag{7.108}$$

The normalisation condition on the probability distribution becomes

$$P_{m0}(n_m) = 1. \tag{7.109}$$

The normalised probability distribution can be obtained by a simple addition to the recursion relation:

$$P_{mk}(n_m) = \frac{P_{m-1,k}(n_{m-1}) - \tfrac{1}{2}e^{-i(\Phi_m-u_m\pi)}P_{m-1,k-1}(n_{m-1}) - \tfrac{1}{2}e^{i(\Phi_m-u_m\pi)}P_{m-1,k+1}(n_{m-1})}{P_{m0}(n_m)}. \tag{7.110}$$

I will consider the same variation in the system phase as in the case of dyne measurements,

$$\varphi(t+dt) = \varphi(t) + \kappa dW(t). \tag{7.111}$$

When the phase varies in time, the time between detections is important. For a photon flux of N, the probability of a photodetection in time dt is Ndt. The probability distribution for the time between detections is given by

$$P_{\rm P}(t)dt = Ne^{-Nt}dt. \tag{7.112}$$

In the results that will be presented here, the time between detections, Δt, was determined according to this probability distribution.

Now in order to determine the effect of this phase diffusion on the probability distribution between detections, we must first consider the effect over some very small time interval δt. This is necessary because the probability distribution for the change in the system phase over time Δt does not go to zero for $\Delta\varphi = \pm\pi$. This means that the probability distribution will not be exactly Gaussian, due to the overlap. In contrast, if we look at a very small time interval δt, the change in the phase will have a normal distribution with a variance of $\kappa^2\delta t$. Explicitly the probability distribution is

$$P_{\rm G}(\Delta\varphi)d(\Delta\varphi) = \frac{1}{\kappa\sqrt{2\pi\delta t}}e^{-\Delta\varphi^2/(2\kappa^2\delta t)}d(\Delta\varphi). \tag{7.113}$$

The probability distribution for the phase after time δt will be the convolution of the initial probability distribution with the Gaussian described by Eq. (7.113). Evaluating this convolution gives

$$\begin{aligned} P'(\varphi|n_m) &= \int_{-\pi}^{\pi} P(\varphi-\theta|n_m) P_{\mathrm{G}}(\theta) d\theta \\ &= \int_{-\pi}^{\pi} \sum_{k=-m/2}^{m/2} P_{mk}(n_m) e^{ik(\varphi-\theta)} P_{\mathrm{G}}(\theta) d\theta \\ &= \sum_{k=-m/2}^{m/2} P_{mk}(n_m) e^{ik\varphi} \int_{-\pi}^{\pi} e^{-ik\theta} P_{\mathrm{G}}(\theta) d\theta. \end{aligned} \tag{7.114}$$

As δt is assumed to be small, $\kappa^2 \delta t \ll 1$, so

$$\int_{-\pi}^{\pi} e^{-ik\theta} P_{\mathrm{G}}(\theta) d\theta = e^{-k^2\kappa^2\delta t/2}. \tag{7.115}$$

The effect of the variation of the system phase on the probability distribution is therefore

$$P^{\delta t}(\varphi|n_m) = \sum_{k=-m/2}^{m/2} P_{mk}(n_m) e^{-k^2\kappa^2\delta t/2} e^{ik\varphi}. \tag{7.116}$$

This shows that the coefficients for the probability distribution are just multiplied by a Gaussian:

$$P^{\delta t}_{mk}(n_m) = P_{mk}(n_m) e^{-k^2\kappa^2\delta t/2}. \tag{7.117}$$

This result is related to the usual result for convolutions and Fourier transforms.

In order to take account of the effect of the phase diffusion on the probability distribution over some significant time interval Δt, this time interval can be divided into M small time intervals δt. Then we find

$$\begin{aligned} P^{\Delta t}_{mk}(n_m) &= P_{mk}(n_m) \prod_{j=1}^{M} e^{-k^2\kappa^2\delta t/2} \\ &= P_{mk}(n_m) e^{-k^2\kappa^2 \sum_{j=1}^{M} \delta t/2} \\ &= P_{mk}(n_m) e^{-k^2\kappa^2\Delta t/2}. \end{aligned} \tag{7.118}$$

This result can be used to take account of the variation of the system phase very simply in the probability distribution.

As time passes the effect of Eq. (7.110) is to broaden the distribution of probability coefficients in k, corresponding to a smaller variance in the phase distribution. In contrast the Gaussian term in Eq. (7.118) tends to narrow the distribution of probability coefficients, corresponding to a greater phase variance. The initially broad phase distribution narrows until an approximate equilibrium is reached, where the two effects cancel each other out.

The derivation for the optimum phase estimate for the single-shot case given in Eq. (6.19) is general enough to hold in this case also. Therefore the optimal phase estimate is given by

$$\begin{aligned} \hat{\varphi} &= \arg \left\langle e^{i\varphi} \right\rangle \\ &= \arg P_{m,-1}(n_m). \end{aligned} \tag{7.119}$$

In addition, much of the reasoning for the feedback phase still holds. The phase variance after the next detection can be minimised by minimising the value of

$$M(\Phi_m) = \sum_{u_m=0,1} \left| \int_{-\pi}^{\pi} P(n_m|\varphi) e^{i\varphi} d\varphi \right|. \tag{7.120}$$

The values of $P(n_m|\varphi)$ can be obtained, except for a normalising constant that is common to $u_m = 0$ and 1, by using Eq. (7.108). This means that we can express $M(\Phi_m)$ as in Eq. (6.33) with the parameters a, b and c given by

$$\begin{aligned} a &= P_{m-1,-1}(n_{m-1}) \\ b &= \tfrac{1}{2} P_{m-1,-2}(n_{m-1}) \\ c &= \tfrac{1}{2} P_{m-1,0}(n_{m-1}). \end{aligned} \tag{7.121}$$

These values of a, b and c can be used to determine the feedback phase as in Sec. 6.3.

The phase uncertainty at equilibrium can be estimated using a similar approach as was used for the single mode case. Let us assume that the equilibrium variance in the best estimate for the system phase is $\Delta\Theta^2$. After time Δt, the variance in this phase estimate with respect to the new system phase, $\varphi(t + \Delta t)$, will be $\Delta\Theta^2 + \kappa^2 \Delta t$. In the equilibrium case this increase in the variance should, on average, be balanced by the decrease due to the next detection.

We now wish to estimate the equilibrium variance based on a weighted average with the previous best phase estimate, and a phase estimate from the new detection. If we use the actual variance for a phase estimate based on a single detection, then we do not get accurate results. This is because the variance for a single detection is large, so the weighted average does not accurately correspond to the exact theory. In order to make the theory based on weighted averages accurate, we need to assume an *effective* variance for the single detection, that is different from the actual variance.

In the case where there is no variation in the system phase, the phase variance after N detections is approximately $1/N$ (see the results given in Secs 5.3 and 6.6). Denoting the variance after N detections as $\Delta\Theta_N^2$, and the effective phase variance from a new detection as $\Delta\Theta_1^2$, the weighted average gives

$$\begin{aligned} \frac{1}{\Delta\Theta_N^2} + \frac{1}{\Delta\Theta_1^2} &= \frac{1}{\Delta\Theta_{N+1}^2} \\ N + \frac{1}{\Delta\Theta_1^2} &= N + 1. \end{aligned} \tag{7.122}$$

Therefore, in order to take account of a single detection via a weighted average, it can be assumed to be equivalent to a phase estimate with a variance of 1. This is, in fact, equal to the variance as estimated using $\langle 2(1 - \cos\varphi) \rangle$.

Applying this to the case with a varying system phase gives

$$\frac{1}{\Delta\Theta^2 + \kappa^2 \Delta t} + 1 = \frac{1}{\Delta\Theta^2} \tag{7.123}$$

Simplifying this to solve for $\Delta\Theta^2$, we find

$$\begin{aligned} \frac{1}{1 + \kappa^2 \Delta t / \Delta\Theta^2} + \Delta\Theta^2 &= 1 \\ 1 - \kappa^2 \Delta t / \Delta\Theta^2 + \Delta\Theta^2 &\approx 1 \\ \kappa^2 \Delta t / \Delta\Theta^2 &\approx \Delta\Theta^2. \end{aligned} \tag{7.124}$$

Thus the approximate value of the variance is

$$\Delta\Theta^2 \approx \kappa \sqrt{\Delta t}. \tag{7.125}$$

On average the time between detections is $1/N$, so the approximate value of the variance should be

$$\Delta\Theta^2 \approx \kappa / \sqrt{N}. \tag{7.126}$$

7.7 Results for Continuous Interferometric Measurements

In order to verify this result, the equilibrium phase variance was determined numerically for a variety of parameters. In this case there are only two parameters, κ and N. In the case of dyne measurements there was the additional parameter χ describing how the latest results were weighted as compared to the previous results. In this case we do not have that parameter, as the phase estimates are not determined in that way.

Similarly to the case of dyne measurements, these two parameters are related. Here the average time between detections is $1/N$, and the mean squared change in the system phase after time Δt is $\kappa^2 \Delta t$. In the above theory the parameters κ and Δt only appear through the product $\kappa^2 \Delta t$. It is therefore the value of κ^2/N that matters, and if κ and N are varied while keeping this product the same, the results will not change. It is therefore convenient to define

$$\mathrm{K} = \frac{\kappa}{\sqrt{N}}, \tag{7.127}$$

similarly to the case for continuous dyne measurements.

The calculations were run for 10^5 detections (or 2×10^5 for the minimum value of K), and the phase error was sampled every detection after 10/K detections. This was done 100 times for each value of K. The equilibrium phase variance was determined in this way for the nearly optimum feedback scheme and in addition for a nonadaptive feedback scheme given by

$$\Phi_m = \Phi_0 + m\mathrm{K}\pi, \tag{7.128}$$

where Φ_0 is a random initial phase. When the value of K was 1 or more this was modified to

$$\Phi_m = \Phi_0 + m\pi/2, \tag{7.129}$$

to prevent Φ_m being constant (modulo π). This is equivalent to the non-adaptive feedback in the single-shot case given by Eq. (6.57), and is analogous to heterodyne feedback for dyne measurements. The reason for the factor of K is that the effective number of detections used for the phase estimate is 1/K. To see this, note that the phase variance is approximately K.

A minor problem with continuous adaptive measurements is that the number of probability coefficients $P_{mk}(n_m)$ needed to determine the probability distribution for the phase rises indefinitely with the number of detections. The narrowing effect of the varying system phase, however, means that the probability coefficients fall approximately exponentially with k. The probability distribution can therefore be approximated very accurately by keeping only a certain number of coefficients. For the results presented here all probability coefficients with a magnitude above about 10^{-20} were used.

All phase variances for the two feedback schemes are plotted in Fig. 7.10. Note that, from the previous section, the analytically predicted value of the phase variance is K. As can be seen, the results for both cases are very close to this analytic result for the smaller values of K. For values of K close to 1 the results for the nonadaptive scheme are noticeably above the analytically predicted values. For large values of K above 1 the variance converges to 3 for both the feedback schemes. This is what can be expected, as the system phase is randomised between detections. This means that the measurements are equivalent to phase measurements with a single photon, for which the Holevo phase variance is 3. The feedback has no effect, as there is no information on which to base it on.

To see the differences more clearly, the phase variances are plotted as ratios to the analytic results in Fig. 7.11. As can be seen, the results for both measurement schemes are very close to the analytic result for very small values of K, but differ increasingly for larger values. The adaptive scheme gives phase variances that are very close to, and slightly below, the analytically predicted values for smaller values of K. In contrast the results for nonadaptive measurements are all above the analytically predicted values (for $\mathrm{K} \leq 1$). For large values of K the variance for both schemes is below K, as the variance is converging to 3.

These results show that there is a small improvement in using an adaptive scheme over a non-adaptive scheme; however, when the system phase is varying slowly there will be very little improvement. This can be expected from the results in the single shot case, where there was very little

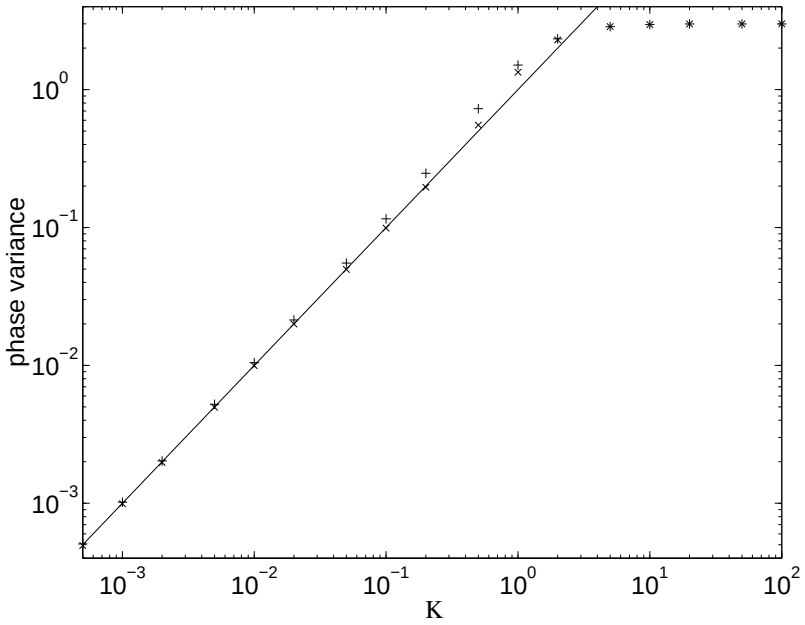

Figure 7.10: The phase variance as a function of K. The numerical results for adaptive and non-adaptive measurements are shown as the crosses and pluses respectively and the analytic result is shown as the continuous line.

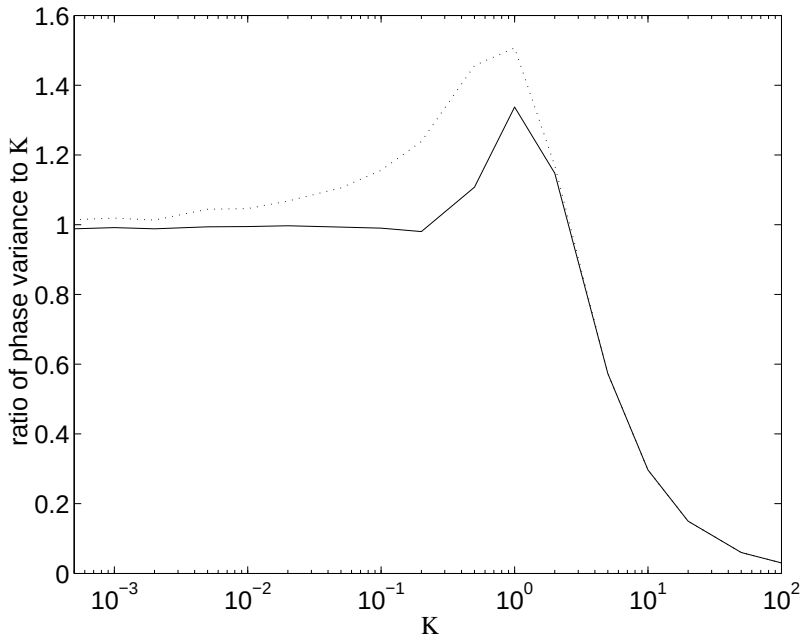

Figure 7.11: The phase variance as a ratio to the approximate analytic result of K. The results for adaptive and nonadaptive measurements are shown as the continuous line and dotted line respectively.

difference between the results for the adaptive and non-adaptive measurements when all photons were in one port. This means that we do not even get an improvement by a factor of $\sqrt{2}$, as we do in the case of continuous dyne measurements on a coherent state. If the system phase is fluctuating rapidly, we can get a significant improvement in the accuracy of the phase measurements. The maximum improvement is about 24% for $\mathrm{K} = 0.5$.

Chapter 8

Conclusions

The major results of this study can be described very succinctly: I have found the optimal input states and adaptive measurement schemes for both dyne and interferometric phase measurements.

8.1 Input States for Dyne Measurements

In Ch. 2 the problem of optimum input states for dyne measurements was considered. Rather than just considering optimisation for minimum intrinsic phase variance, the problem of optimisation for minimum phase variance under more general measurements where $h(n) = cn^{-p}$ was considered. There are three different types of constraints on the states that were considered: an upper limit on the photon number, a fixed mean photon number, and squeezed states with fixed mean photon number. The optimum states for canonical measurements with these three constraints on the states have been determined previous to this study, as has the result for general measurements with an upper limit on the photon number. In Ch. 2 the set of results was completed with analytic results for fixed mean photon number and optimum squeezed states under general measurements. The complete set of results are summarised in Table 8.1.

All of these results (except the exact analytic result for canonical measurements with an upper limit on the photon number) were verified by exact numerical calculation of the optimum states. For the case of canonical measurements on squeezed states it was found that the exact results did not agree with the approximate analytic result to order $\overline{n}^{-2}$. In addition, corrections to this analytic expression derived from the results in [45] did not produce any better agreement.

For the states with an upper limit on the photon number optimised for minimum phase variance under general measurements, it was found that although there was good agreement with the analytic results, larger photon numbers were required for good agreement than was claimed in [35]. This is because in [35] the photon numbers required for good agreement were estimated from the approximation made in omitting a boundary condition. The main approximation made is actually the linearisation of the equations, which means that the photon numbers were underestimated in [35].

An additional correction term of $3c^2N^{-2p}$ was found for this case. This term is specific to the

Table 8.1: The asymptotic formulae for minimum Holevo phase variance for the canonical distribution and general measurements under three different constraints on the states. The results in bold are those that are original to this study.

	Canonical	General measurements
Maximum photon number	$\frac{\pi^2}{N^2}$	$2cN^{-p} + \lvert z_1\rvert\,(2cp)^{2/3}N^{-2(1+p)/3}$+$\mathbf{3c^2N^{-2p}}$
Mean photon number	$\frac{1.893606}{\overline{n}^2}$	$\mathbf{2c\bar{n}^{-p}} + \mathbf{\sqrt{cp(p+1)}\bar{n}^{-p/2-1}} + \mathbf{3c^2\overline{n}^{-2p}}$
Squeezed states	$\frac{\log\bar{n}+\Delta}{4\bar{n}^2}$	$\mathbf{2c\bar{n}^{-p}} + \mathbf{\sqrt{cp(p+1)}\bar{n}^{-p/2-1}} + \mathbf{3c^2\overline{n}^{-2p}}$

Holevo phase variance, and is only significant for mark I measurements. For mark I measurements this term is required in order to obtain agreement to second order with the numerical results. Similar terms were found for the other two constraints upon the state.

For the case of general measurements on states with fixed mean photon number, it was found that the results for both the general case and the squeezed state case agreed well with the analytic result. In addition the results for these two cases agreed very closely with each other; more closely than with the analytic approximation. In fact, when the states were compared, it was found that they were extremely close, although there were small systematic differences, particularly in the tails of the distribution.

These results mean that it should be possible to experimentally produce very good approximations of the states that are optimised for minimum phase variance under the various measurement schemes by using squeezed states. This is an advantage in theoretical work also, as squeezed states are more easily treated numerically. Only the two squeezing parameters need be kept track of, rather than the entire state, resulting in much faster calculation times. The exception to this is canonical measurements, for which the results for general and squeezed states have different scalings. The ratio between the two results only scales as $\log \overline{n}$, however, so the difference becomes significant only for very large photon numbers.

8.2 Optimum Dyne Measurements

The next issue considered was the problem of how to make optimum measurements of these states. Prior to this study it was known that the minimum phase variance introduced by an arbitrary dyne measurement scheme was that of an optimised squeezed state, with variance scaling as $\log \overline{n}/\overline{n}^2$. The best adaptive dyne measurement scheme previously considered, the mark II scheme, only gave scaling of $\overline{n}^{-1.5}$. In this study I have found that it is possible to obtain results that are very close to the theoretical limit using a feedback phase estimate of

$$\hat{\varphi}_v = \arg C_v^{1-\varepsilon(v)} A_v^{\varepsilon(v)}, \tag{8.1}$$

with

$$\varepsilon(v) = \frac{v^2 - |B_v|^2}{|C_v|} \sqrt{\frac{v}{1-v}}. \tag{8.2}$$

If this is used without corrections, the variance as estimated from the variance of the samples is very close to the theoretical limit. Unfortunately this method of estimating the variance does not take account of very low probability results with large phase error. This contribution can be estimated from the values of A and B at the end of the measurement, and it is quite significant for this uncorrected phase feedback scheme.

It is possible to correct the feedback scheme in order to minimise this problem. It was found that very good results were obtained by using a correction that brings the value of B_v directly towards the estimated optimum value. This is used only at the very end of the measurement, and then only if $|B_v|$ is too far above optimum. Using this correction, the introduced phase variance as estimated from the values of A and B is only about 4% above the optimum value, even for the largest photon number considered.

It was found that the phase variance of the samples when this corrected feedback was used was slightly worse than that estimated from the values of A and B. This appears to be because of the large scatter in the photon number of the squeezed states in the POM. The data points with smaller photon number give a disproportionately large contribution to the phase variance, as the phase variance for optimum squeezed states rises very rapidly as the photon number is reduced. This is unavoidable, as any state with a small phase variance will have a large uncertainty in the photon number. This is the case both for the input state and the squeezed state in the POM. It is therefore reasonable to claim that the corrected phase measurement scheme is about 4% above what is theoretically possible.

Under some circumstances, it is actually possible to surpass the theoretical limit based on the POM. This is possible if the input state is known accurately. For coherent states, it is possible to

reduce the introduced phase variance to about a quarter of the theoretical limit. Unfortunately, for states with reduced phase uncertainty, this level of improvement does not appear to be possible. The improvement found for squeezed states was only around 10%.

8.3 Time Delays

All the adaptive feedback schemes considered are adversely affected by time delays. This was considered previous to this study in Ref. [35], which predicted the excess phase variance for mark I and II measurements in a very approximate way. In this study the excess phase variance due to time delays was determined in a far more rigorous way, and verified by numerical calculations.

In Ref. [35] it was predicted that the excess phase variance for mark I measurements was $\tau/2$. In this study it was found that this is the correct result, provided the phase estimate at the end of the measurement is the final value of the running phase estimate. This was found by repeating the derivation of Ref. [35] in a more rigorous way, and the same result was obtained. Fitting to the numerical results gave scaling constants close to the theoretical value of $\frac{1}{2}$, particularly for the larger photon numbers. There was some discrepancy for the smallest photon number considered, which can be expected as the approximation is in the limit of large α.

If, rather than using the final value of the intermediate phase estimate as the final phase estimate, $\arg A$ is used, the phase variance actually decreases with the time delay. This is because the intermediate phase estimates have a larger variance, resulting in the mark I measurements being closer to heterodyne measurements.

It was found that the scaling of the excess phase variance for mark II measurements predicted in [35] was correct, though the scaling constant was incorrect. When the derivation of Ref. [35] was repeated more rigorously, it was found that completely different terms were obtained, casting some doubt on the method used in Ref. [35]. The terms found here are unusable, as they depend explicitly on the initial conditions of the integration.

In order to avoid this problem, an alternative derivation was considered, that gives a lower limit to the total introduced phase variance with a given time delay, rather than the approximate excess phase variance due to the time delay. This result was approximately $\tau/(8\overline{n})$, as opposed to the introduced variance of $\tau/(2\overline{n})$ found in Ref. [35]. This result can be expected to give the introduced phase variance accurately if both the photon number and time delay are relatively large. In contrast, the result found in Ref. [35] was for the limit of small $\alpha\tau$.

Numerically it was found that the introduced phase variance converges to this theoretical limit in the three different cases with unsimplified feedback considered. For the case of simplified feedback, however, the phase variance was far higher. For larger time delays, rather than converging to the theoretical lower limit, the phase variance converged to the variance for heterodyne measurements. It was found to be possible to correct the feedback to improve on these results, but the variance was still much higher than for unsimplified feedback, and did not converge to the theoretical limit.

This means that, if there is any significant time delay in the system, the simplified analog feedback will give far worse results than the exact, unsimplified feedback. This makes the more sophisticated feedback schemes considered in Ch. 3 more attractive, as one of the main advantages of the mark II measurement scheme was that it allowed the simplified feedback to be used.

8.4 Input States for Interferometry

In Ch. 5 optimum input states for interferometry were considered. When the total photon number is fixed, this problem is equivalent to the dyne case with an upper limit on the photon number. This problem has a simple solution, and there is a correspondingly simple solution in the case of interferometry. This state has a canonical phase variance scaling as N^{-2}, as compared to N^{-1} when all the photons are incident on one port.

This state requires contributions from a large number of input photon eigenstates; however, it requires significant contributions from only about 10. This means that although the complete state will be very difficult to produce, it should be less difficult to produce close approximations of it.

Unfortunately, as the photon number is increased, more and more of these eigenstates are required to give a variance that is close to that for the exact state.

An alternative state that has been considered in previous work, is one where equal photon numbers are incident on each input port. This state is the photon number eigenstate that has the highest contribution to the optimum state, and can therefore be considered to be an approximation of it. Unfortunately this state has a canonical phase variance that scales as $N^{-1/2}$. This indicates that the phase uncertainty should scale as $N^{-1/4}$. This is an extraordinary result considering that previous work indicated that this state should give a phase uncertainty scaling as N^{-1}.

This discrepancy was shown to be due to the tails of the phase distribution, which give the main contribution to the phase variance. The previous work, in contrast, only considered the central peak of the distribution. To compare these two states more thoroughly, these states were each evaluated under several different measures of the uncertainty. It was found that the phase uncertainty for the $|j0\rangle_z$ state scaled as N^{-1} for all of the measures of the phase uncertainty except the square root of the variance (standard and Holevo).

The optimal states gave smaller scaling constants under all of the measures except two: the inverse-of-maximal-value and the Fisher length. There are reasons to consider the results given by these measures as being slightly misleading, however. In practice, the $|j0\rangle_z$ state will give results with small errors most of the time, but a small number of results with very large errors. This means that results obtained using this state must be carefully analysed to avoid problems due to the results with large error. For example, if we simply take the mean of a number of results obtained using this state, the uncertainty in this mean will scale as $N^{-1/4}$ rather than N^{-1}.

8.5 Optimum Interferometric Measurements

In order to make interferometric phase measurements with the minimum possible phase uncertainty, the other two areas to optimise are the feedback phases used and the final phase estimate. It was shown in Ch. 6 that it is reasonably simple to determine the final phase estimate that gives the minimum uncertainty. Determining the optimum feedback phases to use is a far more difficult problem.

Solving this problem for N photons is a minimisation problem with approximately 2^{N-2} independent variables. It is possible to solve it numerically; however, the rapidly increasing number of variables means that this is only feasible for photon numbers up to about 12. Even then there is a small possibility of a better result for a significantly different combination of numbers, as the numerical technique finds a local minimum, but does not prove that it is a global minimum.

This numerical minimisation was performed for input states optimised for minimum phase variance under canonical measurements. It is also possible to solve simultaneously for the feedback phases and the input state. This is only a slightly more difficult problem, as the number of state coefficients increases linearly with N. Improvements over the optimum input states were found for every photon number above 2; however, these improvements were only very small, less than 1%.

An alternative approach is to, rather than trying to determine the feedback phase that will minimise the final phase estimate, determine the phase that will simply minimise the phase uncertainty after the next detection. This approach was used in Ch. 6, and this feedback scheme gave results that were the same as optimum for up to 4 photons with optimum input states. For more than 4 photons the phase variances obtained were only slightly higher than for the numerically optimised feedback. For the maximum photon number for which the optimum feedback was determined, 12, the increase was only about 3.6%.

This feedback scheme was tested on optimum input states for higher photon numbers up to 1600, and it was found that the phase variance was only slightly higher than the canonical phase variance. That means that this feedback scheme must be very close to optimum for these larger photon numbers as well. Unfortunately, the ratio of the phase variance under this feedback scheme to the canonical phase variance increases fairly systematically with photon number, indicating that the introduced phase variance is not quite scaling as N^{-2}. It is quite likely that the scaling is $\log N/N^2$, as for dyne measurements, though it would require calculations at much larger photon

numbers to confirm this.

For input states other than the optimum input state, this feedback scheme gave variances almost indistinguishable from the canonical phase variance. Of particular interest is the $|j0\rangle_z$ state, for which the canonical phase uncertainty scales as N^{-1} for measures other than the square root of the variance. For this state, the 2/3 confidence interval scales approximately as $N^{-0.69}$ under this measurement scheme. This is not as good as the N^{-1} scaling for the canonical distribution, but it is an improvement on the result if all photons are in one port. Note, however, that this measurement scheme was designed with the aim of minimising the phase variance, not any other measure of phase uncertainty. This means that the feedback phases and phase estimates are not necessarily close to optimum for minimising the confidence interval. Minimising the confidence interval would require quite a different approach.

Alternative phase feedback schemes were also considered in Ch. 6. A nonadaptive scheme was considered, where rather than using information from the detections to determine the feedback phase, the feedback phase was merely varied linearly. This is equivalent to heterodyne measurements for the single mode case, in that all values of the feedback phase are used with roughly equal probability. This scheme gave a phase variance scaling as N^{-1} for optimum input states, similarly to the heterodyne scheme for single mode measurements. When used on the state with all photons in one port (which has a canonical phase variance scaling as N^{-1}) there was only a marginal increase in the phase variance over the adaptive scheme.

An alternative adaptive feedback scheme based on phase estimates was found to give far poorer results than the feedback scheme of Sec. 6.3. Although most of the results had small error, it gave a small proportion of results with very large error, resulting in a high phase variance. On average the phase variance was between that for the feedback scheme of Sec. 6.3 and the nonadaptive scheme.

For 1 or 2 photons, the phase variances obtained by the feedback scheme of Sec. 6.3 were identical to the phase variances for canonical measurements, for any input state. This means that, provided we have no more than 2 photons, it is possible to perform measurements as good as canonical. For 3 or 4 photons, however, the phase variance was higher than canonical. As the feedback phases are optimum for these photon numbers, this means that it is not possible to make canonical measurements in general, even with perfect photodetectors and the best possible feedback phases. This result was checked by evaluating the entire range of feedback phases, in order to demonstrate that the feedback phases used gave a global minimum to the phase variance.

If arbitrary states are considered, this feedback scheme is no longer optimum for 3 or 4 photons. In addition, it is possible to obtain phase variances that are below canonical. This is not a contradiction, as the variance is not necessarily minimised by having all the H_{nm} equal to 1. For some states the optimal POM should have complex values of H_{nm}. Using this corrected canonical POM, the feedback scheme of Sec. 6.3 no longer gives variances as small as canonical for all 2 photon states.

8.6 Continuous Feedback

The last area considered in this study was the problem of continuous phase measurements, where the phase is being varied and the aim is to follow this variation with the minimum possible excess uncertainty. For this problem, we cannot use the optimum input states considered in other chapters, as these are based on single-shot measurements with a limited number of photons. Instead, I considered a continuous coherent state for dyne measurements, and a state with all photons in one port for the interferometric case. These two cases are extremely similar, with both having canonical phase variances proportional to N^{-1}. In addition, continuous squeezed states for dyne measurements were considered. In all cases the phase variation considered was Gaussian diffusion.

In the case of dyne measurements it was found that good results were obtained using $\arg A_t$ feedback, similarly to mark I and II single-shot measurements. In the continuous case, the feedback simplifies to a very simple form, where the feedback phase is adjusted proportional to the photocurrent. This form is even simpler than for the single-shot case. When the correct proportionality constant is selected, a minimum equilibrium phase variance is found that is proportional to α^{-1}. This is much poorer scaling than the case where the system phase is constant, where the variance

scales as α^{-2} (or $\overline{n}^{-1}$). This is because, as the photon number is increased, data from a shorter time is used (to reduce the contribution to the variance from the varying system phase), with the result that the effective photon number that is used increases proportionally to α rather than α^2.

When heterodyne feedback is used, rather than adaptive feedback, the phase variance is increased by a factor of $\sqrt{2}$. This is less than the difference for the single-shot case for a similar reason as the scaling is different: when a more accurate measurement can be made, the time interval from which data is used is reduced. It was also found that the numerical results matched these theoretical predictions very accurately, particularly for the results with smaller phase variances.

For the case of dyne measurements on continuous squeezed states, the situation is considerably more complicated. Rather than just a single constant that must be varied to find the minimum phase variance there are three. Nevertheless, it is still possible to obtain an approximate analytic result. It was found that the minimum phase variance should vary as $\alpha^{-4/3}$. This is only slightly better scaling than the case for coherent states.

Numerically it was found that the scaling of the minimum phase variance was very close to the $\alpha^{-4/3}$ scaling found analytically. The numerical results were well above the analytic result, however, on average more than twice. This appears to be due to the contribution from large phase error results, which were ignored in the analytic treatment. Numerically it was found that when the parameters were varied beyond the optimal values found, the variance increased due to these large error results.

The case for interferometry is more difficult to treat, as it does not give a simple result for the feedback. The feedback used was based on minimising the variance after the next detection, similarly to the single-shot case. Nevertheless, it was found that it is possible to determine an approximate theory that agrees reasonably well the numerical results. Similarly to the dyne case with a coherent state, the phase variance is proportional to $N^{-1/2}$, where N is the photon flux. When a linearly changing feedback phase was used (analogous to the heterodyne scheme), it was found that the phase variance is above that for the adaptive feedback, but the difference is only small, particularly for the smaller phase variances. This is as can be expected, as the difference is very small for large photon numbers in the single-shot case.

8.7 Questions for Future Research

Although the main aims of this project, finding the optimum states and measurement schemes for phase measurements, have been achieved, this project has also raised a number of unanswered questions that are possible future directions for research.

8.7.1 Optimum Dyne Measurements

In Ch. 3 a feedback scheme was found that produced dyne measurements extremely close to the theoretical limit. Although there are good qualitative reasons for expecting a feedback scheme of this type to be close to optimal, there is no rigorous justification for why this scheme works. This question is only of theoretical interest, as it would only tell us why a feedback scheme that has already been found works, rather than leading to any better phase measurements.

Another question is raised by the results showing that the theoretical limit does not hold when different final phase estimates are used. An alternative limit was found in the case of coherent states, but there is no corresponding result for the case of squeezed states or more general states. When the feedback that gives results close to the theoretical limit for the usual $\arg C$ phase estimates was used, it was found that it is only possible to improve the results slightly using better phase estimates. It is conceivable, though unlikely, that it is possible to do better using some other feedback scheme.

A more promising way of improving upon the theoretical limit would be to use nonlinear elements. The basic reason why there is an introduced phase variance is because the phase is not measured directly. The best we can do is to measure a quadrature of the phase, which is proportional to the sine of the phase. The introduced phase variance is due to the fact that the sine function is nonlinear. In principle, it should be possible to compensate for this nonlinearity using nonlinear optics.

Another possible area for further research is to consider alternative measures of the phase uncertainty, as was done for interferometric measurements. For the Holevo phase variance, it is necessary to take account of very low probability, large error results in order to obtain an accurate estimate of the variance. This is unnecessary for other measures of the uncertainty with less emphasis on the tails. Nevertheless, I do not anticipate that this would give qualitatively different results.

8.7.2 Optimum Interferometry

There are a number of promising areas for future study in the area of interferometry. One is the question of whether the theoretical limit for dyne measurements also holds for interferometric measurements. There is a great deal of similarity between interferometric measurements and dyne measurements, as in both cases we are measuring the phase difference between two modes combined at a beam splitter. In principle it should be possible to take the large amplitude limit in a similar way as for dyne measurements, in order to get a similar theoretical limit.

The main problem with taking this limit is that we cannot assume that one mode is of much larger amplitude than the other, as we can for dyne measurements. Some other complications are that there will be quantum correlations between the modes for interferometry, and the final phase estimate that is being used is not $\arg C$. These factors mean that a theoretical limit would be very difficult to find, and it need not be the same as for the case of dyne measurements.

Another area that can be considered is feedback where the feedback phase is selected so as to minimise the phase variance two or more detections in advance, rather than just one. This should improve on the feedback scheme that minimises the variance after the next detection. In addition, if only very minor improvements were obtained for feedback that minimises two or three detections ahead, then this would indicate that the feedback is very close to optimum.

Unfortunately this approach would be very computationally intensive, as there does not appear to be any analytic solution for the feedback phase. The feedback phases would have to be determined numerically, making the calculations far more time consuming, though not as much as for the case where the feedback phases are chosen to minimise the final phase variance.

Another promising direction is using other measures of the phase uncertainty. In particular it would be interesting to optimise the measurements for minimum entropic length, as this would correspond to maximum information. It is conceivable that minimising the entropic length one detection in advance would also minimise the final entropic length. This problem would also be very computationally intensive, as it would require numerical integrals to evaluate the entropic length.

A more difficult problem is performing measurements that have zero error probability, as in Ref. [59]. The measurements considered in Ref. [59] are not physically possible, but it may be possible to perform measurements with zero error probability using feedback. The states considered in Ref. [59] do not have a fixed total photon number in the two modes, as was assumed in this study, so this theory would have to be substantially modified in order to consider these states.

I will lastly mention that it should be possible to generalise the theory considered in this study to phase measurements with an arbitrary number of modes. These are discussed in Ref. [60], and here it is the phase differences between the modes that we wish to measure. This is a difficult problem, as for example it does not seem to be possible to generalise the optimal states for two-mode interferometry to this case in a simple way.

Appendix A

Longer Derivations

In this appendix I give some of the longer derivations that are too lengthy to present in the main text.

A.1 Perturbation Theory for Optimum Dyne States

We wish to solve Eq. (2.91) of Sec. 2.2.2 by perturbation theory, with the unperturbed Hamiltonian and perturbation term given by Eqs (2.94) and (2.95) respectively. The unperturbed solution is

$$\psi_j^{(0)}(n) = \frac{f_2^{1/8}}{\sqrt{\pi^{1/2} 2^j j!}} \exp\left[-\sqrt{f_2}\,(n-n_0)^2/2\right] H_j\left[f_2^{1/4}(n-n_0)\right], \tag{A.1}$$

where H_j are Hermite polynomials. This is the standard result for the harmonic potential, and is easily derived from the properties of Hermite polynomials given in [46]. Similarly to the other derivations given in Ch. 2, the boundary condition is $\psi(0) = 0$. The unperturbed solution, and the more accurate perturbed solutions below, will approximately obey this boundary condition when $\sqrt{f_2}\, n_0^2 \gg 1$. For this to be the case, $p < 2$ is required, in addition to $n_0 \gg 1$.

The energy eigenvalues are

$$E_j^{(0)} = (2j+1)\sqrt{f_2}. \tag{A.2}$$

From perturbation theory we then have

$$E_j \approx E_j^{(0)} + \langle j|\hat{H}_1|j\rangle + \sum_{k\neq j} \frac{\left|\langle k|\hat{H}_1|j\rangle\right|^2}{E_j^{(0)} - E_k^{(0)}}, \tag{A.3}$$

$$\psi_j(n) \approx \psi_j^{(0)}(n) + \sum_{k\neq j} \frac{\langle k|\hat{H}_1|j\rangle}{E_j^{(0)} - E_k^{(0)}} \psi_k^{(0)}(n), \tag{A.4}$$

where $|j\rangle$ is the state corresponding to $\psi_j^{(0)}(n)$. In terms of number states

$$|j\rangle = \sum_{n=0}^{\infty} \psi_j^{(0)}(n)|n\rangle. \tag{A.5}$$

We can rewrite the perturbation as $\hat{H}_1 = b\xi^3$, where

$$b = -\frac{p+2}{3}\left[cp(p+1)\right]^{1/4} n_0^{-p/4-3/2}, \tag{A.6}$$

$$\xi = f_2^{1/4}(n-n_0). \tag{A.7}$$

In terms of ξ the unperturbed eigenstates are

$$\psi_j^{(0)}(\xi) = \frac{1}{\sqrt{\pi^{1/2}2^j j!}} e^{-\xi^2/2} H_j(\xi). \tag{A.8}$$

The factor of $f_2^{1/8}$ has been omitted here, so these states are normalised when integrated with respect to ξ.

The lowest energy eigenvalue and eigenstate, corresponding to the maximum value of ν and therefore $\langle\cos\phi\rangle$, can be expressed as

$$E_0 \approx E_0^{(0)} + b\langle 0|\xi^3|0\rangle - \frac{b^2}{2\sqrt{f_2}} \sum_{k=1}^{\infty} \frac{\left|\langle k|\hat{\xi}^3|0\rangle\right|^2}{k}, \tag{A.9}$$

$$\psi_0(n) \approx \psi_0^{(0)}(n) - \frac{b}{2\sqrt{f_2}} \sum_{k=1}^{\infty} \frac{\langle k|\hat{\xi}^3|0\rangle}{k} \psi_k^{(0)}(n). \tag{A.10}$$

Here $\hat{\xi}$ is the operator that transforms the state $|j\rangle$ in the same way as ξ transforms the function $\psi_j^{(0)}(\xi)$.

The first four Hermite polynomials are

$$H_0(\xi) = 1, \tag{A.11}$$
$$H_1(\xi) = 2\xi, \tag{A.12}$$
$$H_2(\xi) = 4\xi^2 - 2, \tag{A.13}$$
$$H_3(\xi) = 8\xi^3 - 12\xi. \tag{A.14}$$

Therefore the first four unperturbed eigenstates are

$$\psi_0^{(0)}(\xi) = \pi^{-1/4} e^{-\xi^2/2}, \tag{A.15}$$
$$\psi_1^{(0)}(\xi) = \pi^{-1/4} e^{-\xi^2/2} \sqrt{2}\xi, \tag{A.16}$$
$$\psi_2^{(0)}(\xi) = \pi^{-1/4} e^{-\xi^2/2} \left(\sqrt{2}\xi^2 - \frac{1}{\sqrt{2}}\right), \tag{A.17}$$
$$\psi_3^{(0)}(\xi) = \pi^{-1/4} e^{-\xi^2/2} \left(\frac{2}{\sqrt{3}}\xi^3 - \sqrt{3}\xi\right). \tag{A.18}$$

It is therefore easy to show that

$$\xi^3 \psi_0^{(0)}(\xi) = \frac{3}{2\sqrt{2}} \psi_1^{(0)}(\xi) + \frac{\sqrt{3}}{2} \psi_3^{(0)}(\xi), \tag{A.19}$$

or, in the alternative notation

$$\hat{\xi}^3|0\rangle = \frac{3}{2\sqrt{2}}|1\rangle + \frac{\sqrt{3}}{2}|3\rangle. \tag{A.20}$$

From this it is evident that the only non-zero terms in the sums in (A.9) are $\langle 1|\xi^3|0\rangle = 3/(2\sqrt{2})$ and $\langle 3|\xi^3|0\rangle = \sqrt{3}/2$. This then gives the lowest energy eigenvalue and eigenstate as

$$E_0 \approx E_0^{(0)} - b^2 \frac{11}{16} f_2^{-1/2}, \tag{A.21}$$

$$\psi_0(n) \approx \psi_0^{(0)}(n) - \frac{b}{4\sqrt{2f_2}} \left[3\psi_1^{(0)} + \sqrt{\frac{2}{3}}\psi_3^{(0)}\right]. \tag{A.22}$$

The correction term to the eigenvalue is of order n_0^{-2}, which is sufficiently small to be omitted. The eigenstate can alternatively be expressed as

$$|\psi\rangle \approx |0\rangle - \frac{3b}{4\sqrt{2f_2}}|1\rangle - \frac{b}{4\sqrt{3f_2}}|3\rangle. \tag{A.23}$$

Therefore the expectation value of the photon number is

$$\langle n\rangle = \left(\langle 0| - \frac{3b}{4\sqrt{2f_2}}\langle 1| - \frac{b}{4\sqrt{3f_2}}\langle 3|\right)\left[\frac{\hat{\xi}}{(f_2)^{1/4}} + n_0\right]\left(|0\rangle - \frac{3b}{4\sqrt{2f_2}}|1\rangle - \frac{b}{4\sqrt{3f_2}}|3\rangle\right) \tag{A.24}$$

In order to evaluate this, note first that we wish to keep terms only up to first order in the perturbation, so we can simplify this to

$$\begin{aligned}\bar{n} &\approx n_0 + \left(\langle 0| - \frac{3b}{4\sqrt{2f_2}}\langle 1| - \frac{b}{4\sqrt{3f_2}}\langle 3|\right)\left[\frac{\hat{\xi}}{(f_2)^{1/4}}\right]|0\rangle + \langle 0|\left[\frac{\hat{\xi}}{(f_2)^{1/4}}\right]\left(-\frac{3b}{4\sqrt{2f_2}}|1\rangle - \frac{b}{4\sqrt{3f_2}}|3\rangle\right)\\ &= n_0 + \left(\langle 0| - \frac{3b}{2\sqrt{2f_2}}\langle 1| - \frac{b}{2\sqrt{3f_2}}\langle 3|\right)\left[\frac{\hat{\xi}}{(f_2)^{1/4}}\right]|0\rangle.\end{aligned} \tag{A.25}$$

From the above listing of the eigenstates it can be seen that

$$\xi\psi_0^{(0)}(\xi) = \frac{1}{\sqrt{2}}\psi_1^{(0)}(\xi), \tag{A.26}$$

or

$$\hat{\xi}|0\rangle = \frac{1}{\sqrt{2}}|1\rangle. \tag{A.27}$$

Therefore the only non-zero matrix element above is $\langle 1|\hat{\xi}|0\rangle$. Using this gives

$$\begin{aligned}\bar{n} &\approx n_0 - \frac{3b}{4(f_2)^{3/4}}\\ &= n_0 + \frac{p+2}{4\sqrt{cp(p+1)}}n_0^{p/2}.\end{aligned} \tag{A.28}$$

As we are assuming $p < 2$,

$$n_0 \approx \bar{n}\left[1 - \frac{p+2}{4\sqrt{cp(p+1)}}\bar{n}^{p/2-1}\right], \tag{A.29}$$

so the mean photon number is close to n_0, justifying the expansion around n_0.

A.2 Derivation for Optimum Squeezed States

Next the result (2.149) of Sec. 2.3.3 will be derived in a more rigorous way. Recall that the number state coefficients for squeezed states are given by an expression that depends on Hermite polynomials (2.121). Hermite polynomials satisfy the recursion relation [46]

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0. \tag{A.30}$$

This can be used to derive the recursion relation for number state coefficients:

$$\langle n+1|\alpha,\zeta\rangle\mu\sqrt{n+1} - \langle n|\alpha,\zeta\rangle\beta + \langle n-1|\alpha,\zeta\rangle\nu\sqrt{n} = 0. \tag{A.31}$$

Rearranging this and squaring gives

$$|\langle n+1|\alpha,\zeta\rangle|^2\mu^2(n+1) = |\langle n|\alpha,\zeta\rangle|^2\beta^2 + |\langle n-1|\alpha,\zeta\rangle|^2\nu^2 n - 2\langle\alpha,\zeta|n\rangle\langle n-1|\alpha,\zeta\rangle\beta\nu\sqrt{n}. \tag{A.32}$$

Multiplying this by n^k and summing gives

$$\sum_{n=1}^{\infty}|\langle n+1|\alpha,\zeta\rangle|^2\mu^2(n+1)n^k = \sum_{n=1}^{\infty}|\langle n|\alpha,\zeta\rangle|^2\beta^2 n^k + \sum_{n=1}^{\infty}|\langle n-1|\alpha,\zeta\rangle|^2\nu^2 n^{k+1}$$

$$
\begin{aligned}
&&&-\sum_{n=1}^{\infty}2\langle\alpha,\zeta|n\rangle\langle n-1|\alpha,\zeta\rangle\beta\nu n^{k+1/2}\\
2\beta\nu\sum_{n=1}^{\infty}n^{k+1/2}\langle\alpha,\zeta|n\rangle\langle n-1|\alpha,\zeta\rangle &= \beta^2\sum_{n=1}^{\infty}n^k|\langle n|\alpha,\zeta\rangle|^2+\nu^2\sum_{n=1}^{\infty}n^{k+1}|\langle n-1|\alpha,\zeta\rangle|^2\\
&&&-\mu^2\sum_{n=1}^{\infty}(n+1)n^k|\langle n+1|\alpha,\zeta\rangle|^2\\
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{k+1/2}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle &= \beta^2\sum_{n=1}^{\infty}n^k|\langle n|\alpha,\zeta\rangle|^2+\nu^2\sum_{n=0}^{\infty}(n+1)^{k+1}|\langle n|\alpha,\zeta\rangle|^2\\
&&&-\mu^2\sum_{n=2}^{\infty}n(n-1)^k|\langle n|\alpha,\zeta\rangle|^2 \qquad \text{(A.33)}
\end{aligned}
$$

If $k>0$, then the sums in the first and third terms on the right hand side can be extended to $n=0$, giving

$$
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{k+1/2}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle=\beta^2\left\langle n^k\right\rangle+\nu^2\left\langle(n+1)^{k+1}\right\rangle-\mu^2\left\langle n(n-1)^k\right\rangle. \qquad \text{(A.34)}
$$

In the following we wish to take $k=\frac{1}{2}-p$, so $k\leq 0$. In this case the sums cannot be extended to zero, and in fact the additional terms would be infinite. In the following expansions, however, only the behaviour of the state near $n=\bar{n}$ is considered, and the contribution from $n\sim 0$ is negligible. It is therefore reasonable to use Eq. (A.34) as a basis for the approximate expansions, despite the fact that some of the terms would be infinite if they were worked out exactly.

Taking $k=\frac{1}{2}-p$ and considering the deviation from the mean photon number gives

$$
\begin{aligned}
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle &\approx \beta^2\left\langle(\overline{n}+\Delta n)^{-(p+1/2)}\right\rangle+\nu^2\left\langle(\overline{n}+1+\Delta n)^{-(p-1/2)}\right\rangle\\
&-\mu^2\left\langle(\overline{n}-1+\Delta n)^{-(p-1/2)}\right\rangle-\mu^2\left\langle(\overline{n}-1+\Delta n)^{-(p+1/2)}\right\rangle. \qquad \text{(A.35)}
\end{aligned}
$$

Expanding this in a series in Δn gives

$$
\begin{aligned}
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle &\approx \sum_{j=0}^{\infty}\frac{(-1)^j\left\langle\Delta n^j\right\rangle}{j!}\left\{\frac{(p+j-1/2)!}{(p-1/2)!}\left(\beta^2\overline{n}^{-(p+j+1/2)}\right.\right.\\
\left.\left.-\mu^2(\overline{n}-1)^{-(p+j+1/2)}\right)+\frac{(p+j-3/2)!}{(p-3/2)!}\left[\nu^2(\overline{n}+1)^{-(p+j-1/2)}-\mu^2(\overline{n}-1)^{-(p+j-1/2)}\right]\right\}. \qquad \text{(A.36)}
\end{aligned}
$$

Now we have an expression that can be used to evaluate Eq. (2.145). Here the approximation that will be used for $h(n)$ is $c(n+1)^{-p}$ rather than cn^{-p}. This is reasonable, as the difference is of order $\overline{n}^{-(p+1)}$, which is of higher order than will be considered here. Using the first three terms of the sum in Eq. (A.36) gives

$$
\begin{aligned}
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle &\approx \beta^2\overline{n}^{-(p+1/2)}-\mu^2(\overline{n}-1)^{-(p+1/2)}+\nu^2(\overline{n}+1)^{-(p-1/2)}\\
-\mu^2(\overline{n}-1)^{-(p-1/2)}+\frac{\left\langle\Delta n^2\right\rangle}{2}&\left\{(p+3/2)(p+1/2)\left(\beta^2\overline{n}^{-(p+5/2)}-\mu^2(\overline{n}-1)^{-(p+5/2)}\right)\right.\\
\left.+(p+1/2)(p-1/2)\left[\nu^2(\overline{n}+1)^{-(p+3/2)}-\mu^2(\overline{n}-1)^{-(p+3/2)}\right]\right\} &\qquad \text{(A.37)}
\end{aligned}
$$

Note that the $j=1$ term of the sum is zero, as $\langle\Delta n\rangle=0$. Using $\left\langle\Delta n^2\right\rangle=\alpha^2(\mu-\nu)^2+2\mu^2\nu^2$ this equation becomes

$$
2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle\approx\beta^2\overline{n}^{-(p+1/2)}-\mu^2(\overline{n}-1)^{-(p+1/2)}+\nu^2(\overline{n}+1)^{-(p-1/2)}
$$

$$-\mu^2(\overline{n}-1)^{-(p-1/2)}+\left[\frac{\alpha^2(\mu-\nu)^2}{2}+\mu^2\nu^2\right]\left\{\left(p^2+2p+\frac{3}{4}\right)\left(\beta^2\overline{n}^{-(p+5/2)}\right.\right.$$
$$\left.\left.-\mu^2(\overline{n}-1)^{-(p+5/2)}\right)+\left(p^2-\frac{1}{4}\right)\left[\nu^2(\overline{n}+1)^{-(p+3/2)}-\mu^2(\overline{n}-1)^{-(p+3/2)}\right]\right\}. \quad \text{(A.38)}$$

At this stage the main problem is to determine which terms should be kept. This depends on how n_0 scales with $\bar{n}$. Recall that if the state is optimised for minimum intrinsic phase uncertainty, then $n_0 \propto \log(\overline{n})$. That scaling cannot be assumed in this case; however, it is possible to make some general assumptions about the scaling of n_0. Firstly n_0 should increase with $\overline{n}$, so in order to obtain two orders in the approximation we will keep terms up to leading order divided by n_0.

Secondly n_0 should not increase as rapidly as $\overline{n}$, as the squeezing should increase as the photon number is increased. This means that $n_0/\overline{n}^2$ should be higher order than $1/n_0$, and we can therefore omit terms of leading order times $n_0/\overline{n}^2$. We will, however, include terms of leading order times $n_0/\overline{n}$, as we cannot assume that these are higher order. We will not obtain any terms of leading order times $(n_0/\overline{n})^2$ or any higher power, so we do not need to consider these terms.

The third and fourth terms in (A.38) appear to be of order $\overline{n}^{-(p-3/2)}/n_0$. We can partially cancel these terms, however, by making the expansion

$$(\overline{n}\pm 1)^{-(p-1/2)}=\overline{n}^{-(p-1/2)}\left[1\mp\frac{(p-\frac{1}{2})}{\overline{n}}+\frac{(p+\frac{1}{2})(p-\frac{1}{2})}{2\overline{n}^2}+O(\overline{n}^{-3})\right]. \quad \text{(A.39)}$$

This means that

$$\nu^2(\overline{n}+1)^{-(p-1/2)}-\mu^2(\overline{n}-1)^{-(p-1/2)}\approx(\nu^2-\mu^2)\overline{n}^{-(p-1/2)}\left[1+\frac{(p+\frac{1}{2})(p-\frac{1}{2})}{2\overline{n}^2}\right]$$
$$-(\nu^2+\mu^2)(p-\tfrac{1}{2})\overline{n}^{-(p+1/2)}. \quad \text{(A.40)}$$

Now $\nu^2-\mu^2=-1$, so together these terms are of order $\overline{n}^{-(p-1/2)}$. This is the leading order, and we will now omit terms known to be higher order than $\overline{n}^{-(p-1/2)}/n_0$. This means that we will use

$$\nu^2(\overline{n}+1)^{-(p-1/2)}-\mu^2(\overline{n}-1)^{-(p-1/2)}\approx-\overline{n}^{-(p-1/2)}-(\nu^2+\mu^2)(p-\tfrac{1}{2})\overline{n}^{-(p+1/2)}. \quad \text{(A.41)}$$

Similarly for the terms on the third line of (A.38) there is the expansion

$$(\overline{n}\pm 1)^{-(p+3/2)}=\overline{n}^{-(p+3/2)}\left[1\mp\frac{(p+\frac{3}{2})}{\overline{n}}+\frac{(p+\frac{5}{2})(p+\frac{3}{2})}{2\overline{n}^2}+O(\overline{n}^{-3})\right], \quad \text{(A.42)}$$

so

$$\nu^2(\overline{n}+1)^{-(p+3/2)}-\mu^2(\overline{n}-1)^{-(p+3/2)}\approx-\overline{n}^{-(p+3/2)}-(\nu^2+\mu^2)(p+\tfrac{3}{2})\overline{n}^{-(p+5/2)}. \quad \text{(A.43)}$$

Using this result, and omitting all terms known to be of higher order than $\overline{n}^{-(p-1/2)}/n_0$, Eq. (A.38) simplifies to

$$2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle\approx\beta^2\overline{n}^{-(p+1/2)}-\mu^2(\overline{n}-1)^{-(p+1/2)}-\overline{n}^{-(p-1/2)}$$
$$-(\nu^2+\mu^2)(p-\tfrac{1}{2})\overline{n}^{-(p+1/2)}+\left[\frac{\alpha^2(\mu-\nu)^2}{2}\right]\left(p^2-\frac{1}{4}\right)\left[-\overline{n}^{-(p+3/2)}\right]. \quad \text{(A.44)}$$

Now using $\beta^2\approx n_0$, $\mu^2\approx\nu^2\approx\overline{n}/(4n_0)$ and $\alpha(\mu-\nu)^2\approx\overline{n}^2/n_0$, this simplifies to

$$2\beta\nu\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle\approx-\overline{n}^{-(p-1/2)}\left\{1-\frac{n_0}{\overline{n}}+\frac{1}{2n_0}\left[p(p+1)-\frac{1}{4}\right]\right\}. \quad \text{(A.45)}$$

Now we can expand $2\beta\nu$ to give

$$\begin{aligned} 2\beta\nu &= -\bar{n}^{1/2}\sqrt{1-\frac{\nu^2}{\overline{n}}}\left(1-\frac{n_0}{\bar{n}}\right) \\ &\approx -\bar{n}^{1/2}\sqrt{1-\frac{1}{4n_0}}\left(1-\frac{n_0}{\overline{n}}\right). \end{aligned} \tag{A.46}$$

Using this gives

$$\begin{aligned} \left(1-\frac{n_0}{\bar{n}}\right)\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle &\approx \overline{n}^{-p}\frac{\left\{1-\frac{n_0}{\overline{n}}+\frac{1}{2n_0}\left[p(p+1)-\frac{1}{4}\right]\right\}}{\sqrt{1-\frac{1}{4n_0}}} \\ &\approx \overline{n}^{-p}\left\{1-\frac{n_0}{\overline{n}}+\frac{p(p+1)}{2n_0}\right\}. \end{aligned} \tag{A.47}$$

The two terms of $n_0/\overline{n}$ just cancel, giving the simple result

$$\sum_{n=0}^{\infty}(n+1)^{-p}\langle\alpha,\zeta|n\rangle\langle n+1|\alpha,\zeta\rangle \approx \overline{n}^{-p}\left[1+\frac{p(p+1)}{2n_0}\right]. \tag{A.48}$$

This result is identical to that obtained simply using $\langle n^{-p}\rangle$.

A.3 Perturbation Theory for Mark II Measurements

From Sec. 4.2.2, the phase variance for mark II measurements with a time delay is, according to perturbation theory

$$\left\langle\phi_{\mathrm{II}}^2\right\rangle = \int_{v_1}^1 dt\int_{v_1}^1 dt'\left\langle\hat{\varphi}_t^{(0)}\hat{\varphi}_{t'}^{(0)}\right\rangle + 2\alpha\tau\int_{v_1}^1 dt\int_{v_1}^1 dt'\left\langle\hat{\varphi}_t^{(0)}\hat{\varphi}_{t'}^{(1)}\right\rangle. \tag{A.49}$$

Using the result for $\hat{\varphi}_t^{(0)}$ given in Eq. (4.8), the first term can be evaluated as

$$\begin{aligned} &\int_{v_1}^1 dt\int_{v_1}^1 dt'\left\langle\left[e^{4\alpha(\sqrt{v_1}-\sqrt{t})}\hat{\varphi}_{v_1}^{(0)}+\int_{v_1}^t\frac{e^{4\alpha(\sqrt{u}-\sqrt{t})}}{\sqrt{u}}dW(u)\right]\right. \\ &\qquad\left.\times\left[e^{4\alpha(\sqrt{v_1}-\sqrt{t'})}\hat{\varphi}_{v_1}^{(0)}+\int_{v_1}^{t'}\frac{e^{4\alpha(\sqrt{u}-\sqrt{t'})}}{\sqrt{u}}dW(u)\right]\right\rangle \\ &= \left[\int_{v_1}^1 e^{4\alpha(\sqrt{v_1}-\sqrt{t})}dt\right]^2\left\langle(\hat{\varphi}_{v_1}^{(0)})^2\right\rangle+\int_{v_1}^1 dt\int_{v_1}^1 dt'\int_{v_1}^{\min(t,t')}du\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u} \\ &= \left[\int_{v_1}^1 e^{4\alpha(\sqrt{v_1}-\sqrt{t})}dt\right]^2\left\langle(\hat{\varphi}_{v_1}^{(0)})^2\right\rangle+2\int_{v_1}^1 dt\int_{v_1}^t dt'\int_{v_1}^{t'}du\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}. \end{aligned} \tag{A.50}$$

Considering the first term, this simplifies to

$$\begin{aligned} &\left[2\int_0^{1-\sqrt{v_1}}(\sqrt{v_1}+s)e^{-4\alpha s}ds\right]^2\left\langle(\hat{\varphi}_{v_1}^{(0)})^2\right\rangle \\ &= \left[\frac{\sqrt{v_1}}{2\alpha}+\frac{1}{8\alpha^2}\right]^2\left\langle(\hat{\varphi}_{v_1}^{(0)})^2\right\rangle \\ &= \left[\frac{v_1}{4\alpha^2}+\frac{\sqrt{v_1}}{8\alpha^3}+\frac{1}{64\alpha^4}\right]\left\langle(\hat{\varphi}_{v_1}^{(0)})^2\right\rangle. \end{aligned} \tag{A.51}$$

The term that we take to be leading order here depends on how we take the limits. If we consider the limit of large α with a fixed value of v_1, then the first term is largest. On the other hand, if we

consider the limit of small v_1 the third term is the largest. As the variance should be small at time v_1, a large number of detections should have been made by this time. This means that we should have $\alpha^2 v_1 \gg 1$. In this limit, the first term above is largest.

Now consider the second term in Eq. (A.50). Changing variables to $s = \sqrt{t'} - \sqrt{u}$ gives

$$\begin{aligned} &2\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \int_{\sqrt{t'}-\sqrt{v_1}}^{0} [-2(\sqrt{t'}-s)ds] e^{4\alpha(\sqrt{t'}-\sqrt{t})} \frac{e^{-8\alpha s}}{(\sqrt{t'}-s)^2} \\ &= 4\int_{v_1}^{1} dt \int_{v_1}^{t} dt' e^{4\alpha(\sqrt{t'}-\sqrt{t})} \int_{0}^{\sqrt{t'}-\sqrt{v_1}} ds \frac{e^{-8\alpha s}}{\sqrt{t'}-s}. \end{aligned} \tag{A.52}$$

Performing an expansion in s,

$$\begin{aligned} &4\int_{v_1}^{1} dt \int_{v_1}^{t} dt' e^{4\alpha(\sqrt{t'}-\sqrt{t})} \int_{0}^{\sqrt{t'}-\sqrt{v_1}} \left(\frac{1}{\sqrt{t'}} + \frac{s}{t'} + O(s^2)\right) e^{-8\alpha s} ds \\ &= 4\int_{v_1}^{1} dt \int_{v_1}^{t} dt' e^{4\alpha(\sqrt{t'}-\sqrt{t})} \left(\frac{1}{\sqrt{t'}8\alpha} + \frac{1}{t'(8\alpha)^2} + O(\alpha^{-3})\right). \end{aligned} \tag{A.53}$$

Next we substitute $s = \sqrt{t} - \sqrt{t'}$ so that $dt' = -2(\sqrt{t} - s)ds$. This gives

$$8\int_{v_1}^{1} dt \int_{0}^{\sqrt{t}-\sqrt{v_1}} e^{-4\alpha s} \left(\frac{1}{8\alpha} + \frac{1}{(\sqrt{t}-s)(8\alpha)^2} + O(\alpha^{-3})\right) ds. \tag{A.54}$$

Again expanding in a series we obtain

$$\begin{aligned} &8\int_{v_1}^{1} dt \int_{0}^{\sqrt{t}-\sqrt{v_1}} e^{-4\alpha s} \left(\frac{1}{8\alpha} + \frac{1}{\sqrt{t}(8\alpha)^2} + \frac{s}{t(8\alpha)^2} + O(s^2\alpha^2) + O(\alpha^{-3})\right) ds \\ &= 8\int_{v_1}^{1} \left(\frac{1}{(8\alpha)(4\alpha)} + O(\alpha^{-3})\right) dt \\ &= \frac{1-v_1}{4\alpha^2} + O(\alpha^{-3}). \end{aligned} \tag{A.55}$$

In the limit of small v_1 this simplifies to $1/(4\alpha^2)$. Note that the higher order terms diverge if the limit $v_1 \to 0$ is taken for fixed α. However, as we are taking v_1 such that $\alpha^2 v_1 \gg 1$, they will be smaller than the term given here.

Thus we find that the first term of Eq. (A.49) is

$$\int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(0)} \right\rangle = \frac{1}{4\alpha^2} + \frac{v_1}{4\alpha^2} \left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle + O(\alpha^{-3}). \tag{A.56}$$

Next consider the second term in Eq. (A.49). Evaluating this gives

$$\begin{aligned} &\int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \left[e^{4\alpha(2\sqrt{v_1}-\sqrt{t}-\sqrt{t'})} \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} - \int_{v_1}^{t'} \frac{4\alpha}{s} e^{4\alpha(2\sqrt{v_1}-\sqrt{t}-\sqrt{t'})} (\hat{\varphi}_{v_1}^{(0)})^2 ds \right.\right. \\ &\quad - \int_{v_1}^{t'} \frac{4\alpha}{s} ds \int_{v_1}^{t} \int_{v_1}^{s} \frac{e^{4\alpha(\sqrt{v}-\sqrt{t})}}{\sqrt{v}} dW(v) \frac{e^{4\alpha(\sqrt{u}-\sqrt{t'})}}{\sqrt{u}} dW(u) \\ &\quad \left.\left. + \int_{v_1}^{t} \int_{v_1}^{t'} \frac{e^{4\alpha(\sqrt{v}-\sqrt{t})}}{\sqrt{v}} dW(v) \frac{2}{u} e^{4\alpha(\sqrt{u}-\sqrt{t'})} dW(u) \right] \right\rangle \\ &= \int_{v_1}^{1} dt \int_{v_1}^{1} dt' e^{4\alpha(2\sqrt{v_1}-\sqrt{t}-\sqrt{t'})} \left\langle \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} \right\rangle \\ &\quad + \int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left[4\alpha(\log v_1 - \log t') e^{4\alpha(2\sqrt{v_1}-\sqrt{t}-\sqrt{t'})} \left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle \right. \\ &\quad \left. - \int_{v_1}^{t'} \frac{4\alpha}{s} ds \int_{v_1}^{\min(t,s)} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u} du + 2\int_{v_1}^{\min(t,t')} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u^{1.5}} du \right]. \end{aligned} \tag{A.57}$$

The first term simplifies to

$$\left[\frac{v_1}{4\alpha^2}+\frac{\sqrt{v_1}}{8\alpha^3}+\frac{1}{64\alpha^4}\right]\left\langle\hat{\varphi}^{(0)}_{v_1}\hat{\varphi}^{(1)}_{v_1}\right\rangle. \tag{A.58}$$

As was discussed above, the first term here will be largest when we take v_1 such that $\alpha^2 v_1 \gg 1$.

Now consider the second term in Eq. (A.57). Taking the integral over t this simplifies to

$$\left(2\sqrt{v_1}+\frac{1}{2\alpha}\right)\int_{v_1}^{1}(\log v_1-\log t')e^{4\alpha(\sqrt{v_1}-\sqrt{t'})}\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle dt'. \tag{A.59}$$

Now we make the substitution $s=\sqrt{t'}-\sqrt{v_1}$ so $dt'=2(s+\sqrt{v_1})ds$, giving

$$\begin{aligned}
&-\left(2\sqrt{v_1}+\frac{1}{2\alpha}\right)\int_0^{1-\sqrt{v_1}}\log\left(\frac{(s+\sqrt{v_1})^2}{v_1}\right)e^{-4\alpha s}\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle[2(s+\sqrt{v_1})ds]\\
&=-\left(8\sqrt{v_1}+\frac{2}{\alpha}\right)\int_0^{1-\sqrt{v_1}}\log\left(1+\frac{s}{\sqrt{v_1}}\right)e^{-4\alpha s}\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle[(s+\sqrt{v_1})ds].
\end{aligned} \tag{A.60}$$

Again, there is a slight problem, as the result obtained depends on how the limits are taken. The contributions to this integral will only be significant when s is less than about $1/\alpha$. As we should have $\alpha^2 v_1 \gg 1$, this means we should have $s/\sqrt{v_1} \ll 1$. In this limit, we obtain

$$\begin{aligned}
&-\left(8\sqrt{v_1}+\frac{2}{\alpha}\right)\int_0^{1-\sqrt{v_1}}\left(s+\frac{s^2}{2\sqrt{v_1}}+O(s^3)\right)e^{-4\alpha s}\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle ds\\
&=-\left(8\sqrt{v_1}+\frac{2}{\alpha}\right)\left(\frac{1}{(4\alpha)^2}+\frac{1}{\sqrt{v_1}(4\alpha)^3}+O(\alpha^{-4})\right)\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle\\
&=-\left(\frac{\sqrt{v_1}}{2\alpha^2}+\frac{1}{4\alpha^3}+O(\alpha^{-4})\right)\left\langle(\hat{\varphi}^{(0)}_{v_1})^2\right\rangle.
\end{aligned} \tag{A.61}$$

Next we will consider the third term. To treat the minimum we must split the integral into three parts:

$$\begin{aligned}
&-\int_{v_1}^1 dt\int_{v_1}^1 dt'\int_{v_1}^{t'}\frac{4\alpha}{v}dv\int_{v_1}^{\min(t,v)}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du\\
&=-\int_{v_1}^1 dt\int_{v_1}^t dt'\int_{v_1}^{t'}\frac{4\alpha}{v}dv\int_{v_1}^{v}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du\\
&\quad-\int_{v_1}^1 dt\int_{t}^1 dt'\int_{t}^{t'}\frac{4\alpha}{v}dv\int_{v_1}^{t}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du\\
&\quad-\int_{v_1}^1 dt\int_{t}^1 dt'\int_{v_1}^{t}\frac{4\alpha}{v}dv\int_{v_1}^{v}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du\\
&=-2\int_{v_1}^1 dt\int_{v_1}^t dt'\int_{v_1}^{t'}\frac{4\alpha}{v}dv\int_{v_1}^{v}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du \qquad \text{(a)}\\
&\quad-\int_{v_1}^1 dt\int_{t}^1 dt'\int_{t}^{t'}\frac{4\alpha}{v}dv\int_{v_1}^{t}\frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u}du. \qquad \text{(b)}
\end{aligned} \tag{A.62}$$

Considering term (a), this simplifies to

$$\begin{aligned}
&-2\int_{v_1}^1 dt\int_{v_1}^t dt'\int_{v_1}^{t'}\frac{4\alpha}{v}dv\int_0^{\sqrt{v}-\sqrt{v_1}}\frac{e^{-8\alpha s}}{\sqrt{v}-s}e^{4\alpha(2\sqrt{v}-\sqrt{t}-\sqrt{t'})}ds\\
&=-2\int_{v_1}^1 dt\int_{v_1}^t dt'\int_{v_1}^{t'}\frac{8\alpha}{v}dv\int_0^{\sqrt{v}-\sqrt{v_1}}\left(\frac{1}{\sqrt{v}}+\frac{s}{v}+O(s^2)\right)e^{-8\alpha s}e^{4\alpha(2\sqrt{v}-\sqrt{t}-\sqrt{t'})}ds
\end{aligned}$$

$$
\begin{aligned}
&= -2\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \int_{v_1}^{t'} \frac{8\alpha}{v} dv \left(\frac{1}{8\alpha\sqrt{v}} + \frac{s}{(8\alpha)^2 v} + O(\alpha^{-3}) \right) e^{4\alpha(2\sqrt{v}-\sqrt{t}-\sqrt{t'})} \\
&= -4\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \int_{0}^{\sqrt{t'}-\sqrt{v_1}} \left(\frac{1}{(\sqrt{t'}-s)^2} + \frac{1}{(8\alpha)(\sqrt{t'}-s)^3} + O(\alpha^{-3}) \right) e^{4\alpha(\sqrt{t'}-\sqrt{t})} e^{-8\alpha s} ds \\
&= -4\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \int_{0}^{\sqrt{t'}-\sqrt{v_1}} \left(\frac{1}{t'} + \frac{2s}{t'^{1.5}} + \frac{1}{(8\alpha)t'^{1.5}} + O(\alpha^{-3}) \right) e^{4\alpha(\sqrt{t'}-\sqrt{t})} e^{-8\alpha s} ds \\
&= -\frac{1}{2\alpha}\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \left(\frac{1}{t'} + \frac{3}{(8\alpha)t'^{1.5}} + O(\alpha^{-3}) \right) e^{4\alpha(\sqrt{t'}-\sqrt{t})} \\
&= -\frac{1}{\alpha}\int_{v_1}^{1} dt \int_{0}^{\sqrt{t}-\sqrt{v_1}} \left(\frac{1}{\sqrt{t}-s} + \frac{3}{(8\alpha)(\sqrt{t}-s)^2} + O(\alpha^{-2}) \right) e^{-4\alpha s} ds \\
&= -\frac{1}{\alpha}\int_{v_1}^{1} dt \int_{0}^{\sqrt{t}-\sqrt{v_1}} \left(\frac{1}{\sqrt{t}} + \frac{s}{t} + \frac{3}{8\alpha t} + O(\alpha^{-2}) \right) e^{-4\alpha s} ds \\
&= -\frac{1}{4\alpha^2}\int_{v_1}^{1} dt \left(\frac{1}{\sqrt{t}} + \frac{5}{8\alpha t} + O(\alpha^{-2}) \right).
\end{aligned} \tag{A.63}
$$

Term (b) simplifies to

$$
\begin{aligned}
&-4\alpha\int_{v_1}^{1} dt \int_{t}^{1} dt' [\log t' - \log t] \int_{v_1}^{t} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u} du \\
&= -8\alpha\int_{v_1}^{1} dt \int_{t}^{1} dt' [\log t' - \log t] \int_{0}^{\sqrt{t}-\sqrt{v_1}} \frac{e^{-8\alpha s}}{\sqrt{t}-s} e^{4\alpha(\sqrt{t}-\sqrt{t'})} ds \\
&= -8\alpha\int_{v_1}^{1} dt \int_{t}^{1} dt' [\log t' - \log t] \int_{0}^{\sqrt{t}-\sqrt{v_1}} \left(\frac{1}{\sqrt{t}} + \frac{s}{t} \right) e^{-8\alpha s} e^{4\alpha(\sqrt{t}-\sqrt{t'})} ds \\
&= -\int_{v_1}^{1} dt \int_{t}^{1} dt' [\log t' - \log t] \left(\frac{1}{\sqrt{t}} + \frac{1}{8\alpha t} + O(s^2) \right) e^{4\alpha(\sqrt{t}-\sqrt{t'})} \\
&= -4\int_{v_1}^{1} dt \left(\frac{1}{\sqrt{t}} + \frac{1}{8\alpha t} + O(\alpha^{-2}) \right) \int_{0}^{1+\sqrt{t}} \log\left(1 + \frac{s}{\sqrt{t}} + O(\alpha^{-2} \right) (\sqrt{t}+s) e^{-4\alpha s} ds \\
&= -4\int_{v_1}^{1} dt \left(\frac{1}{\sqrt{t}} + \frac{1}{8\alpha t} + O(\alpha^{-2}) \right) \int_{0}^{1+\sqrt{t}} \left(s + \frac{s^2}{2\sqrt{t}} + O(s^3) \right) e^{-4\alpha s} ds \\
&= -4\int_{v_1}^{1} dt \left(\frac{1}{\sqrt{t}} + \frac{1}{8\alpha t} + O(\alpha^{-2}) \right) \left(\frac{1}{(4\alpha)^2} + \frac{1}{(4\alpha)^3\sqrt{t}} + O(\alpha^{-4}) \right) \\
&= -\frac{1}{4\alpha^2}\int_{v_1}^{1} \left(\frac{1}{\sqrt{t}} + \frac{3}{8\alpha t} + O(\alpha^{-2}) \right) dt.
\end{aligned} \tag{A.64}
$$

Therefore the third term of Eq. (A.57) simplifies to

$$
-\frac{1}{2\alpha^2}\int_{v_1}^{1} \left(\frac{1}{\sqrt{t}} + \frac{1}{2\alpha t} + O(\alpha^{-2}) \right) dt. \tag{A.65}
$$

Next considering the fourth term of Eq. (A.57), we obtain

$$
\begin{aligned}
&2\int_{v_1}^{1} dt \int_{v_1}^{t} dt' \int_{v_1}^{t'} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u^{1.5}} du + 2\int_{v_1}^{1} dt \int_{t}^{1} dt' \int_{v_1}^{t} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u^{1.5}} du \\
&= 4\int_{v_1}^{1} dt \int_{t}^{1} dt' \int_{v_1}^{t} \frac{e^{4\alpha(2\sqrt{u}-\sqrt{t}-\sqrt{t'})}}{u^{1.5}} du \\
&= 4\int_{v_1}^{1} dt \left(\frac{\sqrt{t}}{2\alpha} + \frac{1}{8\alpha^2} \right) \int_{v_1}^{t} \frac{e^{8\alpha(\sqrt{u}-\sqrt{t})}}{u^{1.5}} du
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{4}{\alpha}\int_{v_1}^{1} dt \left(\sqrt{t}+\frac{1}{4\alpha}\right)\int_0^{\sqrt{t}-\sqrt{v_1}} \frac{e^{-8\alpha s}}{(\sqrt{t}-s)^2} ds \\
&= \frac{4}{\alpha}\int_{v_1}^{1} dt \left(\sqrt{t}+\frac{1}{2\alpha}\right)\int_0^{\sqrt{t}-\sqrt{v_1}} e^{-8\alpha s}\left(\frac{1}{t}+\frac{2s}{t^{1.5}}+O(s^2)\right) ds \\
&= \frac{4}{\alpha}\int_{v_1}^{1} dt \left(\sqrt{t}+\frac{1}{4\alpha}\right)\left(\frac{1}{8\alpha t}+\frac{2}{(8\alpha)^2 t^{1.5}}+O(\alpha^{-3})\right) \\
&= \frac{1}{2\alpha^2}\int_{v_1}^{1} dt \left(\frac{1}{\sqrt{t}}+\frac{1}{2\alpha t}+O(\alpha^{-2})\right). \qquad \text{(A.66)}
\end{aligned}
$$

The third and fourth terms therefore cancel to order α^{-3}. This means that

$$
\int_{v_1}^{1} dt \int_{v_1}^{1} dt' \left\langle \hat{\varphi}_t^{(0)} \hat{\varphi}_{t'}^{(1)} \right\rangle \approx \frac{v_1}{4\alpha^2}\left\langle \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} \right\rangle - \frac{\sqrt{v_1}}{2\alpha^2}\left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle. \qquad \text{(A.67)}
$$

Therefore we find that Eq. (A.49) simplifies to

$$
\left\langle \phi_{\rm II}^2 \right\rangle \approx \frac{1}{4\alpha^2} + \frac{v_1}{4\alpha^2}\left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle + \frac{\tau}{\alpha}\left(\frac{v_1}{2}\left\langle \hat{\varphi}_{v_1}^{(0)} \hat{\varphi}_{v_1}^{(1)} \right\rangle - \sqrt{v_1}\left\langle (\hat{\varphi}_{v_1}^{(0)})^2 \right\rangle\right). \qquad \text{(A.68)}
$$

A.4 Derivations for Continuous Measurements

Evaluating Eq. (7.57) of Sec. 7.1.2 for the phase variance under continuous measurements, we find

$$
\begin{aligned}
\left\langle \Theta^2(t) \right\rangle &= 2\chi^3 \left\langle \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} dW(u_1) \int_{-\infty}^{v_2} dW(u_2) e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \right\rangle \\
&+8\chi^3 |\alpha|^2 \kappa^2 \left\langle \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{u_1}^{t} dW'(s_1) \int_{u_2}^{t} dW'(s_2) \right\rangle \\
&= 2\chi^3 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{\min(v_1,v_2)} du\, e^{2|\alpha|\sqrt{2\chi}(2u-v_1-v_2)} \\
&+8\chi^3 |\alpha|^2 \kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{\max(u_1,u_2)}^{t} ds. \qquad \text{(A.69)}
\end{aligned}
$$

Considering the first term first, we have

$$
\begin{aligned}
&2\chi^3 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{\min(v_1,v_2)} du e^{2|\alpha|\sqrt{2\chi}(2u-v_1-v_2)} \\
&= 2\chi^3 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du e^{2|\alpha|\sqrt{2\chi}(2u-v_1-v_2)} \\
&\quad +2\chi^3 \int_{-\infty}^{t} dv_1 \int_{v_1}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du e^{2|\alpha|\sqrt{2\chi}(2u-v_1-v_2)} \\
&= 4\chi^3 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du e^{2|\alpha|\sqrt{2\chi}(2u-v_1-v_2)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\chi^3}{|\alpha|\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} e^{2|\alpha|\sqrt{2\chi}(v_2-v_1)} \\
&= \frac{\chi^3}{|\alpha|\sqrt{2\chi}} \int_{-\infty}^{t} \frac{e^{2\chi(v_1-t)}}{\chi + 2|\alpha|\sqrt{2\chi}} dv_1 \\
&= \frac{\chi^3}{|\alpha|(2\chi)^{3/2}(\chi + 2|\alpha|\sqrt{2\chi})} \\
&\approx \frac{\chi}{8|\alpha|^2}.
\end{aligned} \tag{A.70}
$$

For the approximation in the last line, it has been assumed that $|\alpha| \gg \sqrt{\chi}$.

Now considering the second term in (A.69), we obtain

$$
\begin{aligned}
&8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{\max(u_1,u_2)}^{t} ds \\
&= 8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{\min(v_2,u_1)} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{u_1}^{t} ds \\
&+8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{\min(v_1,v_2)} du_1 \int_{u_1}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)} \int_{u_2}^{t} ds \\
&= 8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{u_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_1) \quad \text{(a)} \\
&+8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{u_1}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{u_1} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_1) \quad \text{(b)} \\
&+8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_1 \int_{u_1}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_2) \quad \text{(c)} \\
&+8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{v_1}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_1} du_1 \int_{u_1}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_2) \quad \text{(d)}
\end{aligned} \tag{A.71}
$$

Due to the difficult nature of the integration limits, we have obtained four terms. Evaluating each of these terms in turn, we firstly find for (a)

$$
\begin{aligned}
&8\chi^3|\alpha|^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{u_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_1) \\
&= \frac{8\chi^3|\alpha|^2\kappa^2}{2|\alpha|\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{-\infty}^{u_1} dv_2 e^{\chi(v_1+v_2-2t)} e^{2|\alpha|\sqrt{2\chi}(u_1-v_1)}(t-u_1) \\
&= \frac{4\chi^2|\alpha|\kappa^2}{\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 e^{\chi(v_1+u_1-2t)} e^{2|\alpha|\sqrt{2\chi}(u_1-v_1)}(t-u_1)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{4\chi^2 |\alpha| \kappa^2}{\sqrt{2\chi}\left(\chi + 2|\alpha|\sqrt{2\chi}\right)} \int_{-\infty}^{t} dv_1 e^{2\chi(v_1-t)} \left(t + \frac{1}{\chi + 2|\alpha|\sqrt{2\chi}} - v_1\right) \\
&= \frac{2\chi |\alpha| \kappa^2}{\sqrt{2\chi}\left(\chi + 2|\alpha|\sqrt{2\chi}\right)} \left(\frac{1}{\chi + 2|\alpha|\sqrt{2\chi}} + \frac{1}{2\chi}\right) \\
&\approx \frac{\kappa^2}{4\chi}.
\end{aligned} \tag{A.72}
$$

The approximation in the last line is accurate in the limit $|\alpha| \gg \sqrt{\chi}$. Turning to the next term, (b), we obtain

$$
\begin{aligned}
&8\chi^3 |\alpha|^2 \kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{u_1}^{t} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{u_1} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_1) \\
&= \frac{8\chi^3 |\alpha|^2 \kappa^2}{2|\alpha|\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 \int_{u_1}^{t} dv_2 e^{\chi(v_1+v_2-2t)} e^{2|\alpha|\sqrt{2\chi}(2u_1-v_1-v_2)}(t-u_1) \\
&= \frac{4\chi^3 |\alpha| \kappa^2}{\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} du_1 e^{\chi(v_1-2t)} e^{2|\alpha|\sqrt{2\chi}(2u_1-v_1)} \frac{\left(e^{(\chi-2|\alpha|\sqrt{2\chi})t} - e^{(\chi-2|\alpha|\sqrt{2\chi})u_1}\right)}{\chi - 2|\alpha|\sqrt{2\chi}}(t-u_1) \\
&= \frac{\chi^3\kappa^2}{2\chi(\chi - 2|\alpha|\sqrt{2\chi})} \int_{-\infty}^{t} dv_1 e^{(\chi+2|\alpha|\sqrt{2\chi})(v_1-t)} \left(t - v_1 + \frac{1}{4|\alpha|\sqrt{2\chi}}\right) \\
&- \frac{4\chi^3 |\alpha| \kappa^2}{\sqrt{2\chi}(\chi^2 - 8|\alpha|^2\chi)} \int_{-\infty}^{t} dv_1 e^{2\chi(v_1-t)} \left(t - v_1 + \frac{1}{\chi + 2|\alpha|\sqrt{2\chi}}\right) \\
&= \frac{\chi^3\kappa^2}{2\chi(\chi^2 - 8|\alpha|^2\chi)} \left(\frac{1}{4|\alpha|\sqrt{2\chi}} + \frac{1}{\chi + 2|\alpha|\sqrt{2\chi}}\right) \\
&- \frac{2\chi^2 |\alpha| \kappa^2}{\sqrt{2\chi}(\chi^2 - 8|\alpha|^2\chi)} \left(\frac{1}{\chi + 2|\alpha|\sqrt{2\chi}} + \frac{1}{2\chi}\right) \\
&\approx \frac{\kappa^2}{8|\alpha|\sqrt{2\chi}}.
\end{aligned} \tag{A.73}
$$

Again this is accurate in the limit $|\alpha| \gg \sqrt{\chi}$. This term is higher order than the result obtained for (a), and so can be ignored.

Now looking at the third term (c),

$$
\begin{aligned}
&8\chi^3 |\alpha|^2 \kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_1 \int_{u_1}^{v_2} du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_2) \\
&= \frac{4\chi^3 |\alpha| \kappa^2}{\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_1 e^{2|\alpha|\sqrt{2\chi}(u_1-v_1)} \left(t - v_2 + \frac{1}{2|\alpha|\sqrt{2\chi}}\right) \\
&- \frac{4\chi^3 |\alpha| \kappa^2}{\sqrt{2\chi}} \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} \int_{-\infty}^{v_2} du_1 e^{2|\alpha|\sqrt{2\chi}(2u_1-v_1-v_2)} \left(t - u_1 + \frac{1}{2|\alpha|\sqrt{2\chi}}\right) \\
&= \chi^2\kappa^2 \int_{-\infty}^{t} dv_1 \int_{-\infty}^{v_1} dv_2 e^{\chi(v_1+v_2-2t)} e^{2|\alpha|\sqrt{2\chi}(v_2-v_1)} \left(t - v_2 + \frac{1}{2|\alpha|\sqrt{2\chi}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{\chi^2\kappa^2}{2}\int_{-\infty}^{t}dv_1\int_{-\infty}^{v_1}dv_2 e^{\chi(v_1+v_2-2t)}e^{2|\alpha|\sqrt{2\chi}(v_2-v_1)}\left(t-v_2+\frac{3}{4|\alpha|\sqrt{2\chi}}\right)\\
&=\frac{\chi^2\kappa^2}{2}\int_{-\infty}^{t}dv_1\int_{-\infty}^{v_1}dv_2 e^{\chi(v_1+v_2-2t)}e^{2|\alpha|\sqrt{2\chi}(v_2-v_1)}\left(t-v_2+\frac{1}{4|\alpha|\sqrt{2\chi}}\right)\\
&=\frac{\chi^2\kappa^2}{2\left(\chi+2|\alpha|\sqrt{2\chi}\right)}\int_{-\infty}^{t}dv_1 e^{2\chi(v_1-t)}\left(t-v_1+\frac{1}{4|\alpha|\sqrt{2\chi}}+\frac{1}{\chi+2|\alpha|\sqrt{2\chi}}\right)\\
&=\frac{\chi\kappa^2}{4\left(\chi+2|\alpha|\sqrt{2\chi}\right)}\left(\frac{1}{4|\alpha|\sqrt{2\chi}}+\frac{1}{\chi+2|\alpha|\sqrt{2\chi}}+\frac{1}{2\chi}\right)\\
&\approx\frac{\kappa^2}{16|\alpha|\sqrt{2\chi}}.
\end{aligned}
\tag{A.74}
$$

This is again higher order than term (a), and so can be ignored. Now looking at the final term, (d), we find

$$
\begin{aligned}
&8\chi^3|\alpha|^2\kappa^2\int_{-\infty}^{t}dv_1\int_{v_1}^{t}dv_2 e^{\chi(v_1+v_2-2t)}\int_{-\infty}^{v_1}du_1\int_{u_1}^{v_2}du_2 e^{2|\alpha|\sqrt{2\chi}(u_1+u_2-v_1-v_2)}(t-u_2)\\
&=\frac{4\chi^3|\alpha|\kappa^2}{\sqrt{2\chi}}\int_{-\infty}^{t}dv_1\int_{v_1}^{t}dv_2 e^{\chi(v_1+v_2-2t)}\int_{-\infty}^{v_1}du_1 e^{2|\alpha|\sqrt{2\chi}(u_1-v_1)}\left(t-v_2+\frac{1}{2|\alpha|\sqrt{2\chi}}\right)\\
&-\frac{4\chi^3|\alpha|\kappa^2}{\sqrt{2\chi}}\int_{-\infty}^{t}dv_1\int_{v_1}^{t}dv_2 e^{\chi(v_1+v_2-2t)}\int_{-\infty}^{v_1}du_1 e^{2|\alpha|\sqrt{2\chi}(2u_1-v_1-v_2)}\left(t-u_1+\frac{1}{2|\alpha|\sqrt{2\chi}}\right)\\
&=\chi^2\kappa^2\int_{-\infty}^{t}dv_1\int_{v_1}^{t}dv_2 e^{\chi(v_1+v_2-2t)}\left(t-v_2+\frac{1}{2|\alpha|\sqrt{2\chi}}\right)\\
&-\frac{\chi^2\kappa^2}{2}\int_{-\infty}^{t}dv_1\int_{v_1}^{t}dv_2 e^{\chi(v_1+v_2-2t)}e^{2|\alpha|\sqrt{2\chi}(v_1-v_2)}\left(t-v_1+\frac{3}{4|\alpha|\sqrt{2\chi}}\right)\\
&=\chi\kappa^2\int_{-\infty}^{t}dv_1 e^{\chi(v_1-t)}\left(\frac{1}{2|\alpha|\sqrt{2\chi}}+\frac{1}{\chi}\right)-\chi\kappa^2\int_{-\infty}^{t}dv_1 e^{2\chi(v_1-t)}\left(t-v_1+\frac{1}{2|\alpha|\sqrt{2\chi}}+\frac{1}{\chi}\right)\\
&-\frac{\chi^2\kappa^2}{2\left(\chi-2|\alpha|\sqrt{2\chi}\right)}\int_{-\infty}^{t}dv_1 e^{\left(\chi+2|\alpha|\sqrt{2\chi}\right)(v_1-t)}\left(t-v_1+\frac{3}{4|\alpha|\sqrt{2\chi}}\right)\\
&+\frac{\chi^2\kappa^2}{2\left(\chi-2|\alpha|\sqrt{2\chi}\right)}\int_{-\infty}^{t}dv_1 e^{2\chi(v_2-t)}\left(t-v_1+\frac{3}{4|\alpha|\sqrt{2\chi}}\right)\\
&=\frac{\kappa^2}{4\chi}+\frac{\kappa^2}{4|\alpha|\sqrt{2\chi}}-\frac{\chi\kappa^2}{2(\chi-8|\alpha|^2)}\left(\frac{3}{4|\alpha|\sqrt{2\chi}}+\frac{1}{\chi+2|\alpha|\sqrt{2\chi}}\right)\\
&+\frac{\chi\kappa^2}{4\left(\chi-2|\alpha|\sqrt{2\chi}\right)}\left(\frac{3}{4|\alpha|\sqrt{2\chi}}+\frac{1}{2\chi}\right).
\end{aligned}
\tag{A.75}
$$

For $|\alpha|\gg\sqrt{\chi}$, the only significant term here is

$$
\frac{\kappa^2}{4\chi}. \tag{A.76}
$$

Using this, together with the results from Eqs (A.70) and (A.72), the total phase variance is

$$\left\langle \Theta^2(t) \right\rangle \approx \frac{\chi}{8\left|\alpha\right|^2} + \frac{\kappa^2}{2\chi}. \tag{A.77}$$

Bibliography

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

[2] F. London, "Über die Jacobischen Transformationen der Quantenmechanik" Z. Phys. **37**, 915 (1926).

[3] F. London, "Winkelvariable und kanonische Transformationen in der Undulationmechanik" Z. Phys. **40**, 193 (1927).

[4] P. A. M. Dirac, "The quantum theory of the emission and absorption of radiation" Proc. R. Soc. London Ser. A **114**, 243 (1927).

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

[6] W. H. Louisell, "Amplitude and phase uncertainty relations" Phys. Lett. **7**, 60 (1963).

[7] L. Susskind and J. Glogower, "Quantum mechanical phase and time operator" Physics **1**, 49 (1964).

[8] P. Carruthers and M. M. Nieto, "Phase angle variables in quantum mechanics" Rev. Mod. Phys **40**, 411 (1968).

[9] D. T. Pegg and S. M. Barnett, "Unitary phase operator in quantum mechanics" Europhys. Lett. **6**, 483 (1988).

[10] S. M. Barnett and D. T. Pegg, "On the Hermitian optical phase operator" J. Mod. Opt. **36**, 7 (1989).

[11] D. T. Pegg and S. M. Barnett, "Phase properties of the quantized single-mode electromagnetic field" Phys. Rev. A **39**, 1665 (1989).

[12] J. C. Garrison and J. Wong, "Canonically conjugate pairs, uncertainty relations, and phase operators" J. Math. Phys. **11**, 2242 (1970).

[13] V. N. Popov and V. S. Yarunin, "On the problem of phase operator of linear harmonic oscillator" Vest. Leningrad. Univ. Ser. Fiz., No. 22, 7 (1973).

[14] V. N. Popov and V. S. Yarunin, "Quantum and quasi-classical states of the photon phase operator" J. Mod. Opt. **39**, 1525 (1992).

[15] J. A. Vaccaro and D. T. Pegg, "Consistency of quantum descriptions of phase" Phys. Scr. **T48**, 22 (1993).

[16] S. M. Barnett and D. T. Pegg, "Limiting procedures for the optical phase operator" J. Mod. Opt. **39**, 2121 (1992).

[17] T. Gantsog, A. Miranowicz, and R. Tanas, "Phase properties of real field states: The Garrison-Wong versus Pegg-Barnett predictions" Phys. Rev. A **46**, 2870 (1992).

[18] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).

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

[20] J. H. Shapiro, S. R. Shepard, and N. C. Wong, "Ultimate quantum limits on phase measurement" Phys. Rev. Lett. **62**, 2377 (1989).

[21] J. H. Shapiro and S. R. Shepard, "Quantum phase measurement: A system-theory perspective" Phys. Rev. A **43**, 3795 (1991).

[22] U. Leonhardt, J. A. Vaccaro, B. Böhmer, and H. Paul, "Canonical and measured phase distributions" Phys. Rev. A **51**, 84 (1995).

[23] M. J. W. Hall, "The quantum description of optical phase" Quantum Opt. **3**, 7 (1991).

[24] S. M. Barnett and D. T. Pegg, "Phase in quantum optics" J. Phys. A **19**, 3849 (1986).

[25] J. W. Noh, A. Fougères, and L. Mandel, "Measurement of the quantum phase by photon counting" Phys. Rev. Lett. **67**, 1426 (1991).

[26] J. W. Noh, A. Fougères, and L. Mandel, "Operational approach to the phase of a quantum field" Phys. Rev. A **45**, 424 (1992).

[27] A. Fougères, J. W. Noh, T. P. Grayson, and L. Mandel, "Measurement of phase differences between two partially coherent fields" Phys. Rev. A **49**, 530 (1994).

[28] J. W. Noh, A. Fougères, and L. Mandel, "Measurements of the probability distribution of the operationally defined quantum phase difference" Phys. Rev. Lett. **71**, 2579 (1993).

[29] M. G. Raymer, J. Cooper, and M. Beck, "Many-port homodyne detection of an optical phase" Phys. Rev. A **48**, 4617 (1993).

[30] D. T. Pegg and S. M. Barnett, "Tutorial review: Quantum optical phase" J. Mod. Opt. **44**, 225 (1997).

[31] M. M. Nieto, "Quantum phase and quantum phase operators: Some physics and some history" Phys. Scr. **T48**, 5 (1993).

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

[33] H. M. Wiseman, "Quantum trajectories and quantum measurement theory" Quantum Semiclass. Opt. **8**, 205 (1996).

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

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

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

[37] G. S. Summy and D. T. Pegg, "Phase optimized quantum states of light" Opt. Comm. **77**, 75 (1990).

[38] R. S. Bondurant and J. H. Shapiro, "Squeezed states in phase-sensing interferometers" Phys. Rev. D **30**, 2548 (1984).

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

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

[41] B. C. Sanders and G. J. Milburn, "Optimal quantum measurements for phase estimation" Phys. Rev. Lett. **75**, 2944 (1995).

[42] B. C. Sanders, G. J. Milburn, and Z. Zhang, "Optimal quantum measurements for phase-shift estimation in optical interferometry" J. Mod. Opt. **44**, 1309 (1997).

[43] E. Polzik, J. Carri, and H. J. Kimble, "Spectroscopy with squeezed light" Phys. Rev. Lett. **68**, 3020 (1992).

[44] S. Takeuchi, J. Kim, Y. Yamamoto, and H. H. Hogue, "Development or a high-quantum-efficiency single-photon counting system" Appl. Phys. Lett. **74**, 1063 (1999).

[45] M. J. Collett, "Phase noise in a squeezed state" Phys. Scr. **T48**, 124 (1993).

[46] A. Erdélyi (editor), *Higher Transcendental Functions* (McGraw-Hill, New York, 1953).

[47] D. W. Berry, H. M. Wiseman, and Zhong-Xi Zhang, "Heterodyne and adaptive phase measurements on states of fixed mean photon number" Phys. Rev. A **60**, 2458 (1999).

[48] G. M. D'Ariano and M. G. A. Paris, "Lower bounds on phase sensitivity in ideal and feasible measurements" Phys. Rev. A **49**, 3022 (1994).

[49] H. P. Yuen, "Two-photon coherent states of the radiation field" Phys. Rev. A **13**, 2226 (1976).

[50] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer, Berlin, 1994).

[51] D. W. Berry and H. M. Wiseman, "Phase measurements at the theoretical limit" Phys. Rev. A **63**, 013813 (2000).

[52] M. Rigo, F. Mota-Furtado, and P. F. O'Mahony, "Continuous stochastic Schrödinger equations and localization" J. Phys. A: Math. Gen. **30**, 7557 (1997).

[53] D. W. Berry and H. M. Wiseman, "The effects of time delays in adaptive phase measurements" J. Mod. Opt. **48**, 797 (2001).

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

[55] M. J. W. Hall, "Quantum properties of classical Fisher information" Phys. Rev. A **62**, 012107 (2000).

[56] C. E. Weatherburn, *Mathematical Statistics* (Cambridge University Press, Cambridge, 1946).

[57] I. Bialynicki-Birula, M. Freyberger, and W. Schleich, "Various measures of quantum phase uncertainty: A comparative study" Phys. Scr. **T48**, 113 (1993).

[58] Z. Hradil, R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch, "Quantum phase in interferometry" Phys. Rev. Lett. **76**, 4295 (1996).

[59] J. H. Shapiro, "Phase conjugate quantum communication with zero error probability at finite average photon number" Phys. Scr. **T48**, 105 (1993).

[60] B. C. Sanders, H. de Guise, D. J. Rowe, and A. Mann, "Vector phase measurement in multipath quantum interferometry" J. Phys A: Math. Gen. **32**, 7791 (1999).