Multi-parameter Optical Metrology: Quantum and Classical

Walker Larson  
*University of Central Florida*

Find similar works at: [https://stars.library.ucf.edu/etd2020](https://stars.library.ucf.edu/etd2020)  
University of Central Florida Libraries [http://library.ucf.edu](http://library.ucf.edu)

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

**STARS Citation**

[https://stars.library.ucf.edu/etd2020/84](https://stars.library.ucf.edu/etd2020/84)
PARTIAL COHERENCE AND ANCILLARY DEGREES OF FREEDOM IN CLASSICAL AND QUANTUM OPTICAL METROLOGY

by

WALKER DOUGLAS LARSON
B.A. Physics, Whitman College, 2014
M.S. Optics, University of Central Florida, 2020

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the College of Optics & Photonics at the University of Central Florida Orlando, Florida

Spring Term
2020

Major Professor: Bahaa E.A. Saleh
The insights offered by quantum mechanics to the field of optical metrology are many-fold, with non-classical states of light themselves used to make sensors that surpass the sensitivity of sensors using classical states of light. Unfortunately, this advantage, referred to often as "super-sensitivity" is notoriously fragile, and even the slightest experimental imperfections may greatly reduce the efficacy of the non-classical sensors, sometimes completely obviating their advantage. In my thesis I have shown that the performance of an otherwise ideal two-photon interferometer, which exploits entanglement between photons to make super-sensitive measurements of phase, is crippled by the slightest introduction of decoherence between modes of the interferometer. I have shown further that such drastic reduction in sensitivity can also appear in classical measurement problems, specifically that the recently developed methods of estimating the separation between a pair of point sources are rendered less effective when the ideal assumption of complete spatial incoherence is relaxed. Towards overcoming these and other issues, I have designed new configurations that use ancillary optical degrees of freedom, a tool-set that has recently garnered much interest in the field of quantum optics. In the context of two-photon interferometry, I have shown that by coupling polarization to the spatial-structure of the two photon state used to probe phase it is possible to obviate the need for a reference phase, even in the context of decoherence and imperfections in the interferometer. In the context of two-point resolution, I have developed an anisotropic imaging system that performs the function of an image-inversion interferometer and is inherently stable, offering an attractive implementation of recently developed methods of sub-Rayleigh imaging. I have further shown both theoretically and experimentally that the same anisotropic image-inversion interferometer is useful in measuring spatially encoded phases, both in the context of classical illumination as well as quantum-aided two-photon
super-sensing. In both cases, the ability to perform interferometric measurements of the spatial structure of an electric field without splitting beam paths forms a bridge between conception and implementation of precision-sensing measurement strategies. Finally, I have shown that binary interferometric method based on the common-path anisotropic imaging system that I introduced, are able to measure both phase gradients and transverse beam tilts with a sensitivity beating conventional systems that are used both commercially and in research laboratories.
Dedicated to Sarah, Douglas, Charlotte and Cristela, the support of whom has made this possible.
I would like to first thank my advisor, Dr. Bahaa Saleh, for all of the guidance, advice, and help he has offered me throughout my years hear. Without his support, I would never have become the scientist I am today, and would have a much less interesting body of work to present to readers here. Thank you as well to my committee members.

I would like to thank all friends I have made at CREOL, who have elevated my time as a graduate student from being something deeply satisfying to being something that I could not imagine even wanting to pursue any alternatives to. Without this community as the backbone of day-to-day life, I would not have been able to accomplish what I have. Special thanks to the administrative staff, without whom I surely would not have filed the proper paperwork needed to graduate on time. Special thanks as well to Dr. Ayman Abouraddy, whose early guidance in my career was extremely formative. Finally, thanks to Dr. Mark Beck of Reed College for convincing me to pursue graduate studies and providing me with the experience to do so successfully.

I would also like to acknowledge the financial support provided by the Air Force Research Laboratory grant FA8651-19-2-0001, which funded some of the work presented in this thesis.
TABLE OF CONTENTS

LIST OF FIGURES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . x

CHAPTER 1: INTRODUCTION . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

CHAPTER 2: THEORY . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
  2.1 Phase in Optics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
  2.2 Classical Parameter Estimation . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.3 Quantum Mechanical State Representation . . . . . . . . . . . . . . . . . . 11
  2.4 Quantum Parameter Estimation . . . . . . . . . . . . . . . . . . . . . . . . 18
  2.5 Multi-Parameter Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . 20

CHAPTER 3: RESURGENCE OF RAYLEIGH’S CURSE IN THE PRESENCE OF
PARTIAL COHERENCE . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
  3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
  3.2 Classical and Quantum Models . . . . . . . . . . . . . . . . . . . . . . . . . 27
  3.3 Quantum Fisher Information Resolution . . . . . . . . . . . . . . . . . . . . 34
  3.4 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
CHAPTER 4: SUPER-SENSITIVE ANCILLA-BASED ADAPTIVE PHASE ESTIMATION

4.1 Introduction ..................................................... 46

4.2 Effect of Decoherence on Phase Sensitivity .......................... 49

4.3 Ancilla Fortification ............................................. 55

4.4 Adaptive Phase Estimation .................................... 59

4.5 Discussion ....................................................... 62

CHAPTER 5: ANCILLARY POLARIZATION USE IN A COLLINEAR IMAGE INVERSION INTERFEROMETER .............................................. 64

5.1 Introduction ..................................................... 64

5.2 Theory ........................................................ 67

5.3 The common-path interferometer ................................. 74

5.4 Experimental verification ....................................... 76

5.5 Discussion ....................................................... 81

CHAPTER 6: Quantum-limited estimation of the phase gradient ............ 83

6.1 Introduction ..................................................... 83

6.2 Quantum Fisher information for single- and N-photon states ........ 85

viii
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Mach-Zehnder interferometer</td>
<td>7</td>
</tr>
<tr>
<td>3.1</td>
<td>Two-point direct-measurement Fisher information</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Concurrent two-point direct-measurement Fisher information</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Quantum Fisher information for two-point resolution</td>
<td>40</td>
</tr>
<tr>
<td>3.4</td>
<td>Quantum Fisher information for two-point resolution</td>
<td>41</td>
</tr>
<tr>
<td>3.5</td>
<td>Two-point Fisher information improvement</td>
<td>42</td>
</tr>
<tr>
<td>3.6</td>
<td>Concurrent two-point Fisher information improvement</td>
<td>44</td>
</tr>
<tr>
<td>4.1</td>
<td>Mach-Zehnder Fisher information</td>
<td>54</td>
</tr>
<tr>
<td>4.2</td>
<td>Ancilla fortified Mach-Zehnder setup</td>
<td>56</td>
</tr>
<tr>
<td>4.3</td>
<td>Fisher information for ancilla fortified interferometer</td>
<td>59</td>
</tr>
<tr>
<td>4.4</td>
<td>Adaptive estimation bias</td>
<td>61</td>
</tr>
<tr>
<td>4.5</td>
<td>Adaptive estimation variance</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Conventional image inversion interferometer</td>
<td>69</td>
</tr>
<tr>
<td>5.2</td>
<td>Interferometer impulse response function</td>
<td>72</td>
</tr>
<tr>
<td>5.3</td>
<td>Image inversion interferometer diagram</td>
<td>75</td>
</tr>
</tbody>
</table>
Figure 5.4: Binary phase measurement ........................................... 79
Figure 5.5: OAM measurement ..................................................... 81
Figure 6.1: Phase gradient Fisher information .................................. 89
Figure 6.2: Phase measurement via an image inverting interferometer ....... 90
Figure 6.3: Phase gradient Fisher information .................................. 94
Figure 6.4: Multi-photon image inverting interferometer ....................... 95
Figure 7.1: Two-photon anisotropic image inversion measurement ........... 105
CHAPTER 1: INTRODUCTION

The insights offered by quantum quantum optics to the field of optical metrology is two-fold: first, the mathematical formalism developed to treat non-classical optical problems has led way to numerous developments in the field of classical optical metrology, and second, non-classical states of light themselves can be used to make sensors that surpass the sensitivity sensors using classical states of light. My thesis focuses on both insights, addressing modern measurement problems with theoretical contributions as well as designing novel implementations of sensors that are aided by the application of non-classical states of light.

My contributions to the field have revolved around the multi-parameter treatment of quantum optical problems. On the less optimistic side, I have shown that measurement strategies must be reconsidered when realistic measurement conditions, where ideal operation parameters will necessarily have small perturbations, are introduced. On the more optimistic side, I show how careful control of ancillary optical degrees of freedom otherwise unused in sensors can offer significant advantages in practical applications, often obviating obstacles that were previously thought to be fundamental.

Optical metrology is the study of using or measuring optical fields to estimate the value of a parameter that characterizes a physical system. From interferometers measuring phase shifts to telescopes measuring the distance between distant stars, the optical field is richly encoded with information about the physical system that it probes or from which it originates. Nonetheless, random fluctuations and noise in the optical probe, system of interest, and measurement apparatus introduce inaccuracies in the measurement that necessitate the use of statistical techniques of parameter estimation to process the data. When all classical sources of noise are eliminated, the unavoidable quantum uncertainty of the optical field re-
mains. Although this may have once seemed to be the ultimate ceiling of precision sensors, innovations provided by using non-classical states of light have been shown to reduce this uncertainty even farther.

A massive body of work has been done to show that the use of optical fields in nonclassical states, including sub-Poissonian, quadrature squeezed, and entangled photon-number states, can offer an advantage in parameter estimation. While these innovations are already being used in commercial applications, the advantage gained by quantum states is notoriously fragile; just as quantum states are more sensitive to the parameters we wish to measure, the advantage afforded to them over classical states is especially hurt by decoherence, loss, experimental imperfections, or any other effect that modifies the otherwise ideal measurement system.

In my work, I focus on the affect that partial coherence has on optical measurement systems: both classical and quantum. Partial coherence, a feature of any real optical system, can be damaging in many ways. For interferometers, the slightest introduction of decoherence from an otherwise completely coherent system can vastly reduce the phase sensitivity of the system. For incoherent imaging systems, the slightest introduction of coherence can drastically stifle the precision of phase estimates inferred from measuring the power distribution at the output of the interferometer. Accepting these obstacles as facts of life, I have investigated ways that ancillary optical resources can help in these metrological scenarios.

The formalism of ancillary resources is relatively new, and is sometimes referred to as "classical entanglement." Just as quantum optics seeks to correlate, or entangle, the state of multiple photons, my approach here is to create optical states — classical or non-classical — where the state of different degrees of freedom are correlated. Both photons and classical electromagnetic fields carry multiple degrees of freedom, be it polarization, frequency, spatial
mode, propagation mode, to name those commonly employed. Metrological implementations typically assume photons to be in plane-wave linearly polarized modes of the electric field. Spurred in part by the mathematical development of quantum information technology, the notion of exploiting for metrological purposes correlations between individual degrees of freedom of an optical beam – for example, polarization and spatial distribution – has recently generated great interest. My research has followed this lead.

The use of ancillary quantum resources, an insight that has been addressed largely in the context of Quantum Computation, has recently found application in the field of quantum-enhanced metrology. As a simple example of a computational application of ancillary quantum resources, consider the task of error correction. The most basic application of a quantum error-correcting-code is to encode a logical qubit on multiple physical qubits, rather than a single system. This is possible whether the physical qubits are encoded on photon polarizations, electron spins, or otherwise.

Looking specifically at the physical platform of optics, it turns out that a single physical system -a photon- can carry multiple pieces of information. Logical bits can be encoded on any of the physical degrees of a photon: polarization, spatial mode, wavelength, propagation direction. Moreover, bits encoded on these degrees of freedom can be engineered to be correlated, completely independent, or somewhere in-between. The characterization of the coherence shared between optical DoFs has been the topic of much recent work [1, 2, 3].

Since our logical qubit now exists on many systems, if a noisy channel causes an error in one of the physical systems comprising the qubit, it is possible to both detect and correct for this error by manipulating and measuring the other physical qubits.

It is simple to see that this may also find application in metrology, as similar procedures could be used to infer deleterious effects such as depolarization or photon-loss within an
interferometer, and a number of papers have been published addressing quantum-enhanced sensitivity when ancillary resources are allowed. In the context of metrology, however, it is almost always better to devote every physical system used for estimation to probe the parameter to be estimated, as this enhances the benefit of using quantum resources in the first place. This begs the question then, as to what use ancillary resources can offer in the context of metrology.

There are a number of ways that utilizing extra DoFs has been shown to be useful in the paradigms of both classical and quantum metrology. [4] The context when this arises usually comes when probe states exhibit completely-correlated joint statistics between different DoFs, similar to the statistics shared between entangled photons in other contexts. Consider the example of metrological applications of optical states containing orbital-angular-momentum (OAM): it is well known that beams with high-OAM are extremely sensitive to rotations in reference frame, but the holographic measurements required to extract and infer the information gained by the probes are involved and often infeasible. It was shown by [5] that the use of vector-beams, where polarization is coupled to the spatial distribution of an optical beam, can overcome this measurement requirement. Vector beams have incoherent statistics when looking at a single DoF alone, but are completely coherent when looking at both DoFs simultaneously. It was shown that by creating these vector beams using liquid crystal optics, it was possible to infer the information contained in the OAM DoF by making a simple polarization projection measurement in the rotated reference frame. The experimenters showed likewise that this procedure can be used in concert with rotation estimation that is already benefiting from a quantum enhancement, and serves as a strong example of the utility of ancillary DoFs in both classical and quantum metrology. Other authors have shown similar utility on other contexts as well [6, 7, 8].

Developments in the engineering of correlations between optical degrees of freedom is further
made crucial by their applications in communication and computation. Since extra informa-
tion can be multiplexed into extra degrees of freedom, utility has been found in both these
fields. As such, recent research has pursued these ends in both the classical and quantum
sensing paradigms. [9, 10].
CHAPTER 2: THEORY

2.1 Phase in Optics

Many problems in optical metrology can be described as a measurement of some phase of an optical field. Consider an electromagnetic field describing a plane-wave expressed mathematically as

\[ \mathcal{E}(t; \omega, \phi) = E_0 e^{i(\omega t - \phi)}, \]  

(2.1)

where \( \omega \) is the oscillation frequency of the field, \( t \) is the time of observation, \( E_0 \) is the real-valued magnitude of the electric field, and \( \phi \) is an unknown phase of the field. Our task then, is to measure the unknown value of \( \phi \).

Suppose we have an identical copy of this field with known phase \( \phi_r \):

\[ \mathcal{E}_r(t; \omega, \phi_r) = E_0 e^{i(\omega t - \phi_r)}. \]  

(2.2)

If we are to combine our original field with this "reference field" we are left with

\[ \mathcal{E}_+(t; \omega, \phi, \phi_r) = \mathcal{E}(t; \omega, \phi) + \mathcal{E}_r(t; \omega, \phi_r) = E_0 e^{i(\omega t - \phi)} + E_0 e^{i(\omega t - \phi_r)}. \]  

(2.3)

We can measure the modulus squared of this combined field using intensity detection. This quantity is given by

\[ |\mathcal{E}_+(t; \omega, \phi, \phi_r)|^2 = 2E_0^2 [1 + \cos(\phi - \phi_r)]. \]  

(2.4)

Hence, if we make a measurement of the intensity, we observe sinusoidal behavior that depends on the difference between the phase \( \phi \) we wish to measure and our known phase
\( \phi_r \). In the ideal case, if we perfectly know \( E_0 \) and \( \phi_r \), then a measurement of the intensity allows us to infer – with certainty – the value of \( \phi \). This procedure is commonly referred to as optical homodyning, and is the crux of many problems in optics.

Figure 2.1: A Mach-Zehnder interferometer uses beam splitters to measure the difference between \( \phi \) and \( \phi_r \), the optical phases induced by transmission through the upper and lower arms of the interferometer.

The most common implementation of this method in the context of metrology comes in the form of an interferometer. First, a beam is split into two possible paths by a beam splitter, with one path probing the known reference phase, and the other path probing the optical phase or object that is to be measured. Examples of phase objects include transparent or reflective objects and images, holograms, isotropic or anisotropic media, or lateral path-length variations, to name a few. After transmission through their respective paths, the beams are recombined at a second beam splitter. Interference between the beams will lead to output fields that are the sum \( \mathcal{E}_+ \) and difference \( \mathcal{E}_- \) of the fields that have passed through the object and reference, allowing for a measurement of the quantity \( \phi - \phi_r \). This geometry can be rearranged, allowing for a wide range of measurement and sensing applications.

2.2 Classical Parameter Estimation

Unfortunately, in real systems, perfect determination of phase is impossible, as all measurement systems will have some degree of noise – both intrinsic and otherwise. Because of this, it is necessary to treat our measurement systems statistically. Fortunately, a large body of
work has been developed to answer the following two questions: What is the best precision a sensor can provide, and what are the best practices of operation of the sensor towards that goal? We begin by applying the Bayesian method of inference to the general problem of parameter estimation.

**Bayesian Inference**

The Bayesian method assumes the phase $\phi$ that we wish to estimate is a random variable distributed according to a prior probability distribution function (PDF) $p(\phi)$, representing whatever information about $\phi$ is available before measurement takes place. In general, however, this method works just as well when we assume that no information exists; that is, a situation where each value of $\phi$ is equally likely. The goal of Bayesian analysis is to choose an estimator $\tilde{\phi}(x)$ that provides the best estimate of $\phi$ available, conditioned on the prior information and measurements that have been obtained. When a measurement - such as the determination of the intensity distribution at the output of an interferometer - is performed, the collected data set $\{x\}$ is dependent on the true value of the variable $\phi$. These measurement results have a conditional probability $p(x|\phi)$ of occurring. Inferring from $\{x\}$, $p(x|\phi)$, and the assumed prior PDF $p(\phi)$, we can find the posterior PDF $p(\phi|x)$ using Bayes theorem, which states

$$p(\phi)p(x|\phi) = p(\phi|x)p(x) = p(x : \phi).$$

(2.5)

Furthermore, we can interpret the joint PDF $p(\tilde{\phi}(x) : \phi) = p(\phi)p(\tilde{\phi}|\phi)$ as the probability that the variable takes the true value $\phi$ and that we subsequently estimate it to have value $\tilde{\phi}$. To calculate this, however, we must pick a method for inferring our estimate from the collected data; thus far, these probabilities have told us nothing about how to infer $\tilde{\phi}(x)$ from the data set $\{x\}$. We could choose any way of inference, but for the task of estimating
a parameter, some inference methods will be subjectively better than others.

To address this, we can define a cost function $C(\tilde{\phi}(x), \phi)$ that quantifies the subjective cost of inferring a value of $\tilde{\phi}$ from the data set $\{x\}$. For an example of how this might change different estimation scenarios, consider the cost function of Seattle Seahawks quarterback Russell Wilson estimating the distance between himself and an intended receiver; comparing this to the cost function of a distant military base estimating the voltage between terminals attached to any large-red-buttons in the oval office shows that there is indeed a degree of subjectivity to how much an estimate "costs." Suspending a discussion of how to pick a cost function, we can look at its statistical mean:

$$\langle C \rangle = \int \int d\phi \, dx \, p(\phi)p(\tilde{\phi}(x)|\phi)C(\tilde{\phi}(x), \phi). \quad (2.6)$$

Here we see that the mean cost is a function of the as-of-yet undetermined joint PDF $p(\tilde{\phi}(x) : \phi) = p(\phi)p(\tilde{\phi}|\phi)$. With this quantity, we can define an optimal inference strategy: for any given cost function $C(\tilde{\phi}(x), \phi)$, we call an estimator $\tilde{\phi}(x)$ that minimizes the mean cost $\langle C \rangle$ an efficient estimator corresponding to that cost function. In the paradigm of phase estimation, we will use the cost function to quantify the distance between $\phi$ and the estimate $\tilde{\phi}$. A common choice of cost function is the mean squared error (MSE), $\Delta^2 \tilde{\phi} = (\tilde{\phi}(x) - \phi)^2$. This leads to

$$\left\langle (\tilde{\phi} - \phi)^2 \right\rangle = \int \int d\phi \, dx \, p(\phi)p(\tilde{\phi}(x)|\phi)(\tilde{\phi}(x) - \phi)^2. \quad (2.7)$$

In this case, we can solve for the minimum MSE estimator $\tilde{\phi}^{MSE}(x)$ by minimizing $\left\langle \Delta^2 \tilde{\phi} \right\rangle$, which leaves us with [11]

$$\tilde{\phi}^{MSE}(\{x\}) = \int d\phi \, p(\phi | \{x\}) \phi. \quad (2.8)$$
which corresponds to the average value of the variable with respect to the posterior PDF.

Along with the MSE of an estimator, it is also important to consider the bias of an estimator. Typically, we require that on average, our estimator provides a 'true' estimate of the parameter that we are trying to estimate, such that \( \langle \hat{\phi} \rangle = \phi \). However, there could also exist estimators for which \( \langle \hat{\phi} \rangle \neq \phi \). If instead, we find that \( \langle \hat{\phi} \rangle = \phi + \phi_b \), we then refer to \( \phi_b = \hat{\phi} - \phi \) as the bias of the estimator \( \hat{\phi} \), given a true value of \( \phi \). Looking at the MSE of a such a biased estimator, we find that

\[
\langle (\hat{\phi} - \phi)^2 \rangle = \langle (\hat{\phi} - \phi_b + \phi_b - \phi)^2 \rangle = \text{Var}(\hat{\phi}) + \phi_b^2
\]

(2.9)

Clearly, any bias will reduce the precision of measurements, a characteristic that would prove undesirable for the purposes investigated here.

**Fisher Information**

Having framed the task of how we will estimate a parameter, we will next address the issue of quantifying the usefulness of our estimators. We have qualified that we wish to pick an unbiased estimator that achieves the minimum MSE, yet we have not described what measurement process will be used to provide information from which such estimates can be inferred. Fortunately, the calculation of the Fisher information lets us quantify the sensitivity that a set of measurement outcomes can provide for estimation purposes. Consider a set \( \{x\} \) of possible measurement outcomes; these outcomes could be anything from the result of a coin toss to a signal measured at one of many photo-detectors placed in an array. If the system that is being measured is parameterized in some way by the parameter \( \phi \) we are trying to estimate, then the probabilities of observing any of the measurement outcomes in the set \( \{x\} \) are conditioned on \( \phi \), leaving us with the set of probabilities \( \{p(x|\phi)\} \) of observing any of the measurement outcomes in \( \{x\} \). From these probabilities we can define the Fisher
information $F(\phi)$ by

$$ F(\phi) = \left\langle \left( \frac{\partial}{\partial \phi} \log p(x|\phi) \right)^2 \right\rangle \quad (2.10) $$

With this metric, we can now define the Cramér-Rao bound, which is the lower limit on the statistical variance of unbiased estimators $\tilde{\phi}$ of

$$ \text{Var}(\phi)_{\text{CRB}} \geq \frac{1}{F(\phi)}. \quad (2.11) $$

This provides a useful way of quantifying the sensitivity of any measurement apparatus itself, allowing us to derive the best-in-principle sensitivity that an unbiased estimator can achieve, regardless of the form of the estimator itself. Throughout this dissertation, I will be using the Cramér-Rao bound, and hence the Fisher information,

**Maximum Likelihood Estimation**

A commonly used method of inference is found in the strategy of maximum likelihood estimation (MLE). In MLE, the quantity of interest after a single experimental observation $x$ is $\mathcal{L}(\phi|x)$, which quantifies the likelihood that a true value of $\phi$ led to experimental observation $x$. As one might guess, this quantity is equal to $p(x|\phi)$, the conditional probability of observing $x$ given a true value of $\phi$.

### 2.3 Quantum Mechanical State Representation

**States**

Pure quantum states can be represented mathematically using vectors existing in linear vector spaces. Of specific interest in quantum optics is the vector space describing the bosonic Fock states, used to represent the presence of a photon. Occupation of a photon is
represented as a creation operator $\hat{a}^\dagger$ acting on the vacuum state.

$$|\psi\rangle = \hat{a}^\dagger |0\rangle = |1_a\rangle$$  \hspace{1cm} (2.12)

The above state indicates the presence of a photon in the optical mode $a$. The specific mode described by $a$ could be a mode of a binary space — a polarization, a path in the arm of an interferometer — or a mode of a continuous space — a frequency, a position along the $x$-axis. In the case of finite vector spaces, we can treat states in these spaces as vectors. Take as two examples a horizontally polarized photon and a diagonally polarized photon, both represented in the basis of horizontal and vertical (HV) polarization:

$$|H\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |D\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$  \hspace{1cm} (2.13)

**Density Operators**

States can also be represented as density operators. The density operator for a given quantum state is represented by

$$\rho = |\psi\rangle \langle \psi|.$$  \hspace{1cm} (2.14)

By definition, density operators are positive (the operators have non-negative, real spectrum of eigenvalues) and have unity trace (a requirement that probabilities add to one).

For finite vector spaces, density operators are represented mathematically as matrices. Density operators allow us to treat probabilistic mixtures of pure quantum states. A state that is found in pure state $\rho_1$ with probability $p_1$ and pure state $\rho_2$ with probability $p_2$ is represented
\[ \rho = p_1 \rho_1 + p_2 \rho_2. \quad (2.15) \]

Note that this state cannot be represented using pure states and their corresponding column vectors alone. States of such a form that cannot be described using pure states alone are said to be mixed states. To this end, it is useful to define a state’s purity \( \gamma \)

\[ \gamma = Tr(\rho^2). \quad (2.16) \]

The purity of a state ranges from a value of 1 (for a pure state) to \( \frac{1}{d} \) (for a maximally mixed state) where \( d \) is the dimension of the density matrix. This can be seen easily for the pure case by noting that all density matrices have unity trace and by exploiting the impotency of pure states: \( (|\psi\rangle \langle \psi|)(|\psi\rangle \langle \psi|) = |\psi\rangle \langle \psi| \). \( \therefore \) \( Tr(\rho_{\text{pure}}^2) = Tr(\rho_{\text{pure}}) = 1. \) Maximally mixed states are defined as the states that have an equal probability of being found in any of the states living in the Hilbert space under question; they are necessarily written in any basis as an identity matrix divided by the dimension of the matrix. This implies that

\[ Tr(\rho^2) = Tr \left( \frac{\hat{1}}{d} \right) = Tr \left( \frac{\hat{1}}{d^2} \right) = \frac{1}{d} \quad (2.17) \]

**Unitary State Evolution**

Pure states evolve by application of unitary operators,

\[ |\psi'\rangle = \hat{U} |\psi\rangle, \quad (2.18) \]
which have the property

\[ UU^\dagger = \hat{1}, \quad (2.19) \]

where \( \hat{1} \) is the identity operator. As one might expect, density operators evolve as

\[ \rho' = U\rho U^\dagger. \quad (2.20) \]

This is obviously the case for pure states, and turns out to be the case for mixed states as well [12]. As a useful example, it is illustrative to look at the unitary transformation representing the action of a wave plate, which acts on the polarization state of a photon.

As an operator, we can use the Pauli operators to represent the unitary operators \( \hat{H} \) and \( \hat{Q} \) used for half-wave plates (HWPs) and quarter-wave plates (QWPs), respectively. They are given by

\[ \hat{\sigma}_x = \cos 2\theta \sigma_z + \sin 2\theta \sigma_x \quad (2.21) \]

and

\[ \hat{Q} = (1 - i \cos^2 \theta) \frac{\hat{I}}{2} + (1 + i \cos^2 \theta) \frac{\sigma_z}{2} + (1 - i) \sin \theta \cos \theta \sigma_x, \quad (2.22) \]

where, in both cases, \( \theta \) is the angle of the optic axis of the wave plates with respect to the horizontal, and \( \hat{I} \) is the identity operator. The Pauli operators can be represented as Jones matrices by

\[
\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}
\]
Hence, the wave plate transformations have Jones matrix representations given by

\[
\hat{H} = \begin{pmatrix}
\cos 2\theta & \sin 2\theta \\
\sin 2\theta & -\cos 2\theta
\end{pmatrix}
\]  \hspace{1cm} (2.23)

and

\[
\hat{Q} = \begin{pmatrix}
\cos^2 \theta + i \sin^2 \theta & (1 - i) \sin \theta \cos \theta \\
(1 - i) \sin \theta \cos \theta & \sin^2 \theta + i \cos^2 \theta
\end{pmatrix},
\]  \hspace{1cm} (2.24)

respectively.

**Multi-Mode States**

To describe a composite quantum system with multiple modes, be it a pair of photons with one optical degree of freedom, or a single photon with multiple degrees of freedom, we will need to deal with the larger-dimensional Hilbert space spanned by the direct product between the sub spaces describing the physics of either mode alone. A pure state describing sub-states \(A\) and \(B\) is written as

\[
|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle,
\]  \hspace{1cm} (2.25)

and the density operator describing sub-states \(A\) and \(B\) is written as

\[
\rho_{AB} = \rho_A \otimes \rho_B.
\]  \hspace{1cm} (2.26)

Product states evolve according to the direct product between the transformations acting
on the individual states

\[ U_{AB} = U_A \otimes U_B. \]  \hfill (2.27)

Quantum Channels

There are many operations that can be carried out on a quantum state that cannot simply be represented using a single unitary operation. A very important set of models that require a more general treatment are many of the models that treat decoherence of quantum states. It is with this in mind that we consider the more general form of quantum operation known as a quantum channel.

A common choice of representation is the operator-sum representation. The action of a quantum operation \( \mathcal{E} \) on a system described by the density operator \( \rho \) is defined as

\[ \mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger, \]  \hfill (2.28)

where \( E_i \) are operation elements.

A commonly used model is the so-called depolarizing channel: acting on a binary quantum state, the operation elements are given by

\begin{align*}
E_1 &= \sqrt{1 - \frac{3p}{4}} I, \\
E_2 &= \sqrt{\frac{p}{4}} \sigma_x, \\
E_3 &= \sqrt{\frac{p}{4}} \sigma_y, \\
E_4 &= \sqrt{\frac{p}{4}} \sigma_z, \end{align*}  \hfill (2.29)
where $\hat{\sigma}_i$ are the Pauli matrices, and $I$ is the identity matrix.

For composite systems, where systems $A$ and $B$ can describe different DoFs of a single photon, or the state of multiple photons, operations can affect individual subsystems. Consider $\rho^{AB} = \rho^A \otimes \rho^B$, with a quantum operation acting on subsystem $A$ only. The operation elements for $\mathcal{E}(\rho^{AB})$ are given by the operation elements for $\mathcal{E}(\rho^A)$:

$$E_i^{AB} = E_i^A \otimes I^B,$$

where superscripts $A$ and $B$ denote transformations acting on respective subsystems $A$ and $B$.

Collective noise is the special case wherein noise acts equally on both subsystems of $\rho^{AB}$:

$$E_i^{AB} = E_i^{\otimes 2}.$$

**Measurement**

Measurement in quantum systems is inherently probabilistic. Hence, we will treat the outputs of a measurement process as the probabilities of observing a given outcome when measuring a state $\rho$. Put formally, the probability that we find our state $\rho$ in the state $|x\rangle$ is given by $P_x = p(x|\rho) = Tr[|x\rangle \langle x| \rho]$. When such a result is observed, the quantum state is left in the state $|x\rangle$: barring any further evolution of the system, it stays in that state for all times after the measurement. In this language, we refer to the operator $\Pi_x = |x\rangle \langle x|$ as the measurement operator corresponding to an observation of $x$.

It is often useful to define a set of all possible measurement outcomes. A typical treatment is to define a positive operator valued measure (POVM): this is a set $\{\Pi_i\}$ such that $\sum_i \Pi_i = I$,
where $I$ is the identity operator. This construction contains the intuitive characteristic that all of the probabilities of measurement outcomes defined by the POVM $P_i = p(\Pi_i|\rho)$ add to one.

### 2.4 Quantum Parameter Estimation

#### Bayesian Analysis

For quantum mechanical systems, we begin with a density matrix $\rho$ and evolve it according to the variable $\phi$ to be estimated: $\rho_{\phi} = U_\phi \rho U_\phi^\dagger$. Furthermore, we must define the way in which we measure the evolved state — this is done using a POVM $P_x$. Each measurement $P_x$ provides a corresponding estimate $\tilde{\phi}(x)$ of $\phi$.

The problem of choosing an optimal operator now includes optimization over both the input state $\rho$ and the measurement operator $P_x$, as well as the estimator $\tilde{\phi}$ itself. Towards this end, we write

$$P_{\tilde{\phi}} = \int dx P_x \delta(\tilde{\phi} - \tilde{\phi}(x)),$$

(2.32)

labeling the measurements with their corresponding estimates. We now have the conditional probability $p(\tilde{\phi}|\phi)$, the probability that measurement $P_{\tilde{\phi}}$ provides data $\{x\}$ given a variable value of $\phi$, is given by

$$p(\tilde{\phi}|\phi) = Tr \left[ U_\phi \rho U_\phi^\dagger P_{\tilde{\phi}} \right].$$

(2.33)

leaving an expected cost function

$$\langle C \rangle = \int \int d\phi \ d\tilde{\phi} \ p(\phi) Tr \left[ U_\phi \rho U_\phi^\dagger P_{\tilde{\phi}} \right] C(\tilde{\phi}, \phi).$$

(2.34)
\( P_\phi \) is constrained by normalization:

\[
\int d\phi P_\phi = I. \tag{2.35}
\]

With this construction, we can apply \( p(\tilde{\phi}|\phi) \) to Eq. (2.10), and calculate the classical Fisher information for the measurement \( P_x \). But, since we have access to the state \( \rho(\phi) \) itself, there is something more powerful we can invoke.

**Quantum Fisher Information**

If we know the state \( \rho \) of a system, we can calculate the quantum Fisher information (QFI) \( F_Q(\phi) \), which generalizes the fisher information and provides the minimum variance of estimators, minimized over all possible measurement probabilities \( p(\phi) \) from which the value of \( \phi \) can be inferred. By this, I mean that the calculation of the QFI give the best possible precision of any measurement strategy, limited only by the inherent indeterminacy of the quantum state itself. The variance bound of this in-principle optimized precision is known as the quantum Cramèr-Rao bound (QCRB):

\[
\text{Var}_Q(\tilde{\phi}) \geq \frac{1}{\sqrt{Q(\phi)}}. \tag{2.36}
\]

Here,

\[
Q(\phi) = Tr(\rho \mathcal{L}^2), \tag{2.37}
\]

where \( \mathcal{L} \) is the symmetric logarithmic derivative (SLD) defined by the relationship

\[
\dot{\rho} = \frac{1}{2}(\mathcal{L}\rho + \rho\mathcal{L}), \tag{2.38}
\]

and \( \rho \) is the density operator describing the output of the system that is sampling the variable \( \phi \), and \( \dot{\rho} \) is its derivative with respect to \( \phi \). To solve for \( \mathcal{L} \), we use the spectral decomposition
of $\rho$:\[
\rho = \sum_i \lambda_i |i\rangle \langle i| ,
\] (2.39)
where $|i\rangle$ are the eigenvectors of $\rho$ with eigenvalues $\lambda_i$. Rearranging we find that
\[
\mathcal{L} = \sum_{i,j} \frac{\langle i| \hat{\rho} |j\rangle}{\lambda_i + \lambda_j} |i\rangle \langle j| .
\] (2.40)

Thus, given an output density operator $\rho$, it is possible to find the minimum estimator variance, optimized over all possible measurements on $\rho$. Furthermore, the measurement operators describing the optimal measurement that achieves this bound is given by the eigenvectors of $\mathcal{L}$.

Not only does the calculation of the quantum Fisher information tell us what the best precision achievable is, it tells us what measurement to use to attain it.

### 2.5 Multi-Parameter Estimation

We have developed a significant amount of artillery in our goal of finding the best sensors. We can calculate the sensitivity of a chosen measurement system, using the classical Fisher information, or we can derive the best possible measurements using the quantum Fisher information. These calculations apply when we have a single unknown parameter of interest, assuming that every other parameter of both the system we are measuring are fixed and known. When there are multiple parameters that are unknown, our task becomes that of a multi-parameter estimation problem: whether or not we intend to estimate every unknown parameter, the imprecision in our knowledge of one parameter will often cause us to be unable to estimate other parameters as precisely. Instead of using the classical and quantum Fisher information measures, we will use the classical (FIM) and quantum (QFIM) Fisher information.
information matrices. In the case of a continuous set of measurement outcomes \( x \in \{x\} \) with \( N \) unknown parameters, the FIM \( F \) has elements

\[
F_{jk} = \int dx \left( \frac{\partial}{\partial j} p(x|j,k) \right) \left( \frac{\partial}{\partial k} p(x|j,k) \right) \frac{1}{p(x|j,k)},
\]

where \( j, k \) are unknown parameters of the system. Hence for an \( N \) parameter estimation problem, \( F \) is an \( N \times N \) matrix. Similarly, the QFIM \( Q \) has elements

\[
Q_{i,j} = \frac{1}{2} \text{Tr} \left[ (L_i L_j + L_j L_i) \hat{\rho} \right],
\]

where \( L_i \) is the symmetric logarithmic derivative (SLD) operator for the parameter \( i \), which is solution to

\[
\frac{1}{2} (\hat{L}_i \hat{\rho} + \hat{\rho} \hat{L}_i) = \frac{\partial}{\partial i} \hat{\rho}
\]

and \( \hat{\rho} \) is the density operator describing the quantum system that encodes the \( N \) parameters.

The multi-parameter version of the Cramér Rao bound sets the bound on the covariance matrix \( C \), whose element \( C_{j,k} \) defines the covariance between estimates of the \( j^{th} \) and \( k^{th} \) unknown parameters of the system. For a measurement \( x \) or a quantum state described by \( \tilde{\rho} \), the Cramér Rao bounds are given by

\[
C^{(x)} \geq \frac{1}{F^{(x)}},
\]

and

\[
C^{(\rho)} \geq \frac{1}{Q^{(\rho)}}.
\]

While the covariances between elements is interesting in many contexts, we will concern ourselves only with the diagonal elements of \( C \), which are themselves \( C_{ii} \) equal to the variances
of each individual parameter $V_i$. Hence the variance of each parameter is bounded by the corresponding diagonal element of the inverse of the FIM or QFIM.

It is elucidating to see the repercussions of the multi-parameter treatment in the case of two parameters $\theta_1$ and $\theta_2$. Suppose, for example, that we want to estimate $\theta_1$, but do not know the value of $\theta_2$, which the evolution of our system relies on. For this case, the precision bound for parameter $\theta_1$ is given by

$$\frac{1}{\text{Var}_Q(\theta_1)} \leq Q_{\theta_1\theta_1} - \frac{Q_{\theta_1\theta_2}^2}{Q_{\theta_2\theta_2}}.$$  \hfill (2.46)

We can see clearly that the maximum attainable precision is reduced by the factor $\frac{Q_{\theta_1\theta_2}^2}{Q_{\theta_2\theta_2}}$, hence any non-diagonal QFIM of 2 parameters will reduce the measurement precision of either of the two parameters.
CHAPTER 3: RESURGENCE OF RAYLEIGH’S CURSE IN
THE PRESENCE OF PARTIAL COHERENCE

The two-point resolution of an optical system is the minimum distance between two point sources that can be estimated with a prescribed precision from measurements in the image plane. When the sources are incoherent, then direct measurements of the optical intensity provides resolution limited by Rayleigh’s curse, i.e., the precision diminishes to zero as the separation is reduced to zero. Using quantum Fisher information bounds on the precision, it was shown recently that estimates based on optimal quantum measurements of the optical field can break Rayleigh’s curse and provide estimates with finite precision even at very small separations. In [13], I showed that if the point sources are partially coherent with an unknown real degree of coherence, no matter how small it is, then the curse resurges. Rayleigh’s curse endures as a fundamental dictum in the face of any finite spatial coherence between sources, and shows that even optimal sensors can fall victim to realistic experimental considerations.

3.1 Introduction

Resolution continues to be a central issue in optical imaging. Earlier notions of resolution were concerned with the minimum separation necessary for two incoherent point sources of equal intensities to be discerned from measurement of the intensity of their image through a diffraction-limited imaging system. Different definitions and criteria were based on various measures of discernibility [14, 15, 16, 17]. Mathematical methods based on analytic continuation were later shown to formally solve the inverse problem associated with diffraction-limited imaging, i.e., restore details finer than the resolution limit [18, 19, 20, 21]. Although such
solutions were deemed to offer superresolution, it was noted that the precision of such solutions deteriorates rapidly in the presence of measurement noise or uncertainty since the inverse problem is ill-posed. More recently, superresolution was demonstrated under different imaging conditions, namely scanning systems with single points or subwavelength areas emitting one at a time [22, 23, 24].

With the theoretical development of statistical signal and image processing tools, image restoration was cast as a formal parameter estimation problem [25, 26]. Two-point resolution was defined as the minimum separation between two point sources that can be estimated from measurement of the optical intensity in the imaging plane with a predefined precision, under the assumption of a given probability model for the measurement noise. The Fisher information (and corresponding Cramér-Rao bound) [27] was used for this purpose [28], and it was shown that the precision of optimal estimators diminishes as the separation is reduced, a problem that has become known as Rayleigh's curse. The realization that reducing uncertainty is a requirement for enhancing precision has also led to investigations of the use of nonclassical states of light, such as squeezed, entangled, and sub-Poisson quantum states [29, 30, 31, 32].

The recent strong interest in quantum information science and the development of new versatile quantum tools have revived interest in the venerable subject of two-point resolution. Also, the earlier formulation of a general quantum estimation theory [33, 34] has provided a theoretical foundation based on the concept of quantum Fisher information for establishing rigorous fundamental bounds on the precision of estimates of parameters of an optical source achievable by employing optimal quantum measurements on the optical field in the image plane. Applying these principles to the problem of estimation of the separation between two incoherent point sources, it was recently shown that the quantum Fisher information bound offers precision significantly greater than that afforded by direct intensity measure-
ment. Moreover, the quantum bound remains constant as the separation diminishes, making it possible in principle to resolve infinitesimally small separations \[35, 36, 37, 38\]. Consequently, it was implied that the optimal resolution is limitless, and hence it was declared that Rayleighs curse is not a fundamental obstacle to imaging.

With insights from quantum metrology, it was further shown that, together with appropriate measurements in the image plane, a specific simple processing scheme based on linear projections of the optical field onto an even and an odd spatial mode offers estimates of the two-point separation with precision approaching the ultimate quantum bound \[35, 39, 36, 37, 38, 40, 41, 42\]. Experimental verification of these relatively simple schemes has confirmed that Rayleighs curse can be practically overcome.

Inherent in the formulation of the two-point separation quantum estimation problem is the assumption that the centroid of the two points is precisely known and their intensities are equal. If this is not the case, then the problem can be formulated as a multiparameter quantum estimation problem \[43\]. It was shown, for example, that for point sources with unknown intensity ratio, the precision bound on estimates of the separation falls to zero as the separation diminishes \[44\]. Nevertheless, enhancement of the precision can be gained by use of optimal quantum measurement. Other multiparameter estimation problems were considered in this context, including estimations of the cartesian components of the separation as a vector in the source plane, and estimation of moments of the spatial distribution of an extended source \[40, 45\].

In this paper, we expand the scope of this quantum estimation problem further by considering the effect of partial coherence on the two-point resolution. In the earlier years of the development of the theory of optical coherence, it was shown that both the magnitude and the phase of the degree of partial coherence of the source play important roles in determining
the distribution of the image and its statistical properties [46, 47]. It was found that, based on direct measurement of the image intensity, a greater degree of coherence corresponds to greater resolvability of pairs of point-sources, i.e., greater accuracy of the binary decision on whether the illumination originates from one or two point sources. However, in the context of classical estimation of the separation between a pair of point sources, coherence has the opposite effect, namely lowering the estimation precision. The question arises as to whether the quantum precision bound on estimates of the two-point separation is also lowered under conditions of partial coherence. We show here that this is indeed the case. This aspect of the resolution problem is important since the optical fields produced by two point sources are correlated when they share a common origin, e.g., when point scatterers are illuminated by a common extended source [48].

By formulating the quantum estimation problem as one of estimating two parameters – the separation and the degree of coherence of the two-point source – we show that the quantum bound on the precision of the separation estimator drops to zero at small separations, even for a very small, but nonzero, degree of coherence. This is remarkable since it means that the noted success of quantum measurements in breaking Rayleigh’s curse for incoherent sources is vulnerable to the smallest correlation between the two sources. Although the optimal quantum measurement offers some benefit over direct intensity measurement, even when an unknown degree of partial coherence is present, it ultimately falls victim to Rayleigh’s curse. It is evident that the limitation caused by partial coherence is more fundamental than the limitation imposed by direct image intensity detection.
## 3.2 Classical and Quantum Models

**Coherent Imaging**

We first compare the imaging of two point sources for coherent classical light and pure-state single-photon light. Consider a coherent shift-invariant imaging system with a symmetric amplitude point spread function (PSF) $h(x)$ that satisfies the condition $\int dx |h(x)|^2 = 1$. For a source comprised of two emitters located at $x = \pm \frac{s}{2}$ in the object plane and having equal amplitudes $E_0$ and equal (in-phase) or opposite (out-of-phase) phases, the optical field in the image plane is the superposition

$$E(x) = E_0 \left[ h_+(x) \pm h_-(x) \right] \quad (3.1)$$

and the optical intensity is

$$I(x) = |E(x)|^2 = I_0 |h_+(x) \pm h_-(x)|^2, \quad (3.2)$$

where $I_0 = |E_0|^2$, and $h_{\pm}(x) = h(x \pm \frac{s}{2})$.

Now, consider another source generating a single photon in a pure quantum state

$$|\psi_c\rangle = \frac{1}{\sqrt{2(1 \pm d)}} \left[ |\psi_+\rangle \pm |\psi_-\rangle \right], \quad (3.3)$$

where $|\psi_{\pm}\rangle = \int dx \, h_{\pm}(x) |x\rangle$ are vectors with unit norm and $|\psi_c\rangle = \int dx \psi_c(x) |x\rangle$ defines the photon wave function

$$\psi_c(x) = \frac{1}{\sqrt{2(1 \pm d)}} \left[ h_+(x) \pm h_-(x) \right]. \quad (3.4)$$
The normalization constant in Eq. (3.3) guarantees that $\langle \psi \mid \psi \rangle = 1$, where

$$d = \text{Re} (\langle \psi_- \mid \psi_+ \rangle) = \text{Re} \left( \int dx \ h^* \left(x + \frac{s}{2}\right) h \left(x - \frac{s}{2}\right) \right)$$

(3.5)

is the real part of the inner product between the displaced states. The parameter $d$ is a function of the displacement $s$ with values in the $[0, 1]$ range. Typically, $d$ decreases as $s$ increases and vanishes for large $s$. The probability density of detecting the photon at position $x$ is

$$f(x) = |\psi_c(x)|^2 = \frac{1}{2(1 \pm d)} |h_+(x) \pm h_-(x)|^2.$$  

(3.6)

Since Eqs. (3.2) and (3.6) become identical if we assume that the total power of the classical source is unity, i.e., $I_0 = 1/[2(1 \pm d)]$, the analogy between the classical (coherent) and the single-photon (pure state) cases is evident. It is important, however, to note the difference in the physical interpretation: for classical imaging $I(x)$ is the optical intensity, while for single-photon imaging $f(x)$ is the probability density function of the position at which the photon is detected.

**Partially Coherent Imaging**

We now generalize this paradigm to a partially coherent classical source and an analogous single-photon source in a mixed state. In the classical case, we assume that the amplitudes of the two emitters are random variables $E_+$ and $E_-$ so that the total optical field

$$E(x) = E_+ h_+(x) + E_- h_-(x)$$

(3.7)
is random. Assuming that $\langle |E_+|^2 \rangle = \langle |E_-|^2 \rangle = I_0$, the average optical intensity is

$$I(x) = \langle |E(x)|^2 \rangle = I_0[|h_+(x)|^2 + |h_-(x)|^2] + 2I_0 \text{Re} \left[ \gamma h_+^*(x)h_-(x) \right],$$

(3.8)

where $\gamma = \langle E_+^* E_- \rangle / I_0$ is the correlation coefficient or the complex degree of coherence of the field at the two points of the source. For incoherent light $\gamma = 0$, and for coherent light $|\gamma| = 1$. If $I_0 = 1/[2 + 2d\text{Re}(\gamma)]$, then the average power is unity.

An analogous quantum source is a single photon in a mixed state assumed to be a statistical combination of a coherent component — the pure state $|\psi_c\rangle$ in Eq. (3.3) — and a maximally mixed state with density operator

$$\hat{\rho}_i = \frac{1}{2} (|\psi_+\rangle \langle \psi_+| + |\psi_-\rangle \langle \psi_-|).$$

(3.9)

The overall density operator is

$$\hat{\rho} = p\hat{\rho}_c + (1 - p)\hat{\rho}_i,$$

(3.10)

where $\hat{\rho}_c = |\psi_c\rangle \langle \psi_c|$ and $p$ is a probability parameter. The formulation of $\hat{\rho}$ and the normalization of the state vectors implies that $\text{Tr}[\hat{\rho}] = 1$ for all values of $s$, and hence the probability of measuring a photon in the image plane is always unity.

For this quantum state, the probability density of detecting the photon at position $x$ is

$$f(x) = \langle x | \hat{\rho} | x \rangle,$$

$$f(x) = I_0 \left( |h_+(x)|^2 + |h_-(x)|^2 \right) + 2I_0 \text{Re} \left( \gamma h_+^*(x)h_-(x) \right),$$

(3.11)
where

\[
I_0 = \frac{1}{2} \left[ \frac{p}{1 \pm d} + (1 - p) \right]
\]

(3.12)

and

\[
\gamma = \pm \frac{p}{1 \pm (1 - p)d}.
\]

(3.13)

Hence, Eq. (3.11) is identical to Eq. (3.8) for a real-valued \( \gamma \). In the limit \( p = 1, \gamma = \pm 1 \), and \( I_0 = \frac{1}{2(1 \pm d)} \), corresponding to the coherent case for which \( \hat{\rho} = \hat{\rho}_c \), discussed before. In the other limit \( p = 0, \gamma = 0 \), and \( I_0 = \frac{1}{2} \), corresponding to the incoherent case \( \hat{\rho} = \hat{\rho}_i \), which was previously considered in the literature in the context of two-point resolution [35]. Note that in our model, \( \gamma \) is real. The \( \pm \) signs denote the in-phase and out-of-phase cases, which correspond to positive and negative degrees of coherence, hereafter called correlated and anti-correlated, respectively. A key assumption in both the classical model and the quantum, single-photon model treated here is that the optical power is fixed at the sources, and equal to that in the image plane. Physically, this means that there is an assumed known rate of emission of the point sources, and that all power generated by the point sources reaches and is measured in the image plane.

**Two-Point Resolution**

We are concerned here with the resolution of the system, viz. the accuracy of estimating the separation parameter \( s \), and the role of coherence in this process. For classical imaging, the optical intensity \( I(x) \) is typically measured and used to estimate \( s \). To assess the accuracy of such estimation, a model for the measurement noise is necessary. For an ideal detector, the accuracy is ultimately limited by the inherent Poisson noise (or shot noise) in the detection process, which depends on the intensity level. For quantum imaging with a single photon, the location of the photon in the image plane is measured, repeatedly, and the probability density \( f(x) = |\psi(x)|^2 \) is determined, from which the separation \( s \) is estimated. Since \( f(x) \)
is a probability density function depending on the unknown parameter \( s \), we may directly
determine bounds on the estimation accuracy by calculating the classical Fisher information
(CFI) and its inverse the Cramèr-Rao bound (CRB) [33].

When \( p \) is known and we wish to estimate \( s \), the Cramèr-Rao theory states that the variance
of estimates based on measurement of a probability distribution \( f(x) \) is bounded by the
inverse of the Fisher Information, \( \text{Var}(s) \geq 1/F \), where

\[
F = \int dx \left( \frac{\partial}{\partial s} f(x) \right)^2 \frac{1}{f(x)}. \tag{3.14}
\]

We will henceforth define the precision as the inverse of the variance bound, \( H_s = F \geq
1/\text{Var}(s) \). When a measurement process is repeated \( N \) times [49], the variance becomes
bounded by \( H_s^{(N)} = NH_s \). For our calculation, we assume that every photon is collected and
measured perfectly.

As an example, we consider a Gaussian PSF with width \( \sigma \), given by

\[
h^2(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \tag{3.15}
\]

for which the inner product \( d = e^{-s^2/8\sigma^2} \). The precision \( H_s \) of estimates of \( s \) is plotted
in Fig. 3.1 as a function of \( s \) for several values of \( p \) for the correlated and anti-correlated
cases. Based on these plots, the following observations can be made: (i) For any fixed
\( s < 3\sigma \), the precision \( H_s \) is a monotonic decreasing function of the coherence parameter \( p \)
so that the highest precision is attained in the incoherent case \( (p = 0) \). (ii) The correlated case
corresponds to precision greater than that afforded by the anti-correlated case, for the same \( s \)
and \( p \). (iii) In all cases, the precision \( H_s \) drops to 0 (i.e., the variance \( \text{Var}(s) \) becomes infinite,
as \( s \) approaches 0, in which case finding an unbiased estimate of \( s \) becomes impossible. This
is Rayleigh’s curse.

Figure 3.1: Classical Fisher-information precision $H_s \geq 1/\text{Var}(s)$ for estimates of the two-point separation $s$ (normalized to the width $\sigma$ of the PSF), based on direct intensity measurement in the image plane, assuming that the coherence parameter $p$ is known, in the correlated case (solid) and the anti-correlated case (dashed). Here, $p$ takes values between 0 and 1, with intermediate values $1/3$ and $2/3$. For $p = 0$ the correlated and anti-correlated cases are identical.

This problem is further compounded when $p$ is not known since we must instead look at the multi-parameter Cramér-Rao bound. To derive this bound, we determine elements of the classical Fisher information matrix (CFIM),

$$F_{jk} = \int dx \left( \frac{\partial}{\partial j} f(x) \right) \left( \frac{\partial}{\partial k} f(x) \right) \frac{1}{f(x)}, \quad (3.16)$$

where $j, k = s, p$. The multi-parameter Cramér-Rao bound states that the covariance matrix of estimates of $s$ and $p$ is bounded by the inverse of the CFIM, $\mathbf{C} \geq \mathbf{F}^{-1}$. This matrix inequality means that for any matrix $\mathbf{G}$, $\text{Tr} [\mathbf{GC}] \geq \text{Tr} [\mathbf{GF}^{-1}]$. Since the diagonal elements of the covariance matrix are equal to the variances of the parameters $s$ and $p$, the variances are $\text{Var}(s) = C_{ss}$ and $\text{Var}(p) = C_{pp}$. Using the multi-parameter Cramér-Rao bound, the precision bounds for either parameter are defined as $H_s \geq 1/C_{ss}$ and $H_p \geq 1/C_{pp}$. The
precision bound for estimates of $s$ is plotted in Fig. 3.2 as functions of $s$ for several values of $p$ in the correlated and anti-correlated cases. The following observations are noted: (i) $H_s$ is a monotonic increasing function of $s$ exhibiting Rayleigh’s curse. (ii) Greater $H_s$ is attained in the anti-correlated case.

Figure 3.2: Same as in Fig. 1, except that the coherence parameter $p$ is estimated concurrently.

These direct detection schemes, both classical and quantum, may be improved if the optical field in the image plane is processed by some prescribed system before the measurement of the optical intensity or the photon probability density is taken. Certain systems may offer enhanced resolution. As an example, it has been shown both theoretically and experimentally that projection onto a subset of spatial modes of the optical field offers measurements that are less susceptible to Rayleigh’s curse than direct imaging [36, 41, 38, 50].

A third paradigm is to seek the best possible measurement system. Although it is not generally possible to find such system, the quantum Cramér Rao bound (QCRB) always allows us to determine a bound on the precision for that optimal quantum measurement for a given quantum state $\hat{\rho}$. Such calculations have been made [39, 35] for the incoherent
system described by $\hat{\rho}_i$ in Eq. (3.9). We determine these resolution bounds here for light in the partially coherent state described by the density operator $\hat{\rho}$ in Eq. (3.10). As was the case for the calculation of the classical Cramér Rao bound, if $p$ is unknown, then we are faced with a multi-parameter estimation problem and its associated QCRB. This calculation proceeds in the next section.

### 3.3 Quantum Fisher Information Resolution

To determine the ultimate precision that a measurement in the imaging plane can obtain, we calculate the multi-parameter quantum Fisher information matrices (QFIM) corresponding to the states $\hat{\rho} = \hat{\rho}^{(\pm)}(x; s, p)$ for the correlated ($+$) and anti-correlated ($-$) cases [51]. For estimation of either the separation $s$ or the coherence parameter $p$ when neither is known, the elements of the Quantum Fisher Information Matrix (QFIM), $Q$, are given by

$$Q_{i,j} = \frac{1}{2} \text{Tr} \left[ (\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i) \hat{\rho} \right], \quad (3.17)$$

where $\mathcal{L}_i$ is the symmetric logarithmic derivative (SLD) operator for the parameter $i = s, p$, which are solutions to the operator equations

$$\frac{1}{2} (\hat{\mathcal{L}}_i \hat{\rho} + \hat{\rho} \hat{\mathcal{L}}_i) = \frac{\partial}{\partial i} \hat{\rho} \quad (3.18)$$

and $\hat{\rho}$ stands for either $\hat{\rho}^+(x; s, p)$ or $\hat{\rho}^-(x; s, p)$. Once elements of the QFIM are determined, as was the case for calculation of the multi-parameter Cramér-Rao bound, the variance of either parameter is bounded by the corresponding diagonal entry of the inverted QFIM, i.e.,

$$\text{Var}(s) \geq Q_{ss}^{-1}(s) = \frac{Q_{pp}}{Q_{pp}Q_{ss} - Q_{sp}^2} \quad (3.19)$$
\[ \text{Var}(p) \geq Q_{pp}^{-1} = \frac{Q_{ss}}{Q_{pp}^2 - Q_{sp}^2}. \quad (3.20) \]

The corresponding precisions \( H_s \geq 1/\text{Var}(s) \) and \( H_p \geq 1/\text{Var}(p) \) are therefore given by

\[ H_s = Q_{ss} - \frac{Q_{sp}^2}{Q_{pp}} \quad (3.21) \]

\[ H_p = Q_{pp} - \frac{Q_{sp}^2}{Q_{ss}}. \quad (3.22) \]

We therefore need to determine \( Q_{i,j} \) in terms of the system parameters. For this purpose, we have to solve Eq. (3.18) for the SLD operators \( \hat{L}_s \) and \( \hat{L}_p \). As shown in Supplement 1, these operators have a commutator that has an expectation value of zero, so that

\[ \text{Tr} [\hat{\rho} [\hat{L}_s, \hat{L}_p]] = 0, \quad (3.23) \]

and it is thus possible to find a single measurement for simultaneous optimal estimation of both \( s \) and \( p \) [52, 53]. However, since the off-diagonal terms \( Q_{s,p} \) and \( Q_{p,s} \) do not necessarily vanish, Eqs. (3.21) and (3.22) indicate that although \( s \) and \( p \) are simultaneously estimatable, the lack of precise knowledge of one parameter will degrade the precision of the estimate of the other [43, 54].

To determine the SLD operators, we decompose the density operator \( \hat{\rho}^{(\pm)} \), defined by Eqs. (3.10), (3.9) and (3.3), in terms of its own orthogonal eigenvectors

\[ |e_1\rangle = \frac{1}{\sqrt{2(1-d)}} (|\psi_+\rangle - |\psi_-\rangle) \quad (3.24) \]
\(|e_2\rangle = \frac{1}{\sqrt{2(1 + d)}}(|\psi_+\rangle + |\psi_-\rangle) \quad (3.25)\)

and corresponding eigenvalues

\[
\lambda_1^\pm = \frac{1}{2} [(1 - p)(1 - d) + p \mp p] \quad (3.26)
\]

\[
\lambda_2^\pm = \frac{1}{2} [(1 - p)(1 + d) + p \pm p]. \quad (3.27)
\]

so that

\[
\hat{\rho}^\pm = \lambda_1 |e_1\rangle \langle e_1| + \lambda_2 |e_2\rangle \langle e_2|. \quad (3.28)
\]

The derivatives of \(\hat{\rho}^\pm\) are given by

\[
\frac{\partial}{\partial p} \hat{\rho}^\pm = \frac{\partial \lambda_1}{\partial p} |e_1\rangle \langle e_1| + \frac{\partial \lambda_2}{\partial p} |e_2\rangle \langle e_2| \quad (3.29)
\]

and

\[
\frac{\partial}{\partial s} \hat{\rho}^\pm = \frac{\partial \lambda_1}{\partial s} |e_1\rangle \langle e_1| + \frac{\partial \lambda_2}{\partial s} |e_2\rangle \langle e_2| + 2 [\lambda_1 \alpha_3 |e_3\rangle \langle e_1| + \lambda_2 \alpha_4 |e_4\rangle \langle e_2| + \text{H.C.}], \quad (3.30)
\]

where we have defined two new vectors \(|e_3\rangle\) and \(|e_4\rangle\) of unit norm by the relations

\[
\alpha_3 |e_3\rangle = \frac{\partial}{\partial s} |e_1\rangle \quad (3.31)
\]

and

\[
\alpha_4 |e_4\rangle = \frac{\partial}{\partial s} |e_2\rangle, \quad (3.32)
\]
with H.C. denoting the Hermitian Conjugate. As shown in Supplement 1, the four vectors $|e_1\rangle$, $|e_2\rangle$, $|e_3\rangle$, and $|e_4\rangle$ form an orthonormal basis that spans the support of both $\hat{\rho}^\pm$ and $\frac{\partial}{\partial i}\hat{\rho}^\pm$ for both $i = s, p$. When these operators, represented in this four-dimensional basis, are used in Eq. (3.18) the following expressions of the SLDs are obtained for $i = s, p$,

$$
L_i^\pm = \sum_{i,k=1}^{4} \frac{2}{\lambda_k + \lambda_l} \left( \langle e_k | \frac{\partial}{\partial i} \hat{\rho}^{(\pm)} | e_l \rangle \right) |e_k\rangle \langle e_l|.
$$

(3.33)

It follows that the only non-zero matrix elements of $L_s^\pm$ and $L_p^\pm$ are

$$
\begin{align*}
[L_s^\pm]_{11} &= \frac{1}{\lambda_1^s} \frac{\partial}{\partial s} \lambda_1^s = \frac{-(1 - p)}{\lambda_1^s} \Gamma \\
[L_s^\pm]_{22} &= \frac{1}{\lambda_2^s} \frac{\partial}{\partial s} \lambda_2^s = \frac{(1 - p)}{2\lambda_2^s} \Gamma \\
[L_s^\pm]_{31} &= [L_s^\pm]_{13} = 2\alpha_3 \\
[L_s^\pm]_{42} &= [L_s^\pm]_{24} = 2\alpha_4 \\
[L_p^\pm]_{11} &= \frac{1}{\lambda_1^p} \frac{\partial}{\partial p} \lambda_1^p = \frac{(d \mp 1)}{2\lambda_1^p} \\
[L_p^\pm]_{22} &= \frac{1}{\lambda_2^p} \frac{\partial}{\partial p} \lambda_2^p = -\frac{(d \mp 1)}{2\lambda_2^p}.
\end{align*}
$$

(3.34)

The parameters used in these equations are related to the PSF $h(x)$, its derivative $h'(x) =$
\( \frac{dh}{dx}, \) and the displacement \( s \) by the following equations:

\[
\alpha_3 = \frac{1}{2} \sqrt{\frac{a^2 + b^2}{1 - d} - \frac{\Gamma^2}{(1 - d)^2}} \quad (3.35)
\]

\[
\alpha_4 = \frac{1}{2} \sqrt{\frac{a^2 - b^2}{1 + d} - \frac{\Gamma^2}{(1 + d)^2}} \quad (3.36)
\]

\[
a^2 = \int dx \left[ h'(x) \right]^2 \quad (3.37)
\]

\[
b^2 = \int dx \ h' \left( x - \frac{s}{2} \right) h' \left( x + \frac{s}{2} \right) \quad (3.38)
\]

\[
\Gamma = \frac{\partial d}{\partial s} = \int dx \ h'(x) h(x - s). \quad (3.39)
\]

Using Eq. (3.17) and accounting for the zero matrix elements of \( \mathcal{L}_s^\pm \) and \( \mathcal{L}_p^\pm \), we obtain the following expressions for elements of the QFIM,

\[
Q_{ss}^\pm = \lambda_1^\pm \left( [\mathcal{L}^\pm_{s}]_{11}^2 + [\mathcal{L}^\pm_{s}]_{13}^2 \right) + \lambda_2^\pm \left( [\mathcal{L}^\pm_{s}]_{22}^2 + [\mathcal{L}^\pm_{s}]_{13}^2 \right) \quad (3.40)
\]

\[
Q_{pp}^\pm = \lambda_1^\pm \left[ \mathcal{L}^\pm_{s} \right]_{11}^2 + \lambda_2^\pm \left[ \mathcal{L}^\pm_{s} \right]_{22}^2 \quad (3.41)
\]

\[
Q_{sp}^\pm = \lambda_1^\pm \left[ \mathcal{L}^\pm_{s} \right]_{11} \left[ \mathcal{L}^\pm_{s} \right]_{11} + \lambda_2^\pm \left[ \mathcal{L}^\pm_{s} \right]_{22} \left[ \mathcal{L}^\pm_{s} \right]_{22}. \quad (3.42)
\]

These equations can be used together with Eqs. (3.21) and (3.22) to calculate the precisions \( H_s \) and \( H_p \) for any PSF \( h(x) \).

We now use the Gaussian PSF described by Eq. (3.15) as an example to determine the dependence of \( H_s \) and \( H_p \) on the normalized separation \( s/\sigma \) and the coherence parameter \( p \). In this example, the parameters in Eqs. (3.35)–(3.39) are

\[
d = e^{-s^2/8\sigma^2}, \quad a^2 = \frac{1}{4\sigma^2},
\]

\[
\Gamma = -d \frac{s}{4\sigma^2}, \quad b^2 = \frac{d}{4\sigma^2} \left( \frac{s^2}{4\sigma^2} - 1 \right).
\]
We now consider two cases: i) estimation of $s$ when $p$ is perfectly known, and ii) concurrent estimation of both $s$ and $p$.

**Estimation of separation with known coherence parameter**

If $p$ is perfectly known, i.e., $s$ is the only unknown parameter, then the variance of the estimate of $s$ is simply bounded by the QFIM element $Q_{ss}$ so that $H_s = Q_{ss}$. This precision bound is plotted in Fig. 3.3 as a function of $s/\sigma$ for several values of $p$. In the limit $p = 0$, which corresponds to the incoherent case, $Q_{ss}$ has no functional dependence on $s$, and hence the precision bound $H_s$ has a constant value extending to the limit $s = 0$, so that it is in principle always possible to make a measurement that gives an unbiased estimate with non-zero precision. This limiting result has recently led to the announcement that Rayleigh’s curse has been broken [39]. The graphs in Fig. 3.3 show that for a source with positive degree of coherence, the precision bound $H_s$ at $s = 0$ drops as $p$ (or $\gamma$) increases, and ultimately vanishes when $p = \gamma = 1$, resurrecting Rayleigh’s curse. It is interesting, however that for a source with negative degree of coherence, with even the smallest magnitude, the curse is revived for any $p \neq 0$ (or $|\gamma| \neq 0$).
Figure 3.3: Quantum-Fisher-information precision bound $H_s = 1/\text{Var}(s)$ on estimates of the separation $s$ between two point sources by use of optimal measurement in the image plane. The coherence parameter $p$ is assumed to be perfectly known. Correlated (solid) and anti-correlated (dashed) coherence are assumed. Here, $p$ takes values between 0 and 1, with intermediate values 1/3 and 2/3. For $p = 0$ the correlated and anti-correlated cases are identical.

**Concurrent estimation of separation and coherence parameter**

In this case, the dependence of the precision bounds $H_s$ and $H_p$ on $s/\sigma$ and $p$ are shown in Fig. 3.4. Remarkably, for both the correlated and anti-correlated cases, $H_s = 0$ as $s \to 0$, for any $p > 0$ so that if there is any degree of correlation between the point sources, no matter how small, Rayleigh’s curse resurges.
Figure 3.4: Same as in Fig.3.3, but assuming that the coherence parameter $p$ is unknown and is concurrently estimated with the separation $s$ using the multi-parameter QFIM. When $p$ is not precisely known, Rayleigh’s curse persists.

**Comparison between direct imaging and quantum-optimal measurement**

It is revealing to compare the precision afforded by estimation based on optimal field measurement together with quantum Fisher information to the precision bound based on direct intensity measurements together with classical Fisher information. This comparison is depicted in Fig. 3.5, showing that the optimal quantum measurement offers significant improvement in precision over direct intensity measurement although neither estimator beats Rayleigh’s curse when $p > 0$. 
Figure 3.5: Ratio of the precision bounds for the optimal strategy against the precision bounds afforded to intensity imaging for the correlated (solid) and anti-correlated (dashed) cases, and $p = 0, \ 1/3, \ 2/3, \ 1$. For $p = 0$ the correlated and anti-correlated cases are identical.

**Precision bounds on the coherence parameter**

While the principal focus of this paper is on the precision of estimates of the separation $s$ when the coherence parameter $p$ is either known or estimated, an important byproduct of the analysis is bounds on the precision of estimates of $p$ when $s$ is either known or estimated. Estimating the degree of coherence can be useful in applications for which the correlation between two emitting or scattering sources is to be assessed. The results are displayed in Fig. 3.6 for estimates based on direct intensity measurement and optimal quantum measurement.

In all cases, higher values of $p$ are estimated with greater precision. Of crucial importance here is whether the sources are correlated or anti-correlated. The precision $H_p$ is always higher for anti-correlated sources and also has a stronger dependence on $p$. This is to be expected for direct intensity measurement since close anti-correlated sources create an intensity distribution with a visible dip at the center, and the depth of the dip is greater for larger $p$. No such dip exists for correlated sources, and if the separation is small, it is difficult
to discern the effect of coherence, so that the precision is low and practically independent of $p$. Also, for both optimal direct intensity measurement and optimal quantum measurement, the precision of estimates of $p$ drops to zero in the correlated case as the separation $s \to 0$, while it remains constant and finite in the anti-correlated case. In both cases, however, the precision $H_p$ is orders of magnitude greater for optimal quantum measurement compared to optimal direct intensity measurement, and this is particularly so for the correlated case at small separations, where direct intensity measurement is not precise.
Figure 3.6: Precision bound $H_p = 1/\text{Var}(p)$ on estimates of the coherence parameter $p$ based on direct intensity measurement (top) and quantum optimal measurement (bottom) for concurrent estimates of $p$ and the two-point separation $s$ in the correlated case (solid) and the anti-correlated case (dashed), for $p = 0, 1/3, 2/3, 1$.

3.4 Discussion

Conventional imaging systems rely on direct measurement of the image-plane optical intensity, which provides only a portion of the information about the object that is carried by
the optical field. For such systems, the two-point resolution is limited by diffraction, which diminishes the precision of estimating two-point separation as the separation is reduced, and the system succumbs to Rayleigh's curse. With optimal quantum measurement of the optical field, the curse is broken and the separation may be estimated with finite precision no matter how small the separation is. What we have shown in this paper is that this is true only when the emissions from the two points are completely uncorrelated, or incoherent. The introduction of any correlation, positive or negative, between the emissions has a detrimental effect on the precision of estimates of the separation, and this effect is particularly strong for small separations so that even optimal quantum measurements, which offer unsurpassable precision, fail to defeat Rayleigh's curse. This effect is similar to ill-posed inverse problems for which solutions exist, but are highly sensitive to the slightest uncertainty in the measured data.

One reason for the resurgence of the curse in the presence of correlation may be attributed to the fact that the degree of coherence is a new unknown parameter that must be estimated jointly with the separation. But we have shown that even if this parameter is known \textit{a priori}, if the correlation is negative, then the curse remains. The presence of known positive correlation does break the curse, however. It should be noted that these findings are applicable when the degree of partial coherence $\gamma$ is real, positive or negative. For a complex $\gamma = |\gamma|e^{i\theta}$, the analysis is more involved, and it is possible that at certain values of $\theta$, the curse is avoided, much like when $|\gamma| = 0$.

Source correlations cannot be ignored, particularly at small separations, which is exactly the region for which Rayleigh's curse is manifest. A Lambertian source is not strictly incoherent, having a correlation width of the order of a wavelength with positive correlation at small separations, and light gains transverse coherence as it propagates [48]. Consequently, Rayleigh's curse endures as a fundamental dictum.
CHAPTER 4: SUPER-SENSITIVE ANCILLA-BASED ADAPTIVE PHASE ESTIMATION

The notorious delicacy of the super-sensitivity attained in quantum phase estimation is demonstrated notably at blind spots – phase values at which sensitivity is completely lost when there is any amount of decoherence added to the quantum probe states. The most common remedy is to use a precisely known reference phase to shift the operation point to a less vulnerable phase value. Since this is not always feasible, we present in [55] an alternative approach based on combining the probe with an ancillary degree of freedom containing adjustable parameters to create an entangled quantum state of higher dimension. We validate this concept by simulating a configuration of a Mach-Zehnder interferometer with a two-photon probe and a polarization ancilla of adjustable parameters, entangled at a polarizing beam splitter. At the interferometer output, the photons are measured after an adjustable unitary transformation in the polarization subspace. Through calculation of the Fisher information and simulation of an estimation procedure, we show that optimizing the adjustable polarization parameters using an adaptive measurement process provides globally super-sensitive unbiased phase estimates for a range of decoherence levels, without prior information or a reference phase.

4.1 Introduction

The maximum measurement sensitivity of optical measurements based on classical optical probes can be surpassed by the use of non-classical light [56, 57, 49, 58, 59, 60]. In many interferometric contexts, measurement of phase is bounded by the shot-noise or classical limit (CL) in classical sensing strategies and by the Heisenberg-limit (HL) [61, 62, 63] in
non-classical sensing strategies. As dictated by the Cramér-Rao bound [11, 31], the variance of estimates employing an average of \( N \) photons can scale at best as \( \frac{1}{\sqrt{N}} \) for classical probes, while the variance of estimates employing exactly \( N \) photons can scale at best as \( \frac{1}{N} \); an estimate that achieves a variance between these ranges is commonly referred to as super-sensitive, and is the hallmark of precision quantum sensors.

The precision attained in entangled-photon phase estimation is compromised in the presence of any finite source of imperfection or decoherence [64, 65, 66, 67] of the states used or device measured, making it rather challenging to reach the goal of super-sensitive estimation. Even worse, for certain values of phase, which we refer to as blind spots, the measurement fails to provide any sensitivity [68] [69]. Traditional adaptive phase estimation overcomes this issue by employing a reference phase and iteratively adjusting the operation parameters of the interferometer (often taking the form of a reference phase) to the range for which it is most sensitive. To our knowledge, in every demonstration of the use of two-photon interferometry, a reference phase has been required to observe super-sensitivity.

Lately, much work has been done to investigate how ancillary photons or degrees of freedom (DoFs) can be used to aid quantum estimation strategies against such deleterious effects [70, 71, 72]. More specifically, recent work [4] has shown that in the presence of partial two-photon spectral distinguishability, an effect that degrades two-photon interference while leaving single-photon interference unhindered, it is possible to employ an ancillary DoF to fortify super-sensitive two-photon states against the total loss of sensitivity at the blind spots. By coupling an ancillary DoF (ancilla) to the probe DoF it was possible to theoretically model and experimentally measure sensitivity above the CL using coincidence measurements at a blind spot.

Here, we consider ancilla-based phase estimation with a two-photon quantum state impaired
by decoherence described by the ‘depolarizing-channel’ model, which is one of the most general models of decoherence in two-photon systems. This effect degrades both two-photon and single-photon interference. We use a configuration of a Mach-Zehnder interferometer with a two-photon probe and a polarization ancilla of adjustable parameters, entangled at a polarizing beam splitter. At the interferometer output, the photons are measured after an adjustable unitary transformation in the polarization subspace. Through calculation of the Fisher information we show that fortification through an ancillary DoF protects the quantum advantage afforded to two-photon measurements for a range of the depolarization probabilities (decoherence levels). Within this range, it is possible to use the ancillary DoF rather than a reference phase, which must be placed within the interferometer itself, to retain the sensitivity of the interferometer.

We also show that adaptive phase estimation can be performed in this paradigm by tuning the polarization (ancilla) of the input two-photon state and the two-photon polarization measurements that are made at the output of the system, rather than tuning the optical system itself by altering a reference phase. Previous experimental and theoretical treatments of this topic have only considered the case where precise prior information of the phase exists, and this is, to our knowledge, the first theoretical work that considers the entire adaptive process. Our simulations suggest that, just as in the case of using a reference phase, adaptive tuning of the ancillary degree of freedom provides unbiased estimates that are super-sensitive for a range of decoherence probabilities. The techniques developed here can therefore play a critical role in phase estimation tasks where introduction of a reference phase is not feasible.
4.2 Effect of Decoherence on Phase Sensitivity

Two-photon probe in a pure state

The phase $\phi$ introduced, for example, by transmission through an optical element is typically measured by placing the element in one arm of a MZI and using an optical probe at the input ports together with an appropriate measurement at the output ports. An unbiased phase estimate $\tilde{\phi}$ inferred from a single measurement outcome $M_i$ from a set of $k$ possible measurement outcomes $\{1, ..., i, ..., k\}$ has a statistical variance satisfying the Cramér-Rao bound, $\text{Var}(\tilde{\phi}) \geq \frac{1}{F(\phi)}$, where

$$F(\phi) = \sum_{i} F_i(\phi)$$  \hspace{1cm} (4.1)

is the total Fisher information given by summing the contributions $F_i(\phi)$ for each possible measurement outcome. The contributions are given by

$$F_i(\phi) = P(M_i|\phi) \left[ \frac{\partial}{\partial \phi} \ln P(M_i|\phi) \right]^2,$$  \hspace{1cm} (4.2)

where $P(M_i|\phi)$ is the conditional probability distribution of measuring $M_i$ given $\phi$. This variance, which defines the sensitivity of the measurement, clearly depends on the choice of the probe and the possible measurement outcomes. If a measurement is repeated $n$ times, the variance of an estimate is bounded by $\text{Var}(\tilde{\phi}) \geq \frac{1}{nF(\phi)}$.

Here, we limit ourselves to two-photon optical probe states and two-photon measurements — either through coincidence between single-photon detection, photon-number resolving detection, or both. When the probe state is a pure state with one photon in each of the interferometer input ports, it turns out that measuring two-photon coincidence and double counts at the output ports of the interferometer is in fact the optimal measurement, with $F(\phi) = 4, \forall \phi$, achieving the HL that corresponds to the highest sensitivity achievable using.
two photons. For comparison, a classical optical probe in a coherent state with an average of two photons obtains a maximum value of $F(\phi) = 2$. In this case, the quantum two-photon probe offers a factor of 2 advantage in the variance of estimates over a classical probe with the same mean number of photons [73, 74].

**Two-photon probe with decoherence**

One would expect that the phase sensitivity achievable with an optical probe in a two-photon state subjected to decoherence would deteriorate gradually as the strength of decoherence increases. It turns out that the effect of decoherence also depends significantly on the actual value of the phase $\phi$. The depolarizing channel model converts a qubit describing the path of a photon in a pure state into a mixed state characterized by a probability parameter $p$. The model, as described in Chapter 2, can be used to characterize decoherence acting on multi-photon or multi-degree-of-freedom states. Here, we consider a two-photon state with two DoFs, path and polarization, with polarization playing the role of the ancilla. Decoherence acts equally on both photons in the path DoF, leaving the polarization DoF unaltered. The operation is

$$\mathcal{E}_p(\rho) = \mathcal{E}_p^{(1)}(\mathcal{E}_p^{(2)}(\rho)) = \mathcal{E}_p^{(2)}(\mathcal{E}_p^{(1)}(\rho)),$$

(4.3)

where $\mathcal{E}_p^{(i)}$ denotes the decoherence channel acting on the interferometer-path degree of freedom of the $i^{th}$ photon with probability $p$. Since this model is responsible for the reduction of the visibility of both two-photon and single-photon interference, we choose to use it in lieu of an ad hoc model representing a generic loss of visibility. Additionally, the operator representing this model commutes with the unitary operator describing transmission through a beam splitter, making it a good model for both noise originating from the source, or from within the interferometer itself.

To investigate how decoherence affects the sensitivity of our system, we calculated the Fisher
information $F(\phi; p)$ using the depolarized input state. We write a pure input state describing the interferometer-path-modes (probe) of a photon pair as superpositions of vectors of the form $|P_1\rangle \otimes |P_2\rangle$, where $P$ denotes the binary probe DoF, and 1, 2 refer to the first and second photon – hence, every mode describes the location of one photon. To create the optimal probe \cite{75}, we use an input state given by,

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} [ |l\rangle |u\rangle + |u\rangle |l\rangle ],$$

(4.4)

where $|u\rangle$ and $|l\rangle$ correspond to the upper and lower input ports of the MZI. In the absence of decoherence, this state is transformed by the beam-splitter into the well-known two photon $N00N$ state $\frac{1}{\sqrt{2}} [ |l\rangle |l\rangle + |u\rangle |u\rangle ]$. In the presence of decoherence, the state $\rho_0 = |\psi_0\rangle \langle \psi_0|$ of the pure input state is altered, becoming a mixed state with density operator

$$\rho(p) = \mathcal{E}_p(|\psi_0\rangle \langle \psi_0|),$$

(4.5)

where $\mathcal{E}_p$ represents the depolarization operation, which is characterized by a probability $p$, as described in Chapter 1. This state is then evolved unitarily by the interferometer, encoding information about the phase difference $\phi$ into the state, leaving the output state

$$\rho(\phi; p) = U_{MZI}(\phi) \rho(p) U_{MZI}^\dagger(\phi),$$

(4.6)

where $U(\phi)_{MZI} = \hat{B} e^{-i\phi \hat{n}_u} \hat{B}$, $\hat{B} = \frac{1}{2}(\mathcal{I} + i\sigma_x)^{\otimes 2}$ is the beam-splitter transformation, $\sigma_x$ is the Pauli-X operation, $\mathcal{I}$ is the identity operation, and $\hat{n}_u$ is the photon-number operator corresponding to the photon-number in the upper-arm of the interferometer.
The probability of measuring a coincidence count is then

\[ P_c(\phi; p) = \text{Tr}[\rho(\phi; p)\Pi_c] = \frac{1}{2} - \frac{1}{2} (p - 1)^2 \cos 2\phi, \quad (4.7) \]

and the probability of measuring a double count is

\[ P_d(\phi; p) = \text{Tr}[\rho(\phi; p)\Pi_d] = 1 - P_c(\phi; p), \quad (4.8) \]

where

\[ \Pi_c = |u\rangle \langle l| + |l\rangle \langle u| , \quad (4.9) \]

and

\[ \Pi_d = |u\rangle \langle u| + |l\rangle \langle l| , \quad (4.10) \]

are the operators representing coincidence and double counts, respectively. From these probabilities, we find

\[ F_c(\phi; p) = \frac{1}{P_c(\phi; p)} \left( \frac{\partial}{\partial \phi} P_c \right)^2 = \frac{2(1 - p)^4 \sin^2 2\phi}{1 + \cos 2\phi(1 - p)^2} \quad (4.11) \]

\[ F_d(\phi; p) = \frac{1}{P_d(\phi; p)} \left( \frac{\partial}{\partial \phi} P_d \right)^2 = \frac{2(1 - p)^4 \sin^2 2\phi}{1 - \cos 2\phi(1 - p)^2} . \quad (4.12) \]

The total fisher information is given by the sum of contributions from each possible mea-
surement outcome, hence

\[ F(\phi; p) = \sum_i F_i(\phi; p). \] (4.13)

Given that the possible measurements are coincidence and double counts, \( i \in \{c, d\} \), we find the total Fisher information to be

\[ F(\phi; p) = F_c(\phi; p) + F_d(\phi; p) = \frac{4(1 - p)^4 \sin^2 2\phi}{1 - (1 - p)^4 \cos^2 2\phi}. \] (4.14)

This quantity is plotted in Figure 4.1.

When the depolarization operation contaminates the input state with any non-zero probability \( p \), there is a drastic change in \( F(\phi; p) \), and we see the emergence of blind spots — phases for which the sensitivity afforded by both two-photon measurements drops sharply to zero. Specifically, for \( \phi \in \{\phi_{BS}\} = \{0, \frac{\pi}{2}, \pi\} \), we find that \( F(\phi; p) = 0 \). This behavior is plotted in Figure 4.1. With this drastic change, it is clear that the once-optimal interferometer will now be completely insensitive to phase values in the neighborhood of any of these blind spots. At these spots, finding an unbiased estimator of \( \phi \) becomes impossible as \( F(\phi; p) \) approaches zero.
Figure 4.1: Fisher information $F(\phi; p)$ as a function of the phase $\phi$ for two cases. (a) When a Mach-Zehnder Interferometer (MZI) is fed with the pure two-photon input state, the sum of the Fisher information provided by contributions from coincidence (blue) and double-count (red) measurements leads to the total (purple) $F(\phi; 0)$ that can be gained from two-photon measurements at the output of an ideal MZI. (b) As in (a), but with non-zero decoherence probability $p$ (shown for $p = 0.005$). Blind spots appear at $\phi = 0, \frac{\pi}{2}, \pi$.

Reference Phase

As a simple remedy to the loss of sensitivity at a blind spot, one could add a precisely known, tunable reference phase $\phi_r$ to a path of the interferometer [74]. To show this, we calculate the Fisher information $F(\phi - \phi_r; p)$ that results from measuring the probabilities of coincidence and double counts when a reference phase is used. The probability of measuring coincidence or double counts are given by $P_c(\phi - \phi_r; p) = \text{Tr}[\rho(\phi - \phi_r; p)\Pi_c]$ or $P_d(\phi - \phi_r; p) = \text{Tr}[\rho(\phi - \phi_r; p)\Pi_d]$, respectively, and give a sensitivity described by $F(\phi - \phi_r; p)$.

The optimal sensitivity afforded to this strategy is then obtained by maximizing $F(\phi - \phi_r; p)$ over $\phi_r$, for which the optimal operating point will be found at $\phi - \phi_r = \phi_{\text{peak}} = \frac{\pi}{4}$. While the introduction of a phase reference conveniently obviates the repercussions of the phase dependence on sensitivity, we are left with a question: could we perform some other modifi-
cation to our system that does not require physical changes inside the interferometer itself?

To answer this question affirmatively, we introduce a second degree of freedom, an ancilla, which we implement by means of the polarization of the photon pair.

### 4.3 Ancilla Fortification

A probe in a two-photon binary pure state is augmented by a binary ancillary DoF – in this case, the polarization of the photon pair – creating a pure two-photon state that is represented as a linear combination of vectors of the product form \(|P_1, A_1 \rangle \otimes |P_2, A_2 \rangle\). Here, \(P, A\) denote the binary probe (interferometer paths |\(u\rangle, |l\rangle\)) and ancillary (polarizations |\(H\rangle, |V\rangle\)) DoFs, respectively, and 1, 2 refer to the first and second photon.

In order to generate a path-polarization entangled state we start with a single-spatial-mode two-photon beam in the pure polarization state \(\frac{1}{\sqrt{2}} [|H\rangle_1 |V\rangle_2 + |V\rangle_1 |H\rangle_2]\), and use a polarizing beam splitter (PBS) to create the two paths. The PBS enacts the transformation |\(H\rangle \rightarrow |l, H\rangle\) and |\(V\rangle \rightarrow |u, V\rangle\). The polarization-path correlations of the outgoing beams are manipulated by use of two half-wave plates (HWPs) with optical axes at controllable angles \(\alpha_1\) and \(\alpha_2\) (see chapter 2), the former placed before the PBS and the later placed in one of the outgoing paths of the PBS (see Figure 4.2). The result is a pure probe state described by

\[
\psi^{(A)}_{in}(\alpha_1, \alpha_2) = \sin \alpha_1 (|H, l\rangle |H, l\rangle - |\alpha_2, u\rangle |\alpha_2, u\rangle) \\
+ \cos \alpha_1 (|H, l\rangle |\alpha_2, u\rangle + |\alpha_2, u\rangle |H, l\rangle),
\]

(4.15)

where the polarization state \(|\alpha_2\rangle = \cos \frac{\alpha_2}{2} |H\rangle - \sin \frac{\alpha_2}{2} |V\rangle\). The state created by this process is a state that is, in general, not optimal when \(p = 0\). This may not be surprising; for a
large number of estimation tasks, the optimal quantum probe states are rarely optimal once decoherence into introduced to the system [76, 77, 78, 79]. Furthermore, while this state is not the most general path-polarization state, we found that it provides the best sensitivity of all the states that we have studied. The states studied include the polarization-path states that could occur from either Type 1 or Type 2, collinear or non-collinear down-conversion and wave plate transformations occurring both before and after path-splitting by the PBS.

Figure 4.2: A path-polarization two-photon entangled state is created by transmitting photon-pairs generated by a source S in the state \( \frac{1}{\sqrt{2}} (|H\rangle |V\rangle + |V\rangle |H\rangle) \) through a half-wave plate (HWP) with optic axis at an angle \( \alpha_1 \), separating the polarization components with a polarizing beam splitter (PBS), and passing one component through a second HWP at angle \( \alpha_2 \). The entangled state created is subjected to the depolarizing channel as it enters an interferometer with the measured phase \( \phi \) in one arm. At each of the output ports of the interferometer, polarization measurements are tuned by placing a HWP at angle \( \beta_1 \), a quarter-wave plate at angle \( 2\beta_1 \), and a PBS with two photon detectors at each output port.

After the preparation stage, this state is then acted upon by the depolarizing channel and transformed by transmission through the interferometer. A pair of HWPs and quarter-wave plates (QWP) are placed into each of the output arms of the interferometer (see Chapter 2). The HWPs have optics axes at angles \( \beta_1 \) (upper arm) and \( \beta_2 \) (lower arm), while the QWPs have optic axes at angles \( 2\beta_1 \) (upper arm) and \( 2\beta_2 \) (lower arm). The angles \( \beta_1 \) and \( \beta_2 \) are independent. Two-photon measurements are then made at the output ports of a PBS placed after each pair of wave plates. While these are not the most general settings for the wave plates at the output, we have found them – through exhaustive simulation – to contain...
the optimal sensitivity for each value of $\phi$, allowing us to reduce the runtime of numerical simulations. As a result, the output state before the final pair of PBSs is a density operator of dimension $16 \times 16$ characterized by the parameter set $\Theta = \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$, and given by

$$
\rho_{\text{out}}^{(a)}(\phi; p, \Theta) = U_\beta(\beta_1, \beta_2)U_{\text{MZI}}(\phi)E_p\left(\rho_{\text{in}}^{(a)}(\alpha_1, \alpha_2)\right)U_\beta^\dagger(\beta_1, \beta_2),
$$

(4.16)

where $U_{\text{MZI}}(\phi)$ is the interferometer transformation acting on the path degree of freedom as in Eq.4.6, $U_\beta^\dagger(\beta_1, \beta_2)$ is the transformation of the wave plates at the output of the interferometer acting on the polarization degree of freedom, $E_p$ is the depolarizing channel acting on the path degree of freedom, and $\rho_{\text{in}}^{(a)}(\alpha_1, \alpha_2) = |\psi^{(A)}_{\text{in}}(\alpha_1, \alpha_2)\rangle \langle \psi^{(A)}_{\text{in}}(\alpha_1, \alpha_2)|$ is the density operator describing the input state.

Now, the calculation of the optimal Fisher information for a given value of phase $\phi$, denoted $F_{\text{opt}}^{(a)}(\phi; p)$, becomes an optimization over the parameter set $\Theta$, as opposed to a reference phase $\phi_r$. The probabilities $P$ of measuring the two photon state in output path modes \{\(P_1, P_2 \in |u\), |l\}\} and polarization modes \{\(A_1, A_2 \in |H\), |V\}\} needed to calculate $F_{\text{opt}}^{(a)}(\phi; p)$ are given by

$$
P_{P_1, A_1, P_2, A_2} = \text{Tr}[\rho_{\text{out}}^{(a)}(\phi; p, \Theta)\Pi_{P_1, A_1, P_2, A_2}],
$$

(4.17)

where

$$
\Pi_{P_1, A_1, P_2, A_2} = \begin{cases} 
|P_1, A_1\rangle \langle P_2, A_2|& \text{for } P_1, A_1 = P_2, A_2, \\
\frac{1}{2} |P_1, A_1\rangle \langle P_2, A_2| & + \\
\frac{1}{2} |P_2, A_2\rangle \langle P_1, A_1| & , \text{else.}
\end{cases}
$$

(4.18)

As was the case when there were only two possible measurement outcomes in Eq. 4.13, the
total Fisher information with the included ancilla is given by

\[
F^{(a)}(\phi, p, \Theta) = \sum_{P_{1,A_1},P_{2,A_2}} \mathcal{P}_{P_1,A_1,P_2,A_2}(\phi; p, \Theta)^{-1} \left( \frac{\partial}{\partial \phi} \mathcal{P}_{P_1,A_1,P_2,A_2}(\phi; p, \Theta) \right)^2,
\]

(4.19)

which now includes 16 terms. For each value of \( \phi \), there will be a subset of the possible values \( \Theta \) that provides the optimal sensitivity \( F^{(a)}_{\text{opt}}(\phi; p) = \max_{\Theta} \{ F^{(a)}(\phi; p, \Theta) \} \).

Calculation of any probability \( \mathcal{P}_{P_1,A_1,P_2,A_2} \) requires multiplication of a number of 16 \( \times \) 16 matrices, the set of which are functions of \( \phi, p, \alpha_1, \alpha_2, \beta_1, \) and \( \beta_2 \). The trace, as well as its derivative, of the resulting probability of measurement is then calculated to compute \( F^{(a)}(\phi, p, \Theta) \). Evidently, analytically determining the sensitivity \( F^{(a)}(\phi, p, \Theta) \) is an infeasible task given the size of this calculation. We have therefore relied on numerical computations to determine \( F^{(a)}_{\text{opt}}(\phi; p) \). For a given value of \( \phi \) and \( p \), we take a set of values for each of the parameters in \( \Theta = \{ \alpha_1, \alpha_2, \beta_1, \beta_2 \} \), defined at points of a four-dimensional discrete grid with dimension 20 \( \times \) 20 \( \times \) 20 \( \times \) 20, and calculate numerically the sensitivity \( F^{(a)}(\phi, p, \Theta) \) at each point and determine the peak value \( F^{(a)}_{\text{opt}}(\phi; p) \) and the corresponding values of \( \Theta \) at which the peak occurs. We then repeat this process for other values of \( \phi \) and \( p \).

We report the results of this exhaustive numerical search for an array of phase values between 0 and \( \frac{\pi}{2} \) in Fig. 3 for \( p=0.05 \), a probability leading to a more severe degree of decoherence than that plotted in Figure 1. While operation at the blind spots may not be as sensitive as the operation that could be attained using a reference phase, it is clear that it makes possible the quantum super-sensitive advantage in a sensing regime where no sensitivity was possible prior. We find that super-sensitivity is globally attainable for all phases for decoherence probabilities less than \( p = 0.072 \).
Figure 4.3: Fisher information for a two-photon interferometer with (purple) and without (orange data points) the employment of an ancillary degree of freedom for $p = 0.05$. Ancilla fortification allows for measurements that retain sensitivity at the blind spots 0 and $\pi/2$. Super-sensitivity is retained for $p < 0.072$ when an ancilla is used.

4.4 Adaptive Phase Estimation

To show that this methodology can find application in the general setting of quantum phase estimation, we have conducted simulations demonstrating how the ancilla-aided strategy provides a full platform for super-sensitive adaptive phase estimation. Simulations of adaptive phase estimation were conducted by supplementing simulations of maximum likelihood estimation with feedback based on measurement results. Traditional adaptive phase estimation uses feedback from measurements to update the value of the reference phase $\phi_r$. In that case, $\phi_r$ is set to the value that maximizes the sensitivity of the optical system, assuming that the true value of $\phi$ is equal to the most recent estimate $\tilde{\phi}$ [80, 81, 82, 83, 84]. Likewise, when using ancilla fortified states without a reference phase, feedback updates the parameter set $\Theta$. After each measurement, $\Theta$ is set to the calculated optimal values (determined by our numerical exhaustive search for each value of $\phi$, $p$) in order to maximize the sensitivity of the system for the assumed value of $\phi = \tilde{\phi}$. 

59
Specifically, performing an \( n^{th} \) two-photon measurement provides a result \( M_n \) that corresponds to the measurement operator element \( \Pi_n \in \Pi_{P_1,A_1,P_2,A_2} \). The likelihood of the result is \( \mathcal{L}_n(\phi) = \text{Tr}[\rho(\phi)\Pi_n]\mathcal{L}_{n-1}(\phi) \), from which an \( n^{th} \) estimate of \( \phi \), \( \tilde{\phi}_n = \arg\max_{\phi} \mathcal{L}_n(\phi) \), is determined. With the estimate \( \tilde{\phi}_n \), the settings of the parameters \( \Theta \) that maximize \( F^{(a)}(\tilde{\phi}_n; p, \Theta) \) are updated \[85\]. To ensure that the final estimate \( \tilde{\phi}_N \) after \( N \) measurements is unbiased (i.e. \( \langle \tilde{\phi}_N \rangle = \phi \)), the initial estimate \( \tilde{\phi}_0 \) is chosen at random. Finally, to avoid degeneracy in the final likelihood function (which would lead to equally likely estimates spaced by a period), a final, non-optimal series of measurements must be made. A pair of measurement settings can be chosen to have different periodicities in the likelihood as a function of \( \phi \): these periodicities are chosen to differ from each other such that at least one differs from the optimal measurement. Hence, by making a set of non-optimal measurements at the end of the adaptive procedure, we are able to resolve any indistinguishability in the likelihood function that came from the larger set of optimal measurements. By experimentation, we have found that setting \( \Theta \) to the values that are optimal at \( \phi = 0 \), then \( \phi = \frac{\pi}{2} \), for the last 3% of adaptive iterations is sufficient to avoid this discrepancy between phases that are degenerate in the likelihood function.

For our numerical testing of this procedure for each phase \( \phi \), we set \( p = 0.01 \) and performed \( S = 5000 \) simulation measurement processes, each simulating adaptive detection of a total of \( N = 1500 \) two-photon states. From these simulated measurement processes, \( S \) final estimates \( \tilde{\phi}_N \) were collected, and statistics were calculated on the ensemble of these estimates. The mean and variance of estimates as a function of the true value of \( \phi \) for each simulation is plotted in Figures 4.4 and 4.5. For the given sample size, we find that the strategy is unbiased, and the functional dependence of the variance on \( \phi \) generally follows the trend predicted by our numerical calculation of the Fisher information. Most importantly, for this probability of decoherence, our simulations show estimates that are super-sensitive for all
values of $\phi$.

Figure 4.4: Mean of final phase estimates after $S = 5000$ simulations of $N = 1500$ adaptive two-photon state detections (blue) are plotted for $p = 0.01$ with the 1:1 correspondence expected for unbiased estimators (orange).
Figure 4.5: Variance of final phase estimates after $S = 5000$ simulations of $N = 1500$ adaptive two-photon state detections (blue) for $p = 0.01$ follows the functional form of the Cramér-Rao Bound, as dictated by the Fisher information (orange). For each value of $\phi$, a variance below the classical estimation limit is observed.

4.5 Discussion

The accuracy of a metrological measurement is limited by inherent noise in the probe. While quantum optical probes offer a global advantage over classical probes, they can be more vulnerable to minute imperfections or contamination by weak extraneous noise. A case in point is the interferometric measurement of phase by use of a two-photon probe. Although this quantum probe offers a sensitivity advantage (super-sensitivity) under ideal conditions, if the probe is subjected to weak decoherence, then the sensitivity will be globally reduced and – surprisingly – lost altogether at certain phase values (blind spots). While the quantum advantage may be partially regained by shifting the phase to a less vulnerable value, the insertion of a precisely known reference phase into the interferometer may not be feasible.

We have investigated an alternative approach based on supplementing the path DoF of the
probe with an ancillary DoF (polarization), and creating an entangled quantum state in a Hilbert space of twice the dimensionality. We have demonstrated that the diversity added into the probe can help avoid the blind-spot predicament. This of course requires tweaking polarization parameters of unitary transformations at both the input and output ports of the interferometer before detecting the outgoing photon-pair. These transformations tailor the input state and the corresponding output detection strategy for optimal estimation. In this paper, the number of these adjustable parameters was limited to four, two at the input of the interferometer and two at the output. Since values of the parameters that maximize the sensitivity depend on the unknown phase itself, the optimization must be conducted adaptively. Based on extensive simulation of the adaptive process, we conclude that for a range of decoherence strengths, super-sensitivity is indeed obtained for any phase.

While the example investigated in this paper uses path and polarization as the principal and ancillary DoFs, respectively, other DoFs may also be used, as long as implementation of the prerequisite unitary transformations are practical. Likewise, we expect this approach to find application in the larger domain of optical parameter estimation. In the more general task of estimating a unitary transformation, rather than a phase, it is expected that blind spots will appear in the estimation of any of the parameters that encode the transformation: in this case, a binary DoF offering a secondary channel for enhanced estimation may be the only way to overcome these blind spots.
CHAPTER 5: ANCILLARY POLARIZATION USE IN A COLLINEAR IMAGE INVERSION INTERFEROMETER

In [13], we present a collinear common-path image-inversion interferometer using the polarization channels of a single optical beam. Each of the channels is an imaging system of unit magnification, one positive and the other negative (inverted). Image formation is realized by means of a set of anisotropic lenses, each offering refractive power in one polarization and none in the other. The operation of the interferometer as a spatial-parity analyzer is demonstrated experimentally by separating even- and odd-order orbital angular momentum modes of an optical beam. The common-path configuration overcomes the stability issues present in conventional two-path interferometers, and serves as an example of how the ancillary degree of freedom found in polarization can help in spatial mode-analysis.

5.1 Introduction

From optical communication and information processing to imaging and metrology, there has been a growing need for analysis and synthesis of the spatial structure of the optical field. Information transmission via spatial modes is based on operations such as modulation for multiplexing, projections for demultiplexing and compression, and filtering and correlation for recognition and classification [86, 87, 88, 89, 90]. Interferometry plays a major role in such operations.

Optical Interferometry is based on mixing a reference optical wave with a delayed, displaced, or rotated version of itself; a comparison of the two enables differential measurement with sub-wavelength resolution. The interferometer is typically configured to generate both the
sum and the difference of the reference and modified waves, whose intensities form two complementary interferograms. A common background in these interferograms may be removed by simple subtraction. This principle is utilized in the balanced homodyne detector, which is widely used for classical and quantum measurement [48]. In the context of spatial structures, the integrated intensities of the sum and difference fields provide binary measures, compressing a simple image into two projections.

A more recent entry into the interferometry toolbox is the image-inversion interferometer. This is based on mixing a reference image with an inverted copy of itself, so that the interferometer is sensitive to inversion symmetry. By producing spatial fields equal to the sum and difference between an original distribution and its inverted copy, the interferometer effectively decomposes the field into its even and odd components, thus serving as a spatial parity analyzer [1]. For example, an image-inversion interferometer decomposes an optical beam containing a superposition of orbital-angular-momentum (OAM) carrying modes into the even-order and the odd-order OAM modes [91, 92]. As shown in previous contexts [93], this binary classification can be cascaded to provide finer classification of the OAM modes.

If each arm of the interferometer has a spatial filter with a given amplitude point spread function (APSF), then in the presence of image inversion in one arm, one output of the interferometer is filtered by a system with even APSF, while the other is filtered with odd APSF. The powers in the output beams provide projections on even and odd functions, which form a binary compression of the original field distribution. Such compression offers a useful platform for communication and information processing.

The image-inversion interferometer has also found applications in imaging and microscopy. If applied to an incoherent light field with each arm including an imaging system of narrow shift-invariant APSF, then the difference between the output intensities is itself an image of
the original intensity distribution with shift-variant response function centered at the origin and with narrower width, so that the system can be used as a scanner with improved spatial resolution [94, 95, 96, 97, 98].

Spatial parity analysis is also useful in quantum optical information processing. It has been shown that a single photon with spatial distribution in either an even or an odd one-dimensional spatial mode, or a superposition thereof, forms a qubit [99], which can be used for quantum logic. One such mode can be converted into the other by use of either a phase plate or spatial light modulator (SLM), and the image-inversion interferometer may be used as a modal analyzer [100, 2, 1].

Additionally, it has been shown that estimation of parameters of a spatial distribution, such as the separation of the diffraction-limited image of a two-point source, by means of projections onto a complete set of spatial modes can provide accuracies unobtainable through standard imaging [50, 38, 35]. Moreover, the projection onto a complete set of spatial modes is unnecessary, since projection onto just two spatial-modes (one even, one odd) offers significant enhancement over traditional measurements [45, 41, 36].

There is no question that the image-inversion interferometer is an important asset for both classical and quantum information processing.

In previous implementations of the image-inversion interferometer, conventional two-path configurations have been adopted with either an extra reflector in one of the arms, or with imaging systems offering upright imaging in one arm and inverted imaging in the other [101, 102, 36, 1]. In this configuration, the path lengths must be kept stable to within a small fraction of a wavelength, and this is not trivial. Stabilization has been addressed by use of post-selection of data based on long scans from which projection fidelity is inferred. Clearly, these methods require a surplus of auxiliary photonic resources, either through
longer integration time or stronger illumination. In either case, especially in the context of quantum systems, the overhead required to implement the device may prove to outweigh the benefits gained by its application in the first place. Additionally, the projections necessary for modal analysis have relied on holographic methods that are often inefficient, especially in the context of quantum applications of OAM [92, 103, 104, 105, 106].

This paper introduces an alternative: a common-path image-inversion interferometer utilizing the two polarization channels. The beam splitters in the conventional Michelson interferometer are replaced with polarization analyzers and combiners. Image inversion in one of the channels is implemented by use of a set of anisotropic lenses providing upright imaging for one polarization and inverted imaging for the other. This collinear configuration obviates the stability limitations inherent in the conventional two-path interferometer.

The paper begins with an overview of the theory underlying the various applications of the image-inversion interferometer and proceeds to describe the polarization-based implementation and its experimental verification.

5.2 Theory

A conventional optical interferometer is a four-port system that mixes an optical field with a phase-shifted, time-delayed, or spatially-translated version of itself and generates the sum and difference fields at its output ports. Likewise, an image-inversion interferometer mixes a spatial field \( f(x) \) with an inverted version of itself \( f(-x) \), and generates at its output ports the fields \( \frac{i}{2}[f(x) + f(-x)] \) and \( \frac{1}{2}[f(x) - f(-x)] \). Here, \( x = (x, y) \) are coordinates in the transverse plane. The inversion operation \( f(x) \rightarrow f(-x) \) is easily implemented by reflection from a simple mirror, but other configurations involving production of an inverted image by
use of a lens may be used. It is also possible to include in the two interferometer arms other linear transformations $L_1$ and $L_2$ and a $(-\pi/2)$ phase shift, as illustrated in Fig. 1, so that the outputs are the fields

$$g_{\pm}(x) = \frac{1}{2}[L_1 f(x) \pm L_2 f(-x)].$$

(5.1)

As summarized below, a number of useful applications may be implemented by means of various choices of $L_1$ and $L_2$. For simplicity, we will present this summary for one-dimensional images $f(x)$, but the results are readily applicable to the two-dimensional case.
Parity analysis

In the simplest image-inversion interferometer, \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are identity operators so that the system generates from the input field \( f(x) \) two output fields with spatial distributions

\[
g_{\pm}(x) = \frac{1}{2} \left[ f(x) \pm f(-x) \right]
\]  

that are proportional to the even and odd components of \( f(x) \). The two arms of the interferometer will have to be perfect imagers, implemented, e.g., by 4-\( f \) optical systems. The interferometer serves as a parity analyzer that separates even and odd spatial distributions. It can be used as a classifier, or demultiplexer, separating a superposition of spatial modes into its constituent even and odd modes, as will be demonstrated experimentally in the next section. This type of parity-based modal analysis can be crucial in the context of the measurement of quantum-optical states that employ even and odd spatial modes in lieu of spin states [1, 2, 99, 107]. Also, by virtue of reciprocity, the interferometer can be used in reverse as a multiplexer, combining even and even modes into a single spatial pattern [98].

Projections onto even and odd functions

If the operators \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) represent multiplication by a prescribed function \( h(x) \) created, e.g., by an SLM, then the areas under the outputs \( g_{\pm}(x) \) provide the projections

\[
\alpha_{\pm} = \frac{1}{2} \int dx h(x) \left[ f(x) \pm f(-x) \right] = \int dx h_{\pm}(x) f(x),
\]  

Figure 5.1: Schematic of a conventional image inversion interferometer. An image inverter (INV) is placed in one branch and \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are linear systems used for various applications.
where \( h_\pm(x) = \frac{1}{2}[h(x) \pm h(-x)] \) are proportional to the even and odd components of \( h(x) \). Thus, \( \alpha_\pm \) represent the projections of the original function \( f(x) \) onto the even and odd functions \( h_\pm(x) \). For example, if \( h(x) \) is a function with support in the interval \([0, d]\), then \( h_\pm(x) \) are even and odd functions with support in the \([-d, d]\) interval.

Projections on even and odd spatial modes have been used in measurements optimized to estimate the separation between two incoherent point sources by measurement on the optical field they emit [36, 41].

**Fourier cosine and sine transforms**

If the operators \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) represent the spatial Fourier transform, as can be implemented by a single-lens in the 2-\( f \) configuration [47], then the interferometer outputs provide the Fourier cosine and sine transforms of the input:

\[
\begin{align*}
g_+(x) &= \frac{1}{2} \int dx \left[ f(x) + f(-x) \right] e^{-ikx} = \int dx \, f(x) \cos(kx), \\
g_-(x) &= \frac{1}{2} \int dx \left[ f(x) - f(-x) \right] e^{-ikx} = i \int dx \, f(x) \sin(kx).
\end{align*}
\]

If \( f(x) \) is a real function, then \( g_\pm(x) \) provide separately the real and imaginary parts of the Fourier transform, from which the phase can be calculated. Also, the Hartley transform \( \int dx \, f(x)[\cos(kx) + \sin(kx)] \) may be determined by adding up these transforms.

**Incoherent image-inversion interferometry**

If the operators \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) represent linear systems of APSF \( h_1(x; x') \) and \( h_2(x; x') \), then the interferometer produces the fields

\[
\begin{align*}
g_\pm(x) &= \frac{1}{2} \int dx' \left[ h_1(x; x') f(x') \pm h_2(x; x') f(-x') \right] \\
&= \int dx' \, h_\pm(x; x') f(x'),
\end{align*}
\]

70
where
\[ h_{\pm}(x, x') = \frac{1}{2}[h_1(x; x') \pm h_2(x; -x')]. \tag{5.7} \]

If \( f(x) \) is a random function representing a spatially incoherent optical field for which \( \langle f^*(x)f(x') \rangle = I_i(x)\delta(x-x') \), where \( I_i(x) \) is the optical intensity, then the average intensities at the outputs of the two branches of the interferometer are
\[ I_{\pm}(x) = \langle |g_{\pm}(x)|^2 \rangle = \int dx' |h_{\pm}(x; x')|^2 I_i(x'). \tag{5.8} \]

Here, \( |h_{\pm}(x; x')|^2 \) represent point spread functions (PSF) of the incoherent imaging systems in the two branches. In view of Eq. (5.7), the two images \( I_{\pm}(x) \) exhibit interference, and when subtracted, a new image \( I_o(x) = I_+(x) - I_-(x) \) is created for which the background terms are canceled out, and the cross terms remain,
\[ I_o(x) = \int dx' h_i(x; x') I_i(x'), \tag{5.9} \]
where
\[ h_i(x; x') = \text{Re} \{h_1^*(x; x')h_2(x; -x')\}. \tag{5.10} \]

The PSF in Eq. (5.10) is a product of a conjugated version of the impulse response function of the first system and an inverted version of that of the second system. If these impulse response functions are shift-invariant, i.e., functions of their coordinate differences, and assuming that they are identical functions \( h(x - x') \), then the overall system is shift variant with PSF
\[ h_i(x; x') = \text{Re} \{h^*(x - x')h(x + x')\}. \tag{5.11} \]

If \( h(x) \) is a narrow 1D function centered about \( x = 0 \), then \( h_i(x; x') \) will be a narrow 2D function centered near the origin, \( x = x' = 0 \), as illustrated in Fig. 5.2. To demonstrate that
this PSF offers enhanced resolution in a scanning configuration, consider as an example the Gaussian APSF \( h(x) = e^{(-x^2/\sigma^2)} \) for which the overall PSF in Eq. (5.11) is

\[
h_i(x; x') = e^{-2x^2/\sigma^2} e^{-2x'^2/\sigma^2}.
\]  

For a point at \( x' = 0 \) in the input plane, the response in the image plane is a Gaussian function \( e^{-2x^2/\sigma^2} \) with width smaller than that of \( h(x) \) by a factor of \( \sqrt{2} \). Additionally, if the point in the object plane is offset from the center of inversion by a distance \( \Delta \), then the measured image is again a Gaussian function \( e^{-2x^2/\sigma^2} \) but with amplitude reduced by a factor \( e^{-2\Delta^2/\sigma^2} \).

It has been shown that integrating over the interferometric transfer function’s image plane coordinate, which is accomplished simply by means of an integrating detector, offers a factor of 2 resolution improvement over direct imaging, and cancels radially symmetric aberrations [94, 95, 96, 101].

**Fourier transform with incoherent light**

A special case of the image-inversion interferometer described by Eq (5.10) is that for which \( h(x, x') = e^{i\pi(x-x')^2/\lambda d} \), which is the kernel for the Fresnel transform (propagation of light of wavelength \( \lambda \) through a distance \( d \) in free space in the paraxial approximation). In this case,
Eq. (5.10) gives
\[ h_i(x; x') = \cos(2\pi x x'/\lambda d), \] (5.13)
which is the kernel for the Fourier cosine transform. The image-inversion interferometer then produces an image \( I_o(x) \) that is the cosine transform of the intensity of the original image \( I_i(x) \). This is remarkable since the original image is incoherent. This lensless system may of course be implemented by use of a Fourier transform lens. This type of configuration has been previously used in two-path interferometers using corner cube mirrors to implement the image inversion [108, 109, 110, 111].

**Fourier transform with partially coherent light**

If the optical field \( f(x) \) is partially coherent with coherence function \( \Gamma(x', x'') = \langle f^*(x') f(x'') \rangle \), then
\[
I_o(x) = \text{Re} \int \int dx' dx'' h_1(x; x') h_2(x; -x'') \Gamma(x', x'').
\] (5.14)

For a system implementing the Fourier transform \( h_1(x, x') = h_2(x, x') = e^{i2\pi xx'/\lambda d} \) using a lens of focal length \( d \),
\[
I_o(x) = \text{Re} \int \int dx' dx'' e^{i2\pi (x' + x'')/\lambda d} \Gamma(x', x'').
\] (5.15)

so that \( I_o(x) \) is proportional to a one-dimensional (1D) Fourier transform of the two-dimensional function \( \Gamma(x', x'') \) calculated along the 135° direction in the \( (x', x'') \) plane. If \( \Gamma(x', x'') \) is sufficiently narrow so that \( \Gamma(x', x'') \approx I_i[1/2(x' + x'')]\gamma(x' - x'') \), then the output of the interferometer \( I_o(x) \) is a cosine Fourier transform of the input intensity \( I_i(x) \), as in Eq. (5.13), regardless of the actual shape of the complex degree of coherence \( \gamma(x) \).

By contrast, the output of a conventional interferometer is proportional to the 1D Fourier
transform of $\Gamma(x', x'')$ along the 45° direction, and $I_0(x)$ is a cosine Fourier transform of the input degree of coherence $\gamma(x)$.

5.3 The common-path interferometer

A single-path image-inversion interferometer may be implemented by using the two polarization components of the propagating field to carry the two interfering images, with one polarization producing an un-inverted image and the other an inverted image. The design requires use of anisotropic imaging components—refractive and/or diffractive. In this work we use an anisotropic optical element made of a combination of a polarization-sensitive diffractive waveplate (DW) and a conventional refractive lens.

The DW is designed to act on incident light in two ways. First, upon transmission, left-circularly polarized light experiences wavefront curvature identical to the effect of a positive lens of focal length $f_{dw}$, while right-circularly polarized light experiences the effect of a negative lens of focal length $-f_{dw}$. Second, right-circularly polarized light becomes left-circularly polarized, and vice-versa. The DW is placed in contact with a conventional lens with focal length $f_{cl}$. If $f_{cl} = f_{dw}$, then the effective focal length of the combined anisotropic lens pair (doublet) is $f_{db} = \frac{1}{2} f_{dw}$ for left-circular polarization, and $\infty$ for right-circular polarization. The anisotropic doublet thus has the effect of a lens acting on a single polarization, leaving the other unchanged.

The interferometer uses a cascade of 6 anisotropic doublets, 4 with focal length $f_{db} = 2''$ and 2 with focal length $f_{db} = 4''$, in the configuration shown in Fig. [3]. For this system, a left-circularly polarized image will see a cascade of two Fourier-transform imaging systems. Hence, in the output plane, the result will be an inverted image with unit magnification of
the spatial distribution in the input plane. Conversely, a right-circularly polarized image will see a cascade of four Fourier-transform imaging systems, resulting in an image with unit magnification and no inversion. Since the handedness of the polarization switches upon transmission through each doublet, a half wave-plate is placed in the middle of the system.

![Diagram of anisotropic doublets](image)

Figure 5.3: A set of six anisotropic doublets used as a common-path polarization-based interferometric spatial parity analyzer. Right-circular polarization sees a cascade of four Fourier-transforming imaging systems made of the lenses labeled "R", creating an uninverted image in the output plane. Meanwhile, left-circular polarization sees a cascade of two Fourier-transforming imaging systems made of the lenses labeled "L", creating an inverted image in the output plane. A polarization analyzer (not shown) generates the sum and difference of the two images, thereby separating the even and odd spatial parities of the input image.

The system is operated as an image-inversion interferometer by using a horizontally polarized optical field with spatial distribution \( f(x) \) in the input plane:

\[
E_1(x) = f(x) \frac{1}{\sqrt{2}} (\hat{e}_R + \hat{e}_L),
\]

(5.16)

where \( \hat{e}_R \) and \( \hat{e}_L \) are unit vectors corresponding to right and left circular polarization, respectively. Transmission through the imaging system transforms this field into

\[
E_2(x) = \frac{1}{\sqrt{2}} [f(x)\hat{e}_R + f(-x)\hat{e}_L],
\]

(5.17)

In a linear polarization basis with horizontal and vertical unit vectors \( e_H, e_V \), this is equal to

\[
E_2(x) = f_0(x)\hat{e}_H + f_e(x)\hat{e}_V,
\]

(5.18)
where \( f_0(x) = \frac{1}{2}[f(x) - f(-x)] \) and \( f_e(x) = \frac{1}{2}[f(x) + f(-x)] \) are the odd and even parts of \( f(x) \), respectively. Hence, simply analyzing the polarization components of \( E_2(x) \) by use of a polarizing beam splitter will also provide separation of the parity components of the input spatial distribution \( f(x) \).

Since this entire process occurs collinearly, acting on a single beam, the issues of alignment and path-length stabilization that have plagued previous implementations of image-inversion interferometers can be virtually alleviated.

### 5.4 Experimental verification

**Diffractive waveplate**

The diffractive waveplate (DW) lens (Pancharatnam-Berry phase lens) is constructed by depositing a liquid crystal (LC) polymer on a plastic substrate and exposing the polymer to polarized UV light with a special spatial pattern. This is accomplished in a multi-step process. First, the substrate is coated with photoalignment layer PAAD-72 (BEAM Co.) [112, 113, 114]. Next, the spatial pattern is created by exposing the photoalignment layer for 30 minutes with a power density of 8.8 mW/cm\(^2\). Finally, a liquid-crystal monomer solution is spin coated onto the photoalignment layer, and is photopolymerized using unpolarized ultraviolet light of 365-nm wavelength generated by a He-Cd laser. To form the lens, a parabolic phase distribution is created by giving each liquid crystal an orientation rotation that is proportional to the square of the radial coordinate, as measured from the center of the lens.

For a DW with diameter \( D \), and grating period \( \Delta \) defined as the period of the parabolic phase modulation at the edge of the DW, light of wavelength \( \lambda \) will experience the effect of
a lens with focal length \( f = D \Delta / 2 \lambda \). The final thickness of the DW is typically on the order of 1 micron. The lenses we use are designed for 100 mm and 200 mm focal length for 532 nm wavelength. The diffraction efficiency spectrum of the DW lenses can typically reach 99.7% efficiency at 532 nm.

The ease of this process has led to a number of manufacturing breakthroughs, allowing for broadband, highly efficient, and tunable DW lenses [115, 116]. Most importantly in the context of future implementation of our image-inversion interferometer, it is possible to coat the DW directly onto a refractive element, allowing for more robust, compact performance. The compactness of the DW offers an attractive advantage over implementing this system with bulkier optics such as spatial-light-modulators.

**One-dimensional parity analysis**

In the first experiment, we test the operation of the common-path image-inversion interferometer and assess its ability to faithfully separate the even and odd components of an optical beam with a one-dimensional spatial distribution. We use a spatial pattern

\[
 f(x) = \cos \left( \frac{\phi}{2} \right) \Phi_e(x) + i \sin \left( \frac{\phi}{2} \right) \Phi_o(x) \tag{5.19}
\]

that is a linear superposition of even and odd spatial functions, \( \Phi_e(x) \) and \( \Phi_o(x) \), of equal power. The optical powers \( P_V \) and \( P_H \) measured at the outputs of the polarization analyzer should be proportional to \( \cos^2 \frac{\phi}{2} \) and \( \sin^2 \frac{\phi}{2} \), respectively.

We found it convenient to use odd and even functions related by the equation \( \Phi_o(x) = [2H(x) - 1] \Phi_e(x) \), where \( H(x) \) is the Heaviside step function centered at \( x = 0 \), i.e., the odd function is simply obtained from the even function by introducing a phase shift of \( \phi = \pi \) for \( x < 0 \). It follows that \( f(x) = [e^{i\frac{\phi}{2}} H(x) + e^{-i\frac{\phi}{2}} (1 - H(x))] \Phi_e(x) \), i.e. \( f(x) \) is obtained from
the even function by adding a phase shift of $\frac{\phi}{2}$ in the right-half plane and $-\frac{\phi}{2}$ in the left-half plane.

The experiment was conducted by the use of a coherent Gaussian beam generated by 532-nm diode laser. Hence, $\Phi_e(x)$ is a Gaussian envelope. Phase modulation was implemented by reflecting the beam off the surface of a liquid-crystal-on-Silicon spatial light modulator (SLM). By having a fixed gray scale value on the left half of the SLM face and a tunable gray scale value on the right half, we were able to tune the phase $\phi$ and thereby vary the even-odd composition of the spatial distribution.

Measured values of $P_V$ and $P_H$ are plotted in Fig. 5.4 as functions of $\phi$ in the $[0, \pi]$ range. The visibility of this interferogram, which is a measure of the efficacy of the interferometer as a spatial parity analyzer, is 0.86. We believe that the imperfect visibility comes from the diffraction efficiency of our anisotropic doublets. Each DW acts on circularly polarized light as either a positive or negative lens, but the process is not perfect. Along with the focused and diffracted beam paths, there is a small amount of light that remains unmodulated. Although the fraction of light that is not diffracted is small, this effect occurs at each of the six DWs used in our system. We believe that this effect can be improved in the manufacturing process, as depositing the DW substrate directly onto the refractive element would allow for the use of DWs with better diffraction efficiency.
Figure 5.4: Powers $P_V$ and $P_H$ of the vertically and horizontally polarized outputs of the polarization analyzer for an input spatial distribution containing a superposition of even and odd functions with amplitudes $\cos^2 \frac{\phi}{2}$ and $\sin^2 \frac{\phi}{2}$ as a function of $\phi$, the phase difference between the faces of the SLM. At each value of $\phi$, $P_V$ and $P_H$ are divided by their sum to normalize for the effect of variations of the SLM reflectance at different values of $\phi$.

**Two-dimensional parity analysis**

As an example of the operation of the system for two-dimensional spatial fields, we consider an optical beam in a superposition of orbital-angular-momentum (OAM) distributions (Laguerre-Gauss modes). The mode of order $\ell$ has the distribution $A^{(\ell)}(r) \exp(-i\ell\phi)$, where $(r, \phi)$ are the polar coordinates and $A^{(\ell)}(r)$ is the radial distribution, so that the image inversion operation $f(x) \rightarrow f(-x)$ in this case is equivalent to multiplication by a phase factor $e^{i\ell\pi} = (-1)^\ell$. Thus, the even-order OAM modes have even distributions and the odd-order modes have odd distributions. The interferometer will therefore serve as an OAM-parity analyzer.

We have experimentally tested the operation of the interferometer in this context by modulating our Gaussian beam $\Phi_i(x)$ with a phase profile $\exp(-i\ell\phi)$ by use of a spatial light modulator (SLM) with vortex phase patterns. We used phase profiles with OAM values
ranging between $\ell = -4$ and $\ell = 4$. The optical powers $P_V$ and $P_H$ at the outputs of the interferometer should hence be proportional to the powers in the even and odd modes of the input beam, respectively. Ideally, $P_V$ should alternate between a high value and zero as $\ell$ alternates from even to odd, and vice-versa for $P_H$.

The results of this measurement are plotted in Fig. 5.5. The ratio $P_V/P_H$ for the even modes, or $P_H/P_V$ for the odd modes, is a measure of selectivity of this mode-parity classifier. This ratio equals 6.66 for $\ell = 0$ and drops to 3.51 and 3.11 for $\ell = \pm 4$, respectively. The fact that this ‘eye diagram’ is less open for higher order modes may be attributed to either the creation of the spatial distributions, or the alignment and diffraction efficiency of the interferometer. For higher order modes, misalignment in either the SLM center or the image-inversion interferometer’s optic axis greatly affects the measurement, since the variation in phase due to small deflections off the beam’s center becomes greater.
Figure 5.5: Powers $P_V$ and $P_H$ of the vertically and horizontally polarized outputs of the polarization analyzer for OAM modes of order $\ell$. As $\ell$ alternates between even and odd, the power switches between the vertical- and horizontal-polarization channels. For each mode, the sum of these powers have been normalized to unity in order to account for variation in the strength of the vortex singularity on the face of the SLM that generates these modes.

5.5 Discussion

Image-inversion interferometry is a versatile tool with potential utility in several areas of optics, including spatial-mode analysis, high-resolution microscopy, and optical image processing. The lack of its widespread use in research and commercial applications may be attributed to the stringent requirements on balancing and pathlength stabilization of the conventional two-path interferometer, which often makes the advantages offered by the device not worth the trouble. The new configuration presented in this paper obviates these challenging requirements since it uses interference of the polarization modes of a single-path optical beam.
The principal challenge in the polarization-based interferometer is the implementation of image inversion in one polarization mode and not the other. We addressed this requirement by means of an anisotropic imaging system that provides upright imaging for one polarization, and inverted imaging for the orthogonal polarization, at the same magnification. This required the use of anisotropic lenses, which we implemented by use of conventional lenses coated with specially designed anisotropic diffractive waveplates. Our design employed a set of six such lenses in two groups: two arranged as a 4-f system acting on one polarization to produce an inverted image; and four, each of half the refractive power, arranged as an 8-f system acting on the other polarization to provide an uninverted image. The lenses are concatenated such that the two images coincide. Evidently, other configurations, possibly with fewer lenses, may also be used to implement the requisite anisotropic imaging. Also, reflective optical elements may substitute refractive components.

We have tested the operation of the common-path interferometer as a spatial-parity analyzer—a demultiplexer of one-dimensional even and odd spatial modes. We have also demonstrated the use of the interferometer to classify orbital angular momentum modes of even and odd order up to ±4. Based on this demonstration, we expect the system to be developed further and to find home in support of many optical toolboxes.
CHAPTER 6: Quantum-limited estimation of the phase gradient

In work currently in peer-review, I showed that the quantum Cramér-Rao bound on the error of measurement of the optical phase gradient with a beam of finite width (or the wavefront tilt within a finite aperture) is consistent with the Heisenberg uncertainty principle for a single-photon state, and is a factor of $N$ lower for the maximally entangled $N$-photon state. This fundamental bound therefore governs the trade-off between quantum sensitivity and spatial resolution. Error bounds for a structured configuration using binary projective-field measurements implemented by an image-inversion (I-I) interferometer are higher, and the factor of $N$ advantage attained by the $N$-photon entangled state is reduced and eventually washed out as the beam width or the phase gradient increase. This reduction is more rapid for larger $N$, so that the quantum advantage is more vulnerable. The precision of the I-I interferometer is greater than that based on a split detector placed in the focal plane of a lens.

6.1 Introduction

The performance of classical optical metrology is limited by standard classical limits on the precision of measurement of optical phase and amplitude. Quantum metrology uses optical probes in nonclassical states of light, enabling precision superseding those classical limits; nonetheless, new superior limits emerge as the ultimate quantum limits. For example, the standard quantum limit for estimation of the optical phase with an average of $N$ photons is $1/\sqrt{N}$ [48], when the ultimate precision limit for a fixed number of photons $N$ (Heisenberg quantum limit) is $1/N$ [117, 118, 58, 57, 62]. In this paper, we determine the quantum limit on the precision of measurement of the optical phase gradient, which is manifested by a local
tilt of the optical wavefront within a finite aperture, or the angle of deflection of an optical beam of finite width introduced upon transmission through a thin slab with a spatially varying refractive index. We consider single-photon and multi-photon entangled quantum states and determine quantum Cramér-Rao (QCR) precision bounds based on the quantum Fisher information (QFI) [33, 34, 118, 119, 56, 63, 49]. As expected, these limits are inversely proportional to the beam width, in accord with the Fourier-transform uncertainty principle [120] and its generalization to a spatially-coded two-photon state [59, 121, 122] that probes the phase slope in a manner similar to that of states used in ”N00N-state” interferometry [67, 123, 124, 77].

Specific measurement configurations aim at reaching the precision limits set by the QFI, but do not always attain these ultimate limits. One evident configuration is to convert the beam tilt into beam displacement by use of an optical Fourier-transform imaging-system [47], as is usually done in the Shack-Hartmann system used in adaptive optics [125]. Standard measurements of beam displacement employ a split detector measuring detected photons in each of its two halves. Earlier studies of precision limits on estimates of displacement show that such configuration falls short of the quantum optimal precision by a factor of $\sqrt{\pi/2} \approx 1.25$ [126, 127]. Although strategies that address this shortfall have been demonstrated, their implementation involves heterodyning of the displaced beam with an independent local oscillator in a specific spatial mode.

We consider here an alternative configuration: an image-inversion (I-I) interferometer [13, 36, 128, 96] that utilizes interference between the phase-modulated beam and a spatially inverted copy of itself, and detects projections of the even component of the optical field distribution in one output port and that of the odd component in the other port. The Fisher information (FI) for this configuration is greater than that of the split-detector and attains the quantum optimal precision for small phase gradients or narrow beam widths for
both single-photon and multiphoton state implementations. However, as we show in this paper, the quantum advantage is reduced and eventually washed out as the phase-gradient beam-width product increases, and this deterioration is more severe for larger $N$, so that the greater the quantum advantage, the more vulnerable it is to the beam width and tilt.

### 6.2 Quantum Fisher information for single- and $N$-photon states

An optical beam probes a phase object that introduces a phase $\phi(x)$ in the plane orthogonal to the beam direction. The beam with is assumed sufficiently narrow so that $\phi(x) \approx \phi_0 + \theta x$, where $\phi_0 = \phi(0)$ and $\theta = \partial \phi / \partial x \big|_{x=0}$. The phase gradient $\theta$ is to be estimated by use of measurements on the transmitted beam. This study is also applicable to measurement of the direction of an optical wave within a finite aperture.

If the quantum state of the light transmitted through the phase object is described by a pure state $|\psi\rangle$, then the QFI is [129]

$$F_Q(\theta) = 4 \left( \langle \psi' \mid \psi' \rangle - \left| \langle \psi \mid \psi' \rangle \right|^2 \right),$$

(6.1)

where $\psi'$ refers to the derivative of $\psi$ with respect to $\theta$. The QCR bound on the variance of the estimate of $\theta$ is $\sigma_\theta^2 = 1/F_Q(\theta)$. In this section, we determine $F_Q(\theta)$ and the associated error bound $\sigma_\theta$ for light in two cases: single-photon, which corresponds to the limit imposed on any classical state of light, and an $N$-photon pure quantum state, which corresponds to the limit achievable by any state of light that seeks to use entanglement to aid estimation.

**Single-photon state**

If the single-photon state is a pure quantum state $|\psi_0\rangle = \int dx \ \psi_0(x) \mid x\rangle$, where $\psi_0(x)$ is
an arbitrary wavefunction normalized such that $\int dx \ |\psi_0(x)|^2 = 1$, then upon transmission through the phase object the state becomes

$$\langle \psi \rangle = \int dx \ e^{-i\theta x} \psi_0(x) \ |x\rangle.$$

(6.2)

Based on Eq. (6.1), the QFI is

$$F^{(1)}_Q(\theta) = 4 \left\{ \int dx \ x^2 |\psi_0(x)|^2 - \left| \int dx \ x |\psi_0(x)|^2 \right|^2 \right\}. \quad (6.3)$$

If $\psi_0(x)$ is an even function, then the second term in (6.3) vanishes, and

$$F^{(1)}_Q(\theta) = 4\sigma_x^2, \quad (6.4)$$

where $\sigma_x^2 = \int dx \ x^2 |\psi_0(x)|^2$ is the second moment of the probability density function $|\psi_0(x)|^2$ and $\sigma_x$ is a measure of its width. The QCR bound on the variance of the estimate of $\theta$ is

$$\sigma^2_\theta = 1/F^{(1)}_Q(\theta),$$

so that

$$\sigma_\theta \sigma_x = \frac{1}{2}. \quad (6.5)$$

Because the phase gradient $\theta$ equals the transverse component $q$ of the wavevector, this is simply an expression of the bound dictated by the Fourier-transform based uncertainty principle $\sigma_x \sigma_q = \frac{1}{2}$.

**N-photon state**

An $N$-photon pure quantum state is described by the integral $|\psi_0\rangle = \int d\mathbf{x} \ \psi_0(\mathbf{x}) \ |\mathbf{x}\rangle$, where $\mathbf{x} = x_1, x_2, \ldots, x_N$, $d\mathbf{x} = dx_1 dx_2 \ldots dx_N$, and $\psi_0(\mathbf{x})$ is an arbitrary $N$-photon wavefunction normalized such that $\int d\mathbf{x} \ |\psi_0(\mathbf{x})|^2 = 1$. Upon transmission through the phase object, the
state becomes

\[ |\psi\rangle = \int \! dx \, \psi_0(x) e^{-i \Theta (\Sigma_x)} |x\rangle, \]  

(6.6)

where \( \Sigma_x = \sum_{n=1}^{N} x_n \). Using Eq. (6.1), the QFI is

\[ F_Q^{(N)}(\Theta) = 4 \int \! dx \, (\Sigma_x)^2 |\psi_0(x)|^2 - 4 \left| \int \! dx \, (\Sigma_x) |\psi_0(x)|^2 \right|^2. \]  

(6.7)

Assuming a maximally entangled state \( \psi_0(x) = f_0(x_1) \Pi_{n=2}^{N} \delta(x_1 - x_n) \), i.e.,

\[ |\psi\rangle = \int \! dx \, f_0(x) e^{-iN\Theta x} |x\rangle^{\otimes N}, \]  

(6.8)

and if \( f_0(x) \) is an even function, then the second term of Eq.(6.7) vanishes and the QFI for the \( N \)-photon state is

\[ F_Q^{(N)}(\Theta) = 4N^2 \sigma_x^2 = N^2 F_Q^{(1)}, \]  

(6.9)

where \( \sigma_x^2 = \int \! dx \, x^2 |f_0(x)|^2 \) is a measure of the width of \( |f_0(x)|^2 \). Therefore, the minimum uncertainty \( \sigma_\theta \) of estimates of the phase-gradient satisfies the relation

\[ \sigma_\theta \sigma_x = \frac{1}{2N}. \]  

(6.10)

The bound for the spatially entangled \( N \)-photon uncertainty product is therefore smaller than that of the single-photon case by a factor of \( N \), assuming equal widths of the functions \( |\psi_0(x)|^2 \) in the single-photon case and \( |f_0(x)|^2 \) in the \( N \)-photon case.

### 6.3 Fisher Information for Special Measurement Configurations

We now consider specific configurations for measuring the phase gradient and assess their optimal precision in comparison with the ultimate quantum bounds described by (6.9) and
Split Detector in Focal Plane

The split detector is a two-sided detector that measures the lateral displacement of an optical beam by detecting the optical power on each side. When a beam described by a symmetric optical field \( \psi_f(x) \) is centered on the detector, the intensity \( |\psi_f(x)|^2 \) on each half will be equal and the difference of the photon counts will be zero, on average. This changes, however, if the beam is displaced by some distance \( s \). The powers in the two detectors are then

\[
P_+ = \int_0^\infty dx \ |\psi_f(x-s)|^2, \quad P_- = \int_{-\infty}^0 dx \ |\psi_f(x-s)|^2, \tag{6.11}
\]

and the power difference \( P_+ - P_- \) can be used to infer the displacement \( s \). For a monochromatic Gaussian beam \( \psi_0(x) \) with modulus \( |\psi_0(x)|^2 = (1/\sqrt{2\pi}\sigma_x) \exp(-x^2/2\sigma_x^2) \) modulated by a linear phase factor \( e^{-i\alpha x} \), a lens of focal length \( f \) produces in the focal plane a Gaussian field \( \psi_f(x) \) of width \( \sigma_f = \lambda f/(4\pi\sigma_x) \) offset from the center by a distance \( s_\theta = (\lambda f/2\pi)\theta \), where \( \lambda \) is the wavelength. The apparatus used for measurement of beam displacement \( s \) can therefore be readily adapted to measurement of the beam tilt \( \theta \). The Fisher information (FI) for such an arrangement is

\[
F_{SD}^{(1)}(\theta) = \frac{8\sigma_x^2}{\pi \sigma_x^2 / \xi(\theta\sigma_x)}, \tag{6.12}
\]

where \( \xi(y) = e^{4y^2} [1 - \text{erf}^2(\sqrt{2}y)] \) and \( \text{erf}(y) \) is the error function. As illustrated in Fig. 6.1, \( F_{SD}^{(1)}(\theta) \) is a monotonic decreasing function of \( \theta\sigma_x \). It has its maximum value of \( 8\sigma_x^2 / \pi \) for \( \theta\sigma_x \ll \frac{1}{2} \) (or \( s_\theta \ll \sigma_f \)). This is a factor of \( \pi / 2 \) smaller than the standard quantum limit \( F_Q^{(1)}(\theta) = 4\sigma_x^2 \), corresponding to a sensitivity lower by a factor of \( \sqrt{\pi/2} \approx 1.25 \), as also noted in [127].
This shortfall of the split detector, which is applicable to the single-photon state (and hence the coherent state), also extends to other implementations aimed at surpassing the classical estimation limit by use of other nonclassical states [126, 130]. For example, for light in a squeezed state with optimal mean photon number $N$, the FI is $\frac{8}{\pi} \sigma_x^2 N^{3/2}$, as compared to the quantum Fisher information $4\sigma_x^2 N^2$ for the maximally entangled $N$-photon state (cf. Eq. 6.9).

Figure 6.1: Fisher information for estimation of the phase gradient $\theta$ using an optical beam of width $\sigma_x$ in a single-photon state by use of a split-detector configuration (green dotted line) and an image-inversion (I-I) interferometer (solid blue line). For the I-I interferometer, the Fisher information reaches the quantum Fisher information (dashed lines) as $\theta \rightarrow 0$ (dashed line), while it is smaller by a factor $2/\pi$ for a configuration using a split-detector in a lens’ focal plane. The single-photon state corresponds to the maximum sensitivity of any classical illumination.

**Image-Inversion Interferometer**

In the image-inversion interferometer, the beam modulated by the phase object is interrogated by an interferometer with an image inversion element (i.e., a mirror) in one arm, as illustrated conceptually in Fig.1(b) [95, 96]. For an optical beam of amplitude $\psi_0 (x)$ and width $\sigma_x$, the beam transmitted through (or reflected from) the phase object has an amplitude $\psi (x) = \psi_0 (x) e^{i\theta x}$ which is mixed with an inverted copy of itself $\psi (-x)$ to generate amplitudes $\frac{1}{2} [\psi (x) \pm \psi (-x)]$ at the output ports of the interferometer. The interferometer can be
made using spatially-separated paths, as depicted in Fig. 6.2(b) or implemented in another ancillary binary degree of freedom such as polarization [8, 38, 96, 13]. The corresponding intensities \( I_\pm(x) = \frac{1}{2} |\psi(x) \pm \psi(-x)|^2 \) are measured with two detectors of areas greater than the beam cross-section \( \sigma_x \). The result is the two signals \( P_\pm = \frac{1}{2} \pm \frac{1}{2} \text{Re} \int dx \, \psi_0^*(x) \psi_0(-x) e^{i2\theta x} \), where we have assumed that \( \int |\psi_0(x)|^2 dx = 1 \). In essence, this binary measurement represents projections of the spatial distribution onto its even (+) and odd (−) components.

If \( \psi_0(x) \) is an even function, then

\[
P_+ = \int dx \, |\psi_0(x)|^2 \cos^2(\theta x), \\
P_- = \int dx \, |\psi_0(x)|^2 \sin^2(\theta x).
\]

(6.13)

For example, for a Gaussian function \( |\psi_0(x)|^2 = (1/\sqrt{2\pi} \sigma_x) \exp(-x^2/2\sigma_x^2) \),

\[
P_\pm = \frac{1}{2} \left( 1 \pm e^{-2\theta^2 \sigma_x^2} \right),
\]

(6.14)

and the difference \( P_+ - P_- = \exp(-2\theta^2 \sigma_x^2) \) is a monotonic decreasing function of \( \theta \sigma_x \) that can be readily used to calculate the phase gradient \( \theta \).

Figure 6.2: Measurement of the optical phase gradient by use of an image-inversion interferometer.
Single-photon Fisher information. If the probe wave is in a single-photon state, then the above classical analysis is applicable with the signals $P_+$ and $P_-$ interpreted as the probabilities of the photon being detected in the + and − output ports, respectively. The FI associated with such measurement is

$$F^{(1)}(\theta) = \frac{1}{P_+} \left( \frac{dP_+}{d\theta} \right)^2 + \frac{1}{P_-} \left( \frac{dP_-}{d\theta} \right)^2.$$  \hspace{1cm} (6.15)

Using the expressions in (6.14), it follows that the FI is

$$F^{(1)}(\theta) = \frac{4\sigma_x^2}{\zeta^2(\theta\sigma_x)},$$  \hspace{1cm} (6.16)

where $\zeta^2(y) = [\exp(4y^2) - 1]/4y^2$ is a monotonically increasing function of $y$ with value equal to 1 for $y = 0$ and $\approx 1.7$ for $y = \frac{1}{2}$. Therefore, in the limit $\theta\sigma_x \ll 1$, i.e., when the phase varies slowly within the beam width, the factor $\zeta(\sigma_x\theta) = 1$, so that $F^{(1)}(\theta) = F_Q^{(1)}(\theta)$, i.e., the I-I interferometer provides the best possible precision for estimating $\theta$. For a fixed value of $\sigma_x$, as $\theta$ increases, $F^{(1)}(\theta)$ drops as depicted in Fig. 6.1, reaching one half of its maximum value at $\theta \approx 0.56/\sigma_x$, so that the larger the beam width is, the faster $F^{(1)}(\theta)$ drops as a function of $\theta$. Based on Eq. (6.16), for a given value of $\theta$, the FI as a function of $\sigma_x$ rises to a peak value at $\sigma_x \approx 0.632/\theta$ and drops with further increase of $\sigma_x$. The Cramér-Rao estimation error $\sigma_\theta$ corresponding to $F^{(1)}(\theta)$ satisfies the relation

$$\sigma_x\sigma_\theta = \frac{1}{2} \zeta(\theta\sigma_x)$$  \hspace{1cm} (6.17)

so that it rises above the minimum value of $1/2$ as $\theta$ increases.

The FI for the I-I interferometer and the split detector are compared in Fig. 6.1. For small $\theta$ the I-I interferometer is superior to the split detector by the largest factor, but this
advantage diminishes as $\theta$ increases, and the split detector becomes slightly more sensitive for $\theta > 0.74/\sigma_x$.

**N-photon Fisher information.** A generalized image-inversion interferometer acting on a phase modulated optical beam in the maximally entangled $N$-photon state in Eq. (6.8) is conceptualized to operate in three stages. In the first, the state is converted into a generalized $N00N$ state in the basis of the two orthogonal modes of the interferometer $|+\rangle$ and $|-\rangle$ (e.g., the upper and lower paths):

$$|\psi_1\rangle = \int dx f_0(x) e^{-iN\theta x} \frac{1}{\sqrt{2}} \left[ |+, x\rangle \otimes N |-, x\rangle \otimes 0 + |+, x\rangle \otimes 0 |-, x\rangle \otimes N \right].$$

(6.18)

In the second stage, spatial inversion is introduced in the $|-\rangle$ mode, generating the state

$$|\psi_2\rangle = \int dx f_0(x) \frac{1}{\sqrt{2}} \left[ e^{-iN\theta x} |+, x\rangle \otimes N |-, x\rangle \otimes 0 + e^{iN\theta x} |+, x\rangle \otimes 0 |-, x\rangle \otimes N \right],$$

(6.19)

where $f_0(x)$ was assumed to be an even function. In the third stage, the $|+\rangle$ and $|-\rangle$ modes are recombined at a beam splitter and the photon-number parity is measured at either output port [131, 132, 122], a measurement represented by the observable operator

$$\Pi = iN \sum_{k=0}^{N} (-1)^k |k, N-k\rangle \langle N-k, k|.$$  

(6.20)

The result of a parity measurement is $+1$ if the photon number detected in the measured mode is even, and $-1$ if it is odd, and the associated probabilities are $P_+$ and $P_-$, with
\( P_+ + P_- = 1 \) and \( P_+ - P_- = \langle \Pi \rangle \), following the standard approach for parity measurement [131]. Based on Eq.(6.15), the Fisher information can be expressed in terms of \( \langle \Pi \rangle \) as

\[
F^{(N)}(\theta) = \frac{\left| \frac{d\langle \Pi \rangle}{d\theta} \right|^2}{1 - \langle \Pi \rangle^2}.
\]

(6.21)

To determine an expression of \( \langle \Pi \rangle = \langle \psi_2 | \Pi | \psi_2 \rangle \) for the \( N00N \) state in (6.19) in terms of \( \theta \) we note that the only terms in the expression (6.20) of \( \Pi \) that contribute to \( \langle \Pi \rangle \) are \( k = 0 \) and \( k = N \), and assuming that \( N \) is even,

\[
\langle \Pi \rangle = \int dx |f_0(x)|^2 \cos(N\theta x).
\]

(6.22)

For a Gaussian function \(|f_0(x)|^2 = (1/\sqrt{2\pi\sigma_x}) \exp(-x^2/2\sigma_x^2)\),

\[
F^{(N)}(\theta) = 4N^2\sigma_x^2/\zeta^2(N\theta\sigma_x) = F_Q^{(N)}/\zeta^2(N\theta\sigma_x),
\]

(6.23)

where, as before, \( \zeta^2(y) = [\exp(4y^2) - 1]/4y^2 \).

Thus, in the limit \( \theta\sigma_x \rightarrow 0 \), \( F^{(N)}(\theta) = F_Q^{(N)} \), i.e., the quantum Fisher information is attained at small phase-gradient beam-width product. Based on Eqs. (6.23) and (6.16), it follows that

\[
F^{(N)}(\theta) = N^2F^{(1)}(N\theta).
\]

(6.24)

Hence, in comparison with the single-photon case, the maximum achievable Fisher information is greater by a factor of \( N^2 \), but drops from its maximal value with increase of \( \theta\sigma_x \) at a rate \( N \) times greater. This remarkable scaling relation, illustrated in Fig. 6.3, highlights both the precision enhancing power of the quantum advantage and its vulnerability to large beam width or aperture area.
Figure 6.3: Fisher information for estimation of the phase gradient $\theta$ using an optical beam of width $\sigma_x$ in a state with $N = 1, 2, 4$ photons, assuming measurements with an image inversion interferometer. In each case, the Fisher information reaches the quantum Fisher information (dashed lines) as $\theta \sigma_x \to 0$ (dashed lines).

Implementation of the I-I interferometer for an arbitrary $N$-photon state requires replacing the first beam splitter with a device that generates the generalized $N00N$ state $|\psi_1\rangle$ in Eq. (6.18) from the state $|\psi\rangle$ in Eq. (6.8). This may be accomplished by use of a wavefront-division component as depicted in Fig. 6.4. Here, the phase-modulated beam in the state $|\psi\rangle$ of Eq. (6.8) is split into two spatial modes: the positive spatial mode, which has all $N$ photons in the region $x > 0$, and the negative spatial mode, which has all $N$ photons in the $x < 0$ region. These modes are directed to the two arms of the interferometer (using, e.g., prisms or a spatial light modulator) so that the state is $|\psi_1\rangle$. After image inversion, the negative mode is converted into a positive mode, but with phase $\varphi(-x)$, so that the phase difference between the modes becomes $\varphi(x) - \varphi(-x) = \theta x$. After recombination at the second beam splitter, by measuring the parity of the detected photons, the sensitivity described by Eq. (6.24) is achieved.

While methods for creating the $N$-photon spatial entanglement required is an ongoing field of research, for $N = 2$ the state may be readily created by use of a process of collinear downconversion, which exhibits a high degree of spatial entanglement [133]. Also, for $N = 2$,
the parity of the photon number may be readily determined by use of a coincidence circuit. The I-I interferometer itself may also be implemented in polarization modes [13], rather than spatially separated path modes.

Figure 6.4: Wavefront-division image-inversion interferometer generating the generalized $N00N$ state. The detectors measure photon-number parity.

6.4 Conclusion

We have shown that, for a single-photon quantum state, the ultimate quantum bound on the precision of estimates of the phase gradient introduced by an optical element probed by a beam of finite width (or estimates of the tilt of an optical wavefront within a finite aperture) reproduces the Fourier-transform uncertainty principle. For an $N$-photon quantum state that is maximally entangled in the spatial domain, the quantum precision bound is superior by the familiar factor of $N$ and the uncertainty product is tighter by the same factor. Here, uncertainty is defined as a bound on the statistical accuracy – as dictated by the quantum Fisher information – of estimating the phase gradient, which corresponds to the transverse component of the optical field’s wave vector.

We have determined the sensitivity of a specific optical configuration for measurement of the phase gradient, namely the image-inversion (I-I) interferometer, and shown that it meets the
standard quantum limit for small phase-gradient beam-width product and a single-photon state (or a coherent state). It also attains the Heisenberg quantum precision limit for an $N$-photon quantum state with maximum spatial entanglement. The I-I interferometer achieves this super-sensitivity by utilizing interference between the phase-modulated beam and a spatially inverted copy of itself, along with binary projective measurements similar to those used in other recent applications [41, 39, 36, 37]. For the $N$-photon state the system is a $N00N$-like interferometer. Unfortunately, the precision drops rapidly as the phase-gradient beam-width product increases, and the rate of such drop is greater for larger $N$. This is another manifestation of the fragility of quantum super-sensitivity. While the regime of small beam width, which is desirable since it enables greater spatial resolution in scanning systems, preserves the quantum advantage, it corresponds to lower quantum sensitivity since the QFI is proportional to the squared beam width. An optimal beam width is inversely proportional to the phase gradient.

In comparison with a measurement configuration using a split detector placed in the focal plane of a Fourier-transform imaging system, the sensitivity of the I-I interferometer is superior by a factor of at least $\sqrt{\pi}/2 \approx 1.25$ for the single-photon state. Other nonclassical states have been considered for use with the split-detector configuration, but their sensitivity is also limited by the same factor. The performance of the I-I interferometer is actually similar to that of a homodyne detection system designed to measure beam displacement by use of even and odd spatially distributed signal and local oscillator beams [126].
In my most recent work, I am applying the use of an image inversion interferometer in the context of measuring a binary phase object. Just as we can approximate many phase estimation problems as the measurement of a phase gradient, we can likewise investigate the utility of the interferometer in the context of resolving a binary phase object, which has two discrete regions with a phase difference between them. As I show here, it turns out the interferometer is able to resolve this phase object, and to do so while taking advantage of entanglement found in two-photon optical states. Although experimental verification is in process, I present here an outline of the theory and experimental design.

7.1 Introduction

On-demand measurement of a spatially coded phase distribution remains crucial in a number of optical metrology applications. In many cases, it is sufficient to measure the contrast between the phase of two points sampled from a distribution, either assuming that the values at each point are discrete or that there is a linear gradient between the two [48, 47]. This treatment is found in both classical and quantum-aided metrological systems [83].

A typical application takes the form of differential contrast interference, wherein an optical beam is split into two paths through an imaging system and subsequently recombined. Interference between the paths traveled through the system results in a binary power measurement at the output of the imaging system, from which the phase contrast between the
paths can be inferred. This has been demonstrated in both classical metrological contexts as well as quantum meteorological contexts, which seek to exploit the sensitivity advantage offered to quantum states of light that entangle the spatial modes of multiple photons. While the use of this system is sufficient, useful and appropriate in many contexts, there are always a number of issues inherent in the use spatially-separating interferometers. Namely, path length stability and alignment can be problematic in in-situ and scanning applications.

In this work, we aim to obviate these issues by using only a single optical path that probes a phase target in a collinear arrangement. We achieve this by the use of a polarization-sensitive image inversion interferometer, which splits and modulates the polarization of the input, rather than splitting it into two paths. In a similar arrangement to that used in prior demonstration [13], the system uses an anisotropic imaging system that inverts the spatial distribution of one polarization mode and leaves an orthogonal mode untouched. By measuring the polarization at the output of the system, the electric field is separated into spatially symmetric and anti-symmetric beams.

In the context of optical metrology with classical illumination, by using an input that is in a completely symmetric spatial beam, transmission through a binary phase object centered on the beam will introduce an anti-symmetric component into the distribution of the electric field. This allows us to estimate the value of the phase contrast via the binary polarization measurement. Hence, the phase contrast becomes coded into the polarization of the electric field.

Furthermore, in the context of illumination by non-classical photonic states, by entangling the polarization and spatial mode of a multi-photon state, we can use an anisotropic imaging system to extract an estimate of the phase distribution with a sensitivity beating that allowed for classical states of light. This super-resolving advantage is found by looking at...
the correlations in polarization of the photon pair after they has probed the phase target and anisotropic imaging system. Here we demonstrate experimental results showing that the two-photon states provided by spontaneous parametric downconversion (SPDC) can be used to access super-resolution to measurement of the binary phase object while using an anisotropic image-inversion interferometer.

7.2 Theory

Classical estimation sensitivity

The state of the electric field of classical coherent illumination is given by

$$|\phi_0\rangle = \int dx \phi_0(x) |x\rangle$$, \hspace{1cm} (7.1)

where $\phi_0(x)$ is a spatial distribution normalized such that $\int dx |\phi_0(x)|^2 = 1$.

Upon transmission through a phase shift described by $\phi(x) = \theta H(x)$, where $H(x)$ is the Heaviside-step function with value $-1/2$ for $x < 0$ and $1/2$ for for $x > 0$, the state becomes

$$|\phi\rangle = \int dx \phi_0(x)e^{-i\theta H(x)} |x\rangle$$, \hspace{1cm} (7.2)

Assuming this field is horizontally polarized, our field can be written in the left and right circular polarization basis, using the states states $|L\rangle$ and $|R\rangle$, as

$$\int dx e^{-i\theta H(x)} \phi(x) \left( |R\rangle + |L\rangle \right) \frac{2}{2} |x\rangle$$, \hspace{1cm} (7.3)
With the left-circular polarization mode, $|L\rangle$, introducing image inversion, the state becomes

$$\int dx \phi(x) \left[ e^{-i\theta H(x)} |R\rangle + e^{i\theta H(x)} |L\rangle \right] |x\rangle.$$ 

In the H-V polarization basis, this state can be written as

$$|\phi_{HV}\rangle = \int dx \left[ \cos(\theta H(x)) |H\rangle + \sin(\theta H(x)) |V\rangle \right] \phi(x) |x\rangle. \quad (7.4)$$

Noting that $\cos(\theta H(x)) = \cos(\frac{\theta}{2})$ and $\sin(\theta H(x)) = \sin(\frac{\theta}{2})$, the probabilities of measuring a photon to be polarized horizontally or vertically is given by

$$p_H = \int dx |\phi(x)|^2 \cos^2 \frac{\theta}{2} = \cos^2 \frac{\theta}{2},$$

$$p_V = \int dx |\phi(x)|^2 \sin^2 \frac{\theta}{2} = \sin^2 \frac{\theta}{2}. \quad (7.5)$$

The Fisher information can then be calculated as

$$F(\theta) = \left( \frac{\partial}{\partial \theta} p_H \right)^2 \frac{p_H}{p_H} + \left( \frac{\partial}{\partial \theta} p_V \right)^2 \frac{p_V}{p_V} = 1, \quad (7.6)$$

meaning that the Cramér-Rao bound for estimates of $\theta$ are given by $\text{Var}_{CRB} \geq 1$. If this experiment is repeated by further illumination for a total of $N$ collected photons, the best possible estimation variance is thus given by

$$\text{Var}_{CRB}^{(N)} \geq \frac{1}{N}. \quad (7.7)$$
Two-photon state estimation sensitivity

A collinear two-photon pure quantum state is described by the integral

$$|\psi_0\rangle = \int \int dx_1 \, dx_2 \, \psi_0(x_1, x_2) \, |x_1, x_2\rangle. \quad (7.8)$$

Here, we assume $\psi_0(x_1, x_2)$ is an arbitrary two-photon wavefunction normalized such that

$$\int \int dx_1 \, dx_2 \, |\psi_0(x_1, x_2)|^2 = 1.$$ Upon transmission through the phase shift described by $\phi(x)$ the state becomes

$$|\psi\rangle = \int \int dx_1 \, dx_2 \, \psi_0(x_1, x_2) e^{-i\phi(x_1) - \phi(x_2)} \, |x_1, x_2\rangle. \quad (7.9)$$

Assuming a maximally entangled state $\psi_0(x_1, x_2) = f_0(x_1) \delta(x_1 - x_2)$ [133] and that $f_0(x)$ is an even function,

$$|\psi\rangle = \int dx \, f_0(x) e^{-i2\phi(x)} \, |x, x\rangle, \quad (7.10)$$

In order to take advantage of the spatial entanglement to increase our measurement sensitivity, we create a generalized two-photon $N00N$ state in the polarization modes of the interferometer $|R\rangle$ and $|L\rangle$. The result is a superposition state

$$\int dx \left[ e^{-i\theta} \psi_+ (x) |R, R\rangle + e^{i\theta} \psi_- (x) |L, L\rangle \right] |x, x\rangle. \quad (7.11)$$

With the left-circular polarization mode, $|L\rangle$, introducing image inversion, the state becomes

$$\int dx \left[ e^{-i\theta} |R, R\rangle + e^{+i\theta} |L, L\rangle \right] \psi_+ (x) \, |x, x\rangle.$$
The two channels are then combined with a regular beam splitter to produce the state

\[ \int dx \sqrt{2} \left[ \cos(\theta) |c\rangle + i \sin(\theta) |a\rangle \right] \psi_+(x) |x, x\rangle , \]

where \(|c\rangle = \frac{1}{\sqrt{2}} (|R, R\rangle + |L, L\rangle)\) and \(|a\rangle = \frac{1}{\sqrt{2}} (|R, L\rangle + |L, R\rangle)\) are correlated and anticorrelated states, respectively. It follows that the probability \(p_a\) of measuring one photon in each channel (anti-correlated outcome) and the probability \(p_c\) of measuring the two photons together in either channel (correlated outcome) are:

\[
\begin{align*}
    p_c &= \int dx 2|\psi_+(x)|^2 \cos^2 (\theta) = \cos^2 \theta, \\
    p_a &= \int dx 2|\psi_+(x)|^2 \sin^2 (\theta) = \sin^2 \theta.
\end{align*}
\]

These outcomes can then be measured using two-photon coincidence measurements.

The Fisher information for the two-photon measurement outcomes can then be calculated as

\[
F(\theta) = \left( \frac{\partial}{\partial \theta} p_c \right)^2 \frac{p_c}{p_c} + \left( \frac{\partial}{\partial \theta} p_a \right)^2 \frac{p_a}{p_a} = 4,
\]

meaning that the Cramér-Rao bound for estimates of \(\theta\) are given by \(Var^{(2P)}_{CRB} \geq 1/4\). Comparing this sensitivity the sensitivity shown in Eq. (7.7), we see that using two photons in the entangled state we have presented here provides estimates of \(\theta\) with a variance that is a factor of two smaller than the estimates provided by illumination that uses an average of two classically described photons. This improvement is the hallmark of quantum sensors.
7.3 Experimental verification

Anisotropic Image inversion interferometer
We construct our anisotropic image inversion interferometer using a set of 6 diffractive waveplates [116] and 6 refractive lenses. Diffractive waveplates operate as lenses with focal lengths $\pm f$ for right and left circular polarization, respectively. Placing a diffractive waveplate of focal length $\pm f$ in contact with a refractive lens of focal length $f$ results in a doublet with focal length $\frac{f}{2}$ for right-circular polarization, and no focusing power for the left-circular polarization. By using a combination of lenses acting on left and right-polarization (see supplement), we are able to construct imaging systems that affect one of the polarizations but not the other. Since these imaging systems are non interacting, we can overlap the object and image planes of each system, inducing interference between the imaging systems output in the image plane of the combined system. It is this interference that allows us to measure the odd and even components of the electric field distribution in the object plane of the system.

Two-photon state creation
The crucial step in our super-sensitive phase estimation protocol is to create the state given by Eq. (7.11) To achieve this, we use the process of collinear type 1 SPDC. This process is carried out by the use of a 100 mW pump operating at 405 nm (Coherent CUBE 405-100C), pumping a 0.5 mm thick BBO crystal cut for type 1 collinear downconversion. This produces the desired photon pair with polarization state $|H, H\rangle$, due to the strict phase-matching conditions imposed by the SPDC process. The pump is filtered from the downconverted photon pairs by the use of polarizers and band-pass filters with a cutoff wavelength of 780 nm. It is reasonable to assume a large degree of spatial entanglement given the thin length of the crystal.
To emulate a phase target, we use an SLMS with the phase pattern described by $\phi(x) = \theta H(x)$. The SLM has its liquid crystal axis oriented horizontally, so that each photon has its spatial distribution modulated as if the liquid crystals imposed an isotropic phase shift. Hence, the state of the photon pair is now given approximately by Eq. (7.10). To create the generalized $N_{00}N$ state, we rotate the polarization state of each photon by the application of a half-wave plate whose optic axis is oriented at an angle of 22.5$^\circ$ with respect to the horizontal. This creates leads to a state described by

$$|\psi\rangle = \int dx \ f_0(x) e^{-i2\phi(x)} |D, D\rangle |x, x\rangle,$$

(7.14)

where $|D\rangle = \frac{|H\rangle + |V\rangle}{\sqrt{2}}$. We next reflect this beam off of a second SLM with phase distribution given by $\phi(x) = \pi(1/2 + H(x))$, creating the state

$$\int dx \ [e^{-i\theta} \psi_+ (x) |D, D\rangle + e^{i\theta} \psi_- (x) |A, A\rangle] \ |x, x\rangle,$$

(7.15)

where $|A\rangle = \frac{|H\rangle - |V\rangle}{\sqrt{2}}$. After this reflection, we place a quarter-wave plate in the path of the beam, transforming the state into the desired state given by Eq. (7.11). The beam is then directed to the image-inversion interferometer.

In the image plane of the interferometer, a polarizing beam splitter is placed, with both of its outputs are then coupled into fibers, and each fiber is then coupled into a fiber beam-splitter. The output facets of this beam splitter are then sent to a pair of avalanche photo-diodes (Pacer SPCM-ARQH-15FC), and the time-coincidence of the resulting photo-count signals is measured using an FPGA. By measuring the coincidence rates between detectors, the value of $\theta$ can be inferred.
7.4 Progress so far

We have assembled and constructed the image inversion interferometer, but, as opposed to the interferometer constructed in [13], alignment and maximizing of interference visibility prove to be more difficult when using near-infrared wavelength illumination. Nonetheless, iterative improvement is achievable, and we are pursuing the performance of the previously constructed implementation. Once this is satisfactory, we are confident that application of the two-photon state will be straightforward.

Figure 7.1: Implementation of super-sensitive phase measurements. A two-photon state is created using spontaneous parametric downconversion. After probing the phase target, an anisotropic image inversion interferometer is used to measure the spatial distribution of the state, thereby allowing for inference of the distribution of the phase target.
CHAPTER 8: DISCUSSION

My dissertation has shown, that the realistic treatment of metrological, where partial coherence and other deleterious effects are a fact of life, require careful consideration of every photonic resource available to a measurement system. I have presented here that partial coherence, a characteristic of any real optical system, can have crippling effects on the sensitivity of measurement devices. Whether the ideal operation assumes perfect coherence or complete incoherence, deviations from this must be carefully considered in both theory and practice.

Towards building sensors that perform well under these conditions, my dissertation shows that it is possible to employ otherwise unused optical resources to increase the utility and usability of optical sensors. Namely, I have shown that it is possible to use polarization to make measurements of phase and spatial structures that might otherwise be impossible. Not only does this application find home in the context of classical optical sensors, it can be used in quantum-sensing scenarios, opening for attractive new avenues of sensing where entanglement in ancillary degrees of freedom (e.g. polarization) can be used to measure parameters encoded in degrees of freedom (e.g. spatial mode) in which entanglement is difficult to exploit.

Currently, the specific concept of using interferometric polarization measurements, where stability is inherent when optical paths are not split, appears to be a severely under-explored area of research. Fortunately, research and development in the applications of liquid crystal devices are making polarization-dependent wavefront modulation more straightforward than ever. I hope that insights provided by my research can help lead the way to better precision sensors that take advantage of ancillary photonic resources.
LIST OF REFERENCES


vol. 58, no. 4, 1968.


[22] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimu-
lated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics Letters*, 

[23] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, 
M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging Intracellular Fluoro-
rescent Proteins at Nanometer Resolution,” *Science*, vol. 313, no. 5793, pp. 1642–1645, 
2006.

fluorescence photoactivation localization microscopy,” *Biophysical Journal*, vol. 91, 


[26] C. W. Helstrom, “Resolvability of objects from the standpoint of statistical parameter 


