2010

New strategies for phase estimation in quantum optics

Yang Gao
Louisiana State University and Agricultural and Mechanical College, ygao1@tigers.lsu.edu

NEW STRATEGIES FOR PHASE ESTIMATION IN QUANTUM OPTICS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics and Astronomy

by
Yang Gao
B.S., Harbin Institute of Technology, Harbin, 2003
M.S., Peking University, 2006
August, 2010

Acknowledgments

First of all, I would like to thank my advisor Prof. Hwang Lee for his excellent guidance and support. His encouragement and unlimited patience in correcting the drafts of the papers are invaluable. I learnt a lot of the material covered in this dissertation through this way.

I would like to thank Prof. Jonathan P. Dowling for providing me an opportunity to work in his group. I was deeply influenced by his optimistic attitude towards physics and life. I thank him for forcing me give more presentations. I also appreciate his invaluable assistance on my personal problems.

I would like to thank Prof. Robert F. O'Connell for allowing me to work with him on quantum entanglement and quantum Langevin equation. I also thank him for showing me how research should be done.

I would like to thank Prof. Phillip Sprunger for being on my committee. I would also like to thank Prof. Bahadir Gunturk from Department of Electrical and Computer Engineering for being on my committee as deans representative.

At last my words can express my gratitude to my family. Without their lasting understanding and encouragement, I would not have come this far.

Table of Contents

List of Figures

Abstract

One important problem in quantum optics is to resolve an extremely small change of phase shift. The complementarity between photon number and phase sets an ultimate limit, the so-called Heisenberg limit, on the phase measurement sensitivity. The precise phase estimation has many technological applications, such as optical gyroscopes, gravitational wave detection, quantum imaging and sensing.

In this thesis I show that the utilization of the parity measurement in the optical interferometry is actually applicable to a wide range of quantum entangled input states. Comparison of the performance of the various quantum states then can be made within such a unified output measurement scheme. Based on such a universal detection scheme, we present a comparison of the phase sensitivity reduction for various quantum states of light in the presence of photon loss.

I also provide a simple condition that could be used to check whether an arbitrary state can achieve the Heisenberg limit independently of the detection scheme. It implies that the fidelity between the two output states with zero phase and minimal detectable phase applied respectively should significantly different from unity as the minimal phase shift scaling as $1/\langle N\rangle$, whereas the average number of input photons $\langle N\rangle$ goes to infinity.

Next I give several measures to characterize the which-way information in the interference experiment. We define a new distinguishability associated with the fidelity between two density matrices to measure the which-way information. We demonstrate that the changes of mutual entropy as well as the entanglement of formation of the whole system, i.e. the physical system plus the which-way detector can also be used to describe the which-way information. With such quantities, we show that as the fringe visibility of the interference pattern gets larger, the less which-way information is obtained.

Finally, I show that coherent light coupled with photon number resolving detectors can provide a super-resolution much below the Rayleigh diffraction limit, with sensitivity no worse than shot-noise in terms of the detected photon power. This scheme would have applications to laser

radar, given the difficulty in making entangled states of light, as well as their susceptibility to atmospheric absorption.

Chapter 1: Introduction

The birth of quantum mechanics was marked with Planck's assumption of the discrete energy levels of simple harmonic oscillators in the walls of black-body that was necessary to explain the black-body radiation spectrum. The generalization of these ideas led Einstein to explain the photoelectric effect, and to introduce the photon concept by analogy with particle. Then Dirac associated each mode of the radiation field with a quantized simple harmonic oscillator, which is the essence of the quantum theory of radiation or *quantum optics*, to combine the wave- and particle-like aspects of light [1, 2]. Through this way the radiation field or photon field is capable of explaining all interference phenomena as well as showing the excitation of a specific atom along a wave front absorbing one photon of energy. In this new formalism, photon is interpreted as the excitation of an appropriate normal mode of electromagnetic field. The old fashioned content of photon as a real particle like electron is thus given up.

A consequence of quantum optics is the zero-point fluctuations of photon fields. These fluctuations have no classical analog and are responsible for many phenomena in quantum optics. Practically a semiclassical theory of atom-field interaction in which only the atom is quantized and the radiation field is treated classically, with zero-point fluctuations, can account for many phenomena in modern optics, such as spontaneous emission, the Lamb shift, and the Casimir effect. However, quantum beat phenomena, quantum eraser, and the production of two-photon entangled states from a three-level atom in cascade configuration provide simple examples in which the results of quantum optics differ qualitatively from those obtained via a semiclassical theory. Further support of quantum optics comes from the experimental observations of nonclassical states of photon field, such as squeezed states, photon anti-bunching, and number states in cavity quantum electrodynamics [3].

The implementation of quantum optics in real situation is best summarized by Lamb [4]: *Begin by deciding how much of the universe needs to be brought into the discussion. Decide what normal*

modes are needed for an adequate treatment of the problem under consideration. Find a suitable approximation for the normal mode; the simpler, the better. Decide how to model the light sources and work out how they drive the wave function for the system. No more and no less!

All physical predictions thus obtained must be backed by measurements on certain real parameters of the investigated system. In quantum mechanics the description of a state of a physical system is radically different from its classical counterpart. A quantum state has an inherent probability, due to which an ultimate limit, i.e. the so-called *Heisenberg limit* (HL), on the sensitivity of all measurements are imposed. So it is more appropriate to estimate the parameters of the system via the results of measurements. In most cases, the adopted measurement methods fail to reach these limits. Conventional bounds to the sensitivity of measurements such as *shot-noise limit* (SNL) are not as fundamental as HL, and can be beaten through quantum strategies such as squeezing and entanglement [5].

Quantum optical metrology attempts to reach HL using photons. It is the quantum entanglement between the photons in various modes that react to the zero-point fluctuations in a correlated way and beat the SNL. The physical parameter that is of primary concern in the context of quantum optical metrology is the optical phase. Quantum mechanically an observable must be described by a Hermitian operator. However there is no Hermitian phase operator in quantum mechanics. This calls for an estimation scheme of phase through some intermediate methods. An important method that can detect the difference of phase shifts is to use interferometric phenomenon. This is can be achieved with Mach-Zehnder interferometer (MZI), as shown in Fig. 1.1, because the output signal is sensitive to the relative phase shift between two fields traveling down separated paths. Its ability to resolve extremely small relative phase shifts in the two paths finds applications in optical gyroscopes [6], gravitational wave detectors [7], quantum imaging and sensing [8, 9]. It is not only of great importance in such technological advances, but also a fundamental problem in quantum mechanics because phase measurement sensitivity is fundamentally limited by the phase uncertainty introduced by the fluctuations of photon fields in the two input ports. For separated single photons or coherent states, the phase sensitivity subject to the shot-noise limit $\Delta\varphi \sim 1/\langle N\rangle^{1/2}$, where $\langle N\rangle$ is the average photon number used in the

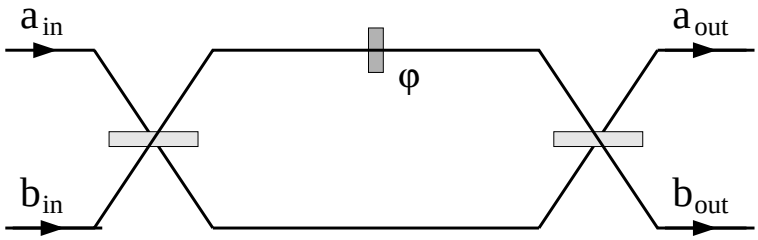

FIGURE 1.1. Schematic of MZI, which consists of two 50-50 beam splitters. The symbols a, b represent the photon annihilation operators. The angle φ denotes the relative phase difference between the two arms. Note that the NOON state is inserted to the right of the first beam splitter, and other states to the left.

given experiment. To measure a small phase shift with high precision, the measurement time or the strength of the incident laser resource has to be increased to a huge level. As we said before, a reasonable way to beat the shot-noise limit is to entangle the photons together, due to the fact that entangled states react to the fluctuations in a correlated way. Caves first showed how to beat SNL by squeezed states [10]. The complementarity between photon number and phase sets a fundamental limit of $\Delta\varphi \sim 1/\langle N\rangle$, i.e. HL, which represents an improvement of $\langle N\rangle^{1/2}$ over the shot-noise limit. For example the states such as $|N, N\rangle$, $\alpha|N, N\rangle + \beta|N+1, N-1\rangle$ with $\alpha^2 + \beta^2 = 1$, and $(|N, 0\rangle + |0, N\rangle)/\sqrt{2}$ with $|N_a, N_b\rangle = |N_a\rangle|N_b\rangle$ and $|N_{a(b)}\rangle$ being the number state of the annihilation operator $a\,(b)$, which are also known as twin- [11, 12], Yurke- [13], and NOON- [14, 15] states respectively, can be inserted into the input ports of MZI to reach HL. The aim of this dissertation is to discuss several strategies and applications for phase estimation in quantum optics to achieve the HL. In Chapter 2, I will briefly review some basics of quantum optics that are pertinent to the subsequent chapters.

The utilization of above quantum correlated input states are accompanied by various output measurement schemes. In some cases the conventional measurement scheme of photon-number difference is used. A certain probability distribution, a specific adaptive measurement [16, 17], and the parity measurement (P), which gives "+" or "−" for even or odd photons at one output port, are used in other cases. In optical interferometry, Gerry first showed the use of the parity measurement for the NOON state to reach the exact HL [18]. Later the parity measurement scheme was suggested to be used for the twin state inputs by comparing the quantum state inside the interferometer with the NOON state [19]. The extension to all the known entangled states of photon then promotes the parity measurement to a universal detection scheme for quantum

interferometry. Furthermore, it leads to a great reduction of the efforts in precise quantum state preparation as well as in various optimization strategies involving quantum state engineering for the HL interferometry. In chapter 3, I will show that the utilization of the parity measurement in HL interferometry is actually applicable to a wide range of quantum entangled input states. Comparison of the performance of the various quantum states then can be made within such a unified output measurement scheme. Then, based on such a universal detection scheme comparisons of performance of various quantum states can be made in a common ground. As an example, we presented a comparison of the phase sensitivity reduction for various quantum states of light in the presence of photon loss. The calculations reveal that in the parity measurement scheme the NOON states show the best performance for phase detection and can still beat SNL if the transmittance of interferometer is not too small and the photon number is not too large.

However, it needs to be pointed out that the correlated states used in [20] only have real coefficients, such as the particular Yurke state $(|N, N\rangle + |N+1, N-1\rangle)/\sqrt{2}$. If the imaginary coefficients are allowed, the phase sensitivity with parity measurement will blow up at some points such as the Yurke state with $\beta/\alpha = i$. It is thus convenient to provide a condition that could be used to check whether a given state can achieve HL in principle. In [21, 22, 23] a lower bound for the minimum detectable phase shift based on the Cramer-Rao bound is given. In Chapter 4, I will show that the condition for one input state to reach HL can be expressed as the overlap between two possible output states with zero phase and the minimal detectable phase $\Delta\varphi$ applied respectively should be less than 1 as $\Delta\varphi \sim 1/\langle N\rangle$ and $\langle N\rangle \to \infty$, i.e.

$$\lim_{\langle N\rangle\to\infty} |\langle\Psi_{\rm out}(0)|\Psi_{\rm out}(\Delta\varphi)\rangle|_{\Delta\varphi\to 1/\langle N\rangle} < 1. \tag{1.1}$$

This condition gives more information than the Cramer-Rao bound, and applies to the single run of detection of phase shift as well as multiple runs. Here the two states should be understood as the output states for the overall measurements. It thus provides us a simple condition to determine whether a given state can reach HL or not.

MZI can also be used to illustrate the complementarity principle [24], which states that quantum systems inherit equally real but mutually exclusive properties. For example, an electron can behave either like a particle or like a wave depending on the experimental situation. The situa-

tions "perfect fringe visibility and no which-way information" and "full which-way information and no fringes" are well known for MZI. The intermediate stages were studied in [25], where a quantity called distinguishability ($\mathcal{D}$) to quantify the which-way information as well as the fringe visibility ($\mathcal{V}$) for interference pattern is defined. An inequality $\mathcal{D}^2 + \mathcal{V}^2 \leq 1$ is proved and thus the complementarity principle is verified for the intermediate cases. To study the intermediate cases from other points of view. In chapter 5, I will use three different measures, i.e. fidelity between two parts of the final density matrix of which-way detector, the change of the mutual entropy, and the entanglement between the particle and the which-way detector, to express the which-way information. With such quantities, I will give new proofs of the above inequality, and reinstate the notion that the complementarity principle is actually one side of the entanglement in the quantum system [26].

There has been much recent interest in quantum optical interferometry for applications to sub-wavelength imaging and remote sensing in laser radar (LADAR) [27]. Ever since the work of [28], it has been realized that by exploiting quantum states of light, such as NOON states, it is possible to beat the Rayleigh diffraction limit in imaging and lithography (super-resolution) while also beating SNL in phase estimation (super-sensitivity). However such quantum states of light are very susceptible to photon loss [29]. Recent work has shown that in the presence of high loss, the optimal phase sensitivity achieved is always $\Delta\varphi = 1/\langle N\rangle^{1/2}$, where $\langle N\rangle$ is the average number of photons to arrive at the detector [30]. These results suggest that, given the difficulty in making quantum states of light, as well as their susceptibility to loss, the most reasonable strategy for quantum LADAR is to transmit coherent states of light to maximize sensitivity [27]. It is well known that such an approach can only achieve at best shot-noise limited sensitivity [10]. However, such a conclusion leaves open the question as to what is the best detection strategy to optimize resolution. Recent experimental results and numerical simulations have indicated that such a coherent-state strategy can still be super-resolving, provided a quantum detection scheme is employed [31]. In chapter 6, I will derive a quantum detection scheme inspired by parity measurement, that is super-resolving, and which can be implemented with photon-number-resolving detectors. The proposed scheme exploits coherent states of light, is shot-noise limited

in sensitivity, and can resolve features by an arbitrary amount below the Rayleigh diffraction limit. This scheme would have applications to quantum optical remote sensing, metrology, and imaging.

In chapter 7, I will discuss some further problems related with the phase estimation in quantum optics.

Chapter 2: Some Basics of Quantum Optics

2.1 A Short Review of Quantum Mechanics

Quantum Mechanics divides the world into two parts, commonly called the system and the observer. Except at specified times the system and the observer do not interact. An interaction at those specified times is called a measurement. Quantum Mechanics predicts all the information that the observer can possibly obtain about the system. This information can be represented in different ways. It is often represented in terms of a wave function. A measurement changes the information an observer has about the system and therefore changes the wave function of the system. The basic assumptions of non-relativistic quantum mechanics are [32]:

✓ The full knowledge of a system that an observer can possibly obtain is characterized by a pure quantum state $|\Psi\rangle$ in the Hilbert space $\mathcal{H}$ with $\langle\Psi|\Psi\rangle = 1$.

✓ If $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are two possible quantum states, then $|\Psi\rangle = c_1|\Psi_1\rangle + c_2|\Psi_2\rangle$ is also a possible one.

✓ The measurement of an arbitrary physical quantity is described by a Hermitian operator A. The result of a measurement belongs to a set of eigenvalues $\{\alpha_n\}$. Each eigenvalue is associated with an eigenstate $|n\rangle$, namely $A|n\rangle = \alpha_n|n\rangle$.

✓ If the measurement corresponding to A is applied on the quantum state $|\Psi\rangle = \sum_n c_n|n\rangle$ associated with $\sum_n |c_n|^2 = 1$, the probability of obtaining result α_n is $|c_n|^2$. Therefore, the average value of A is given by $\langle A\rangle = \langle\Psi|A|\Psi\rangle$. After the measurement the quantum state becomes to be $|n\rangle$ immediately.

✓ The conjugate canonical observables $\{p, q\}$ satisfy the commutation relation $[p, q] = -i\hbar$.

✓ The evolution of quantum state is governed by the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = H|\Psi\rangle.$$

Here H is the Hamiltonian of the system.

✓ For a system with identical Bosons or Fermions, its quantum state is symmetric or anti-symmetric under the exchange of any two particles.

In real cases a system under investigation is usually surrounded by another external system, such as the thermal environment. The two systems as a whole are described by a pure state, but the system to which we can access is not pure any more. Formally the density matrix is introduced to characterized our incomplete information of the investigated system. The density matrix should satisfy the following properties constrained by physical requirements:

(1) ρ is positive, $\rho > 0$.

(2) ρ is Hermitian, $\rho^\dagger = \rho$.

(3) ρ is normalized, $\mathrm{tr}\rho = 1$.

These three properties guarantee that a density matrix can be immediately decomposed as $\rho = \sum_r p_r |r\rangle\langle r|$, where $|r\rangle$ is the eigenvector of ρ with eigenvalue p_r. Hence we may interpret ρ as describing an ensemble of pure quantum states, in which the state $|r\rangle$ occurs with probability p_r. Obviously we have $0 \leq p_r \leq 1$ and $\sum_r p_r = 1$. Moreover one can see $\rho^2 - \rho \leq 0$, where the equal sign is valid iff there is only one term in the above summation, i.e. $\rho = |\Psi\rangle\langle\Psi|$, which corresponds to a pure state. Otherwise the density matrix is a description of a mixed state. The expectation value of any observable $A = \sum_n \alpha_n |n\rangle\langle n|$ on the system can be expressed as

$$\langle A\rangle = \sum_r p_r \langle r|A|r\rangle = \mathrm{tr}A\rho. \tag{2.1}$$

The final state after the measurement is given by

$$\begin{aligned} \rho' &= \sum_r p_r \sum_n |\langle n|r\rangle|^2 |n\rangle\langle n| \\ &= \sum_n |n\rangle\langle n| \left(\sum_r p_r |r\rangle\langle r| \right) |n\rangle\langle n| \\ &= \sum_n E_n \rho E_n^\dagger, \end{aligned} \tag{2.2}$$

where the projector E_n is defined as $E_n = |n\rangle\langle n|$ and $E_n^2 = E_n$. This formalism of projective measurement is called Von Neumann scheme, which is just a specific case of the generalized

measurement theory. In the generalized measurement theory, each outcome has a probability that can be expressed as

$$p_r = \mathrm{tr} F_r \rho, \tag{2.3}$$

where the non-negative operator set $\{F_r = \Omega_r^\dagger \Omega_r\}$, namely *positive operator-value measure* (POVM), forms a partition of unity, $\sum_r F_r = 1$. A general POVM changes a quantum state ρ in the following way,

$$\rho \to \rho' = \sum_r \Omega_r \rho \Omega_r^\dagger. \tag{2.4}$$

On the other hand, according to *Neumark's theorem* [33], a POVM can be realized by extending the Hilbert space to a larger space, and performing projective measurement in the larger space.

As for the time evolution law of density matrix of an open system interacting with environment, the conventional Schrödinger equation is not applicable to the open system alone. Its equation of motion can be obtained by eliminating the variables of environment from the Schrödinger equation of the system and the environment. The resulting equation of motion is described by the so-called master equation of density matrix. The general form of master equation can be deduced from the following physical assumptions [33]:

(1) There exists a linear mapping $\$: \rho \to \rho'$ that takes an initial state ρ to a final state ρ'.

(2) $\$ preserves hermiticity: $\rho' = \rho'^\dagger$.

(3) $\$ is trace preserving: $\mathrm{tr}\rho' = 1$.

(4) $\$ is positive: $\rho' > 0$. More precisely, $\$ should be completely positive, which means that for any extension of $\mathcal{H}_A$ to the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$, the mapping $\$_A \otimes I_B$ is positive for all such extensions.

Under such conditions, Kraus proved that the action of $\$ on ρ has an operator-sum representation,

$$\$(\rho) = \sum_r \Omega_r \rho \Omega_i^\dagger, \qquad \sum_r \Omega_r^\dagger \Omega_r = 1. \tag{2.5}$$

It has the same form as the action of a POVM on state ρ. This observation allows us to interpret the master equation of an open system interacting with environment as how the system changes in time under continuous POVM. In general one can write

$$\begin{aligned} \rho(t+dt) &= \sum_r \Omega_r \rho(t) \Omega_r^\dagger, \\ \Omega_0 &= 1 - \left(\frac{i}{\hbar} H + \frac{1}{2} \sum_{r \neq 0} L_r^\dagger L_r \right) dt, \\ \Omega_r &= L_r \sqrt{dt}, \qquad r \neq 0, \end{aligned} \tag{2.6}$$

where we have used the Markovian approximation (the surrounding environment is nearly memoryless), and the operators $L_{r\neq 0}$ are called jump operators. Under the same condition,

$$\rho(t+dt) \approx \rho(t) + \dot{\rho}(t) dt, \tag{2.7}$$

we obtain the master equation in the Lindblad form,

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] + \sum_{r \neq 0} \left(L_r \rho L_r^\dagger - \frac{1}{2} L_r^\dagger L_r \rho - \frac{1}{2} \rho L_r^\dagger L_r \right). \tag{2.8}$$

In general it is very difficult to solve the master equation for real problems, even numerically. One possible way out is to transform the master equation into a stochastic schrödinger equation (SSE), which is much easier to solve or to simulate [35]. The formulation of SSE is to first define a conditioned state, not necessarily normalized, by

$$|\bar{\Psi}_r(t+dt)\rangle = \frac{\Omega_r(dt)}{\sqrt{\Lambda_r(dt)}} |\bar{\Psi}(t)\rangle, \tag{2.9}$$

where $\Lambda_r(dt)$ is the fictitious probability for result r such that $\sum_r \Lambda_r(dt) = 1$. The actual probability for result r is given by

$$p_r(dt) = \Lambda_r(dt) \langle \bar{\Psi}_r(t+dt) | \bar{\Psi}_r(t+dt) \rangle = \langle \bar{\Psi}(t) | \Omega_r^\dagger \Omega_r | \bar{\Psi}(t) \rangle, \tag{2.10}$$

which is independent of the fictitious probability $\Lambda_r(dt)$. In fact, the actual probability of a particular process or a *quantum trajectory* is given by the product of the fictitious probability of generating this trajectory and the norm of the conditioned state.

Physically the change of conditioned state produced by the jump operator $L_{r\neq 0}$ is not continuous, so it has to be expressed as a stochastic process $dN_r(t)$ with the average $E[dN_r(t)] =$

$\Lambda_r(dt) \sim \mathcal{O}(dt)$ and $dN_r(t)^2 \sim \mathcal{O}(dt)$. Explicitly the time evolution of the conditioned state can be written as

$$d|\bar{\Psi}(t)\rangle = \left[\sum_{r\neq 0} dN_r \left(\frac{\Omega_r(dt)}{\sqrt{\Lambda_r(dt)}} - 1\right) + \left(\frac{\Omega_0(dt)}{\sqrt{\Lambda_0(dt)}} - 1\right)\right] |\bar{\Psi}(t)\rangle. \tag{2.11}$$

Consider an example with

$$\begin{aligned} \Omega_1(dt) &= \sqrt{dt}(a+\gamma), \\ \Lambda_1(dt) &= E[dN(t)] = |\gamma|^2 dt, \qquad (2.12) \\ \Omega_0(dt) &= 1 - dt\left[iH + \frac{1}{2}(a\gamma^* - a^\dagger\gamma) + \frac{1}{2}(a^\dagger + \gamma^*)(a+\gamma)\right], \\ \Lambda_0(dt) &= 1 - |\gamma|^2 dt, \qquad (2.13) \end{aligned}$$

where $dN(t)$ is chosen to be a Poisson process. This amounts to adding a coherent field $|\gamma(t)\rangle$ with $\gamma(t) = |\gamma(t)|e^{i\varphi(t)}$ to the output radiation field a before it is measured by the photodetector. This can be realized by a low-reflectance beam-splitter. In the limit $|\gamma|^2 \gg \langle a^\dagger a\rangle$, this procedure is known as non-balanced homodyne detection. In the limit of $\gamma \to \infty$, Eq. (2.11) becomes

$$\begin{aligned} d|\bar{\Psi}(t)\rangle &= \left[dW(t)e^{-i\varphi(t)}a - dt(iH + a^\dagger a/2)\right]|\bar{\Psi}(t)\rangle, \\ E[dW(t)] &= 0, \qquad dW(t)^2 = dt, \end{aligned} \tag{2.14}$$

where $dW(t)$ describes a Wiener process. This SSE is much easier to solve or simulate and provides a powerful tool for adaptive measurement, i.e. updating the phase $\varphi(t)$ according to the previous results to accelerate or decelerate the evolution of system to some target states.

2.2 Entanglement

Entanglement is not only an essential concept of quantum mechanics, but is an extremely useful resource for quantum information and computation. It describes correlations between two or more degrees of freedom that can not be accounted for classically. A quantum state is separable if it can be written as a convex sum of product states belong to different degrees of freedom, otherwise it is entangled. For the two-photon polarization states, $|\Psi\rangle = |\uparrow\rangle_1|\downarrow\rangle_2$ is separable whereas $|\Psi\rangle = (|\uparrow\rangle_1|\downarrow\rangle_2 + |\downarrow\rangle_1|\uparrow\rangle_2)/\sqrt{2}$ is entangled.

The importance of entanglement can be illustrated by the Young's double slit experiment in Fig. 2.1. It clearly demonstrates that the principle of complementarity is just one aspect of

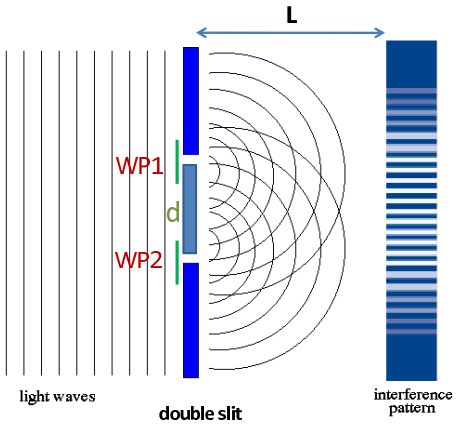

FIGURE 2.1. Schematic of Young's double slit experiment. The notions WP1 and WP2 represent two wave plates that can change the polarization of passing photons.

entanglement. The complementarity principle states that the interference pattern is washed away if the path or which-way information of photon is obtainable. Suppose the screen with double slit is in the plane $z = 0$ with the narrow slits at positions $x_\pm = \pm d/2$. Under the assumption of an impinging plane wave with wave number k and polarization $|\uparrow\rangle$, the photon state after passing the two infinitesimally wide slits on the screen located in the plane $z = L$ is a superposition of eigenstates $|\psi_\pm\rangle|\uparrow\rangle$,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\psi_+\rangle + |\psi_-\rangle)|\uparrow\rangle, \tag{2.15}$$

where

$$\psi_\pm(x) \propto \exp\left[\frac{ik(x - x_\pm)^2}{2L}\right]. \tag{2.16}$$

The probability density for detecting a photon at the position x on the screen is

$$\begin{aligned} |\Psi(x)|^2 &= \frac{1}{2}\left\{|\psi_+(x)|^2 + |\psi_-(x)|^2 + 2\mathrm{Re}[\psi_+^*(x)\psi_-(x)]\right\} \\ &\propto 1 + \cos\frac{kdx}{L}, \end{aligned} \tag{2.17}$$

which is double slit interference pattern with unit visibility.

Now let us insert two wave plates WP1 and WP2 in the front of two slits, which change the polarizations of passing photon from $|\uparrow\rangle$ to $|P_\pm\rangle$, respectively. The new state on the screen then

is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\psi_+\rangle|P_+\rangle + |\psi_-\rangle|P_-\rangle), \tag{2.18}$$

which is in an entangled form and yields the probability density

$$\begin{aligned} |\Psi(x)|^2 &= \frac{1}{2}\left\{|\psi_+(x)|^2 + |\psi_-(x)|^2 + 2\mathrm{Re}[\psi_+^*(x)\psi_-(x)\langle P_+|P_-\rangle]\right\} \\ &\propto 1 + |\langle P_+|P_-\rangle| \cos\left(\frac{kdx}{L} + \delta\right), \end{aligned} \tag{2.19}$$

where the angle δ is the phase of $\langle P_+|P_-\rangle$. The wave plates reduce the visibility by the factor $|\langle P_+|P_-\rangle|$. If the polarization states $|P_+\rangle$ and $|P_-\rangle$ are orthogonal, it becomes possible to distinguish them with certainty and interference pattern is washed away. According to Feynman [36], one should add *amplitudes* associated with *indistinguishable* processes, and add *probabilities* associated with *distinguishable* processes.

The degree of entanglement is well defined for all pure states. According to Schmidt theorem, we can always orthogonally decompose the two-particle state $|\Psi\rangle$ as

$$|\Psi\rangle = \sum_n \lambda_n |\psi_n^{(1)}\rangle|\psi_n^{(2)}\rangle. \tag{2.20}$$

The degree of entanglement for $|\Psi\rangle$ is defined to be the Von Neumann entropy [32],

$$E(\Psi) = -\sum_n \lambda_n \ln \lambda_n. \tag{2.21}$$

As for mixed states, there are no such simple expressions, or even no simple criteria to check whether a given mixed state entangled or not. The only exceptions are for two-qubit system and its continuous version — two-Gaussian system, where it is possible to write down the explicit formulas for the degree of entanglement [37, 38].

In literatures there are many necessary conditions for entangled state. The only sufficient and necessary condition we can find is derived in Ref. [39]. Suppose the spectrum decomposition of two-particle state ρ is

$$\rho = \sum_n \lambda_n |\Psi_n\rangle\langle\Psi_n|. \tag{2.22}$$

The state ρ is separable iff the equation

$$\begin{aligned} \sigma^2 - \sigma &= 0, && (2.23)\\ \sigma &= \mathrm{tr}_2|\Psi\rangle\langle\Psi|, && (2.24)\\ |\Psi\rangle &= \sum_n c_n|\Psi_n\rangle && (2.25) \end{aligned}$$

(tr_2 denotes the partial trace over the second particle) has different solutions (labeled by index m) for $c_n^{(m)}$ such that the equation

$$\begin{aligned} E^\dagger E &= 1, && (2.26)\\ E_{mn} &= \sqrt{\frac{p_m}{\lambda_n}} c_n^{(m)} && (2.27) \end{aligned}$$

have a positive solution for p_m.

2.3 A Short Review of Quantum Theory of Radiation

The classical electrodynamics in free space without any sources such as currents and charges is fully described by Maxwell's equations [40]

$$\begin{aligned} \nabla \times \mathbf{B} &= \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, && (2.28)\\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, && (2.29)\\ \nabla \cdot \mathbf{B} &= 0, && (2.30)\\ \nabla \cdot \mathbf{E} &= 0. && (2.31) \end{aligned}$$

Since Maxwell's equations are gauge invariant when no sources are present, it is convenient to choose the vector potential $\mathbf{A}(\mathbf{r}, t)$ in the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ such that

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A}, && (2.32)\\ \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t}. && (2.33) \end{aligned}$$

In terms of $\mathbf{A}(\mathbf{r}, t)$, we have the wave equation

$$\nabla^2 \mathbf{A}(\mathbf{r}, t) - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial^2 t} = 0. \quad (2.34)$$

Separating the vector potential into two complex terms and restricting the field to a finite volume of space, we can Fourier expand the vector potential in terms of discrete set of orthogonal mode functions $\mathbf{u}_k(\mathbf{r})$,

$$\mathbf{A}(\mathbf{r},t) = \sum_k [c_k(t)\mathbf{u}_k(\mathbf{r}) + c_k^*(t)\mathbf{u}_k^*(\mathbf{r})], \tag{2.35}$$

where the Fourier coefficients $c_k(t)$ and $c_k^*(t)$ satisfy the equation

$$\frac{dc_k}{dt} + i\omega_k c_k = 0, \tag{2.36}$$

$$\frac{dc_k^*}{dt} - i\omega_k c_k^* = 0. \tag{2.37}$$

From the wave equation (2.34) in the Coulomb gauge, we immediately get

$$\left(\nabla^2 + \frac{\omega_k^2}{c^2}\right)\mathbf{u}_k(\mathbf{r}) = 0, \tag{2.38}$$

$$\nabla \cdot \mathbf{u}_k(\mathbf{r}) = 0. \tag{2.39}$$

The mode functions depend on the boundary conditions of the physical volume under consideration. For a cube of volume $V = L^3$ with periodic boundary conditions corresponding to traveling-wave modes, we can write $\mathbf{u}_k(\mathbf{r})$ as

$$\mathbf{u}_k(\mathbf{r}) = \frac{1}{V^{1/2}}\mathbf{e}^{(\lambda)}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.40}$$

where $\mathbf{e}^{(\lambda)}$ is the unit polarization vector with the polarization index $\lambda = 1, 2$ and is required to be perpendicular to $\mathbf{k}$. The components of $\mathbf{k}$ take the values

$$k_x = \frac{2\pi n_x}{L}, \qquad k_y = \frac{2\pi n_y}{L}, \qquad k_z = \frac{2\pi n_z}{L}, \qquad n_x, n_y, n_z = 0, \pm 1, \pm 2, \ldots \tag{2.41}$$

Thus the vector potential may now be written as,

$$\mathbf{A}(\mathbf{r},t) = \sum_k \left(\frac{\hbar}{2\omega_k \epsilon_0}\right)^{1/2} [a_k(t)\mathbf{u}_k(\mathbf{r}) + a_k^*(t)\mathbf{u}_k^*(\mathbf{r})] \tag{2.42}$$

where the normalization factors have been chosen such that the amplitudes a_k and $a_k^\dagger$ are dimensionless and the corresponding electric field is

$$\mathbf{E}(\mathbf{r},t) = i\sum_k \left(\frac{\hbar\omega_k}{2\epsilon_0}\right)^{1/2} [a_k(t)\mathbf{u}_k(\mathbf{r}) - a_k^*(t)\mathbf{u}_k^*(\mathbf{r})] \tag{2.43}$$

In the classical electrodynamics these Fourier amplitudes are complex numbers. Quantization of the electromagnetic field is accomplished by replacing a_k and a_k^* with mutually adjoint operators a_k and $a_k^\dagger$ obeying the commutation relations

$$[a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0, \qquad [a_k, a_{k'}^\dagger] = \delta_{kk'}. \tag{2.44}$$

The electric field thus becomes an operator

$$\mathbf{E}(\mathbf{r}, t) = i \sum_k \left(\frac{\hbar \omega_k}{2\epsilon_0} \right)^{1/2} [a_k(t)\mathbf{u}_k(\mathbf{r}) - a_k^\dagger(t)\mathbf{u}_k^*(\mathbf{r})]. \tag{2.45}$$

The dynamical behavior of the electric-field amplitude then is described by an ensemble of independent harmonic oscillators obeying the above relations describes. The Hamiltonian of the field then is given as,

$$H = \frac{1}{2} \int \left(\epsilon_0 \mathbf{E}^2 + \frac{1}{\mu_o} \mathbf{B}^2 \right) d^3 r = \sum_k \hbar \omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right). \tag{2.46}$$

This represents sum of the number of photons in each mode multiplied by the energy of a photon in that mode, plus $\hbar\omega_k/2$ representing the energy of the zero-point fluctuations in each mode.

We shall now consider three possible states of the electromagnetic fields. The eigenstates of the Hamiltonian H are the photon number states or Fock-states. The photon number operator is given by $N = a^\dagger a$ with eigenvalue $n = 0, 1, 2, \ldots$ and eigenstates $|n\rangle$, i.e.

$$N|n\rangle = n|n\rangle, \qquad |n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n |0\rangle. \tag{2.47}$$

The set $\{|n\rangle\}$ forms a complete orthogonal basis for a Hilbert space,

$$\langle m|n\rangle = \delta_{mn}, \tag{2.48}$$

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = 1. \tag{2.49}$$

Note that

$$a|n\rangle = \sqrt{n}|n-1\rangle, \qquad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \tag{2.50}$$

a and $a^\dagger$ are therefore also named as annihilation and creation operators, and state $|0\rangle$ is called vacuum state. The number states form a useful representation, but they are not the most suitable

representation for optical fields where total number of photons is large, such as the field of laser. The laser field is more appropriately described by the so-called coherent states [41]. They are the closest quantum states to a classical description of the field. The coherent states are eigenstates of the annihilation operator a,

$$a|\alpha\rangle = \alpha|\alpha\rangle. \tag{2.51}$$

They can be expanded in terms of the number states $|n\rangle$ as follows,

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{2.52}$$

So $\langle\alpha|\alpha\rangle = 1$ and the coherent states are normalized. The average photon number of coherent state $|\alpha\rangle$ is $\langle N\rangle = \langle\alpha|N|\alpha\rangle = |\alpha|^2$. An alternative representation of coherent states is to used the unitary displacement operator,

$$D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}. \tag{2.53}$$

Equivalently the coherent state $|\alpha\rangle$ can be generated by operating with $D(\alpha)$ on the vacuum state

$$|\alpha\rangle = D(\alpha)|0\rangle. \tag{2.54}$$

The expectation value of the electric field operator

$$\mathbf{E}(\mathbf{r},t) = i\mathbf{e}^{(\lambda)}\left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2}\left[ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - a^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right] \tag{2.55}$$

in the coherent state $|\alpha\rangle$ with $\alpha = |\alpha|\, e^{i\varphi}$ is given by

$$\langle\alpha|\mathbf{E}(\mathbf{r},t)|\alpha\rangle = 2\mathbf{e}^{(\lambda)}|\alpha|\left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2}\sin(\omega t - \mathbf{k}\cdot\mathbf{r} - \varphi), \tag{2.56}$$

which takes the form of a classical field. The last quantum states that we consider is the so-called squeezed states [42] defined by

$$|\alpha,\xi\rangle = D(\alpha)S(\xi)|0\rangle, \qquad S(\xi) = \exp\left[\frac{1}{2}(\xi^* a^2 - \xi a^{\dagger 2})\right], \tag{2.57}$$

where $\xi = re^{i\theta}$, and r is known as the squeezing parameter. The squeezed states can also be obtained as the eigenstates of the operator $\mu a + \nu a^\dagger$ with $\mu = \cosh r$ and $\nu = e^{i\theta}\sinh r$,

$$(\mu a + \nu a^\dagger)|\alpha,\xi\rangle = \gamma|\alpha,\xi\rangle, \qquad \gamma = \alpha\cosh r + \alpha^* e^{i\theta}\sinh r. \tag{2.58}$$

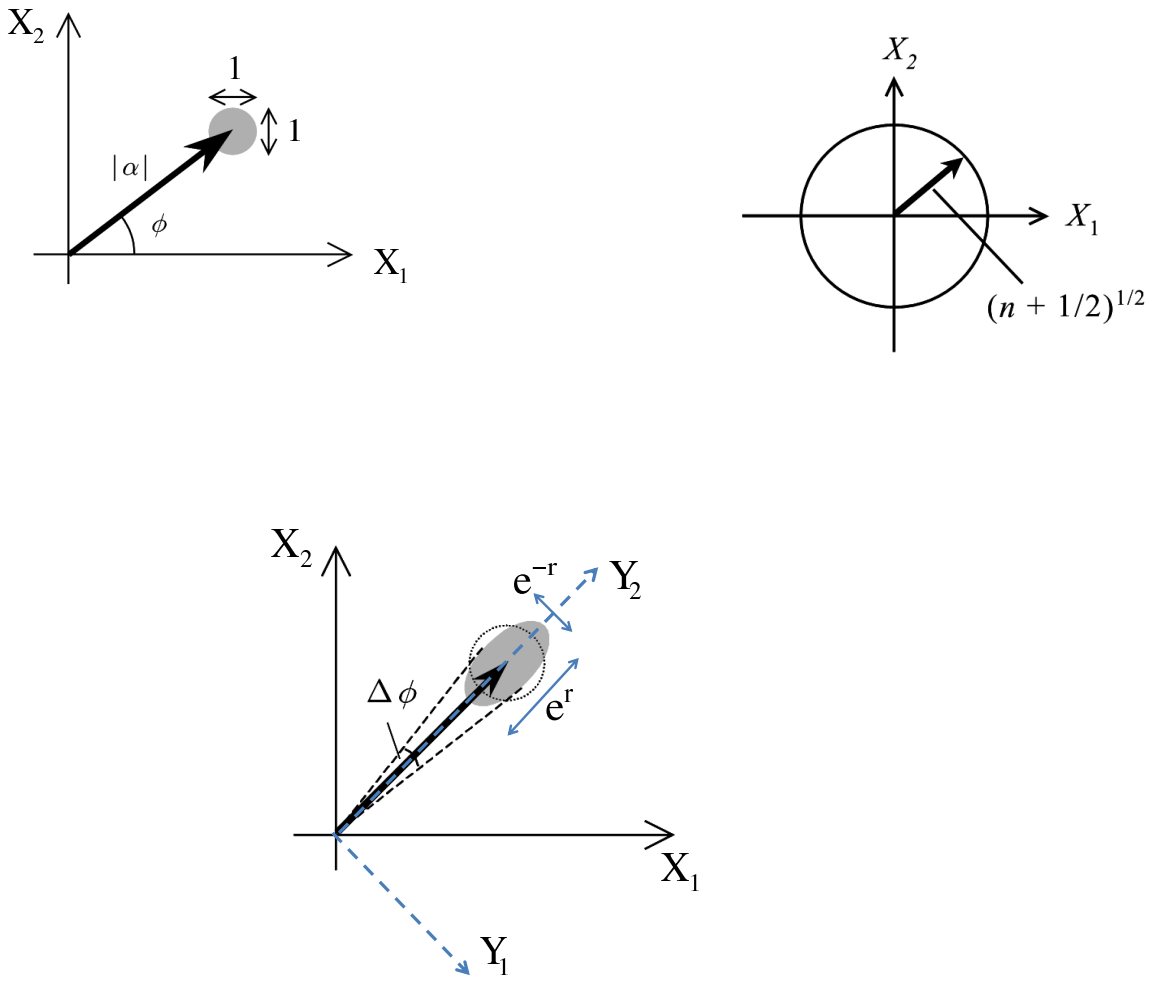

FIGURE 2.2. Schematics of coherent state $|\alpha\rangle$, number state $|n\rangle$, and squeezed state $|\alpha, \xi\rangle$ with $\varphi = -\theta/2$ in phase space.

For $r = 0$ this is just the eigenvalue problem for the coherent state.

One intuitive way to view the above three types of state is to look at their phase space diagrams in the dimensionless position $X_1 = (a + a^\dagger)/2$ and momentum $X_2 = (a - a^\dagger)/(2i)$ coordinates. Because the quadrature operators X_1, X_2 satisfy the commutation relation

$$[X_1, X_2] = \frac{i}{2}, \tag{2.59}$$

the uncertainty relation $\Delta A \Delta B \geq |\langle [A, B] \rangle| /2$ requires that $\Delta X_1 \Delta X_2 \geq 1/4$. Thus the quantum state is not well localized as a point in phase space as it is in classical mechanics, but is distributed over some area. A minimum uncertainty state is a state which minimizes the uncertainty product. The vacuum state $|0\rangle$ is such a minimum uncertainty state. Since $\langle X_1 \rangle = \langle X_2 \rangle = 0$ and $\Delta X_1 =$

$\Delta X_2 = 1/2$, the vacuum is thus depicted as a filled circle at the center with radius 1/2. The coherent state is also a minimum uncertainty state. Note $\langle X_1 \rangle = \mathrm{Re}[\alpha]$, $\langle X_2 \rangle = \mathrm{Im}[\alpha]$ and $\Delta X_1 = \Delta X_2 = 1/2$ for a coherent state $|\alpha\rangle$, it can be pictured as a displaced vacuum to the center $(\mathrm{Re}[\alpha], \mathrm{Im}[\alpha])$. For the number states $|n\rangle$, we have $\langle X_1 \rangle = \langle X_2 \rangle = 0$ but $\langle X_1^2 \rangle = \langle X_2^2 \rangle = (2n+1)/4$. It thus can be represented in phase space as a circle of radius $\sqrt{n+1/2}$. To visualize the squeezed state $|\alpha, \xi\rangle$, we introduce the rotated quadrature operators Y_1, Y_2 as

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \tag{2.60}$$

or equivalently $Y_1 + iY_2 = (X_1 + iX_2)e^{-i\theta/2}$. Note that

$$\begin{aligned} \langle a \rangle &= \alpha, \\ \langle Y_1 + iY_2 \rangle &= \alpha e^{-i\theta/2}, \\ \langle N \rangle &= |\alpha|^2 + \sinh^2 r \\ \Delta Y_1 &= \frac{1}{2}e^{-r}, \\ \Delta Y_2 &= \frac{1}{2}e^{r}, \end{aligned} \tag{2.61}$$

so the squeezed state can be represented as a filled rotated ellipse at the center $(\mathrm{Re}[\alpha], \mathrm{Im}[\alpha])$ in phase space, as shown in Fig. 2.2.

2.3.1 Definition of Phase Parameter

Suppose a portion of dielectric crystal with dielectric constant ϵ and length l in the traveling path of electromagnetic field. Neglecting the reflection effect of the crystal, the electric field after the crystal is

$$\begin{aligned} \mathbf{E}_\lambda(l,t) &= i\mathbf{e}^{(\lambda)} \left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2} \left[a e^{i(\chi kl - \omega t)} - a^\dagger e^{-i(\chi kl - \omega t)}\right] \\ &= i\mathbf{e}^{(\lambda)} \left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2} \left[a e^{i\varphi} e^{i(kl - \omega t)} - a^\dagger e^{-i\varphi} e^{-i(kl - \omega t)}\right], \end{aligned} \tag{2.62}$$

where the index $\chi = \sqrt{\epsilon/\epsilon_0} = 1 + \delta$ and the phase $\varphi = \delta kl$. In the limit of $l \to 0$ and keeping φ fixed, this change of field amounts to a unitary transform, $a \to e^{-iN\varphi} a\, e^{iN\varphi} = a\, e^{i\varphi}$. In other words, the unitary matrix $U(\varphi) = e^{-iN\varphi}$ just realizes a phase shift φ. However, there does not

exist a Hermitian operator corresponding to phase shift φ [43]. Otherwise the conjugate variables N, φ should satisfy the commutation relation

$$[\varphi, N] = -i. \tag{2.63}$$

If we take the matrix element of the commutator for arbitrary number states $|n\rangle$ and $|n'\rangle$, we should have

$$\langle n|[\varphi, N]|n'\rangle = (n' - n)\langle n|\varphi|n'\rangle = -i\delta_{nn'}, \tag{2.64}$$

which contains an obvious contradiction in the case when $n = n'$ giving $0 = -i$! To compare this situation with the usual momentum and position operators $[p, q] = -i\hbar$, this reasoning is not valid, because it gives a singular identity

$$(q' - q)\frac{\partial}{\partial q}\delta(q' - q) = \delta(q' - q). \tag{2.65}$$

The root of the problem is that the operator N has a spectrum bounded from below, unlike the unbounded position operator. This means if we superficially define the number operator by

$$N = i\frac{\partial}{\partial \varphi}, \qquad [\varphi, N] = -i, \tag{2.66}$$

the eigenstates of N have the form $e^{-in\varphi}/\sqrt{2\pi}$ using periodicity. However, we get the wrong spectrum $-\infty < n < \infty$, whereas the true spectrum runs from 0 to ∞ with the number states $|n\rangle$. Similar difficulty is also encountered in the energy-time problem, where the spectrum of the Hamiltonian is bounded from below. Another specific problem of the phase shift is that φ and $\varphi + 2\pi$ makes no difference physically.

The nonexistence of the Hermitian phase shift operator denies the possibility to read it in the sense of projective measurement. But this does not prevent us to detect the phase shift through the generalized measurement with the POVM, or through the measurement of a particular intermediate observable.

2.3.2 Phase Value of Quantum State

It is worth to point out that though there is a rigorous scheme to estimate the phase shift φ in the transform $U = e^{iN\varphi}$, the problem how to assign a phase value to a given quantum

state still remains open. We will consider only one possible solution to this problem. We already know that the quantum state that mimics a classical electric field is given by the coherent state. In the coherent state $|\alpha\rangle$ with $\alpha = |\alpha|e^{i\varphi}$, the expectation value of the electric field is $\mathbf{E}(\mathbf{r}, t) \propto \sin(\omega t - \mathbf{k}\cdot\mathbf{r} - \varphi)$ with a phase value φ. It is thus reasonable to define the phase value of $|\alpha\rangle$ as $\varphi = \arg\alpha$ or in terms of the average values of quadrature operators $\varphi = \tan^{-1}(\langle X_2\rangle/\langle X_1\rangle)$ due to $\langle X_1\rangle = |\alpha|\cos\varphi$, $\langle X_2\rangle = |\alpha|\sin\varphi$. This definition of phase value is only an average quantity over all quantum fluctuations, and can be generalized to any quantum state. Pictorially the phase value of a quantum state is defined as the polar angle of its diagram in the phase space. The distributional nature of a quantum state in the phase space clearly shows that its phase value also subjects to fluctuations.

Let us now consider the examples as shown in Fig. 2.2 with the above definition. For coherent state $|\alpha\rangle$ with $\alpha = |\alpha|\, e^{i\varphi}$, the polar angle of its portrait in the phase space is scattered around $\varphi \pm r/|\alpha|$, where $r = 1/2$ is the radius of the filled circle. The phase uncertainty therefore is given as

$$\Delta\varphi = \frac{1}{2\langle N\rangle^{1/2}}. \tag{2.67}$$

Obviously, the number state $|n\rangle$ has an arbitrary phase since its phase diagram is an empty circle around the origin. For a specific squeezed state $|\alpha, \xi\rangle$ with $\alpha = |\alpha|e^{i\varphi}$ and $\xi = re^{-2i\varphi}$, the phase uncertainty is equal to

$$\Delta\varphi = \frac{\Delta Y_1}{|\alpha|} \approx \frac{e^{-r}}{2\langle N\rangle^{1/2}}, \tag{2.68}$$

where we have assumed $|\alpha| \gg 1$ in order to neglect the $\sinh^2 r$ term in the expression of $\langle N\rangle$ in Eq. (2.61).

Experimentally the above defined phase value can be measured via the so-called heterodyne or homodyne detection [3] as shown in Fig. 2.3. A single-mode quantum state is input at one port of a balanced beam splitter. At the other input port of the beam splitter a strong local oscillator such as a laser in the coherent state, $|\alpha_l\rangle$ where $\alpha_l = |\alpha_l|e^{i\varphi_l}$ and $|\alpha_l| \gg 1$ is incident. When the signal and the local oscillator have the same frequencies, it is called as homodyne

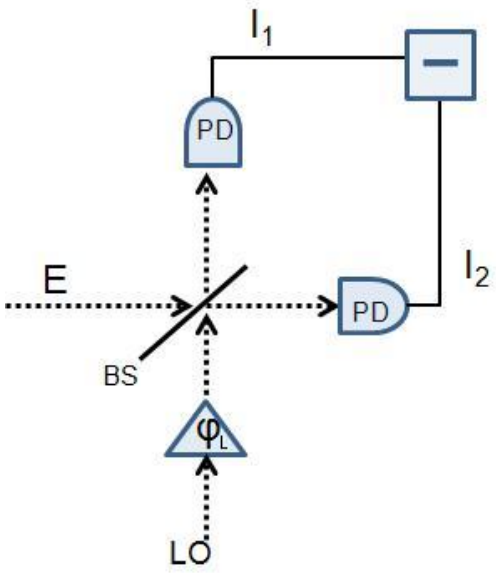

FIGURE 2.3. Schematic of balanced homodyne detection. Here E represents the input field, I_1 and I_2 are measured signals, LO: local oscillator, BS: beam splitter, PD: photodetector.

detection. Otherwise it is called as heterodyne detection. For simplicity we suppose that all the states considered in this thesis have the same frequencies.

The input and the oscillator modes are described by the annihilation operators a and b, respectively. Then the two output modes reaching photon number detectors 1 and 2 by c and d are given by

$$\begin{aligned} c &= (a + ib)/\sqrt{2}, \\ d &= (ic + b)/\sqrt{2}. \end{aligned} \tag{2.69}$$

The signals measured by two detectors are determined by the operators,

$$\begin{aligned} c^\dagger c &= \frac{1}{2}[a^\dagger a + b^\dagger b + i(a^\dagger b - b^\dagger a)], \\ d^\dagger d &= \frac{1}{2}[a^\dagger a + b^\dagger b - i(a^\dagger b - b^\dagger a)]. \end{aligned} \tag{2.70}$$

The difference of the two signal operators gives

$$n_{cd} = c^\dagger c - d^\dagger d = i(a^\dagger b - b^\dagger a). \tag{2.71}$$

The measured signal of n_{cd} then is

$$\langle n_{cd} \rangle = 2\,|\alpha_l|\,\langle X(\varphi_l + \pi/2) \rangle, \tag{2.72}$$

and the associated variance can be found to be

$$\Delta n_{cd} = 2\,|\alpha_l|\,\Delta X(\varphi_l + \pi/2), \tag{2.73}$$

where the quadrature operator $X(\varphi)$ is

$$X(\varphi) = (ae^{-i\varphi} + a^\dagger e^{i\varphi})/2. \tag{2.74}$$

Clearly, by adjusting the laser phase φ_l one can obtain $\langle X(0)\rangle = \langle X_2\rangle$ and $\langle X(-\pi/2)\rangle = \langle X_1\rangle$, and thus the averaged phase value of mode a through the homodyne detection can be expressed as

$$\varphi = \tan^{-1}\frac{\langle X(0)\rangle}{\langle X(-\pi/2)\rangle}. \tag{2.75}$$

The variance of φ can also be calculated from the variance of n_{cd}. By some adaptive techniques of adjusting of the phase of the local oscillator one can improve the variance of φ and reduce the cycles of measurement [44].

The phase value discussed above should not be confused with the geometric phase in quantum mechanics [45]. For example, the phase difference φ_{12} between any two non-orthogonal states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, which is defined via

$$\langle\Psi_1|\Psi_2\rangle = |\langle\Psi_1|\Psi_2\rangle|\,e^{i\varphi_{12}}. \tag{2.76}$$

This definition has no classical analog because it is not transitive: If $|\Psi_1\rangle$ is in phase with $|\Psi_2\rangle$, and $|\Psi_2\rangle$ with $|\Psi_3\rangle$, then $|\Psi_1\rangle$ need not to be in phase with $|\Psi_3\rangle$.

2.3.3 Atom-Field Interaction in Quantum Optics

For completeness, we will briefly review atom-field interactions in quantum optics [40]. Suppose an electron bound to an atom in the absence of external field. The relevant Hamiltonian is given by

$$H_a = \frac{\mathbf{p}^2}{2m} + V(r), \tag{2.77}$$

where $V(r)$ is the usual Coulomb potential. Assume the energy levels are $|a\rangle$, satisfying the eigenvalue equation

$$H_a|a\rangle = E_a|a\rangle. \tag{2.78}$$

We can therefore rewrite H_a and the dipole operator $\mathbf{d} = e\mathbf{r}$, respectively, as

$$H_a = \sum_a E_a |a\rangle\langle a| = \sum_a E_a \sigma_{aa}, \tag{2.79}$$

$$e\mathbf{r} = \sum_{a,b} e|a\rangle\langle a|\mathbf{r}|b\rangle\langle b| = \sum_{ab} \mathbf{d}_{ab}\sigma_{ab}, \tag{2.80}$$

where $\mathbf{d}_{ab} = e\langle a|\mathbf{r}|b\rangle$ and $\sigma_{ab} = |a\rangle\langle b|$. In the presence of external electromagnetic fields the total Hamiltonian of the atom and fields in the dipole approximation, where the field is assumed to be uniform over the whole atom, is

$$\begin{aligned} H &= H_a + H_f - e\mathbf{r}\cdot\mathbf{E}(\mathbf{r},t), \\ H_f &= \sum_k \hbar\omega_k a_k^\dagger a_k, \end{aligned} \tag{2.81}$$

where we have omitted the zero-point energy of field as it does not contribute to the dynamics, $\mathbf{r}$ is the position of the electron, and the electric field $\mathbf{E}(\mathbf{r},t)$ is expressed by Eq. (2.45). For an atom with two energy levels E_a and E_b in a single mode field and in the rotating wave approximation, the Hamiltonian can be further simplified to be

$$H = \hbar\omega_0\sigma_z + \hbar\omega a^\dagger a + \hbar g(a\sigma^+ + a^\dagger\sigma^-), \tag{2.82}$$

where $\hbar\omega_0 = E_b - E_a$, $\sigma^\pm = \sigma_x \pm i\sigma_y$, $\sigma_x, \sigma_y, \sigma_z$ are Pauli spin matrices, and we have taken the coupling constant g to be real. This is the simplest form of the atom-field interaction and is known as the Jaynes-Cummings model [46].

Now let us consider an ideal detector consisting of a single ground-state atom of dimension small compared to the wavelength of the light. This single-atom detector couples to the quantized field via Eq. (2.81). Using the first order perturbation theory in quantum mechanics, the probability of the atom making a transition to an ionized atomic state by absorbing a photon from field is proportional to

$$\sum_f \left|\langle f|\mathbf{E}^+(\mathbf{r},t)|i\rangle\right|^2 = \langle i|\mathbf{E}^{(-)}(\mathbf{r},t)\cdot\mathbf{E}^{(+)}(\mathbf{r},t)|i\rangle, \tag{2.83}$$

where $\mathbf{E}^{(+)}$ is the positive frequency part of $\mathbf{E}$ and $\mathbf{E}^{(-)} = \mathbf{E}^{(+)\,\dagger}$. Because we are really only interested in the final state of the detector, not the field, we have summed over all possible field

transitions from the initial state $|i\rangle$ to the final states $|f\rangle$. For a mixed photon state described by density matrix ρ, the absorbing probability then is proportional to

$$\mathrm{tr}\left[\rho E^{(-)}(x)E^{(+)}(x)\right], \tag{2.84}$$

where $E^{(+)} = \mathbf{e}\cdot\mathbf{E}^{(+)}$, $E^{(-)} = \mathbf{e}^*\cdot\mathbf{E}^{(-)}$, and $x = (\mathbf{r}, t)$. It can be viewed as a specific case of the first order correlation function defined by

$$G^{(1)}(x, x') = \mathrm{tr}\left[\rho E^{(-)}(x)E^{(+)}(x')\right], \tag{2.85}$$

which describes the correlation between field at the point x and the field at x'. To describe delayed coincidence experiments such as Hanbury-Brown and Twiss experiment, it is necessary to define higher-order correlation functions [41] as

$$G^{(n)}(x_1, ..., x_n, x_{n+1}, ..., x_{2n}) = \mathrm{tr}\left[\rho E^{(-)}(x_1)...E^{(-)}(x_n)E^{(+)}(x_{n+1})...E^{(+)}(x_{2n})\right]. \tag{2.86}$$

Such an expression follows from a consideration of an n-atom photon detector by n-order perturbation. The n-fold delayed coincidence rate is

$$\eta^n G^{(n)}(x_1, ..., x_n, x_n, ..., x_1), \tag{2.87}$$

where η is the efficiency of the detector. For the case of $n = 1$ and a single mode field, the rate of detecting one photon is proportional to $\mathrm{tr}[\rho a^\dagger a]$, namely the average photon number of the field. In other words, the single atom detector can work as a photon number resolver of a monochromatic field, as we have already used for homodyne detection in last section.

2.4 Phase Shift Estimation

The estimation of the phase shift φ in the unitary transform $U(\varphi) = e^{iN\varphi}$ is a particular problem in quantum estimation theory, which aims to tangle such problems in a rigorous way [47]. Let us first discuss the general estimation theory in the classical way, and then generalize its results quantum mechanically.

2.4.1 Classical Estimation Theory

If a classical system is observed to obtain the signal v, a decision is to be made about its state. The system may be in any one of M possible states, and the hypothesis H_j, $j = 1, 2, ..., M$,

asserts that "The system is in state j". The signal v is a random variable whose joint probability density function (p.d.f.) is $p_j(v)$ if the system is in state j. Suppose the system is in state j with a prior probability ξ_j from past experience. The normalization requires $\sum_{j=1}^{M} \xi_j = 1$. Let C_{ij} be the cost incurred by choosing hypothesis H_i when H_j is true. They assign relative weights to the various possible error and correct decisions. A physical constrain for the cost is $C_{jj} \geq C_{ij}$, i.e. the chance to get a right answer should win over the chance to get a wrong one. For example, the *Bayes strategy* corresponds to an assignment of costs $C_{ij} = -\delta_{ij}$, which penalizes all errors equally. Under these general conditions, one desires a strategy whose average cost is a minimum.

A decision strategy is represented by the probabilities $\pi_i(v)$, that hypothesis H_i is chosen when v is obtained. These probabilities are subject to the conditions

$$0 \leq \pi_i(v) \leq 1, \qquad \sum_{i=1}^{M} \pi_i(v) = 1. \tag{2.88}$$

Pure guessing among the M hypotheses corresponding to $\pi_i(v) = M^{-1}$ for all i. The probability that the hypothesis H_i is chosen when H_j is true is $p(i|j) = \int \pi_i(v) p_j(v) dv$. With the incurred cost C_{ij}, the average cost of the strategy is

$$\begin{aligned} \overline{C}\left[\{\pi_i\}\right] &= \sum_{i,j=1}^{M} \xi_j C_{ij} p(i|j) = \sum_{i=1}^{M} \int W_i(v) \pi_i(v) dv, \\ W_i(v) &\equiv \sum_{j=1}^{M} \xi_j C_{ij} p_j(v). \end{aligned} \tag{2.89}$$

The function $W_i(v)$ is called "risk function" for H_i.

The average cost $\overline{C}$ will be least if at each point v we choose the hypothesis for which the risk $W_l(v)$ is smallest. That is $\pi_i(v) = \delta_{il}$ for $W_l(v) \leq W_i(v)$. Define $\Upsilon(v) = W_l(v)$. These conditions can be expressed as

$$\begin{aligned} \left[W_i(v) - \Upsilon(v)\right] \pi_i(v) = 0, \\ W_i(v) - \Upsilon(v) \geq 0, \end{aligned} \tag{2.90}$$

for each v and all hypotheses H_i.

If the possible states of the system under observation are represented by a continuous parameter θ, which is to be estimated from the observable v, $\hat{\theta} = \hat{\theta}(v)$. The function $\hat{\theta}(v)$ is often called

estimator of θ. We can introduce the joint p.d.f. $p(v|\theta)$ of v given the parameter θ, the prior p.d.f. $z(\theta)$, the cost function $C(\hat{\theta},\theta)$, and the joint p.d.f. $\pi(\hat{\theta}|v)$ for estimation strategy. A common quadratic cost function is $C(\hat{\theta},\theta) = (\hat{\theta}-\theta)^2$. For the maximum-likelihood estimator, the cost function is $C(\hat{\theta},\theta) = -\delta(\hat{\theta}-\theta)$.

When the true values of the parameters are θ, the probability that the estimates lie in a volume elements $d\hat{\theta}$ about the point $\hat{\theta}$ is

$$q(\hat{\theta}|\theta)d\hat{\theta} = \int \pi(\hat{\theta}|v)p(v|\theta)dvd\hat{\theta}. \tag{2.91}$$

The average cost function thus is

$$\begin{aligned} \overline{C}\,[\pi] &= \int\int z(\theta)C(\hat{\theta}-\theta)q(\hat{\theta}|\theta)d\theta d\hat{\theta} \\ &= \int\int W(\hat{\theta};v)\pi(\hat{\theta}|v)d\hat{\theta}dv \end{aligned} \tag{2.92}$$

by introducing the risk function $W(\hat{\theta};v) = \int z(\theta)C(\hat{\theta}-\theta)p(v|\theta)d\theta$. The optimum strategy is specified by the nonnegative function $\pi(\hat{\theta}|v)$ satisfying $\int \pi(\hat{\theta}|v)d\hat{\theta} = 1$ and

$$\begin{aligned} [W(\hat{\theta};v) - \Upsilon(v)]\pi(\hat{\theta}|v) = 0, \\ W(\hat{\theta};v) - \Upsilon(v) \geq 0, \end{aligned} \tag{2.93}$$

where $\Upsilon(v)$ is given by $\Upsilon(v) = \min_{\hat{\theta}}[W(\hat{\theta};v)]$. The least average cost is thus $\overline{C}_{\min} = \int \Upsilon(v)dv$ and the corresponding strategy is $\pi(\hat{\theta}|v) = \delta(\hat{\theta}-\hat{\theta}(v))$.

For the quadratic cost function, the optimum strategy leads to

$$\hat{\theta}(v) = \int \theta p(\theta|v)d\theta, \tag{2.94}$$

where the posterior p.d.f. is obtained by Bayes rule

$$p(\theta|v) = \frac{z(\theta)p(v|\theta)}{\int z(\theta)p(v|\theta)d\theta}. \tag{2.95}$$

For the delta-function cost, we have the maximum-likelihood estimator

$$\hat{\theta}(v) = \arg\ \max_{\theta}[p(\theta|v)]. \tag{2.96}$$

The optimum estimator of θ of an arbitrary p.d.f. $p(v|\theta)$ is usually complicated, and analytically evaluating the minimum cost is impossible. However, a lower bound for the accuracy of any estimator is provided by the following Cramer-Rao bound. Note the identity

$$\int p(v|\theta)\Delta\hat{\theta}\, dv = 0, \tag{2.97}$$

where $\Delta\hat{\theta} = \hat{\theta}(v) - \overline{\theta}$ and $\overline{\theta} = \int p(v|\theta)\hat{\theta}(v)\, dv$. Taking the derivative of this identity with respect θ, we obtain

$$\int p(v|\theta)\frac{\partial \ln p(v|\theta)}{\partial\theta}\Delta\hat{\theta}\, dv = \frac{d\overline{\theta}}{d\theta}. \tag{2.98}$$

Applying the Schwarz inequality

$$\int |f|^2\, dx \int |g|^2\, dx \geq \left|\int fg dx\right|^2 \tag{2.99}$$

to the functions $f = p^{1/2}\partial \ln p/\partial\theta$ and $g = p^{1/2}\Delta\hat{\theta}$ yields

$$F(\theta)\,\overline{(\Delta\hat{\theta})^2} \geq \left(\frac{d\overline{\theta}}{d\theta}\right)^2,$$

where the Fisher information is defined by

$$F(\theta) = \int p(v|\theta)\left(\frac{\partial \ln p(v|\theta)}{\partial\theta}\right)^2 dv. \tag{2.100}$$

The Cramer-Rao inequality is given by

$$\overline{(\delta\theta)^2} \geq \frac{1}{F(\theta)} + \overline{(\delta\theta)}^2 \geq \frac{1}{F(\theta)} \tag{2.101}$$

with

$$\delta\theta = \frac{\hat{\theta}(v)}{|d\overline{\theta}/d\theta|} - \theta. \tag{2.102}$$

A zero value of $\overline{\delta\theta}$ means the estimator is unbiased, i.e. $\overline{\theta} = \theta$ locally. The well known formula of the sensitivity of θ obtained from an arbitrary estimator A can thus be defined as

$$\Delta\theta = \frac{\Delta A}{|\partial\overline{A}/\partial\theta|}. \tag{2.103}$$

For ν such measurements, the Cramer-Rao bound Eq. (2.101) becomes

$$\overline{(\delta\theta)^2} \geq \frac{1}{\nu F(\theta)}. \tag{2.104}$$

The Cramer-Rao bound only places a lower bound on the uncertainty of an estimator. Fisher's theorem says that asymptotically for large ν, maximum-likelihood estimator is unbiased and achieves the Crame-Rao bound. For a single measurement, this lower bound usually can not be achieved.

2.4.2 Quantum Estimation Theory

Quantum estimation theory seeks the best strategy to estimate the parameter θ of the density matrix $\rho(\theta)$ of a system. This procedure can be realized by a POVM, $dE(\hat{\theta})$ with $\int dE(\hat{\theta}) = 1$. The joint conditional p.d.f. $q(\hat{\theta}|\theta)$ of the estimates is then given by

$$q(\hat{\theta}|\theta)d\hat{\theta} = \mathrm{tr}[\rho(\theta)dE(\hat{\theta})]. \tag{2.105}$$

With the prior p.d.f. $z(\theta)$ and the cost function $C(\hat{\theta}, \theta)$, the average cost incurred is

$$\begin{aligned} \overline{C}\,[E] &= \mathrm{tr}\int\int z(\theta)C(\hat{\theta},\theta)\rho(\theta)dE(\hat{\theta})d\theta \\ &= \mathrm{tr}\int W(\hat{\theta})dE(\hat{\theta}), \\ W(\hat{\theta}) &= \int z(\theta)C(\hat{\theta},\theta)\rho(\theta)d\theta. \end{aligned} \tag{2.106}$$

The least average cost is attained by the optimal strategy, which must satisfy

$$\begin{aligned} [W(\hat{\theta}) - \Upsilon]dE(\hat{\theta}) = 0, \\ W(\hat{\theta}) - \Upsilon \geq 0, \end{aligned} \tag{2.107}$$

where the Hermitian operator Υ is given by

$$\Upsilon = \int W(\hat{\theta})dE(\hat{\theta}) = \int dE(\hat{\theta})W(\hat{\theta}).$$

The minimum cost of error is

$$\min_E \overline{C}\,[E] = \mathrm{tr}\Upsilon. \tag{2.108}$$

To illustrate the quantum optimal strategy, we consider an example of the binary decisions. The system has two density matrix ρ_0 and ρ_1 with the prior probabilities ξ_0 and ξ_1, $\xi_0 + \xi_1 = 1$.

The optimal POVM that can distinguish the two states is described by two operators E_0 and E_1, $E_0 + E_1 = 1$. The costs are given by the Bayes strategy, and the risks are now

$$\begin{aligned} & W_0 = \xi_1\rho_1, \qquad W_1 = \xi_0\rho_0, \\ & \Upsilon = W_0E_0 + W_1E_1 = W_0 + (W_1 - W_0)E_1. \end{aligned} \tag{2.109}$$

The optimal conditions require

$$\begin{aligned} 0 &= (W_0 - \Upsilon)E_0 = (W_0 - W_1)E_1E_0, \\ 0 &\leq W_0 - \Upsilon = (W_0 - W_1)E_1 = (\xi_1\rho_1 - \xi_0\rho_0)E_1, \end{aligned} \tag{2.110}$$

which are satisfied by choosing

$$E_0 = \sum_\lambda \Theta(-\lambda)|\lambda\rangle\langle\lambda|, \qquad E_1 = \sum_\lambda \Theta(\lambda)|\lambda\rangle\langle\lambda|, \tag{2.111}$$

where $|\lambda\rangle$ is the eigenvector of $\Delta = \xi_1\rho_1 - \xi_0\rho_0$ with the eigenvalue λ, and $\Theta(x)$ the unit step function. The average probability of error in such a optimal decision is equal to

$$\begin{aligned} P_{\text{error}} &= \xi_0 \text{tr}[\rho_0E_1] + \xi_1\text{tr}[\rho_1E_0] = \xi_0 + \text{tr}[(\xi_1\rho_1 - \xi_0\rho_0)E_0] \\ &= \xi_0 + \sum_{\lambda\leq 0}\lambda = (1 - \sum_\lambda |\lambda|)/2 \\ &= \left(1 - \text{tr}\sqrt{\Delta^\dagger\Delta}\right)/2 = \frac{1}{2} - D(\xi_0\rho_0, \xi_1\rho_1), \end{aligned} \tag{2.112}$$

where the distance $D(\rho, \sigma)$ is defined by $D(\rho, \sigma) = \text{tr}\,|\rho - \sigma|/2$ and $|A| \equiv \sqrt{A^\dagger A}$. For the two pure states, $\rho_0 = |\Psi_0\rangle\langle\Psi_0|$ and $\rho_1 = |\Psi_1\rangle\langle\Psi_1|$,

$$P_{\text{error}} = \left(1 - \sqrt{1 - 4\xi_0\xi_1\,|\langle\Psi_1|\Psi_0\rangle|^2}\right)/2. \tag{2.113}$$

The second example is to estimate the phase parameter appeared in the states

$$\Psi(\theta) = e^{-iJ\theta}\Psi, \tag{2.114}$$

where $|\Psi\rangle = \sum_j c_j|j\rangle$ and $J|j\rangle = j|j\rangle$. If nothing is known in advance about the true value of θ, i.e.

$$z(\theta) = 1/2\pi, \qquad -\pi < \theta \leq \pi, \tag{2.115}$$

and the delta cost function is used, the optimal POVM is given by (for details see section 2.5)

$$dE(\theta) \;=\; \frac{\mathcal{N}}{2\pi}|\theta\rangle\langle\theta|d\theta, \qquad |\theta\rangle = \frac{1}{\sqrt{\mathcal{N}}}\sum_j \frac{c_j}{|c_j|}e^{-ij\theta}|j\rangle, \tag{2.116}$$

where $\mathcal{N}$ is a normalization factor. For convenience, it is assumed from now on that all coefficients c_j of the initial state are real. In the case of phase shift $J = N$ and $\theta = -\varphi$, the optimal measurement is [47]

$$E(\varphi)d\varphi = \frac{1}{2\pi}|\varphi\rangle\langle\varphi|d\varphi, \qquad |\varphi\rangle = \sum_{n=0}^{\infty} e^{in\varphi}|n\rangle. \tag{2.117}$$

With this optimal POVM, let us now look at the estimated phase distribution of a coherent state $|re^{i\varphi_0}\rangle = e^{iN\varphi_0}|r\rangle$ with real r [48], which is

$$p(\varphi) = \frac{1}{2\pi}|\langle\varphi|re^{i\varphi_0}\rangle|^2 = \frac{1}{2\pi}e^{-r^2}\left|\sum_{n=0}^{\infty} e^{-in(\varphi-\varphi_0)}\frac{r^n}{\sqrt{n!}}\right|^2. \tag{2.118}$$

For large r, the Poisson distribution may be approximated as a Gaussian,

$$e^{-r^2}\frac{r^{2n}}{n!} \approx \frac{1}{\sqrt{2\pi r^2}}\exp\left[-\frac{(n-r^2)^2}{2r^2}\right]. \tag{2.119}$$

Then the sum can be further approximated by a Fourier integral, which is evaluated to be

$$p(\varphi) = \sqrt{\frac{2r^2}{\pi}}\exp[-2r^2(\varphi-\varphi_0)^2]. \tag{2.120}$$

This is a Gaussian peaked at $\varphi = \varphi_0$ with the variance

$$\Delta\varphi = \frac{1}{2r} = \frac{1}{2\langle N\rangle^{1/2}}. \tag{2.121}$$

This is SNL and is in coincidence with the intuitive result of Eq. (2.67).

Next let us seek an input state that gives minimal phase variance. Suppose we have N photons available for this input state $|\Psi\rangle$,

$$|\Psi\rangle = \sum_{n=0}^{N} c_n|n\rangle. \tag{2.122}$$

The phase distribution is expressed as

$$p(\varphi) = \frac{1}{2\pi}\langle\Psi|E(\varphi)|\Psi\rangle, \tag{2.123}$$

where $E(\varphi)$ is given by Eq. (2.117). The use of the Holevo variance defined below, taking account of the cyclic property of phase shift, enables an analytic solution. The Holevo variance is [17]

$$\begin{aligned} V_\varphi &= S_\varphi^{-2} - 1, \\ S_\varphi &= \langle e^{i\varphi} \rangle = \int_{-\pi}^{\pi} e^{i\varphi} p(\varphi) d\varphi, \end{aligned} \tag{2.124}$$

where we have chosen the true phase shift to be zero to make S_φ positive without loss of generality. If the variance is small, i.e. $S_\varphi \approx 1$, its equivalence to the usual definition of variance for well-localized distributions is obvious by observing the fact

$$\Delta\varphi = \int \varphi^2 p(\varphi) d\varphi \approx 2(1 - S_\varphi) \approx (S_\varphi^{-2} - 1). \tag{2.125}$$

From Eqs. (2.123) and (2.124) we get

$$S_\varphi = \langle \Psi | E | \Psi \rangle, \tag{2.126}$$

where the integral over φ is performed and the operator E is a matrix with elements $E_{m,n} = \delta_{m,n\pm1}$ and $m, n = 0, ..., N$. Since the minimal variance corresponds to a maximal S_φ, and the maximum of $\langle \Psi | E | \Psi \rangle$ is achieved when $|\Psi\rangle$ is an eigenvector of matrix S with the largest eigenvalue. This eigenvector is solved to be [43]

$$|\Psi\rangle = \sqrt{\frac{2}{N+2}} \sum_{n=0}^{N} \sin \frac{(n+1)\pi}{N+2} \; |n\rangle \tag{2.127}$$

with eigenvalue $\cos \pi/(N+2)$. Hence the minimal variance is

$$V_\varphi = \tan^2 \frac{\pi}{N+2} \approx \frac{\pi}{N+2}. \tag{2.128}$$

2.4.3 Quantum Cramer-Rao Bound and the Heisenberg Limit

As in the classical estimation theory, there is also a quantum Cramer-Rao bound, which is a variant of uncertainty principle [21, 22]. Suppose an operator $U = \exp(-iJ\theta)$ is applied to a probe state $|\Psi\rangle$ where J is a Hermitian operator and θ is a parameter to be estimated. Then the observable A is measured on the state $|\Psi(\theta)\rangle = U|\Psi\rangle$ to estimate the parameter θ. The uncertainty relation yields

$$\Delta A \Delta J \geq \frac{1}{2} \left| \langle \Psi(\theta) | [J, A] | \Psi(\theta) \rangle \right| = \frac{1}{2} \left| \frac{\partial \langle \Psi(\theta) | A | \Psi(\theta) \rangle}{\partial \theta} \right|, \tag{2.129}$$

where $\Delta A, \Delta J$ denote the variance in A, J. The resulting precision in estimating θ is thus given by

$$\Delta\theta = \frac{\Delta A}{|\partial\langle\Psi(\theta)|A|\Psi(\theta)\rangle/\partial\theta|} \geq \frac{1}{2\Delta J}. \tag{2.130}$$

Once again if we take $J = N$ and $\theta = -\varphi$, this bound tells us

$$\Delta\varphi \geq \frac{1}{2\Delta N}. \tag{2.131}$$

For a coherent state $|\alpha\rangle$, the variance of N is $\Delta N = |\alpha| = \langle N\rangle^{1/2}$ and therefore

$$\Delta\varphi \geq \frac{1}{2\langle N\rangle^{1/2}}, \tag{2.132}$$

which is just the so-called SNL.

For a general state with N available photons,

$$|\Psi\rangle = \sum_{n=0}^{N} c_n|n\rangle, \tag{2.133}$$

intuitively the variance of photon number N should satisfy the energy constrain,

$$\Delta N \sim N. \tag{2.134}$$

Hence, the best phase uncertainty we can expect is

$$\Delta\varphi = \frac{1}{N} \tag{2.135}$$

after canceling out some constants. This is what is called HL for phase estimation. Mathematically, consider a discrete random variable N taking value from the set $\{n\}$ with probability $\{|c_n|^2\}$, subject to the constrain

$$\sum_{n=0}^{N} n|c_n|^2 = \langle N\rangle, \tag{2.136}$$

then the variance of N is bounded by the inequalities

$$0 \leq \Delta N^2 \leq \langle N\rangle(N - \langle N\rangle) \leq \frac{N^2}{4}. \tag{2.137}$$

The last two equality signs hold iff $|c_n|^2 = 0$ for $n = 1, ..., N-1$ and $\langle N\rangle = N/2$ with $|c_0|^2 = |c_N|^2 = 1/2$, respectively. Its proof is based on the following theorem [49]: If $X = E[x]$, where x represents an arbitrary random variable and $E[x]$ the average value of x, then for all a and b,

$$E[(x-X)^2] = E[(x-a)(x-b)] - (a-X)(b-X). \tag{2.138}$$

So the state with the maximal variance of N is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |N\rangle) \tag{2.139}$$

up to some phase factors, and

$$\langle N\rangle = \frac{N}{2}, \qquad \Delta N = \frac{N}{2} = \langle N\rangle. \tag{2.140}$$

Then a more rigorous form of HL from Eq. (2.131) is obtained

$$\Delta\varphi \geq \frac{1}{N} = \frac{1}{2\langle N\rangle}. \tag{2.141}$$

Now let us extend the finite terms in the sum of Eq. (2.133) to be infinite, i.e.

$$|\Psi\rangle = \sum_{n=0}^{\infty} c_n|n\rangle, \tag{2.142}$$

and only keep the average photon number constrain

$$\sum_{n=0}^{\infty} n|c_n|^2 = \langle N\rangle. \tag{2.143}$$

We will show in chapter 4 that to detect a phase shift of order $\Delta\varphi$ with the input state $|\Psi\rangle$ the quantity defined via $F = \left|\langle\Psi|e^{iN\Delta\varphi}|\Psi\rangle\right|$ should be less than 1 by a finite quantity. This gives

$$\begin{aligned} F &= \left|\langle\Psi|e^{iN\Delta\varphi}|\Psi\rangle\right| \\ &= \left|1 - \sum_n |c_n|^2 (1 - e^{in\Delta\varphi})\right| \\ &\geq 1 - 2\left|\sum_n |c_n|^2 e^{in\Delta\varphi/2} \sin(n\Delta\varphi/2)\right| \\ &\geq 1 - 2\sum_n |c_n|^2 \sin(n\Delta\varphi/2) \geq 1 - \langle N\rangle\Delta\varphi, \end{aligned} \tag{2.144}$$

where the relations $|a-b| \geq ||a| - |b||$ and $\sin x \leq x$ are used. Hence we get the result

$$\Delta\varphi \geq \frac{1-F}{\langle N\rangle} \sim \frac{1}{\langle N\rangle} \tag{2.145}$$

in parallel with Eq. (2.141).

2.5 Optimal Phase Estimator

The optimal POVM for phase parameter in the state $\Psi(\theta) = e^{-iJ\theta}\Psi$ is shown in details [47]. Suppose the desired POVM $dE(\hat{\theta}) = \mathcal{E}(\theta)d\theta/(2\pi) = e^{-\mathrm{i}J\theta}\mathcal{E}e^{\mathrm{i}J\theta}d\theta/(2\pi)$. The p.d.f. of this measurement $p(\hat{\theta}|\theta)$ depends only on $\hat{\theta} - \theta$,

$$p(\hat{\theta}|\theta) = \mathrm{Tr}[\rho(\theta)\mathcal{E}(\hat{\theta})] = \mathrm{Tr}[e^{-\mathrm{i}J(\theta-\hat{\theta})}\rho\, e^{\mathrm{i}J(\theta-\hat{\theta})}\mathcal{E}]. \tag{2.146}$$

If $z(\theta) = 1/(2\pi)$ and $C(\hat{\theta}, \theta) = -\delta(\hat{\theta} - \theta)$, the conditions for the optimal POVM are

$$\begin{aligned} [\Upsilon - \rho(\theta)]\mathcal{E}(\theta) &= 0, \\ \Upsilon - \rho(\theta) &\geq 0, \\ \Upsilon = \int e^{-\mathrm{i}J\theta}\rho\,\mathcal{E}\, e^{\mathrm{i}J\theta}\frac{d\theta}{2\pi}&. \end{aligned} \tag{2.147}$$

Suppose $\mathcal{E} = |\Phi\rangle\langle\Phi|$, $J|j\rangle = j|j\rangle$, and the set $\{|j\rangle\}$ forms a complete orthogonal base. Then the identity $\int \mathcal{E}(\theta)d\theta/(2\pi) = 1$ leads to $|\langle j|\Phi\rangle| = 1$. It will be shown below that the choice $\langle j|\Phi\rangle = \gamma_j \equiv c_j/|c_j|$ satisfy the optimal condition. Hence

$$|\Phi\rangle = \sum_j |j\rangle\langle j|\Phi\rangle = \sum_j |j\rangle\gamma_j. \tag{2.148}$$

Note that $e^{-\mathrm{i}J\widetilde{\theta}}\Upsilon e^{\mathrm{i}J\widetilde{\theta}} = \Upsilon$ for any $\widetilde{\theta}$ gives $[\Upsilon, J] = 0$. Therefore, the optimal conditions become

$$\begin{aligned} (\Upsilon - \rho)\mathcal{E} &= 0, \\ \Upsilon - \rho &\geq 0. \end{aligned} \tag{2.149}$$

It remains to check the chosen POVM satisfies these two conditions.

For the proof of the nonnegative $\Upsilon - \rho$, the useful matrix elements in the base $\{|j\rangle\}$ are given by

$$\begin{aligned}
\langle m|\rho|n\rangle &= c_m c_n^*, \\
\langle m|\mathcal{E}|n\rangle &= \gamma_m \gamma_n^*, \\
\langle \Psi|\Phi\rangle &= \sum_j |c_j|, \\
\langle m|\Upsilon|n\rangle &= \int e^{-\mathrm{i}(m-n)\theta} \langle m|\rho\,\mathcal{E}|n\rangle \frac{d\theta}{2\pi} \\
&= \langle m|\rho\,\mathcal{E}|m\rangle \delta_{mn}, \\
\langle m|\rho\,\mathcal{E}|n\rangle &= c_m \gamma_n^* \sum_j |c_j|. \qquad (2.150)
\end{aligned}$$

For any state $|\widetilde{\Psi}\rangle = \sum_j \widetilde{c}_j |j\rangle$, one has

$$\langle \widetilde{\Psi}|\Upsilon - \rho|\widetilde{\Psi}\rangle = \sum_m |\widetilde{c}_m|^2 |c_m| \sum_n |c_n| - |\sum_m \widetilde{c}_m^* c_m|^2 \geq 0, \qquad (2.151)$$

where the last step follows from the Schwarz inequality. The first condition is also fulfilled,

$$\begin{aligned}
\langle m|(\Upsilon - \rho)\mathcal{E}|n\rangle &= \langle m|\rho\,\mathcal{E}|m\rangle\langle m|\mathcal{E}|n\rangle - \langle m|\rho\,\mathcal{E}|n\rangle \\
&= |c_m|\gamma_m \gamma_n^* \sum_j |c_j| - c_m \gamma_n^* \sum_j |c_j| \\
&= 0. \qquad (2.152)
\end{aligned}$$

Chapter 3: Parity as a Unified Measurement Scheme

Phase measurement using a lossless MZI with certain entangled N-photon states can lead to a phase sensitivity of the order of $1/N$, i.e. HL. However, previously considered output measurement schemes are different for different input states to achieve this limit. We show that it is possible to achieve this limit just by the parity measurement for all the commonly proposed entangled states. Based on the parity measurement scheme, the reductions of the phase sensitivity in the presence of photon loss are examined for the various input states.

3.1 Introduction

The notion of quantum entanglement holds great promise for certain computational and communication tasks. It is also at the heart of metrology and precision measurements in extending their capabilities beyond the so-called standard quantum limit [5]. For example, the phase sensitivity of a usual two-port interferometer is SNL as $1/\sqrt{N}$, where N is the number of the photons entering the input port. However, a properly correlated Fock-state input for MZI can lead to an improved phase sensitivity that scales as $1/N$, i.e., HL [13, 42, 51, 52]. In the subsequent development, the dual Fock-state [11] and the so-called intelligent state [53, 54] were proposed to reach a sub-shot-noise sensitivity as well. Recently, much attention has been paid to NOON state to reach the exact HL in interferometry as well as super-resolution imaging [15, 55, 31, 56].

The utilization of those quantum correlated input states are accompanied by various output measurement schemes. In some cases the conventional measurement scheme of photon-number difference is used, whereas a certain probability distribution [57, 12, 58, 59], a specific adaptive measurement [16, 17, 60], and the parity measurement are used for other cases.

Gerry and Campos first showed the use of the parity measurement for the "maximally entangled state"–the NOON state–of light to reach the exact HL [18], following the earlier suggestion of the HL spectroscopy with N two-level atoms [14]. Campos, Gerry, and Benmoussa later suggested that the parity measurement scheme can also be used for the dual Fock state inputs by comparing

the quantum state *inside* the interferometer with the NOON state [19]. In this paper we show that the parity measurement can be used as a detection scheme for sub-shot-noise interferometry with the correlated Fock states first proposed by Yurke, McCall, and Klauder [13], as well as with the intelligent states first suggested by Hillery and Mlodinow [53]. Extension of its use for all these input states then promote the parity measurement to a kind of universal detection scheme for quantum interferometry. Then, based on such a universal detection scheme comparisons of performance of various quantum states can be made in a common ground. As an example, we present a comparison of the phase sensitivity reduction for various quantum states of light in the presence of photon loss.

3.2 The Group-Theoretic Analysis of MZI

In order to describe the notations, we briefly review the group theoretical formalism of MZI. The key point of such a formalism is that any passive lossless four-port optical system can be described by the SU(2) group [13]. First, we use the mode annihilation operators $a_{in(out)}$ and $b_{in(out)}$, which satisfy boson commutation relations, to represent the two light beams entering (leaving) the beam splitter (BS), respectively. Then the action of BS takes the form

$$\begin{pmatrix} a_{out} \\ b_{out} \end{pmatrix} = \begin{pmatrix} e^{i(\alpha+\gamma)/2}\cos\frac{\beta}{2} & e^{-i(\alpha-\gamma)/2}\sin\frac{\beta}{2} \\ -e^{i(\alpha-\gamma)/2}\sin\frac{\beta}{2} & e^{-i(\alpha+\gamma)/2}\cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} a_{in} \\ b_{in} \end{pmatrix}. \tag{3.1}$$

Here α, β, and γ denote the Euler angles parameterizing SU(2), and they are related to the complex transmission and reflection coefficients. Through the angular momentum representation we can construct the operators for the angular momentum and for the occupation number from the mode operators a and b,

$$\mathbf{J} = \begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \frac{1}{2}\begin{pmatrix} ab^\dagger + ba^\dagger \\ i(ab^\dagger - ba^\dagger) \\ aa^\dagger - bb^\dagger \end{pmatrix}, \tag{3.2}$$

and $N = a^\dagger a + b^\dagger b$. The commutation relations $[a, b] = [a, b^\dagger] = 0$ and $[a, a^\dagger] = [b, b^\dagger] = 1$ lead to the relation $\mathbf{J} \times \mathbf{J} = i\mathbf{J}$. The group invariant has the form $J^2 = J_x^2 + J_y^2 + J_z^2 = (N/2)(N/2+1)$ that commutes with J_i and N. Next, it was shown that the operation of the BS is equivalent to

[13]

$$\mathbf{J}_{out} = e^{i\alpha J_z} e^{i\beta J_y} e^{i\gamma J_z} \mathbf{J}_{in} e^{-i\gamma J_z} e^{-i\beta J_y} e^{-i\alpha J_z}, \tag{3.3}$$

in the Heisenberg picture, and to

$$|\text{out}\rangle = e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z} |\text{in}\rangle, \tag{3.4}$$

in the Schrödinger picture. If we use the symbols j and m to indicate the eigenvalues of $N/2$ and J_z, then the theory of angular momentum tells that the representation Hilbert space is spanned by the complete orthonormal basis $|j, m\rangle$ with $m \in [-j, j]$, which can also be labeled by the Fock states of the two modes, $|j, m\rangle = |j + m\rangle_a |j - m\rangle_b$. In terms of this language, we may make the geometrical interpretation of the elements of a MZI. For example, the effect of a 50/50 BS, which leads a $\pm\pi/2$ rotation around the x axis (given by the unitary transformation $e^{\pm i(\pi/2)J_x}$), is equivalent to the transformation

$$\begin{pmatrix} a_{out} \\ b_{out} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \mp i \\ \mp i & 1 \end{pmatrix} \begin{pmatrix} a_{in} \\ b_{in} \end{pmatrix}. \tag{3.5}$$

Similarly, the relative phase shift φ acquired between the two arms of MZI can be expressed by $a_{out} = a_{in}$, $b_{out} = e^{i\varphi} b_{in}$, or by the unitary transformation $e^{-i\varphi J_z}$ equivalently.

MZI can be illustrated schematically in Fig. 1.1, where the two light beams a and b first enter the BS_+, and then acquires a relative phase shift φ, and finally pass through the BS_-. The photons leaving the BS_- are counted by detectors D_a and D_b. Therefore, in the language of the group theory, the input states of BS_+ and the output states of BS_- is connected by a simple unitary transformation $U = e^{i(\pi/2)J_x} e^{-i\varphi J_z} e^{-i(\pi/2)J_x} = e^{-i\varphi J_y}$. Its effect is the equivalent to a BS, which gives a φ rotation around the y axis,

$$\begin{pmatrix} a_{out} \\ b_{out} \end{pmatrix} = \begin{pmatrix} \cos\frac{\varphi}{2} & -\sin\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix} \begin{pmatrix} a_{in} \\ b_{in} \end{pmatrix}. \tag{3.6}$$

The information on the phase shift φ is inferred from the photon statistics of the output beams. There are many statistical methods to extract such information. The most common one is to use

the difference between the number of photons in the two output modes, $N_d = a_{out}^\dagger a_{out} - b_{out}^\dagger b_{out}$, or equivalently, $J_{z,out} = N_d/2$. The minimum detectable phase shift then can be estimated by [3]

$$\Delta\varphi = \frac{\Delta J_{z,out}}{|\partial\langle J_{z,out}\rangle/\partial\varphi|}, \tag{3.7}$$

where $\Delta J_{z,out} = \sqrt{\langle J_{z,out}^2\rangle - \langle J_{z,out}\rangle^2}$. The expectation value of $J_{z,out}$ and $J_{z,out}^2$ are calculated by $\langle J_{z,out}\rangle = \langle \text{in}|J_{z,out}|\text{in}\rangle = \langle \text{in}|U^\dagger J_{z,in} U|\text{in}\rangle$, $\langle J_{z,out}^2\rangle = \langle \text{in}|J_{z,out}^2|\text{in}\rangle = \langle \text{in}|U^\dagger J_{z,in}^2 U|\text{in}\rangle$, and $U^\dagger J_{z,in}^n U = (-\sin\varphi J_{x,in} + \cos\varphi J_{z,in})^n$.

Now the application of the group formalism to analyze the phase sensitivity of the ideal MZI is straightforward. Let us first consider the correlated photon-number states [13, 42, 52]. In particular, the so-called Yurke state has the form $|\text{in}\rangle = [|j,0\rangle + |j,1\rangle]/\sqrt{2}$, which is one of the earliest proposals of utilizing the correlated photon-number states [13]). A simple calculation for the Yurke-state input gives

$$\Delta\varphi = \frac{\left\{[j(j+1)-1]\sin^2\varphi + \cos^2\varphi\right\}^{1/2}}{|\sqrt{j(j+1)}\cos\varphi + \sin\varphi|}, \tag{3.8}$$

which has its minimum value $\Delta\varphi_{min} \approx 1/\sqrt{j(j+1)}$ when $\sin\varphi \approx 0$. Hence, when the Yurke state is fed into the input ports of an interferometer, the minimum of $\Delta\varphi$ has the order of $2/N$ limit since $j = N/2$. We should bear in mind that the minimum phase sensitivity is achieved only at particular values of $\varphi \approx 0$. For other values of φ the phase sensitivity is decreased. However, one can always control the phase shift by a feed-back loop which keeps φ at any particular value.

3.3 Parity as a Unified Measurement Scheme

The parity measurement, represented by the observable $P = (-1)^{b^\dagger b} = e^{i\pi(j-J_z)}$ has an advantage when the simple photon number counting method ceases to be appropriate to infer the phase shift and provides a wider applicability than J_z. The parity measurement scheme was first introduced by Bollinger, Itano, Wineland, and Heinzen for spectroscopy with trapped ions of maximally entangled form [14]. Gerry and Campos adopted such a measurement scheme to the optical interferometry with the NOON state [18]. The NOON state can be formally written as $|\text{NOON}\rangle = [|j,j\rangle + |j,-j\rangle]/\sqrt{2}$. Note that the NOON state is *not* the input state of MZI, but the state *after* the first beam splitter BS_+. Hence the output state is described as $|\text{out}\rangle = e^{i(\pi/2)J_x} e^{-i\varphi J_z}|\text{NOON}\rangle$.

The expectation value for the parity operator is then given by

$$\begin{aligned}\langle P\rangle &= i^N\langle \mathrm{NOON}|e^{i\varphi J_z}e^{i\pi J_y}e^{-i\varphi J_z}|\mathrm{NOON}\rangle \\ &= i^N\left[e^{iN\varphi}+(-1)^N e^{-iN\varphi}\right]/2, \end{aligned} \tag{3.9}$$

so that we have

$$\langle P\rangle = \begin{cases} i^{N+1}\sin N\varphi, & N \ \text{ odd}, \\ i^N\cos N\varphi, & N \ \text{ even}, \end{cases} \tag{3.10}$$

where the identity $e^{-i(\pi/2)J_x}e^{-i\pi J_z}e^{i(\pi/2)J_x}=e^{i\pi J_y}$ is applied and Eq. (3.6) with $\varphi=-\pi$ is used. Since $P^2=1$, the equation (3.10) then immediately leads to the result $\Delta\varphi=1/N$, exactly.

Now, let us consider the dual Fock-state as the input state, $|j,0\rangle=|j\rangle_a|j\rangle_b$. Here, if we still use $J_{z,out}$ as our observable, we have

$$\langle J_{z,out}\rangle = \langle j,0|-\sin\varphi J_x+\cos\varphi J_z|j,0\rangle = 0. \tag{3.11}$$

The expectation value of the difference of the output photon number is now independent of the phase shift. Therefore, in this case the measurement of $J_{z,out}$ contains no information about the phase shift. A method of reconstruction of the probability distribution has been proposed to avoid this phase independence and to reach HL [11, 58, 59]. More recently, Campos, Gerry, and Benmoussa suggested the use of the parity measurement for the dual Fock-state inputs [19].

The expectation value of P can be derived from $\langle P_{out}\rangle=\langle \mathrm{in}|e^{i\varphi J_y}P_{in}e^{-i\varphi J_y}|\mathrm{in}\rangle$ and $\langle P_{out}^2\rangle=\langle \mathrm{in}|\mathrm{in}\rangle=1$. For the dual Fock-state, we have

$$\begin{aligned}\langle P_{out}\rangle^{\text{d-Fock}} &= \langle j,0|e^{i\varphi J_y}(-1)^{j-J_z}e^{-i\varphi J_y}|j,0\rangle \\ &= (-1)^j \ d_{0,0}^{(j)}(2\varphi), \end{aligned} \tag{3.12}$$

where $d_{m,n}^{(j)}$ denotes the rotation matrix element: $e^{-i\varphi J_y}|j,n\rangle=\sum_{m=-j}^{j}d_{m,n}^{(j)}(\varphi)|j,m\rangle$, and

$$\begin{aligned} d_{m,n}^{(j)}(\varphi) &= (-1)^{m-n}2^{-m}\sqrt{\frac{(j-m)!(j+m)!}{(j-n)!(j+n)!}} \\ &\times \ P_{j-m}^{(m-n,m+n)}(\cos\varphi)\ (1-\cos\varphi)^{\frac{m-n}{2}}(1+\cos\varphi)^{\frac{m+n}{2}}, \end{aligned}$$

where $P_n^{(\alpha,\beta)}(x)$ represents the Jacobi polynomial. Thus the phase sensitivity is obtained as

$$\Delta\varphi^{\text{d-Fock}} = \frac{\sqrt{1-[d_{0,0}^{(j)}(2\varphi)]^2}}{|\partial d_{0,0}^{(j)}(2\varphi)/\partial\varphi|} \tag{3.13}$$

for the dual Fock-state, and in the limit of $\varphi \to 0$, we have

$$\Delta\varphi^{\text{d-Fock}} \to \frac{1}{\sqrt{2j(j+1)}} \sim \frac{\sqrt{2}}{N}. \tag{3.14}$$

If we use the parity measurement scheme for the Yurke-state input, we obtain

$$\begin{aligned}
&\langle P_{out}\rangle^{\text{Yurke}} = \langle \text{in}|e^{i\varphi J_y}(-1)^{j-J_z}e^{-i\varphi J_y}|\text{in}\rangle \\
&= \sum_{m=-j}^{j} \frac{(-1)^{j-m}}{2}\left(d_{m,0}^{(j)*} + d_{m,1}^{(j)*}\right)\left(d_{m,0}^{(j)} + d_{m,1}^{(j)}\right) \\
&= \frac{(-1)^j}{2}\left[d_{0,0}^{(j)} + d_{0,1}^{(j)} - d_{1,0}^{(j)} - d_{1,1}^{(j)}\right](2\varphi),
\end{aligned} \tag{3.15}$$

where have used the following properties of the matrix element [61] in the last line of (4.3):

$$\begin{aligned}
d_{m,n}^{(j)*} &= d_{m,n}^{(j)} = (-1)^{m-n} d_{n,m}^{(j)} = d_{-n,-m}^{(j)} \\
\sum_{m=-j}^{j} d_{k,m}^{(j)}(\varphi_1)\ d_{m,n}^{(j)}(\varphi_2) &= d_{k,n}^{(j)}(\varphi_1+\varphi_2).
\end{aligned} \tag{3.16}$$

Again, using

$$\Delta\varphi = \frac{\sqrt{1-[\langle P_{out}\rangle^{\text{Yurke}}]^2}}{|\partial\langle P_{out}\rangle^{\text{Yurke}}/\partial\varphi|}, \tag{3.17}$$

we have

$$\Delta\varphi^{\text{Yurke}} \to \frac{1}{\sqrt{j(j+1)}} \sim \frac{2}{N}, \tag{3.18}$$

in the limit of $\varphi \to 0$. This shows that, for the Yurke state, the parity measurement scheme leads to the same phase sensitivity as the $J_{z,out}$ measurement scheme. The dual-Fock state then performs better than the Yurke-state by a factor of $\sqrt{2}$ within the parity measurement scheme.

We can also use parity observable for the intelligent state entering the first beam splitter BS_+ in Fig. 1.1. The intelligent state is defined as the solution of the equation

$$(J_y + i\eta J_z)\,|j, m_0, \eta\rangle = \beta|j, m_0, \eta\rangle, \tag{3.19}$$

where $\eta^2 = (\Delta J_y)^2/(\Delta J_z)^2$ and m_0 is an integer belonging to $[-j, j]$ [53]. The eigenvalue corresponding to $|j, m_0, \eta\rangle$ is $\beta = im_0\sqrt{\eta^2-1}$ and the eigenvector $|j, m_0, \eta\rangle = \sum_{k=-j}^{j} C_k|j, k\rangle$, where

an explicit form of the expansion coefficient C_k is given in Ref. [54]. The expectation value of the parity operator is then obtained as

$$\langle P_{out}\rangle^{\rm Int} = (-1)^j \sum_{k,n=-j}^{j} C_k^* C_n (-1)^k d_{k,n}^{(j)}(2\varphi). \tag{3.20}$$

It follows that from the explicit form of C_k's the phase sensitivity scales better with a larger η and a smaller $|m_0|$. As $\eta \to \infty$, the phase sensitivity becomes

$$\Delta\varphi^{\rm Int} \to \frac{1}{\sqrt{2(j^2 - m_0^2 + j)}} \sim \frac{\sqrt{2}}{N}. \tag{3.21}$$

On the other hand, as $\eta \to 1$, we have $\Delta\varphi^{\rm Int} \to 1/\sqrt{2j} \sim 1/\sqrt{N}$, which is the standard shot-noise limit. So the minimum value of $\Delta\varphi$ is only accessible for $m_0 = 0$. This limiting behavior is the same as the phase sensitivity with J_z measurement at $\varphi = 0$ [54]. We note that, within the parity measurement scheme, of all states considered here only the NOON state reaches exactly HL.

3.4 The Effect of Photon Loss on Phase Sensitivity

We have seen that we can adopt the parity measurement as a universal detection scheme for all the commonly used entangled states, we will use it as a common ground to compare the effect of photon loss on phase sensitivity, thus we can put each input state on the same footing.

The effect of photon loss has been recently studied for the NOON states. Gilbert and coworkers applied a model for loss as a series of beam splitters in the propagation paths [70]. Rubin and Kaushik applied a single beam-splitter model for loss on the detection operator [29]. Whereas the two approaches are equivalent, we adopt the one given in Ref. [70] by putting the the effect of photon loss in the following form [63]:

$$a_{out} = e^{(-i\eta_a\omega/c - K_a/2)L_a} a_{in} + i\sqrt{K_a}\int_0^{L_a} dz\ e^{(-i\eta_a\omega/c - K_a/2)(L_a - z)} d(z), \tag{3.22}$$

$$b_{out} = e^{(-i\eta_b\omega/c - K_b/2)L_b} b_{in} + i\sqrt{K_b}\int_0^{L_b} dz\ e^{(-i\eta_b\omega/c - K_b/2)(L_b - z)} d(z), \tag{3.23}$$

where η_i is the index of refraction for arm i of the interferometer, K_i is the absorption coefficient, and L_i is the path length. The annihilation operator $d(z)$ is the modes into which photons are scattered. Typically, the photon loss can be modeled by making the substitution $\varphi \to \varphi + i\gamma$, where $\gamma = KL/2$ is the rate of absorption. The effects that it produces on coherent state $|\alpha\rangle$ and

number state $|N\rangle$ are

$$|\alpha\rangle \rightarrow |e^{-\gamma+i\varphi}\alpha\rangle, \tag{3.24}$$

$$|N\rangle \rightarrow e^{-N(\gamma+i\varphi)}|N\rangle, \tag{3.25}$$

respectively. The exponential dependence of the loss in the coherent or classical state is called Beer's law for optical absorption. For number state, we have a super-exponential behavior, which is called super-Beer's law.

The observable used for the output detection schemes in both Refs. [70, 29], namely, $A = |N,0\rangle\langle 0,N| + |0,N\rangle\langle N,0|$, is equivalent to the parity measurement for the NOON state [15]. In addition, if we now only consider the measurement performed in the *N-photon subspace* of the output state, we can ignore the scattering term of the above transformation.

Following Ref. [70], we assume that the losses are present only in the one of the two arms of the interferometer and set $e^{-K_a L_a} = 1$ and $e^{-K_b L_b} \equiv \lambda^2$. The associated operation of the lossy MZI then can be expression as

$$\begin{pmatrix} a_{out} \\ b_{out} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\lambda e^{i\varphi} & -i(1-\lambda e^{i\varphi}) \\ i(1-\lambda e^{i\varphi}) & 1+\lambda e^{i\varphi} \end{pmatrix}\begin{pmatrix} a_{in} \\ b_{in} \end{pmatrix}, \tag{3.26}$$

which is non-unitary unless $\lambda = 1$. In the angular momentum representation, this transformation can be rephrased as

$$\begin{aligned} L(\varphi) &= e^{iJ_x\frac{\pi}{2}}\Lambda e^{-iJ_x\frac{\pi}{2}}e^{iJ_x\frac{\pi}{2}}e^{-iJ_z\varphi}e^{-iJ_x\frac{\pi}{2}} \\ &= e^{iJ_x\frac{\pi}{2}}\Lambda e^{-iJ_x\frac{\pi}{2}}e^{-iJ_y\varphi}, \end{aligned} \tag{3.27}$$

where Λ is a matrix representing the effect of path absorption. Then we get

$$\begin{aligned} L^\dagger(\varphi)P_N L(\varphi) &= e^{iJ_y\varphi}e^{iJ_x\frac{\pi}{2}}\Lambda e^{-iJ_x\frac{\pi}{2}}P_N e^{iJ_x\frac{\pi}{2}}\Lambda e^{-iJ_x\frac{\pi}{2}}e^{-iJ_y\varphi} \\ &= e^{iJ_y\varphi}L(0)P_N L(0)e^{-iJ_y\varphi} \\ &= \lambda^N e^{iJ_y\varphi}P_N e^{-iJ_y\varphi} \equiv \mathcal{Y}_1. \end{aligned} \tag{3.28}$$

with

$$P_N = P \otimes \sum_{m=-j}^{j} |j,m\rangle\langle j,m| \tag{3.29}$$

denoting the N-photon projected parity operator. That is to say, one needs to detect all the N photons, even though that probability decreases exponentially with N. To prove above relation, we note that in the spinor representation, the truncated parity operator P_N can be expressed as

$$P_N = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z \tag{3.30}$$

up to common phase factor. Hence $L(0)P_N L(0)$ in Eq. (3.28) corresponds to

$$\frac{1}{2}\begin{pmatrix} 1+\lambda & -i(1-\lambda) \\ i(1-\lambda) & 1+\lambda \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1+\lambda & -i(1-\lambda) \\ i(1-\lambda) & 1+\lambda \end{pmatrix} = \lambda\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.31}$$

in the spinor representation, which corresponds to $\lambda^N P_N$ in the angular momentum representation.

Similarly, we find

$$\begin{aligned} L^\dagger P_N^2 L &= L^\dagger L = e^{iJ_y\varphi}e^{iJ_x\frac{\pi}{2}}\Lambda^2 e^{-iJ_x\frac{\pi}{2}}e^{-iJ_y\varphi} \\ &= e^{iJ_x\frac{\pi}{2}}\Lambda^2 e^{-iJ_x\frac{\pi}{2}} \equiv \mathcal{Y}_2, \end{aligned} \tag{3.32}$$

where the commutability of $\mathcal{Y}_2$ and $e^{-iJ_y\varphi}$ is applied. It is easy to check in the spinor representation,

$$\begin{aligned} &\frac{1}{2}\begin{pmatrix} 1+\lambda^2 & -i(1-\lambda^2) \\ i(1-\lambda^2) & 1+\lambda^2 \end{pmatrix}\begin{pmatrix} \cos\frac{\varphi}{2} & -\sin\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos\frac{\varphi}{2} & -\sin\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1+\lambda^2 & -i(1-\lambda^2) \\ i(1-\lambda^2) & 1+\lambda^2 \end{pmatrix} \end{aligned} \tag{3.33}$$

by the matrix identity

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} c & d \\ -d & c \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{3.34}$$

Now, for a general input state, $|\text{in}\rangle = \sum_{m=-j}^{j} c_m |j,m\rangle$, we obtain

$$\langle P_N\rangle_{\text{out}} = \langle \mathcal{Y}_1\rangle_{\text{in}} = (-1)^j \lambda^{2j} \sum_{m,n} c_m^* c_n (-1)^m d_{mn}^{(j)}(2\varphi) \tag{3.35}$$

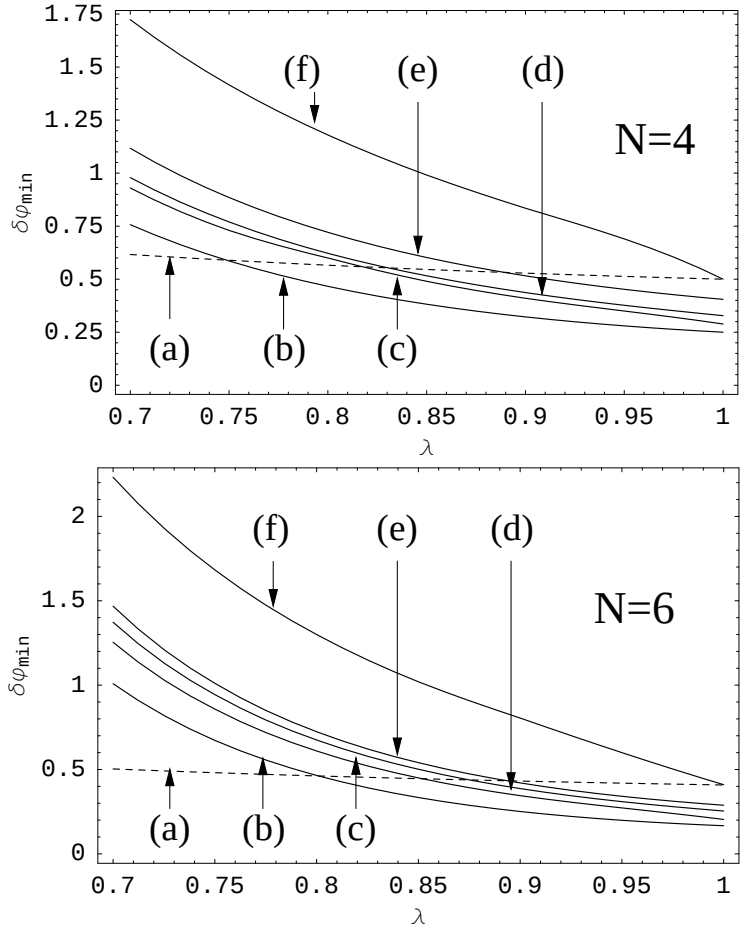

FIGURE 3.1. The minimum phase sensitivity, $\Delta\varphi_{\min}$, for the various entangled states as a function of λ, the transmission coefficient. The upper and lower figures are for $N = 4$ and $N = 6$, respectively. The dotted line (a) represents that of the uncorrelated input state [70]. The solid lines represent (b) the NOON state, (c) the dual Fock state, (d) the intelligent ($\eta = 10$) state, (e) the Yurke, and (f) the intelligent ($\eta = 1$) state, respectively.

and

$$\langle P_N^2 \rangle_{\rm out} = \langle \mathcal{Y}_2 \rangle_{\rm in} = (1/2) \sum_{m,n} c_m^* c_n [Q_{mn} + Q_{nm}](\lambda). \tag{3.36}$$

Here, the polynomial $Q_{mn}(\lambda)$ is defined as the matrix element $\langle j, m|\mathcal{Y}_2|j, n\rangle$ such that

$$\begin{aligned} Q_{mn}(\lambda) &= \frac{i^{-m-n}}{(j-n+1)_{j+n}} \frac{(2j)!\sqrt{(j+n)!}}{\sqrt{(j-m)!(j+m)!(j-n)!}} \\ &\quad \times \left(\frac{x^2-1}{4}\right)^j \left(\frac{1+x}{1-x}\right)^{j+n-m} P_{j+n}^{(-2j-1,m-n)}\left(1 - \frac{8x}{(x+1)^2}\right), \end{aligned} \tag{3.37}$$

where $x \equiv \lambda^2$ and $p_q \equiv \Gamma(p+q)/\Gamma(p)$.

We now compare the phase sensitivity for different entangled states in the presence of photon loss. The plots depicted in Fig. 3.1 show the reduced phase sensitivity due to the photon loss, in this case as a function of λ (the transmission coefficient). All the commonly proposed entangled

states are compared to the lossy-environment shot-noise limit. Among the entangled states, the best possible phase sensitivity can be achieved by the NOON state, and it gets worse in the following order: the dual Fock state, the $\eta = 10$ intelligent state, the Yurke state, and then the $\eta = 1$ intelligent state. Within the restricted parity measurement scheme the NOON states show the best performance for phase detection and can still beat SNL if the transmittance of interferometer is not too small and the photon number is not too large. We see that beating SNL (dotted line, represented by the uncorrelated input state) requires less attenuation as the number of photons increases. For example, the lowest solid line (representing the NOON states) requires 75% transmission for $N = 4$ and 80% for $N = 6$.

3.5 Conclusion

We showed that the utilization of the parity measurement in sub-shot-noise interferometry is applicable to a wide range of quantum entangled input states, so far known entangled states of light. Comparison of the performance of the various quantum states then can be made within such a unified output measurement scheme. Furthermore, it may lead to a great reduction of the efforts in precise quantum state preparation as well as in various optimization strategies involving quantum state engineering for the sub-shot-noise interferometry.

Chapter 4: A Simple Condition for Heisenberg-Limited Optical Interferometry

In this chapter I will present a simple condition to determine whether a given input state can be used to achieve the optical phase measurement at HL. In terms of the fidelity F between the two output states with zero phase and minimal detectable phase shift applied respectively, it can be expressed as $F < 1$, with the minimal phase shift scaling as HL, in the limit where the average number of input photons goes to infinity.

4.1 Introduction

The entangled states used in chapter 3 all have real coefficients. If the imaginary coefficients are allowed, the phase sensitivity with parity measurement will blow up as indicated below. It is thus useful to have a condition to check if a state can achieve HL that is independent of the detection scheme. One such criterion is the quantum Fisher information [47] that provides one lower bound for the phase sensitivity. However, the lower bound given by quantum Fisher information is only asymptotically achievable for a large number of runs of detections. Another lower bound for minimal phase shift based on the uncertainty relation is given in Ref. [22, 23], but whether it can be actually attained for a given state remains open. In this paper we find that the condition for one input state to reach HL can be expressed as the fidelity between two output states with zero phase and the minimal detectable phase $\Delta\varphi$ applied respectively should be less than 1 as $\Delta\varphi \sim 1/\langle N\rangle$ and $\langle N\rangle \to \infty$. We need to point out that this condition applies to the single run of detection of phase shift as well as the multiple runs. For multiple runs, the output state should be understood as the output state for the overall measurements.

4.2 An Example

In this section we calculate the phase sensitivity with an input state to a lossless Mach-Zehder interferometer, $|\Psi\rangle = \alpha|N\rangle_a|N\rangle_b + \beta|N+1\rangle_a|N-1\rangle_b$, which is the superposition of dual Fock state and Yurke state. In the angular momentum representation it can be expressed by $|\Psi\rangle = \alpha|j,0\rangle + \beta|j,1\rangle$ with $j = N$. We have known that the effect of MZI can be represented by a

unitary rotation operator, $U = e^{-i\varphi J_y}$. Then the output state is given by

$$|\Psi(\varphi)\rangle = U|\Psi\rangle = e^{-i\varphi J_y}\left[\,\alpha|j,0\rangle + \beta|j,1\rangle\right]. \tag{4.1}$$

Using the standard rotation matrix $U|j,n\rangle = \sum_{m=-j}^{j} d_{mn}^{(j)}(\varphi)|j,m\rangle$, we have

$$|\Psi(\varphi)\rangle = \sum_{n=-j}^{j}\left[\,\alpha\, d_{n,0}^{(j)} + \beta\, d_{n,1}^{(j)}\right](\varphi)|j,n\rangle. \tag{4.2}$$

To determine the phase sensitivity, we need to choose a detection scheme. Specifically, we use parity operator as our measurement. The parity measurement is given by the operator $P = (-1)^{(j-j_z)}$. We have previously shown that this scheme works well with all the known entangled states achieving HL in chapter 3. For the above states $|\Psi\rangle$, we get

$$\begin{aligned}
\langle P\,\rangle &= \langle\Psi(\varphi)|P\,|\Psi(\varphi)\rangle = \langle\Psi|e^{i\varphi J_y}P\,e^{-i\varphi J_y}|\Psi\rangle \\
&= (-1)^j\left[|\alpha|^2\, d_{0,0}^{(j)} + \beta\alpha^*\, d_{0,1}^{(j)} - \alpha\beta^* d_{1,0}^{(j)} - |\beta|^2\, d_{1,1}^{(j)}\right](2\varphi), \\
\langle P^2\rangle &= \langle\Psi(\varphi)|P^2\,|\Psi(\varphi)\rangle = 1,
\end{aligned} \tag{4.3}$$

and the phase sensitivity is given by

$$\Delta\varphi = \frac{\Delta P}{|\partial\langle P\rangle/\partial\varphi|} = \frac{\sqrt{1-\langle P\rangle^2}}{|\partial\langle P\rangle/\partial\varphi|} \tag{4.4}$$

In the limit of $\varphi \to 0$,

$$\Delta\varphi^2 \to \frac{1-\left(|\alpha|^2-|\beta|^2\right)^2}{\left(\alpha\beta^*+\beta\alpha^*\right)^2 j(j+1)} = \frac{1}{j(j+1)\cos^2\theta}, \tag{4.5}$$

where θ is the phase difference of α and β. If we confine the parameters, α and β to be real, the phase sensitivity goes exactly like $1/\sqrt{j(j+1)}$, which is the same as the pure Yurke state [13]. To achieve HL, the input state does not need to be exactly prepared as pure dual Fock state or Yurke state. We can use their superposition as input freely if the mixing amplitudes are real. When α and β are not in phase, i.e. $\alpha\beta^* + \beta\alpha^* = 0$ and $|\Psi\rangle = [|j,0\rangle + i\,|j,1\rangle]/\sqrt{2}$, $\Delta\varphi$ would go to infinity. Another known example is that the conventional $J_z = (a^\dagger a - b^\dagger b)/2$ measurement does not work with the dual Fock states [58]. However, this superficial infinity can be avoid by other estimation schemes, such as the maximal-likelihood method [64]. To avoid this scheme dependence, it is convenient to present a condition to determine whether a given input state can achieve HL independently of the detection scheme.

4.3 A Condition Based on Fidelity

For a generic normalized input state $|\Psi\rangle$, we apply a unitary operation U, after which it becomes $|\Psi'\rangle = U|\Psi\rangle$. In order to effectively distinguish these two states, we require the fidelity defined by $F = |\langle\Psi|\Psi'\rangle| < 1$. The best distinguishability corresponds to $F = 0$, which means the two states are orthogonal. To prove it, we apply a positive operator-value measurement (POVM) [65] E_m on Ψ and Ψ' to distinguish them. The POVM satisfies the conditions: (i) $\sum_m E_m = 1$, and (ii) the spectrum of E_m is positive for all m. The resulting probability distributions thus are $p_m = \langle\Psi|E_m|\Psi\rangle$ and $p'_m = \langle\Psi'|E_m|\Psi'\rangle$. A faithful quantity that quantifies the difference of the states Ψ and Ψ' is that of the probabilities p_m and p'_m. We therefore can take the overlap $\min_E \left\{\sum_m \sqrt{p_m}\sqrt{p'_m}\right\}$ to quantify this difference. From Schwarz inequality,

$$\sum_m \sqrt{\langle\Psi|E_m|\Psi\rangle\langle\Psi'|E_m|\Psi'\rangle} \geq \sum_m |\langle\Psi|E_m|\Psi'\rangle| \geq |\langle\Psi|\Psi'\rangle|. \tag{4.6}$$

So $\min_E \left\{\sum_m \sqrt{p_m}\sqrt{p'_m}\right\} = |\langle\Psi|\Psi'\rangle|$ if we choose $E_1 = |\Psi\rangle\langle\Psi|$ and $E_2 = 1 - |\Psi\rangle\langle\Psi|$. Hence if the fidelity $F < 1$, there is finite probability to distinguish them by performing the POVM, E_1 and E_2.

We now investigate its implication for our phase sensitivity in MZI. The unitary transformation for MZI is $U = e^{-i\Delta\varphi J_y}$. In order to detect a small phase shift $\Delta\varphi$, we need

$$F = \left|\langle\Psi|e^{-i\Delta\varphi J_y}|\Psi\rangle\right| < 1. \tag{4.7}$$

In particular, the condition for HL state can be stated as

$$\lim_{\langle N\rangle\to\infty} \left|\langle\Psi|e^{-i\Delta\varphi J_y}|\Psi\rangle\right|_{\Delta\varphi\to 1/\langle N\rangle} < 1. \tag{4.8}$$

Here the limiting process is formally used to test HL in the mathematical sense. It does not actually require infinite number of photons.

With this condition, we can also provide a lower bound for the phase sensitivity. The Margolus-Levitin theorem [66] indicates that a quantum system of energy E needs at least a time of $\pi\hbar/2E$ to go from one state to an orthogonal state. By analogy, the minimum value of $\Delta\varphi$ that make $\langle\Psi|e^{-i\Delta\varphi J_y}|\Psi\rangle \to 0$ satisfies

$$\Delta\varphi \geq \frac{\pi}{2(j_{max} + \langle J_y\rangle)}. \tag{4.9}$$

Here j_{max} denotes the maximum angular momentum in the components of $|\Psi\rangle$, which is added to $\langle J_y\rangle$ to make the denominator positive, in analogy with the positivity of E.

On the other hand, the uncertainty relation provides another bound for phase sensitivity [22]. Suppose an observable X is measured on the state $|\Psi(\varphi)\rangle = e^{-i\varphi J_y}|\Psi\rangle$ to estimate the phase. The uncertainty relation yields

$$\Delta X \Delta J_y \geq \frac{1}{2}\left|\langle\Psi(\varphi)|[J_y, X]|\Psi(\varphi)\rangle\right| = \frac{1}{2}\left|\frac{\partial\langle X\rangle}{\partial\varphi}\right|. \tag{4.10}$$

The resulting precision in estimating θ is thus given by

$$\Delta\varphi = \frac{\Delta X}{|\partial\langle X\rangle/\partial\varphi|} \geq \frac{1}{2\Delta J_y}. \tag{4.11}$$

Combining the above two lower bounds (4.9) and (4.11), we have

$$\Delta\varphi \geq \max\left\{\frac{\pi}{2(j_{max} + \langle J_y\rangle)},\ \frac{1}{2\Delta J_y}\right\}. \tag{4.12}$$

This inequality only gives a lower bound for phase sensitivity, whether the lower bound can be achieved should be checked separately with our initial condition (4.7). As an example, the inequality (4.12) implies $\Delta\varphi \geq \pi/N$ for the NOON state, $|\Psi\rangle = \left[|j, j\rangle_y + |j, -j\rangle_y\right]/\sqrt{2}$ with $j = N/2$, $\langle J_y\rangle = 0$, and $\Delta J_y = j$. Note that at $\Delta\varphi = \pi/N$ the output state becomes $|\Psi(\pi/N)\rangle = \left[e^{-i\pi/2}|j, j\rangle_y + e^{i\pi/2}|j, -j\rangle_y\right]/\sqrt{2}$, which is orthogonal to $|\Psi\rangle$. In other words, NOON state can reach HL.

4.4 Relation with Previous Works

For the single-mode unitary phase shift transformation $U = e^{-i\varphi N}$ with N the number operator, we have inequality similar to (4.12),

$$\Delta\varphi \geq \max\left\{\frac{\pi}{\langle N\rangle},\ \frac{1}{2\Delta N}\right\}. \tag{4.13}$$

The necessary condition to achieve HL is $\mathcal{O}(\langle N\rangle) \leq \Delta N$, which is also obtained in [51]. We should point out that (4.12) and (4.13) are not mathematically rigorous since $\Delta\varphi$ is a bounded quantity. If the right hands of (4.12) and (4.13) become larger than 2π, we have to go back to the initial condition (4.7) to find the minimal detectable phase shift.

In Ref. [21] Braunstein and Caves, by minimizing the Fisher information on all possible POVMs, found an inequality for estimating a parameter φ, which states

$$\Delta\varphi^2 \frac{ds^2}{d\varphi^2} \geq \frac{1}{4}, \qquad ds^2 = 2(1 - |\langle\Psi_\varphi|\Psi_{\varphi+d\varphi}\rangle|). \tag{4.14}$$

Here $|\Psi_\varphi\rangle$ is a class of states generated by parameter φ. As long as $|\langle\Psi_\varphi|\Psi_{\varphi+\Delta\varphi}\rangle|$ is less than one, which is equivalent to (4.7), it is possible to tell the difference between the states Ψ_φ and $\Psi_{\varphi+\Delta\varphi}$. A similar quantity has also been proposed by Ou in Ref. [51] from the complementarity principle.

From the perspective of quantum estimation theory [47], the average probability of error of distinguishing two known states ρ_0, ρ_1 is equal to

$$P_{\text{error}} \;=\; \frac{1}{2}[1 - \mathcal{D}(\rho_0, \rho_1)] \leq \frac{1}{2},$$

where $\mathcal{D}(\rho_0, \rho_1) = |\rho_0 - \rho_1|/2$ and $|A| \equiv \text{Tr}\sqrt{A^\dagger A}$ and the equality holds only for $\rho_0 = \rho_1$. In Ref. [65] $\mathcal{D}$ is defined as the distance between states ρ_0 and ρ_1, which characterizes how distinguishable they are, because $\mathcal{D} > 0$ implies $P_{\text{error}} < 1/2$. Note the inequality $\sqrt{1-F} \leq \mathcal{D} \leq \sqrt{1-F^2}$ proved in Ref. [65]. Here the fidelity is $F = \text{Tr}\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}}$, and for pure states $F = |\langle\Psi_1|\Psi_0\rangle|$. We can equally use $\mathcal{D}_F = \sqrt{1-F^2}$ to characterize the distinguishability between ρ_0 and ρ_1. Thus the condition to effectively distinguish ρ_0 and ρ_1 can be expressed as $F < 1$. To determine whether HL is attainable for an input state ρ, we should require

$$\lim_{\langle N\rangle\to\infty} F[\rho(0), \rho(\Delta\varphi)]|_{\Delta\varphi\to 1/\langle N\rangle} < 1, \tag{4.15}$$

where $\rho(0), \rho(\Delta\varphi)$ are output states with null and $\Delta\varphi$ phase shift applied respectively, which is a generalization to the condition (4.7) for mixed states.

In addition there is another way to distinguish two states with zero errors, but the price to pay is to introduce the inconclusive results. For the two pure states Ψ_1, Ψ_2 with equal prior probabilities, the minimal probability of inconclusive results is given by $|\langle\Psi_1|\Psi_2\rangle|$. To prove it, we use a theorem in Ref. [67]: The states Ψ_1, Ψ_2 can be identified with the respective probabilities p_1, p_2 iff the matrix

$$\begin{pmatrix} 1-p_1 & \langle\Psi_1|\Psi_2\rangle \\ \langle\Psi_2|\Psi_1\rangle & 1-p_2 \end{pmatrix} \tag{4.16}$$

is semi-positive. It implies the inequalities

$$1 - p_1 \geq 0, \tag{4.17}$$

$$(1 - p_1)(1 - p_2) - |\langle \Psi_1 | \Psi_2 \rangle|^2 \geq 0, \tag{4.18}$$

and hence

$$(p_1 + p_2)/2 \leq 1 - |\langle \Psi_1 | \Psi_2 \rangle| . \tag{4.19}$$

The inconclusive probability is $1 - (p_1 + p_2)/2 = |\langle \Psi_1 | \Psi_2 \rangle|$. It therefore also qualifies the fidelity F as a meaningful measure for the distinguishability of two states.

4.5 Applications

In this section we illustrate the applications of (4.7) or (4.8) with some examples. First, we take $|\Psi\rangle = |j, j\rangle$ in the angular momentum representation as the input state with total photon number $N = 2j$,

$$\langle \Psi | e^{-i\Delta\varphi J_y} | \Psi \rangle = d^{(j)}_{j,j}(\Delta\varphi) = \left(\frac{1 + \cos\Delta\varphi}{2} \right)^{2j} \approx e^{-j\Delta\varphi^2/2}, \tag{4.20}$$

where the approximations $\cos x \approx 1 - x^2/2$ and $(1 - x)^{1/x} \approx e^{-1}$ for small x are used. Therefore,

$$\lim_{N\to\infty} \left| \langle \Psi | e^{-i\Delta\varphi J_y} | \Psi \rangle \right|_{\Delta\varphi \to 1/N} = \lim_{N\to\infty} e^{-j\Delta\varphi^2/2}|_{\Delta\varphi\to 1/N} = \lim_{N\to\infty} e^{-1/(4N)} = 1, \tag{4.21}$$

i.e. HL can not be achieved. In fact we note that

$$\lim_{N\to\infty} e^{-j\Delta\varphi^2/2}|_{\Delta\varphi\to 1/\sqrt{N}} = e^{-1/2} < 1, \tag{4.22}$$

and thus the minimal measurable phase shift is $\Delta\varphi \sim N^{-1/2}$, which is the standard shot-noise limit.

If the input state is the dual Fock state $|j, 0\rangle$,

$$\langle j, 0 | e^{-i\Delta\varphi J_y} | j, 0 \rangle = d^{(j)}_{0,0}(\Delta\varphi) \to \mathcal{J}_0(j\Delta\varphi) + \mathcal{O}(j^{-3/2}). \tag{4.23}$$

where $\mathcal{J}_\nu(x)$ is Bessel function, and the asymptotic formula for $d^{(j)}_{m,n}(\beta)$ is used, i.e. $d^{(j)}_{m,n}(\beta) = \mathcal{J}_{n-m}(j\beta) + \mathcal{O}(j^{-3/2})$ when $m, n \ll j$. In order to make (4.23) distinctly different from 1, we need

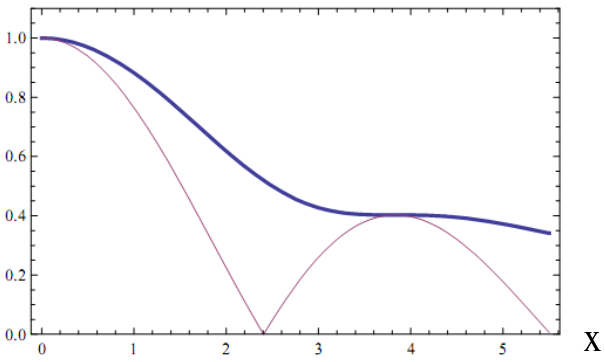

FIGURE 4.1. The plots for $|\mathcal{J}_0(x)|$ (thin) and $\sqrt{\mathcal{J}_0(x)^2 + \mathcal{J}_1(x)^2}$ (thick).

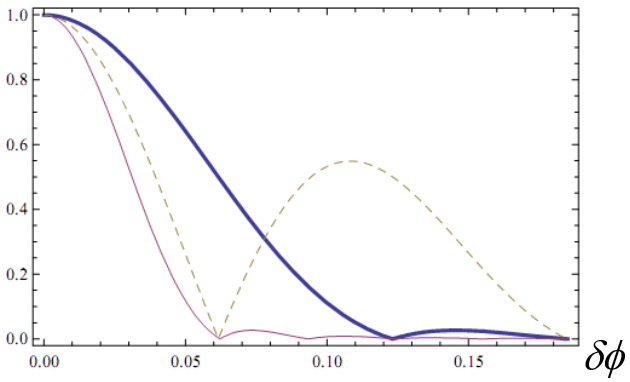

FIGURE 4.2. The plots for the magnitude of (4.30) with $\{N = 100, r = 1\}$ (thick), $\{N = 200, r = 1\}$ (thin), and $\{N = 100, r = 2\}$ (dashed).

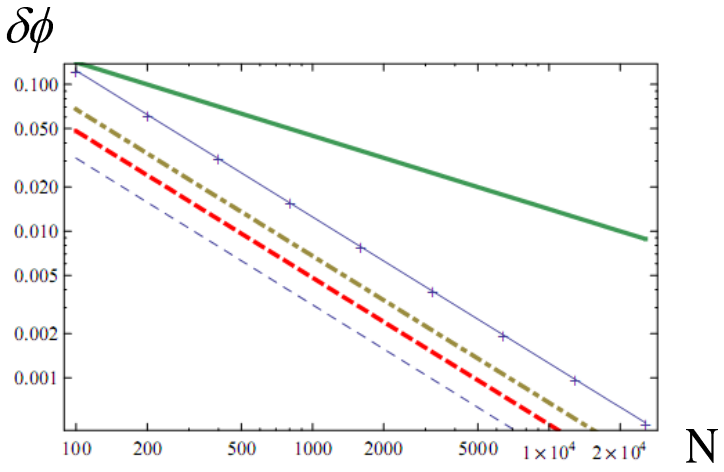

FIGURE 4.3. The scaling behavior of $\Delta\varphi$ versus the total photon number for (4.30) with $r = 1$ (thin solid), NOON state (thin dashed), $|j, j\rangle$ (thick solid), dual Fock-state (thick dashed), and the superposition of dual Fock- and Yurke-states with $|\mathrm{Im}[\alpha\beta^*]| = 1/2$ (thick dotdashed).

to require $j\Delta\varphi \sim x_0 = 2.405$ with x_0 the first root of $\mathcal{J}_0(x)$, see Fig. 4.1. This means that the dual Fock state can achieve HL with a proper detection scheme [58].

The superpostion of dual Fock state and Yurke state can be written as

$$|\Psi\rangle_{FY} = \alpha|j, 0\rangle + \beta|j, 1\rangle \quad \left(|\alpha|^2 + |\beta|^2 = 1\right), \tag{4.24}$$

$$\langle J_y\rangle = \mathrm{Im}[\alpha\beta^*]\sqrt{j(j+1)}, \qquad \langle J_y^2\rangle = \frac{1}{2}\left[j(j+1) - |\beta|^2\right]. \tag{4.25}$$

Note that $|\mathrm{Im}[\alpha\beta^*]| \leq 1/2$. Hence $j + \langle J_y\rangle \sim \Delta J_y \sim \mathcal{O}(N)$, and (4.12) gives $\Delta\varphi \geq \mathcal{O}(N^{-1})$. Now let us check if $\Delta\varphi$ can actually reach this lower bound. For large photon number, we note

$$\begin{aligned} \langle\Psi_{FY}|e^{-i\Delta\varphi J_y}|\Psi_{FY}\rangle &= \left[|\alpha|^2 d_{0,0}^{(j)} + \beta\alpha^* d_{0,1}^{(j)} + \alpha\beta^* d_{1,0}^{(j)} + |\beta|^2 d_{1,1}^{(j)}\right](\Delta\varphi) \\ &\to \mathcal{J}_0(j\Delta\varphi) + (\beta\alpha^* - \alpha\beta^*)\mathcal{J}_1(j\Delta\varphi) + \mathcal{O}(j^{-3/2}). \end{aligned} \tag{4.26}$$

At the worst case $|\mathrm{Im}[\alpha\beta^*]| = 1/2$, the magnitude of it can attain 0.402 at $j\Delta\varphi \approx 3.382$ as shown in Fig. 4.1. This means we can detect phase shift as small as $\mathcal{O}(N^{-1})$ even through the parity measurement fails in this case.

Next we consider the eigenstates of operator $\cos(\hat{\varphi}_a - \hat{\varphi}_b)$ or $\sin(\hat{\varphi}_a - \hat{\varphi}_b)$ [43] inserted between the two beam splitters of MZI,

$$|\cos\theta_{Nr}\rangle = \sqrt{\frac{2}{N+2}} \sum_{n=0}^{N} \sin\left[(n+1)\theta_{Nr}\right] |n\rangle_a |N-n\rangle_b, \tag{4.27}$$

$$|\sin\theta_{Nr}\rangle = \sqrt{\frac{2}{N+2}} \sum_{n=0}^{N} (-i)^n \sin\left[(n+1)\theta_{Nr}\right] |n\rangle_a |N-n\rangle_b, \tag{4.28}$$

$$\langle\cos\theta_{Nr}|\cos\theta_{Ms}\rangle = \langle\sin\theta_{Nr}|\sin\theta_{Ms}\rangle = \Delta_{NM}\Delta_{rs}. \tag{4.29}$$

Here $\theta_{Nr} = \pi r/(N+2)$ and $r = 1, 2, ..., N+1$. Under the action of the unitary transformation $U = e^{-i\Delta\varphi\hat{N}_a}$, we find that if $N \gg 1$,

$$\begin{aligned}
&\langle\cos\theta_{Nr}|\,U\,|\cos\theta_{Nr}\rangle = \langle\sin\theta_{Nr}|\,U\,|\sin\theta_{Nr}\rangle \\
&= \frac{2}{N+2} \sum_{n=0}^{N} e^{-in\Delta\varphi} \sin^2[(n+1)\theta_{Nr}] \\
&\to \frac{2}{\pi} \int_0^{\pi} dx\, e^{-i(N\Delta\varphi)x} \sin^2(rx).
\end{aligned} \tag{4.30}$$

From above expression, we can see that only if the exponential factor $N\Delta\varphi \sim 1$, it is possible to have the fidelity less than 1. Some typical plots of the magnitude of (4.30) are shown in Fig. 4.2. The relation between the minimal detectable phase shift $\Delta\varphi$ and the total photon number N can be derived by demanding the magnitude of (4.30) distinctly away from 1, as plotted in Fig. 4.3. It clearly shows that we can achieve HL using such states.

At last we consider the single-mode unitary phase shift transformation $U = e^{-i\varphi\hat{N}}$ with the input state proposed in Ref. [68], which is given by

$$|\Psi\rangle_{SSW} = \sqrt{\frac{6}{\pi^2}} \sum_{n=0}^{M} \frac{1}{n+1} |n\rangle \ \ (M \gg 1), \tag{4.31}$$

$$\langle N\rangle = \frac{6\,\ln M}{\pi^2}, \qquad \Delta N = \sqrt{\frac{6\,M}{\pi^2}}. \tag{4.32}$$

The relation (4.13) shows that $\Delta\varphi \geq \mathcal{O}(\langle N\rangle^{-1})$, but this bound is not attainable. Note that

$$\begin{aligned}
&\left|\langle\Psi_{SSW}|e^{-i\Delta\varphi\hat{N}}|\Psi_{SSW}\rangle\right| = \frac{6}{\pi^2} \left|\sum_{n=0}^{M} e^{-in\Delta\varphi} \left(\frac{1}{n+1}\right)^2\right| \\
&\to \frac{6}{\pi^2} \left|\mathrm{Li}_2(e^{-i\Delta\varphi})\right| + \mathcal{O}(M^{-2}) \\
&= 1 - \frac{3}{\pi}\Delta\varphi + \mathcal{O}(\Delta\varphi^2) + \mathcal{O}(M^{-2}),
\end{aligned} \tag{4.33}$$

which approaches to 1 as $\Delta\varphi \sim \mathcal{O}(\langle N\rangle^{-1}) \to 0$ for $M \gg 1$. Here $\mathrm{Li}_2(x)$ is the polylogarithm function. Therefore, the SSW state can not achieve HL [51].

4.6 Conclusion

We gave a detection-scheme-independent condition to determine whether an arbitrary input state can achieve the HL. The condition for one input state ρ to reach HL is that the fidelity between two output states with zero phase and the minimal detectable phase $\Delta\varphi$ applied respectively should satisfies the condition $\lim_{\langle N\rangle\to\infty} F|_{\Delta\varphi\to 1/\langle N\rangle} < 1$, which is applicable to a single run of detection as well as multiple runs. This condition serves as sufficient check for the lower bound provided by the uncertainty relation or the Fisher information for the minimal detectable phase, because in principle the possibility of such lower bound achievable for single run of measurement is not guaranteed.

Chapter 5: The Measures for Which-Way Information

We use several physical quantities to characterize the which-way information in the interference experiment. One quantity that can express the which-way information is associated with fidelity between two density matrices. Others are the mutual entropy and the change of the entanglement of the whole system, i.e. physical system plus which-way detector. With such quantities, we find that as the visibility of the interference pattern gets larger, we get less which-way information.

5.1 Introduction

The complementary principle states that quantum systems inherit equally real but mutually exclusive properties. For example, an electron can behave either like a particle or like a wave depending on the experiment situation. The situations "perfect fringe visibility and no which-way information" and "full which-way information and no fringes" are well known for Young's double slit experiment [24]. The intermediate stages are studied by Englert in Ref. [25], where he defined a quantity called distinguishability $\mathcal{D}$ to quantify the which-way information as well as the fringe visibility $\mathcal{V}$ for interference pattern. With $\mathcal{D}$ and $\mathcal{V}$ an inequality $\mathcal{D}^2+\mathcal{V}^2 \leq 1$ is proved so that the complementary principle is verified for the intermediate cases. We will continue to study the intermediate cases from other points of view and give a new and strong form of above inequality.

5.2 Distinguishability Revisited

We will follow the treatment of MZI as that of [25], see Fig. 5.2. The quantum number $+1$ and -1 are used to label the two ways between beam splitter and beam merger. The polarization degree of freedom of photon can be summarized by the density matrix,

$$\rho_P^{(i)} = \frac{1+\mathrm{s}^{(i)}\cdot\sigma}{2} \tag{5.1}$$

with σ Pauli's matrix, and the initial Bloch vector $\mathrm{s}^{(i)} = \mathrm{tr}_P[\rho_P^{(i)}\sigma]$ describes the initial state of photon at the input port. The action of the beam splitter and the beam merger can be represented

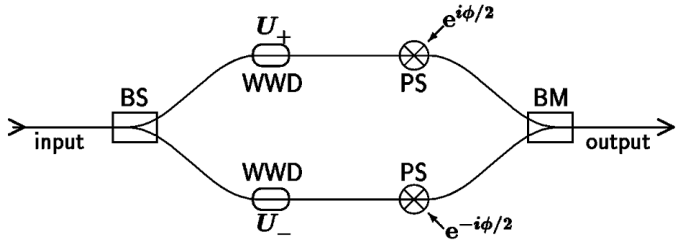

FIGURE 5.1. Schematic two-way interferometer [25]. The beam splitter BS splits the input photon into the two ways, then its which-way information is acquired by the which-way detector WWD, and after passing the phase shifter PS, the beam merger BM produces the output photon.

by

$$\rho_P \rightarrow e^{-i\pi\sigma_y/4}\rho_P\, e^{i\pi\sigma_y/4}, \tag{5.2}$$

whereas the phase shifter at the central stage effects

$$\rho_P \rightarrow e^{-i\varphi\sigma_z/2}\rho_P\, e^{i\varphi\sigma_z/2}. \tag{5.3}$$

The consequent result of MZI is turning $\rho_P^{(i)}$ to the final state

$$\rho_P^{(f)} = \frac{1+\mathrm{s}^{(f)}\cdot\sigma}{2}, \tag{5.4}$$

with

$$\mathrm{s}^{(f)} = (-s_x^{(i)}, s_y^{(i)}\cos\varphi + s_z^{(i)}\sin\varphi, s_y^{(i)}\sin\varphi - s_z^{(i)}\cos\varphi). \tag{5.5}$$

The probabilities for taking either one of the two ways are given as

$$\begin{aligned}\xi_\pm &= \mathrm{tr}\left[\frac{1}{2}(1\pm\sigma_z)e^{-i\pi\sigma_y/4}\rho_P^{(i)}\, e^{i\pi\sigma_y/4}\right] \\ &= \frac{1}{2}(1\mp s_x^{(i)}).\end{aligned} \tag{5.6}$$

Their difference $|\xi_+ - \xi_-| = s_x^{(i)}$ gives the predictability of the ways through MZI. To simplify the following analysis, let us first consider symmetric interferometer with $s_x^{(i)} = 0$ and $s_z^{(i)} + is_y^{(i)} =$

$e^{-i\theta}$, which is coupled to another physical system that serves as a which-way detector. Suppose the detector is in state $\rho_P^{(i)}$ initially so that the density matrix of the whole system has the form $\rho^{(i)} = \rho_P^{(i)} \otimes \rho_D^{(i)}$. The interaction between photon and detector at the central stage of the interferometer now can be described by the unitary operator

$$\mathcal{U} = \frac{1+\sigma_z}{2} e^{i\varphi/2} \otimes U_+ + \frac{1-\sigma_z}{2} e^{-i\varphi/2} \otimes U_-, \tag{5.7}$$

instead of $e^{i\varphi\sigma_z/2}$ of (5.3). Here U_+ and U_- are unitary operators that only affect the detector. The resulted final state of the whole system then becomes

$$\begin{aligned} \rho^{(f)} &= \frac{1+\sigma_x}{4} \otimes U_+^\dagger \rho_D^{(i)} U_+ + \frac{1-\sigma_x}{4} \otimes U_-^\dagger \rho_D^{(i)} U_- \\ &\quad - e^{-i(\phi-\theta)} \frac{\sigma_z - i\sigma_y}{4} \otimes U_+^\dagger \rho_D^{(i)} U_- - e^{i(\phi-\theta)} \frac{\sigma_z + i\sigma_y}{4} \otimes U_-^\dagger \rho_D^{(i)} U_+. \end{aligned} \tag{5.8}$$

Before we move onto the next step, one thing can further simplify our calculation. If the interaction between photon and detector is totally symmetrical, i.e. $U_+ = U_- = U$, the detector obtains no which-way information. Because in this case $\mathcal{U} \to e^{i\varphi\sigma_z/2} \otimes U$, it leads to $\rho^{(f)} = \rho_P^{(f)} \otimes U^\dagger \rho_D^{(i)} U$, which means the detector has no effects on photon. Without loss of generality, let us set $U_+ = U$ and $U_- = 1$ for the rest.

The final states of the photon and the detector are given by tracing over detector's degree of freedom. The photon state takes the same form as Eq. (5.4) with $s_x^{(f)} = 0$ and $s_z^{(f)} + i s_y^{(f)} = -e^{-i(\varphi-\theta)}\mathcal{C}$, where $\mathcal{C} = \mathrm{tr}_D[U^\dagger \rho_D^{(i)}]$ is a complex number. By measuring the observable σ_z after the photon passing the beam merger, the interference pattern is revealed through the relative frequency for finding the value -1,

$$\begin{aligned} p_\varphi &= \mathrm{tr}_P \left[\frac{1}{2}(1-\sigma_z) \rho_P^{(f)} \right] \\ &= \frac{1}{2} \left[1 + \mathcal{C} \cos(\varphi - \theta) \right], \end{aligned} \tag{5.9}$$

so that the magnitude of $\mathcal{C}$ gives the fringe visibility $\mathcal{V} = |\mathcal{C}|$. Similarly, The final detector state equals

$$\rho_D^{(f)} = \mathrm{tr}_P[\rho^{(f)}] = \frac{1}{2}(U^\dagger \rho_D^{(i)} U + \rho_D^{(i)}). \tag{5.10}$$

Based on optimal likelihood guess [24], the distinguishability of the ways is defined by $\mathcal{D} = \mathrm{tr}_D|U^\dagger\rho_D^{(i)}U - \rho_D^{(i)}|/2$. Englert then proved the inequality [25]

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1, \tag{5.11}$$

which incorporates the extreme situations mentioned above.

For asymmetric interferometer $s_x^{(i)} \neq 0$ with a priori fringe visibility $\mathcal{V}_0 = \sqrt{(s_y^{(i)})^2 + (s_z^{(i)})^2}$ and the priori which-way probabilities $\xi_\pm = (1 \mp s_x^{(i)})/2$, the distinguishability is given by

$$\mathcal{D} = \mathrm{tr}_D|\xi_+ U^\dagger\rho_D^{(i)}U - \xi_-\rho_D^{(i)}|. \tag{5.12}$$

Then the relation (5.11) can be generalized to

$$\mathcal{D}^2 + 4\,\xi_+\xi_- \left(\frac{\mathcal{V}}{\mathcal{V}_0}\right)^2 \leq 1, \tag{5.13}$$

where $\mathcal{V} = \mathcal{V}_0|\mathcal{C}|$. In next section, we will define distinguishability from other points of view, which allows us to give a simple proof of this inequality.

5.3 Fidelity and Distinguishability

The distinguishability defined by (5.12) can be applied to any two density matrix ρ_1, ρ_2 given the priori probabilities ξ_1, ξ_2 for choosing one of them with $\xi_1 + \xi_2 = 1$. More explicitly, we define

$$\mathcal{D}(\rho_1, \rho_2) \equiv \mathrm{tr}|\xi_1\rho_1 - \xi_2\rho_2|. \tag{5.14}$$

In the case of ρ_1 and ρ_2 being pure states, i.e. $\rho_1 = |\Psi_1\rangle\langle\Psi_1|$ and $\rho_2 = |\Psi_2\rangle\langle\Psi_2|$, $\mathcal{D}(|\Psi_1\rangle, |\Psi_2\rangle) = \sqrt{1 - 4\,\xi_1\xi_2|\langle\Psi_1|\Psi_2\rangle|^2}$. From this definition, we can prove the following property for $\mathcal{D}(\rho_1, \rho_2)$. Suppose $\mathcal{E}$ is a positive tracing preserving operator, then

$$\mathcal{D}(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)) \leq \mathcal{D}(\rho_1, \rho_2). \tag{5.15}$$

The proof is based on the fact that $w_1\rho_1 - w_2\rho_2$ can be expressed as $\xi_1\rho_1 - \xi_2\rho_2 = Q - S$, where Q and S are positive operators with orthogonal supports. It implies $|\xi_1\rho_1 - \xi_2\rho_2| = Q + S$, so

$$\begin{aligned}
\mathcal{D}(\rho_1, \rho_2) &= \mathrm{tr}Q + \mathrm{tr}S \\
&= \mathrm{tr}[\mathcal{E}(Q)] + \mathrm{tr}[\mathcal{E}(S)] \\
&\geq \mathrm{tr}|\mathcal{E}(Q) - \mathcal{E}(S)| \\
&= \mathcal{D}(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)).
\end{aligned} \tag{5.16}$$

Here the triangle inequality for the trace-class norm for all trace-class operators A, B, $|\mathrm{tr}|A| - \mathrm{tr}|B|| \leq \mathrm{tr}|A \pm B| \leq \mathrm{tr}|A| + \mathrm{tr}|B|$, is applied. Similarly, it is easy to check $1 \geq \mathcal{D}(\rho_1, \rho_2) \geq |\xi_1 - \xi_2|$. Moreover, we can prove that

$$\mathcal{D}(\rho_1, \rho_2) \;\; \geq \;\; \sum_m |\xi_1 p_m - \xi_2 q_m| \tag{5.17}$$

for any positive operator-valued measures (POVM) $\{E_m\}$, where $p_m = \mathrm{tr}[\rho_1 E_m]$ and $q_m = \mathrm{tr}[\rho_2 E_m]$ are the probabilities for obtaining outcome m for the states ρ_1 and ρ_2, respectively.

We now define a new distinguishability in terms of fidelity between two density matrix given the priori probabilities ξ_1, ξ_2 for choosing one of them,

$$\mathcal{D}_F(\rho_1, \rho_2) = \sqrt{1 - 4\,\xi_1 \xi_2 F(\rho_1, \rho_2)^2}, \tag{5.18}$$

with the fidelity

$$F(\rho_1, \rho_2) = \mathrm{tr}\left[\left(\sqrt{\rho_1} \rho_2 \sqrt{\rho_1} \right)^{1/2} \right]. \tag{5.19}$$

If $\mathcal{E}$ is a positive tracing preserving operator, we have the relation

$$F(\rho_1, \rho_2) \leq F(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)). \tag{5.20}$$

Let us establish the relation between these two distinguishabilities following [65]. Let $|\Psi_1\rangle$ and $|\Psi_2\rangle$ be purifications of ρ_1 and ρ_2 chosen such that $F(\rho_1, \rho_2) = |\langle \Psi_1 | \Psi_2 \rangle| = F(|\Psi_1\rangle, |\Psi_2\rangle)$. Using (5.16) by taking $\mathcal{E}$ as a partial trace, we see that

$$\mathcal{D}(\rho_1, \rho_2) \leq \mathcal{D}(|\Psi_1\rangle, |\Psi_2\rangle) = \sqrt{1 - 4\,\xi_1 \xi_2 |\langle \Psi_1 | \Psi_2 \rangle|^2} = \mathcal{D}_F(\rho_1, \rho_2). \tag{5.21}$$

Conversely, let E_m be a POVM, such that $F(\rho_1, \rho_2) = \sum_m \sqrt{p_m q_m}$. Next we observe that

$$\sum_m \left(\sqrt{\xi_1 p_m} - \sqrt{\xi_2 q_m} \right)^2 = 1 - 2\sqrt{\xi_1 \xi_2} F(\rho_1, \rho_2), \tag{5.22}$$

and note that

$$\begin{aligned} \sum_m \left(\sqrt{\xi_1 p_m} - \sqrt{\xi_2 q_m} \right)^2 \;\; &\leq \;\; \sum_m |\sqrt{\xi_1 p_m} - \sqrt{\xi_2 q_m}||\sqrt{\xi_1 p_m} + \sqrt{\xi_2 q_m}| \\ &= \;\; \sum_m |\xi_1 p_m - \xi_2 q_m| \leq \mathcal{D}(\rho_1, \rho_2). \end{aligned} \tag{5.23}$$

Here we have used (5.17) in the last line of (5.23). Finally, we get the relation

$$\begin{aligned} 1 - 2\sqrt{\xi_1\xi_2}F(\rho_1,\rho_2) \le \mathcal{D}(\rho_1,\rho_2) &\le \sqrt{1-4\,\xi_1\xi_2 F(\rho_1,\rho_2)^2} \\ &= \mathcal{D}_F(\rho_1,\rho_2). \end{aligned} \tag{5.24}$$

This means that the two distinguishabilities are qualitatively equivalent measures. So we can use $\mathcal{D}_F(\rho_1,\rho_2)$ as well as $\mathcal{D}(\rho_1,\rho_2)$ to measure the distinguishability of the two ways.

We are ready to give a new proof for (5.12). Let $|\Psi\rangle$ be a purification $\rho_D^{(i)}$, then

$$\begin{aligned} \frac{\mathcal{V}}{\mathcal{V}_0} &= |\mathcal{C}| = \left|\mathrm{tr}_D[U^\dagger \rho_D^{(i)}]\right| = \left|\mathrm{tr}[U^\dagger|\Psi\rangle\langle\Psi|]\right| \\ &= |\langle\Psi|U^\dagger|\Psi\rangle| = F(U^\dagger|\Psi\rangle,|\Psi\rangle) \le F(U^\dagger\rho_D^{(i)}U,\rho_D^{(i)}), \end{aligned} \tag{5.25}$$

where (5.20) is used by taking $\mathcal{E}$ as a partial trace. If we choose $\rho_1 = U^\dagger\rho_D^{(i)}U$ and $\rho_2 = \rho_D^{(i)}$ in (5.24), and consider (5.25), the following inequality is obtained,

$$|\xi_+ - \xi_-| \le \mathcal{D} \le \mathcal{D}_F \le \sqrt{1 - 4\,\xi_+\xi_- \left(\frac{\mathcal{V}}{\mathcal{V}_0}\right)^2}. \tag{5.26}$$

Thus the relation (5.12) is confirmed again. If the distinguishability is measured according to $\mathcal{D}_F = \mathcal{D}_F(\rho_1,\rho_2)$ rather than $\mathcal{D} = \mathcal{D}(\rho_1,\rho_2)$, it is obvious that $\mathcal{D}_F$ may take a larger magnitude.

To illustrate above analysis, let us examine an example where the which-way detector is a two-level quantum system. The fidelity can then be given by

$$F(\rho_1,\rho_2)^2 = \mathrm{tr}[\rho_1\rho_2] + 2\sqrt{\det[\rho_1]\det[\rho_2]}. \tag{5.27}$$

In some particular base, the density matrix $\rho_D^{(i)}$ can be written as

$$\rho_D^{(i)} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}, \qquad 1 \ge p \ge 0. \tag{5.28}$$

The unitary matrix U generally is

$$U = \begin{pmatrix} a+ic & b+id \\ -b+id & a-ic \end{pmatrix}, \qquad a^2+b^2+c^2+d^2 = 1, \tag{5.29}$$

where a, b, c, and d are real numbers. The calculation gives that for symmetric interferometer,

$$\mathcal{D} = \mathcal{D}_F = x\sqrt{1-a^2-c^2}, \qquad \mathcal{V} = \sqrt{a^2+c^2x^2}, \tag{5.30}$$

where the parameter $x = |2p-1| \le 1$. From these expressions, the inequality $\mathcal{D}^2 + \mathcal{V}^2 \le 1$ or $\mathcal{D}_F^2 + \mathcal{V}^2 \le 1$ is straightforward to obtain.

5.4 Mutual Entropy

Using mutual entropy, we will see that it can also measure the which-way information. To be specific, let us still assume that the detector is a two-level system. The von Neumann entropy of a given density state ρ is $S(\rho) = -\mathrm{tr}[\rho \ln \rho]$, which is invariant under any unitary transformation, $\rho \to U^\dagger \rho U$. The mutual entropy for a coupled system AB is defined as [65]

$$S(A{:}B) = S(A) + S(B) - S(AB) \tag{5.31}$$

with $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{tr}_A[\rho_{AB}]$. The physical interpretation of the mutual entropy is that it describes the information shared by A and B. If we apply it to our coupled P, D system, intuitively we would expect the mutual entropy of the final state $\rho^{(f)}$ decreases (increases) as the fringe visibility $\mathcal{V}$ (the distinguishability $\mathcal{D}$ or $\mathcal{D}_F$) gets larger. In other words, a large $\mathcal{V}$ or a small distinguishability means there is little information shared by P and D. To justify this statement, we need to find out the explicit expression for mutual entropy of the final state. First note that

$$S^{(f)}(PD) = -\mathrm{tr}[\rho^{(f)} \ln \rho^{(f)}] = S^{(i)}(PD) = S^{(i)}(P) + S^{(i)}(D), \tag{5.32}$$

for the total final state is related to the total initial state by a unitary transformation, and $\rho^{(i)} = \rho_P^{(i)} \otimes \rho_D^{(i)}$. Since $\rho_P^{(i)}$ is a pure state and $\rho_D^{(i)}$ takes the form of (5.28), we find that

$$\begin{aligned} S^{(f)}(PD) &= -\left(\lambda_+ \ln \lambda_+ + \lambda_- \ln \lambda_-\right), \\ \lambda_\pm &= (1 \pm x)/2. \end{aligned} \tag{5.33}$$

Next, from the expressions for $\rho_P^{(f)}$ and $\rho_D^{(f)}$, the associated entropies are

$$\begin{aligned} S^{(f)}(P) &= -\left(\delta_+ \ln \delta_+ + \delta_- \ln \delta_-\right), \\ \delta_\pm &= (1 \pm x\sqrt{a^2 + c^2})/2, \end{aligned} \tag{5.34}$$

$$\begin{aligned} S^{(f)}(D) &= -\left(\eta_+ \ln \eta_+ + \eta_- \ln \eta_-\right), \\ \eta_\pm &= (1 \pm \sqrt{a^2 + c^2 x^2})/2. \end{aligned} \tag{5.35}$$

In terms of variables $\mathcal{D}, \mathcal{V}$ and x, the final mutual entropy becomes

$$\begin{aligned} \Delta S(P{:}D) &= S^{(f)}(P{:}D) - S^{(i)}(P{:}D) \\ &= H\left(\sqrt{x^2 - \mathcal{D}^2}\right) + H(\mathcal{V}) - H(x), \end{aligned} \tag{5.36}$$

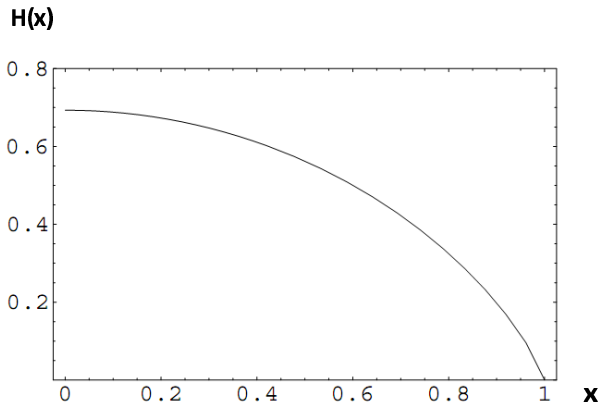

FIGURE 5.2. The plot of the function $H(x)$ versus x.

where $\mathcal{D} \leq \min\left\{x, \sqrt{1-\mathcal{V}^2}\right\}$, and the function

$$H(x) \equiv -\left(\frac{1+x}{2}\ln\frac{1+x}{2} + \frac{1-x}{2}\ln\frac{1-x}{2}\right). \tag{5.37}$$

Because $H(x)$ is a monotonously decreasing function as shown in Fig. 5.4, for fixed x and $\mathcal{D}$ ($\mathcal{V}$), $\Delta S(P\!:\!D)$ is a decreasing (increasing) function of $\mathcal{V}$ ($\mathcal{D}$).

5.5 Entanglement of Formation

Our last measure for the which-way information is related with the entanglement between the whole final state, while the whole initial state has no entanglement. We can also imagine that the entanglement has similar property as that of the mutual entropy from its physical interpretation. If the final state is separable, i.e. $\rho^{(f)} = \rho_P^{(f)} \otimes \rho_D^{(f)}$, the fringe visibility is unaffect but which-way information is null. On the other hand, if the final state is in the maximal entanglement state, i.e. $\rho^{(f)} = |\Psi\rangle\langle\Psi|$ with $\Psi = \left(|+\rangle_P|+\rangle_D + |-\rangle_P|-\rangle_D\right)/\sqrt{2}$, where $|\pm\rangle_{P(D)}$ is the photon (detector) state corresponding to the path of $\sigma_z = \pm$, the which-way information is gained but the fringe visibility is lost. In order to quantify this statement, we calculate the explicit expression for entanglement defined in Ref. [37]. The entanglement of formation of a mixed state of two qubits is given by

$$\begin{aligned} E(\rho) &= H\left(\sqrt{1-C^2(\rho)}\right), \\ C(\rho) &= \max\left\{0,\, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\right\}, \end{aligned} \tag{5.38}$$

where $C(\rho)$ is called concurrence, and the $\{\lambda_i\}$ are the eigenvalues, in decreasing order, of the Hermitian matrix $(\sqrt{\rho}\tilde{\rho}\sqrt{\rho})^{1/2}$ with the spin-flipped state $\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho(\sigma_y \otimes \sigma_y)$. Note that each λ_i is a non-negative real number and $0 \leq C(\rho) \leq 1$. The explicit expression of entanglement of formation for any mixed state has not been found yet. Apply the definition of entanglement to our simple model, it gives that

$$\begin{aligned} \Delta E(PD) &= E(\rho^{(f)}) - E(\rho^{(i)}) \\ &= H\left(\sqrt{1 - C^2(\rho^{(f)})}\right), \\ C^2(\rho^{(f)}) &= \mathcal{D}^2(1 - \mathcal{V}^2). \end{aligned} \tag{5.39}$$

Since $\Delta E(PD)$ is a monotonously increasing function of $C(\rho^{(f)})$, our statement about the entanglement of formation as the measure of the which-way information is justified. This also clearly demonstrates that the complementary principle is independent of uncertainty principle, since it is only related to the entanglement of the state through (5.11).

5.6 Conclusion

We used several physical quantities to characterize the which-way information in the interference experiment. The complementary principle states that the interference pattern and the acquisition of which-way information are mutually exclusive. We defined a new distinguishability associated with fidelity between two density matrix, and then gave a new proof for Englert inequality. It was demonstrated that the changes of the mutual entropy and the entanglement of the whole system, i.e. physical system and which-way detector can be also used to describe the which-way information. With such quantities, it was shown that as the fringe visibility of the interference pattern gets larger, the less which-way information is obtained. Thus the complementary principle in the such experiment is quantified and confirmed.

Chapter 6: Super-Resolution at the Shot-Noise Limit with Coherent States

There has been much recent interest in quantum optical interferometry for applications to metrology, sub-wavelength imaging, and remote sensing, such as in quantum laser radar (LADAR). For quantum LADAR, atmospheric absorption rapidly degrades any quantum state of light, so that for high-photon loss the optimal strategy is to transmit coherent states of light, which suffer no worse loss than the Beer law for classical optical attenuation, and which provides sensitivity at SNL. In this chapter we show that coherent light coupled with photon number resolving detectors can provide a super-resolution much below the Rayleigh diffraction limit, with sensitivity no worse than SNL in terms of the detected photon power.

6.1 Introduction

Ever since the work of Boto *et al.* in 2000, it has been realized that by exploiting quantum states of light, such as NOON states, it is possible to beat the Rayleigh diffraction limit — the minimal scale accessible by a classical light field is limited by its wavelength — in imaging and lithography (super-resolution) while also beating SNL in phase estimation (super-sensitivity) [28, 5, 69, 56, 27]. This is a non-trivial result, because the super-sensitivity and the super-resolution usually are not related each other. However such quantum states of light are very susceptible to photon loss [29, 70, 71]. Recent work has shown that in the presence of high loss, the optimal phase sensitivity achieved is always SNL [30]. These results suggest that, given the difficulty in making quantum states of light, as well as their susceptibility to loss, that the most reasonable strategy for quantum LADAR is to transmit coherent states of light to mitigate a super-Beer's law in loss and maximize sensitivity [27]. It is well known that such an approach can only ever achieve at best shot-noise limited sensitivity [10]. However, such a conclusion leaves open the question as to what is the best detection strategy to optimize resolution. Recent experimental results and numerical simulations have indicated that such a coherent-state strategy can still be super-resolving, provided a quantum detection scheme is

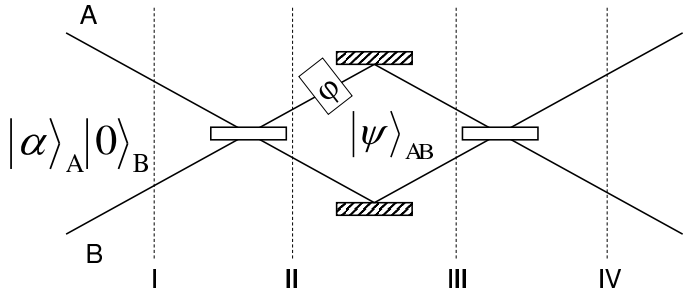

FIGURE 6.1. Here we indicate MZI used in the calculations. The coherent state is incident in mode A and vacuum in mode B at the left at Line I. After the first beam splitter transformation we have the two-mode coherent state of Eq. (6.1), as indicated at Line II. After the phase shifter φ this state becomes the two-mode coherent state of Eq. (6.2). At Line III we also implement the detection schemes corresponding to the operators $\widehat{N}_{AB}$, $\hat{\nu}_{AB}$, and $\hat{\mu}_{AB}$. Finally after the final beam splitter, we implement the parity operator $\widehat{\Pi}_A$ detection of Eq. (6.9) in the upper mode.

employed [31, 72]. In this paper we derive a quantum detection scheme, that is super-resolving, and which can be implemented with photon-number-resolving detectors. Our proposed scheme exploits coherent states of light, is shot-noise limited in sensitivity, and can resolve features by an arbitrary amount below the Rayleigh diffraction limit. Our scheme would have applications to quantum optical remote sensing, metrology, and imaging.

6.2 Three Proposed Photon-Number-Resolving Detectors

In Fig. 6.1 we illustrate schematically a two-mode interferometric quantum LADAR scheme. The source at the left is assumed to contain a laser producing a coherent state $|\alpha\rangle_A$ in upper mode A with vacuum in lower mode B, which is illustrated at line I in Fig. 6.1. The state is incident on a 50-50 beam splitter (BS), which mixes this coherent state with the vacuum state $|0\rangle_B$ in lower mode B. The output of such a mixing is computed by the beam-splitter transformation [48], and is the state,

$$|\alpha/\sqrt{2}, \alpha/\sqrt{2}\rangle = e^{-\bar{n}/2} \sum_{n,m=0}^{\infty} \frac{\left(\alpha/\sqrt{2}\right)^{n+m}}{\sqrt{n!m!}} |n, m\rangle, \tag{6.1}$$

where $\bar{n} = |\alpha|^2$ and without loss of generality, a BS phase factor of $i = e^{i\pi/2}$ has been dropped for clarity. (This physically corresponds to a symmetric 50-50 BS with a quarter-wave plate in one of the output ports.) This is the state at line II in Fig. 6.1. As shown in Fig. 6.1 the upper mode A imparts a phase shift φ on this state, yielding,

$$\begin{aligned} |\Psi\rangle_{AB} &= |\alpha e^{i\varphi}/\sqrt{2}, \alpha/\sqrt{2}\rangle \\ &= e^{-\bar{n}/2} \sum_{n,m=0}^{\infty} \frac{(e^{i\varphi}\sqrt{\bar{n}/2})^n (\sqrt{\bar{n}/2})^m}{\sqrt{n!m!}} |n, m\rangle, \end{aligned} \tag{6.2}$$

where, without loss of generality, we take $\alpha = \sqrt{\bar{n}}$, which amounts to setting an irrelevant overall phase to zero. This is the state at the line III in Fig. 6.1. In the coherent state basis, this state is obviously separable. However, while it is less obviously separable in the number basis, it is clear to see that the double sum contains the NOON states with the contributions from $\{n = N, m = 0\}$ and $\{n = 0, m = N\}$, which are $|N :: 0\rangle^{\varphi} = e^{iN\varphi}|N, 0\rangle + |0, N\rangle$, apart from a common factor $e^{-\bar{n}/2}\sqrt{\bar{n}/2}^N/\sqrt{N!}$, for any N. This NOON-state component can lead to the N-fold super-resolution [28, 31]. Similarly, for the dummy indices being $\{n = M, m = M'\}$ and $\{n = M', m = M\}$, it yields the so-called MM' states: $|M::M'\rangle^{\varphi} = e^{iM\varphi}|M, M'\rangle + e^{iM'\varphi}|M', M\rangle$, which can be $M - M'$ fold super-resolving [71]. Indeed this observation suggests a strategy, similar to that employed by Resch *et al.*, of projecting the state $|\Psi\rangle_{AB}$ of Eq. (6.2) onto the NOON basis through the implementation of the operator $\widehat{N}_{AB} = |N, 0\rangle\langle 0, N| + |0, N\rangle\langle N, 0|$ [31]. This results in an expectation value of ${}_{AB}\langle\Psi|\widehat{N}_{AB}|\Psi\rangle_{AB} = (\bar{n}/2)^N e^{-\bar{n}} 2\cos(N\varphi)/N!$ that is clearly N-fold super-resolving. However the factor of $(\bar{n}/2)^N e^{-\bar{n}}/N!$ indicates that the visibility of this expectation value, as a function of φ, is much less than unity. The trade-offs between the periods of oscillating features and their magnitudes have been noted previously by Berry and Popescu in the context of super-resolution in optical imaging [74].

Let us attempt to optimize this visibility by tuning the return power $\bar{n}$. The maximal visibility occurs when $\bar{n} = N$. For example, with $\bar{n} = N = 10$ we achieve 10-fold super-resolution in ${}_{AB}\langle\Psi|\widehat{N}_{AB}|\Psi\rangle_{AB}$ with a visibility of about 0.024%. In a similar fashion we may now estimate the minimal phase sensitivity variance via the usual Gaussian error propagation formula [27],

$$\Delta\varphi_N^2 = \frac{\Delta\widehat{N}_{AB}^2}{|\partial\langle\widehat{N}_{AB}\rangle/\partial\varphi|^2} = \frac{2^N e^{\bar{n}} N! - 2\bar{n}^N \cos^2 N\varphi}{2\bar{n}^N \sin^2 N\varphi} \frac{1}{N^2}, \tag{6.3}$$

where the factor of $1/N^2$ would provide HLed sensitivity, as it would be in the case of a pure NOON state [27], if it were not for the Poisson weight factors inherited from the coherent state. This expression, Eq. (6.3), can be minimized by inspection, again by taking $N = \bar{n}$ and simultaneously choosing φ (which can be done by putting a phase shifter in arm B) such that $N\varphi = \pi/2$, yielding, $\Delta\varphi_N^2 = 2^{\bar{n}} e^{\bar{n}} \bar{n}!/(2\bar{n}^{\bar{n}}\bar{n}^2) \cong \sqrt{\pi/2}(2^{\bar{n}}/\bar{n}^{3/2})$, where we have used Stirling's approximation, and where again HL behavior of $1/\bar{n}^2$ appears in a tantalizing fashion. However, the exact expression has a minimum at $\bar{n} = 2$ of $\Delta\varphi_N^2|_{\bar{n}=2} \cong 1.85$, and hence for all values of $\bar{n}$ always does worse than the shot-noise, in agreement with the conclusion of Ref. [31].

It is easy to see that the reason this above strategy does worse than SNL, is that the measurement on the operator $\widehat{N}_{AB}$, defined above, throws away very many photon-number amplitudes in the coherent state, keeping only the two terms in Eq. (6.2) with $n = N$ and $m = 0$, or $n = 0$ and $m = N$. Such an observation suggests that we introduce a new operator, which projects onto all of the maximally super-resolving terms, that is, $\hat{\nu}_{AB} = \sum_{N=0}^{\infty} \widehat{N}_{AB}$, which corresponds to the phase-bearing anti-diagonal terms in the two-mode density matrix [71]. We may now carry out the expectation with respect to the dual-mode coherent state of Eq. (6.2) to get, $\langle \hat{\nu}_{AB} \rangle = 2e^{-\bar{n}+(\bar{n}\cos\varphi)/2} \cos\left[(\bar{n}\sin\varphi)/2\right] \cong 2e^{-\bar{n}/2}\cos(\bar{n}\varphi/2)$. Here we have approximated the expression near $\varphi = 0$, where the function is sharply peaked. We can see from this expression that by carefully choosing the return power $\bar{n} = N$, we recover super-resolution. However the exponential pre-factor produces a loss in visibility. Choosing $\bar{n} = N = 20$, to obtain again 10-fold super-resolution, we obtain a visibility of only 0.009%. This low visibility does not bode well for the sensitivity, which we may compute as per Eq. (6.3) to get,

$$\begin{aligned} \Delta\varphi_\nu^2 &= \left\{e^{\bar{n}} + e^{3\bar{n}/2} - e^{\bar{n}\cos\varphi}[1 + \cos(\bar{n}\sin\varphi)]\right\} \\ &\quad \times e^{-\bar{n}\cos\varphi} \csc^2\left(\varphi + \frac{\bar{n}}{2}\sin\varphi\right) \frac{2}{\bar{n}^2}, \end{aligned} \tag{6.4}$$

which again displays the Heisenberg-limiting pre-factor of $1/\bar{n}^2$. This expression is singular at the phase origin, but has a minimum nearby $\varphi = \pi/2\bar{n}$, which for large photon number is approximately at $\bar{n} = N$, which by again choosing this to be a large integer, simplifies Eq. (6.4) to, $\Delta\varphi_\nu^2|_{\varphi=\pi/2\bar{n}} \cong 16\pi e^{\bar{n}/2}/(2\bar{n} + \pi)^2$, which only approaches the SNL and HL near $\bar{n} = 1$ and then rapidly diverges to be much worse than the SNL for large $\bar{n}$. Hence, from these examples, we

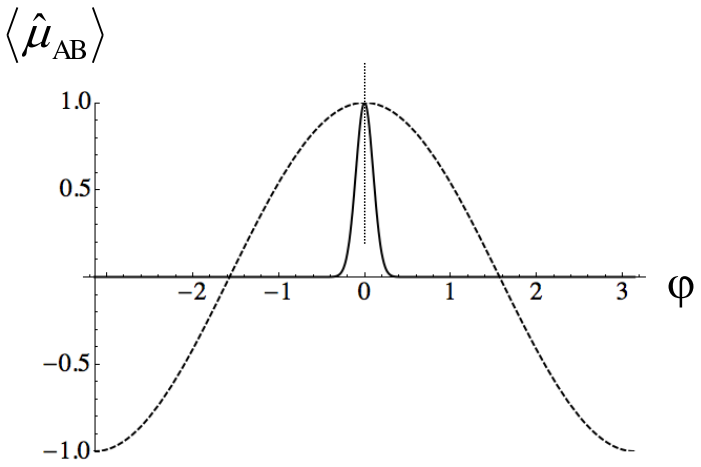

FIGURE 6.2. This plot shows the expectation value $\langle\hat{\mu}_{AB}\rangle$ of Eq. (6.6) plotted as a function of the phase shift φ (solid curve) for a return power of $\bar{n} = 100$. For reference we plot the normalized "classical" two-port difference signal (dashed curve). We see that the plot of the $\langle\hat{\mu}_{AB}\rangle$ is super-resolving and is narrower than the classical curve by a factor of $\Delta\varphi = 1/\sqrt{\bar{n}} = 1/10$, as given in Eq. (6.8).

see there is a high price to pay for NOON-like super-resolution with coherent states — so many photon amplitudes are discarded that we always do far worse than shot-noise. This analysis then suggests our final protocol — what if we choose a measurement scheme that includes all of the phase-carrying off-diagonal terms in the two-mode density matrix? Such a scheme is to consider the operator constructed from all the terms bearing $M - M'$ fold super-resolving capabilities [71], that is,

$$\hat{\mu}_{AB} = \sum_{M,M'=0}^{\infty} |M', M\rangle\langle M, M'|, \tag{6.5}$$

where we note this is evidently not a resolution of the identity operator. It is easy to show that this operator of Eq. (6.5) is both Hermitian and idempotent, that is $\hat{\mu}^{\dagger}_{AB} = \hat{\mu}_{AB}$ and $\hat{\mu}^{2}_{AB} = \hat{I}_{AB}$, respectively, where $\hat{I}_{AB}$ is the two-mode identity operator. Using these properties, with a bit of algebra, we establish, with respect to the two-mode coherent state of Eq. (6.2), the following results,

$$\langle\hat{\mu}_{AB}\rangle = {}_{AB}\langle\Psi|\hat{\mu}_{AB}|\Psi\rangle_{AB} = e^{-2\bar{n}\sin^2(\varphi/2)}, \tag{6.6}$$

$$\Delta\varphi^2_{\mu} = \frac{e^{4\bar{n}\sin^2(\varphi/2)} - 1}{\bar{n}^2\sin^2\varphi}, \tag{6.7}$$

where once again HL term of $1/\bar{n}^2$ appears accompanied by an exponential factor in the sensitivity estimate. We plot as a solid curve the expectation value of Eq. (6.6), for a return power of $\bar{n} = 100$,

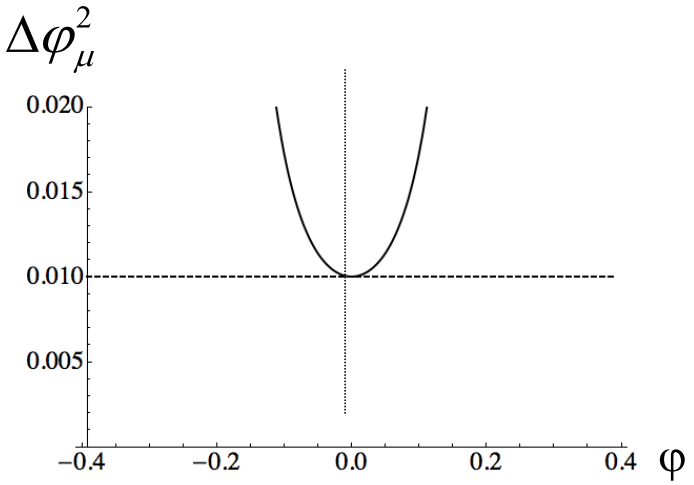

FIGURE 6.3. In this plot we depict the sensitivity expression $\Delta\varphi_\mu^2$ of Eq. (6.7), again for the return power of $\bar{n} = 100$ (solid curve). The horizontal dashed line indicates the shot-noise limit of $\Delta\varphi_{\rm SNL}^2 = 1/\bar{n} = 1/100$. We see that the sensitivity of the super-resolving $\hat{\mu}_{AB}$ detection scheme hits the SNL at $\varphi = 0$, as indicated by expanding Eq. (6.7) in a power series.

in Fig. 6.2, along with the standard photon difference detection interferogram (dashed curve) [27]. Clearly $\langle\hat{\mu}_{AB}\rangle$ has a visibility of 100% now, and is periodic in φ with period 2π, and highly peaked at the phase origin where $\varphi = 0$. This curve is not super-resolving in the usual sense of the word, as there are no multiple narrow peaks as would be the case in a NOON-state scheme, but it is super-resolving in the sense that there is a well defined narrow feature that is clearly sub-Rayleigh limited in resolution. Such a feature would be useful, for example, in LADAR ranging or laser Doppler velocimetry in the small $\bar{n}$ return-power regime, where one would lock onto the side of such a feature and then monitor how it changes in time with a feedback loop in the interferometer.

To estimate the width of this central peak we note that in the small phase angle limit, Eq. (6.6) may be approximated as,

$$\langle\hat{\mu}_{AB}\rangle|_{\varphi\cong 0} \cong e^{-\bar{n}\varphi^2/2}, \tag{6.8}$$

which is clearly a Gaussian of width $\Delta\varphi = 1/\sqrt{\bar{n}}$. Hence by choosing a return power of $\bar{n} = 100$, we are 10-fold super resolving.

Now we check the sensitivity of this scheme. In Fig. 6.3 we plot as a solid curve the variance of Eq. (6.7), again for $\bar{n} = 100$, near the phase origin. We include as a dashed curve the shot-

noise limit. We see that the sensitivity is shot-noise limited about the phase origin. This may be established analytically by first noting that the expression of Eq. (6.7) for the sensitivity has a removable singularity at the origin, and then by expanding it in a power series around $\varphi = 0$ to get, $\Delta\varphi_\mu|_{\varphi\cong 0} \cong 1/\sqrt{\bar{n}}$, which is precisely SNL. Hence by counting all the photons in an off-diagonal fashion, the detection scheme embodied in the operator $\hat{\mu}_{AB}$ of Eq. (6.5) produces a new kind of super-resolution, and performs at the optimal shot-noise limit in sensitivity. Since all of this information is extracted at the detector, and only coherent states are used at the source and in the interferometer, this scheme will be no worse in sensitivity in the presence of absorption or loss than an equivalent classical LADAR scheme, but it will in addition have super-resolving capabilities.

6.3 The Experimental Implementation

It remains to be understood how the observable of Eq. (6.5) might be detected in the laboratory. We note that at this stage of the analysis the operator $\hat{\mu}_{AB}$ is to be carried out with respect to the two-mode coherent state at the line III in the Fig. 6.1. That is we have not yet applied the second beam splitter. To the right of the second beam splitter, at the line IV in Fig. 6.1, we wish to now carry out the so-called parity operator [**?**, 75] on output mode A in the upper arm,

$$\widehat{\Pi}_A = (-1)^{\hat{n}_A} = e^{i\pi\hat{a}^\dagger\hat{a}}, \tag{6.9}$$

which simply indicates whether an even or odd number of photons exits that port. Here $\hat{n}_A = \hat{a}^\dagger\hat{a}$ is the number operator for that mode. Such a detection scheme can be implemented by placing a highly efficient photon-number-resolving detector at this port, and such detectors with 95% efficiency and number resolving capabilities in the tens of photons have been demonstrated [76]. The connection between these two operators may be established via the identity,

$${}_{AB}\langle\Psi|\hat{\mu}_{AB}|\Psi\rangle_{AB} \equiv \langle\alpha, 0|\hat{U}^\dagger(\widehat{\Pi}_A \otimes \hat{I}_B)\hat{U}|\alpha, 0\rangle, \tag{6.10}$$

where $\hat{I}_B$ is the B-mode identity operator, and $\hat{U} = e^{-i\hat{J}_y\varphi}$ with $\hat{J}_y = i(\hat{a}\hat{b}^\dagger - \hat{a}^\dagger\hat{b})/2$ denotes the transformation of the whole MZI, from the line I to IV in Fig. 6.1. That is, the effect of measuring the two-mode coherent state with respect to the operator $\hat{\mu}_{AB}$ to the left of the second BS at

line III is equivalent to measuring it with respect to the parity operator $\widehat{\Pi}_A$ to the right of the second BS at line IV, all indicated in Fig. 6.1. Hence Eq. (6.10) establishes that the two schemes have the same super-resolving and shot-noise limiting properties, with the important point that the parity operator is perhaps far easier to implement in the laboratory, and it has been shown to be a universal detection scheme in quantum interferometry [19, 20]. However, it is yet to be determined whether the parity operator is optimal or not. In the super-sensitivity consideration, for a given input state, whether a particular measurement is optimal or not can be determined by comparing the phase uncertainty given the Cramer-Rao bound and the one obtained directly by that particular measurement [62]. For coherent states, the parity observable is indeed optimal in the phase sensitivity as it saturates the Cramer-Rao bound. We, however, know of no such bound for super-resolution.

6.4 Conclusion

We have provided a super-resolving interferometric metrology strategy, which achieves SNL. The protocol has the appealing feature that it requires only the production and transmission of ordinary laser beams in the form of coherent states of light. Hence, unlike the issues concerning the propagation of non-classical states of light, such as squeezed light or entangled Fock states, this scheme suffers no worse degradation in the presence of absorption and loss than a classical coherent LADAR system. All of the quantum trickery, which provides the super resolution, is carried out in the detection, which can be carried out with current photon number-resolving technology. However, a scheme by which we count the number of photons and then decide if that number is even or odd is overkill. The parity operation only requires that we know the sign — even or odd — independently of the actual number. Hence counting photon number is sufficient but perhaps not necessary, if a general scheme to determine the parity of a photon state could be found that did not require photon number counting. We conjecture such a scheme exists, perhaps through the exploitation of optical nonlinearities [77], or projective measurements, and this is an area of ongoing research. Our protocol for super-resolving phase measurements at SNL has applications to biomedical imaging, metrology, and remote sensing. In particular, for applications such as Doppler velocimetry of rapidly moving objects with few return photons, one does not

have the luxury to integrate the data for long periods of time in order to push down the signal to noise. In such scenarios the signal resolution in the form of the Rayleigh criteria is the usual limit to the system performance.

Chapter 7: Outlook

It needs to point out all the above considerations are confined to the local phase sensitivity according to [62], i.e. the phase shift is known around some specific value in prior. As for the global phase sensitivity, one has no prior phase information and employs a POVM for phase estimation such as Eq. (2.116). It is shown in [78] that the distinctions between the local and the global distinguishabilities merge each other if one uses the multiple-shot measurement to estimate phase shift. For example, in the single-shot measurement the NOON state is locally optimal, but globally suboptimal. If one is allowed to do the adaptive multiple-shot measurement with NOON state, HL can be restored. Moreover, even the separated single photon state can achieve HL within an adaptive multiple-shot measurement scheme, though the whole process takes a much longer period [50]. So the development of a new adaptive protocol, which makes use of easily produced correlated states, in order to find a method that achieves the same HL sensitivity in a much shorter integration time is highly demanded. A more challenging problem will be to develop a formulation of the optimal estimation condition for adaptive multiple-shot measurement.

On the other hand, it is not known how to implement the continuous optimal POVM for MZI, simply because it is impossible to implement it by counting photons in the two output ports. As it was shown in Ref. [17] that it is possible to approximate this optimal POVM by counting output photons if one makes the measurement adaptive. The other possibility is that in finite dimensions the continuous set of POVM can be replaced by a finite set of POVM without error [80, 79]. Such finite POVM can be produced by time reversed measurement probabilistically [81] due to the fact that quantum mechanics permits us to calculate the probabilities of the past preparation events on the basis of the known outcome of a measurement. To improve the efficiency of the measurement, new methods to optimize the probability of success is needed.

Another possible extension of research is to study the spatial correlations of photon field in the interferometer, which has a number of applications in quantum imaging [9].

References

[1] P. A. M. Dirac, Proc. Roy. Soc. London A **114**, 243 (1927).

[2] E. Fermi, Rev. Mod. Phys. **4**, 87 (1932).

[3] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, 1997).

[4] W. E. Lamb, Appl. Phys. B **60**, 77 (1995).

[5] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).

[6] W. W. Chou, Rev. Mod. Phys. **57**, 61 (1985).

[7] K. Thorne, Rev. Mod. Phys. **52**, 285 (1980).

[8] L. A. Lugiato *et al.*, J. Opt. B **4** S176-S183 (2002).

[9] M. I. Kolobov, *Quantum Imaging* (Springer 2005).

[10] C.M. Caves, Phys. Rev. D **23**, 1693 (1981).

[11] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1335 (1993).

[12] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).

[13] R. Yurke, S. L. McCall, and J. R. Klauder, Phys. Rev. A **33**, 4033 (1986).

[14] J. J. Bollinger, *et al.*, Phys. Rev. A **54**, R4649 (1996).

[15] H. Lee, P. Kok, and J. P. Dowling, J. Mod. Opt. **54**, 2325 (2002).

[16] H. M. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995).

[17] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[18] C. C. Gerry and R. A. Campos, Phys. Rev. A **64**, 063814 (2001).

[19] R. A. Campos, C. C. Gerry, and A. Benmoussa, Phys. Rev. A **68**, 023810 (2003).

[20] Y. Gao and H. Lee, J. Mod. Opt. **55**, 3319(2008).

[21] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[22] S. M. Roy and S. L. Braunstein, Phys. Rev. Lett. **100**, 220501 (2008).

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

[24] W. K. Wootters and W. H. Zurek, Phys. Rev. D **19**, 473 (1979).

[25] B.-G. Englert, Phys. Rev. Lett. **77**, 2154 (1996).

[26] B.-G. Englert, M. O. Scully, and H.Walther, Nature **351**, 111 (1991).

[27] J. P. Dowling, Contemp. Phys. **49**, 125 (2008).

[28] A. N. Boto, *et al.*, Phys. Rev. Lett. **85**, 2733 (2000).

[29] M. A. Rubin and S. Kaushik, Phys. Rev. A **75**, 053805 (2007).

[30] U. Dorner, *et al.*, Phys. Rev. Lett. **102**, 040403 (2009).

[31] K. J. Resch, *et al.*, Phys. Rev. Lett. **98**, 223601 (2007).

[32] A. Peres, *Quantum Theory* (Kluwer Academic Publishers, 1993).

[33] J. Preskill, Lecture Note for Physics: *Quantum Information and Computation*

[34] K. Kraus, *States, Effects, and Operations* (Springer, 1983).

[35] H. M. Wiseman, Quantum Semiclass. Opt. **8**, 205 (1996).

[36] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Vol. III (Addison-Wesley, 1989).

[37] W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1998).

[38] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).

[39] S.-J. Wu, X.-M. Chen, and Y.-D. Zhang, Phys. Lett. A **275**, 244 (2000).

[40] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer-Verlag, 1994).

[41] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

[42] H. P. Yuen, Phys. Rev. Lett. **56**, 2176 (1986).

[43] P. Carruthers and M. M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).

[44] H. M. Wiseman, D. W. Berry, and Z.-X. Zhang, Phys. Rev. A **60**, 2458 (1999.

[45] R. Bhandari and J. Samuel, Phys. Rev. Lett. **60**, 2339 (1988).

[46] E. T. Jaynes and F. W. Cummings, Proc. IEEE **51**, 89 (1963).

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

[48] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2005).

[49] J. Muilwijk, *Sankhyā*, Series B **28**, 183 (1966).

[50] B. L. Higgins, *et al.*, Nature **450**, 393 (2007).

[51] Z. Y. Ou, Phys. Rev. Lett. **77**, 2352 (1996).

[52] J. P. Dowling, Phys. Rev. A **57**, 4736 (1998).

[53] M. Hillery and L. Mlodinow, Phys. Rev. A **48**, 1548 (1993).

[54] C. Brif and A. Mann, Phys. Rev. A **54**, 4505 (1996).

[55] P. Walther, *et al.*, Nature **429**, 158 (2004); M.W. Mitchell, J.S. Lundeen, and A.M. Steinberg, *ibid.*, 161 (2004).

[56] T. Nagata, *et al.*, Science **316**, 726 (2007).

[57] Z. Hradil, Phys. Rev. A **51**, 1870 (1995).

[58] T. Kim *et al.*, Phys. Rev. A **57**, 4004 (1998).

[59] R. C. Pooser and O. Pfister, Phys. Rev. A **69**, 043616 (2004).

[60] M. A. Armen, *et al.*, Phys. Rev. Lett. **89**, 133602 (2002).

[61] D. M. Brink and G. R. Satchler, *Angular Momentum* (Clarendon Press, 1993).

[62] G. A. Durkin and J. P. Dowling, Phys. Rev. Lett. **99**, 070801 (2007).

[63] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, Oxford, 2000).

[64] H. Uys and P. Meystre, Phys. Rev. A **76**, 013804 (2007).

[65] M. A. Nielson and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000).

[66] N. Margolus and L. B. Levitin, Physica D **120**, 188 (1998).

[67] L.-M. Duan and G.-C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).

[68] J. H. Shapiro, S. R. Shepard, and N. C. Wong, Phys. Rev. Lett. **62**, 2377 (1989).

[69] P. Kok, S. L. Braunstein, and J. P. Dowling, J. Opt. B **6**, S811 (2004).

[70] G. Gilbert, M. Hamrick, and Y. S. Weinstein, J. Opt. Soc. Am. B **25**, 1336 (2008).

[71] S. D. Huver, C. F. Wildfeuer, and J. P. Dowling, Phys. Rev. A **78**, 063828 (2008).

[72] V. Giovannetti, *et al.*, Phys. Rev. A **79**, 013827 (2009).

[73] R. A. Campos, B. E. A. Saleh, and M. C. Teich, Phys. Rev. A **40**, 1371 (1989).

[74] M. V. Berry and S. Popescu, J. Phys. A **39**, 6965 (2006).

[75] C. C. Gerry, Phys. Rev. A **61**, 043811 (2000).

[76] A. J. Miller, *et al.*, Appl. Phys. Lett. **83**, 791 (2003); G. Khoury, *et al.*, Phys. Rev. Lett. **96**, 203601 (2006); A. E. Lita, *et al.*, Opt. Express **5**, 3032 (2008); C. F. Wildfeuer, *et al.*, `arXiv`: 0905.1085.

[77] C. C. Gerry, A. Benmoussa, and R. A. Campos, Phys. Rev. A **66**, 013804 (2002); *ibid.*, **72**, 053818 (2005).

[78] W. M. Mitchell, Proc. SPIE **5893**, 589310 (2005).

[79] G. Chiribella and G. M. D'Ariano, Phys. Rev. Lett. 98, 190403 (2007).

[80] R. Derka, V. Buzek, and A. K. Ekert, Phys. Rev. Lett. 80, 1571 (1998).

[81] K. L. Pregnell and D. T. Pegg, J. Mod. Opt. **51**, 1613 (2004).

Appendix: Permission to Use the Published Work

The following is copyright section of the Taylor & Francis Group, which can be found at http://www.tandf.co.uk/journals/pdf/copyright-author-rights.pdf. Here I am copying only the relevant part that allows me to use my work published in Journal of Modern Optics **55**, 3319 (2008).

The rights that you personally retain as Author:

10. The right to include an article in a thesis or dissertation that is not to be published commercially, provided the acknowledgement to prior publication in the relevant Taylor & Francis journal is made explicitly.

Vita

Yang Gao is a native of Henan province in People's Republic of China. He was born in November 1982. He finished his undergraduate studies in applied physics at Harbin Institute of Technology in July 2003. He earned a master of science degree in theoretical physics from Peking University in July 2006. In August 2006 he came to Louisiana State University to pursue graduate studies in physics. He is currently a candidate for the degree of Doctor of Philosophy in physics, which will be awarded in August 2010.