Sequential Quantum Measurements In Both the Weak and Strong Regimes

Davor Curic

Department of Physics

University of Ottawa

A thesis completed under the supervision of J. S. Lundeen submitted to the Faculty of Graduate and Postgraduate studies in partial fulfillment of the requirements for the for the degree of

Masters of Science

in Physics

Abstract

Quantum physics provides us with a formalism with which we can try to understand the microscopic world. However, the underpinning concepts of the basic principles of quantum mechanics itself are surprisingly ill understood. Even the ubiquitous concept of the wavefunction collapsing post measurement, lives outside the formal theory of quantum mechanics and must be taken as a postulate. The Schrödinger equation, which governs the evolution of quantum systems, does not describe the system during the measurement process. This is referred to as the measurement problem, and has been debated since the inception of quantum theory. Many different interpretations of quantum mechanics try to elucidate what happens during wavefunction collapse, such as the famous many-worlds interpretation. However, these interpretations either give rise to the same physical outcome or simply are not testable, and so cannot be physically differentiated from one another. We attempt to gain insight into the measurement problem through experiments involving sequential measurements, both in the weak and strong regime. We find surprising outcomes such as the persistence of time-ordering asymmetry in the weak measurement limit, and the survival of coherence in the strong measurement limit.

Acknowledgements

First, I would like to acknowledge the financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the University of Ottawa.

I am grateful to my supervisor, Jeff Lundeen, for his guidance both academically and personally. His enthusiasm towards scientific research is inspiring, but my biggest take away is his passion for scientific communication and presentation. The clarity with which he approaches explaining even seemingly complicated topics is something I plan on taking forward with me to whichever future project I pursue.

I truly doubt that I would have accomplished as much as I did without the help of Lambert Giner, the Lundeen lab post-doc. Lambert provided countless codes which allowed for automation of lab equipment. As well, Lambert was always available for back and forth discussion, both scientific and otherwise.

I would also like to thank the entirety of the Canada Excellence Research Chair (CERC) group. Moving to a new city is always hard, but I can genuinely say that the people in the CERC group will make it hard to leave. Thank you for the fun movie nights (Sharknado night in particular), curling, ice-skating etc. I would in particular like to thank the usual lunch group (of whom there are too many to name) and my colleges in the Lundeen lab.

I would also like to thank the great people I met outside the CERC group. In particular Jen, who was supportive and encouraging throughout my Masters, and my amazing room-mates, David, Candace, and Justin who always served as a great distraction from math. As well as my friends from Calgary who were immensely supportive even though they were far away.

Most importantly, I am grateful for my family's loving support. Leaving them to start my Masters was one of the hardest things I've had to do. But even from across the country, I could not have succeeded without their guidance and support.

Contents

1

The Measurement Problem

1.1 Introduction

Quantum mechanics has moved beyond its original primarily academic setting. Modern commercial devices are either already utilizing quantum mechanics in some way, or are starting to consider quantum mechanical effects as devices become smaller (1). Material scientists utilize complex quantum phenomena to develop materials that could potentially revolutionize future technologies (2). The word "Quantum" has even moved into the mainstream consciousness, with many commercial goods using it as a prefix to make products sound powerful, or revolutionary (e.g. Finish Quantum Max Fresh dishwasher tablets), or politicians giving public lectures about the importance of quantum computing.

Given the ubiquity of quantum mechanics, it might then be surprising to learn how poorly the fundamental aspects of quantum mechanics are understood. Fundamental concepts such as the mechanism of wavefunction collapse, the transition between the quantum and classical worlds, the concept of time, and more, are equally poorly understood. However, what is understood is that the mechanism that underlies these seemingly disparate concepts is measurement. Studying the role of measurement in quantum mechanics provides a promising avenue to studying poorly understood fundamental concepts.

Recent interest has developed around utilizing sequences of measurements as a means to gain deeper insight into measurement. Sequential measurements allow one to study how the wavefunction is disturbed as it propagates through the sequence.

Sequential measurements have already been used to perform interesting quantum experiments such as making measurements of incompatible observables (3, 4, 5). On a more fundamental note, sequential measurements have been used to study paradoxes, and they open the possibility of studying fundamental concepts such as time in quantum mechanics, the role and propagation of coherence (6, 7, 8). Sequential measurements provide us a tool with which we can hope to study measurement experimentally rather than philosophically.

1.2 Summary of Work

This thesis covers the three major projects I worked on during the course of my Masters. These projects all serve as an experimental dissection of the measurement problem. The first major topic is studying the effect of time-ordering of non-commuting weak measurements. Quantum mechanics has a built in time-asymmetry due to the non-reversibility of the measurement process, also known as wavefunction collapse. As a consequence of this collapse, permuting the ordering of a measurement within a sequence of measurements yields a different result (non-commutivity). Weak measurements are a class of measurements that yield little information on a single shot basis, but as a result they do not collapse the quantum wavefunction of the measured state. It is natural to then ask, is time-symmetry restored in a sequence of weak measurements, given that the measurements do not collapse the state. We find, surprisingly, that time-symmetry is in fact *not* restored, even in the limit of zero interaction strength between the system and measurement apparatus. However, this asymmetry manifests in a complicated way. For example, a sequence of two non-commuting weak measurements turn out to be time-symmetric, where as a sequence of three measurements is not.

My second major project was studying the propagation of coherence through sequences of strong measurements. Basic intuition of quantum theory suggests that in a sequence of three, non-commuting measurements, each successive measurement collapses the state, totally randomizing the result. However, collapse only occurs when we have learned of the result. We perform an experiment to show three sequential measurements retain coherence between their possible results. This implies the wavefunction has not in fact collapsed between the first and last strong measurements.

The third project was to develop the tools necessary to measure the results for second project. To check if coherences between strong measurement results has remained, one has to reconstruct the state of an eight dimensional density matrix encoded into the eight possible path modes that result from the sequential measurements. As it turns out, state reconstruction can be challenging. Not only does the number of measurements needed scale as the square of the dimension of the system, a procedure for reconstructing high dimensional path encoded states is in itself not clear. As such, we were required to develop a method of reconstructing an eight dimensional state path encoded state simply. We show that this can be done simply by using a rotating cylindrical lens to interfere states encoded onto transverse spatial path modes. We describe this method, and show its effectiveness at reconstructing states faithfully.

1.3 Context

Quantum measurement theory links the probabilistic world of quantum mechanics and the deterministic classical world. However, before physicists started studying measurement itself, quantum theory was built on the measurement postulates; measurement itself lies outside the formal theory. And this, while unsatisfying perhaps, worked well. Scientists were able to study the structure of atoms and molecules, developed scattering theory, quantized the electromagnetic field, derived rate equations for spontaneous emission, and so on. Never the less, the lack of a formal theory of measurement was a black eye on quantum mechanics. Heisenberg (9) introduced the collapse of the wavefunction, Dirac set the initial foundation for a quantum measurement theory (10), but it was John von Neumann who added rigorous mathematics to the problem (11).

John von Neumann developed many important concepts such as the density matrix formalism, generalizing the quantum state to allow for classical statistical mixtures. This was one of the first major bridges between the quantum and classical world. Perhaps the most significant, at least in the context of measurement theory, was the introduction of what many now call the von Neumann model of measurement. Prior to this, measurement theory was an unmotivated whirlwind of eigenvalues and eigenvectors, and did not take into account the measurement device itself. John von Neumann introduced a mathematical treatment of the measurement apparatus. He also conceived of an interaction Hamiltonian for the measurement process: $\boldsymbol{\mathcal{H}} = g\boldsymbol{A} \otimes \boldsymbol{p}$. The symbol

$\otimes$ denotes the tensor product of operators belonging to separate Hilbert spaces, and may be dropped when the context is clear. Here $\boldsymbol{A}$ is the observable being measured, $\boldsymbol{p}$ is the momentum operator of the measurement apparatus, and g is the interaction parameter between the two systems. He showed that this interaction shifts a dial on the measurement apparatus to point to a result, conditioned on the outcome of measuring $\boldsymbol{A}$. This model was later tweaked by Lüders to include a decohereing environment, producing the theory of measurements that is still used today (12).

After the initial formalization by von Neumann, mathematical physicists generalized quantum measurement theory further to describe non-ideal measurements, leading to the development of positive-operator valued measures (13). The field of quantum optics was an especially fruitful area for researching measurement. Physicists in this field seemed particularly adept at pushing the bounds of what could theoretically be measured. In an ever expanding attempt to beat the fundamental limits set in place by nature, they introduced exotic quantum states, such as squeezed states, and methods such as weak measurement (14, 15). Indeed, in pushing the resolution of measurements (16), they could develop methods that in essence put error bars on their error bars (for example, the beam splitter model of loss in quantum optics (17)). The result of an experiment was no longer based on how well the outcome agreed with the theory, but also how well the noise of the measurement agreed with the theory.

Nevertheless, much work remains to be done in the context of measurement. Fundamental concepts are still poorly understood, despite the progress made elsewhere. Take for example the related concept of time in quantum mechanics. Physicists hold time symmetry, the idea that dynamics look the same whether played forward or in reverse, in high regard. The one exception was entropy, which is always said to increase with time, therefore providing an 'arrow' of forward time. However, the law that entropy always increases is an emergent property of systems in the thermodynamic limit. For an ensemble with only a few particles, entropy is just as likely to decrease as it is to increase. Similarly, CPT (charge, parity, time) symmetry is also thought to provide an arrow of time. CPT symmetry is believed to be a fundamental symmetry of the universe. Experiments have concluded that CP symmetry is violated (i.e. it is not conserved) (18), which implies T symmetry must also be violated in order for CPT symmetry to remain. However, CP symmetry is only very rarely violated, meaning that T symmetry is only weakly broken. Another candidate for an 'arrow of time' was

quantum measurement, which appears to be non-unitary (i.e. irreversible). However, time in quantum mechanics is a surprisingly tricky concept (19, 20, 21). Physicists since the days of Wolfgang Pauli have struggled to come up with a rigorous meaning of time in the context of quantum mechanics. Moreover, recent theoretical and experimental methods such as weak measurement question how fundamental this time asymmetry is. Given that measurement is at the heart of quantum mechanics, more needs to be done in order to elucidate its more obscure features, and help contextualize our classical reality within the weird reality that supposedly forms the foundation of our universe.

1.4 Contributions of the Author

The work presented in this thesis is of course a collective effort. Here I will contextualize my contributions to each project. All data presented was obtained by myself in the lab during the course of my Master's degree. Likewise, all figures were created by my self.

The original theoretical framework for the time-ordering project was developed in (6, 22). The experimental setup was a modification of the one used in (5), which was headed by Guillaume Thekkadath. Guillaume developed the alignment procedure and initial analysis code, off which the time ordering project was developed. My contributions included taking the data as well as adapting the theory from (6, 22) to our experiment. I was also assisted in the data collection by Magdalena Richardson, an undergraduate student working on her research project. As well, in trying to understand the project, I independently discovered an interesting mathematical structure for weak sequential measurements, however, this was later discovered to have been published a year prior (23).

The initial proposal for the second project was detailed in (8). However, the paper offered little clues as to how to obtain the data needed to test it, which led into the third project of tomography of states encoded into a superposition of many possible paths. The experimental design, setup, data collection, analysis code, and theory were done by my self. Once this setup was completed, and shown to work, I returned to the second project to collect the necessary data.

1.5 Outline

The thesis is broken up into five chapters (not including the introductory chapter 1, or the concluding chapter):

Chapter 2: This chapter goes over the background for the time-ordering chapter. Here I introduce the notion a system and pointer. These are fundamental concepts of the von Neumann model of quantum measurement. Using it we derive the weak and strong limits of measurements. We show how weak measurements can still offer information on the average result of a measurement, and setup the concept of sequential measurements.

Chapter 3: This chapter discusses the time-ordering of weak measurements topic. Time-ordering dependence of measurements is shown to remain even in the weak limit.

Chapter 4: This chapter details the background for chapters 5 and 6. The concept of the density matrix is introduced, allowing us to write states as statistical mixtures. We show how the loss of coherence in a pure entangled state produces statistical mixtures. We then consider sequential strong measurements, and show that the end results retain coherence.

Chapter 5: This chapter details the method devised to obtain the data for the project introduced in the previous chapter. We show how using one-dimensional Fourier transforms implemented by a cylindrical lens can allow one to reconstruct the state of a large dimensional density matrix.

Chapter 6: In this chapter, we return to the second major project, now with the tools necessary to measure the large states required. We show that indeed, after a sequence of three strong measurements, the resulting state retains coherence, a result incompatible with the notion of collapse. We conclude and discuss future work.

2

Measurement Disturbance

Preamble

One of the most unconventional features in quantum mechanics is the superposition principle; if $|a_1\rangle$ and $|a_2\rangle$ are two valid states, then the state $\alpha\,|a_1\rangle + \beta\,|a_2\rangle$ $(\alpha, \beta \in \mathbb{C})$, is also a physically possible state. Superpositions are a critical feature in the study of foundational quantum mechanics and the development of quantum information sciences. There is great interest in creating systems which are robust against collapse for use in information sciences. This ranges from photons, to superconducting qubits which can occupy a state of left-spinning current and right-spinning current at the same time (24). However, in our daily life, we do not see objects as being in a superposition of two states at once. If quantum theory underlies our natural world, why do we not see it? In this section we introduce the measurement problem and how its significance to the theory of collapse. This provides the background for the next chapter where we look at how we can 'reduce' this collapse and the consequences of doing so.

2.1 Measurement Disturbance

Unlike other theories of physics, the concept of measurement plays a central role in quantum mechanics. In a naive sense, measurement is how one obtains information about a quantum system. Generally, an experiment is designed to measure a property of a system, which can mathematically be represented by an observable operator $\boldsymbol{A}$. The distinct, orthogonal outcomes of measuring of $\boldsymbol{A}$ are represented by *eigenstates* $\{|a_n\rangle\}$. The measurement returns a real value a_n, known as the *eigenvalue*, that

informs the outcome of measuring $\boldsymbol{A}$. The operator $\boldsymbol{A}$ can be represented in terms of its eigenvectors and eigenvalues as $\boldsymbol{A} = \sum_n a_n |a_n\rangle \langle a_n|$, which is known as the *spectral decomposition* of $\boldsymbol{A}$. The outcome of the measurement will depends on the state, $|\psi\rangle$, of the system that is to be measured. Quantum mechanics tells us that the state $|\psi\rangle$ may be written as a superposition of these eigenstates

$$|\psi\rangle = \sum_n c_n |a_n\rangle .$$

The $c_n = \langle a_n|\psi\rangle$ are the complex probability amplitudes. As an inner product, they represent a measure of the 'amount' that $|\psi\rangle$ is along $|a_n\rangle$. The measurement postulate of quantum mechanics states that the original state, which was in a superposition of the $\{|a_n\rangle\}$ states, is, after having been measured, in only one particular $|a_n\rangle$ states. Measurement has collapsed the state. Which state the system will be in after the measurement is entirely random, but has a probability given by $|c_n|^2 = |\langle a_n|\psi\rangle|^2$.

The collapse of the state into eigenstates of $\boldsymbol{A}$ sets limits on how much can be learned about two properties at the same time. To see why this is, we consider another observable $\boldsymbol{B}$ with eigenstates $\{|b\rangle_n\}$ which are different from the eigenstates of $\boldsymbol{A}$. We will assume that the initial state of the system is the $\boldsymbol{A}$ eigenstate $|a\rangle$, and we measure $\boldsymbol{A}$ followed by $\boldsymbol{B}$. Because the state is in $|a\rangle$, the outcome of the first measurement is always $|a\rangle$. As $|a\rangle$ is not an eigenstate of $\boldsymbol{B}$, the second measurement collapses it to $|b\rangle \in \{|b_n\rangle\}$ with probability $|\langle b|a\rangle|^2$. Consider now the reverse sequence of measurements. Measuring $\boldsymbol{B}$ collapses the initial state $|a\rangle$ to one of the states in $|b\rangle \in \{|b_n\rangle\}$. This state is then collapsed again by the second measurement $\boldsymbol{A}$ as $|b\rangle$ is not an $\boldsymbol{A}$ eigenstate. The resulting state is an eigenstate of $\boldsymbol{A}$, but not necessarily the one we started off with. From this, we can see that $\boldsymbol{AB}|a\rangle \neq \boldsymbol{BA}|a\rangle$. This is known as the *non-commutativity* of observables. Mathematically, this is written as $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{AB} - \boldsymbol{BA} \neq 0$, where $[\cdot, \cdot]$ is called the *commutator*. If $[\boldsymbol{A}, \boldsymbol{B}] = 0$, then $\boldsymbol{AB} = \boldsymbol{BA}$, and the observables are said to commute. This is only the case if both $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same set of eigenstates, as then sequential measurement of one will not disturb the other. Famously, the position, $\boldsymbol{x}$, and momentum, $\boldsymbol{p}$, cannot be simultaneously measured because $[\boldsymbol{x}, \boldsymbol{p}] = i\hbar$.

If one cannot measure two results simultaneously, can accuracy on one be given up in order to measure the other better? This question contributed to the development of the field of quantum metrology, where researchers study how to make higher

resolution measurements of physical parameters using exotic quantum states and methods (15, 25). Predating quantum metrology was *Heisenberg's uncertainty principle*. The uncertainty principle states that the uncertainty in a particle's position, Δx, and uncertainty in a particle's momentum, Δp, have to obey the relation

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{2.1}$$

In order to minimize the uncertainty on position, certainty on momentum needs to be given up. A more general result states that for any two observables $\boldsymbol{A}$ and $\boldsymbol{B}$, the product of their uncertainties has to obey

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\boldsymbol{A}, \boldsymbol{B}]\rangle|. \tag{2.2}$$

Therefore, the outcome of two simultaneous measurements can only be known with absolute certainty when the two measurements commute. While $\boldsymbol{x}$ and $\boldsymbol{p}$ are the most famous example, non-commuting observables are everywhere. Other examples include measuring the horizontal polarization of a photon, $|H\rangle$, followed by measuring diagonal polarization, $|D\rangle$, or number-phase uncertainty in quantum optics[1] (26). The uncertainty principle is a measure of how *fundamental* the uncertainty of a measurement is, and whether or not one can design a device that can beat that limit.

Equation 2.2 is a hard and fast universal law and as such cannot be violated. However, physicists quickly found loop-holes to exploit. One of the most promising avenues is often brought up (and dismissed) in introductory quantum mechanics courses; just perform the measurements very gently so as to not disturb the state. Before we can discuss how to reduce the disturbance, a more realistic model of measurement needs to be introduced so as to be able to quantify what reducing the disturbance means.

2.2 Von Neumann's Model for Measurement

The above treatment of measurement, often called a direct measurement, is mathematically simple. Measuring the observable $\boldsymbol{A}$ collapses the system into an eigenstate $|a\rangle$ of $\boldsymbol{A}$ with a probability of $|\langle\psi|a\rangle|^2$. However this is rather abstract. Students

[1] Despite its ubiquity in the field, this one is strange in that a phase operator is not well defined. Regardless, this uncertainty relation has been well verified in countless experiments. Another example of this is the time-energy uncertainty relation, as a time operator is difficult to consistently define.

learning quantum mechanics may ask, what does it mean for the observable to return a particular eigenvalue? How does this eigenvalue manifest itself physically? Where is the measurement device and how is it related to $\boldsymbol{A}$? How fast does the collapse $|\psi\rangle \to |a\rangle$ happen? Needless to say, a more rigorous definition for this process needed to be developed.

The first mathematically rigorous attempt at describing the measurement process is generally attributed to John von Neumann. In 1932, he provided a more realistic model of measurement, now called the indirect measurement model, or von Neumann model (27). Von Neumann introduced to the theory of measurement a mathematical model of a measurement device. He introduced to the theory a model for the measurement apparatus which is treated quantum mechanically. To do this, the total Hilbert space is subdivided into two pieces; $\mathcal{S}$ is the Hilbert space of the quantum system, and $\mathcal{P}$ contains measurement device, which the experimenter has complete control over. For reasons which will be made apparent in a moment, $\mathcal{P}$ is often called the 'pointer' space.

At the start of an experiment to measure the observable $\boldsymbol{A}$, the quantum system is in an initial state $|\chi\rangle \in \mathcal{S}$. Independently, the measuring device is prepared in an initial state, $|\psi(x)\rangle \in \mathcal{P}$. The mean position of the state represents a pointer on a gauge that one can read off classically (e.g. by looking at it). It is assumed that the state is prepared in such a way so that its initial mean position $\langle \boldsymbol{x} \rangle$ is zero. The total of the two independent states, $|\Psi\rangle = |\chi\rangle \otimes |\psi\rangle$, is called a *bipartite state* (i.e. a state belonging to two separate parties), and belongs to the joint Hilbert space $\mathcal{S} \otimes \mathcal{P}$. As the experiment progresses, the quantum system and the measurement device evolve together according to an interaction Hamiltonian given by $\boldsymbol{\mathcal{H}} = g(t)\boldsymbol{A} \otimes \boldsymbol{p}$. Here $\boldsymbol{p}$ is the momentum operator on the measurement device Hilbert space. g is known as the *measurement strength*[1]. This Hamiltonian corresponds to continuous unitary time evolution given by $\boldsymbol{U} = e^{i\boldsymbol{\mathcal{H}}} = \exp(-ig\boldsymbol{A} \otimes \boldsymbol{p})$. At the end of the experiment the final state is given by $\exp(-ig\boldsymbol{A} \otimes \boldsymbol{p}) |\chi\rangle \otimes |\psi\rangle$.

Before proceeding we discuss how to treat functions of operators. Using the spectral decomposition, we can write $\boldsymbol{A} = \sum_n a_n |a_n\rangle \langle a_n|$. A function $f(\boldsymbol{A})$ can be then be written as $f(\boldsymbol{A}) = \sum_n f(a_n) |a_n\rangle \langle a_n|$. With this in mind, the evolution operator

[1] I am abusing notation here and am assuming that $\boldsymbol{\mathcal{H}}$ is time independent and so g is the interaction integrated over the entire interaction time T, i.e. I redefine $\int_0^T g(t)\mathrm{d}t = gT$ to be equal to g.

becomes

$$\boldsymbol{U} = \exp\left(-ig\boldsymbol{A}\otimes\boldsymbol{p}\right) = \sum_n |a_n\rangle\langle a_n| \otimes \boldsymbol{T}(ga_n), \tag{2.3}$$

where $\boldsymbol{T}(ga) = \exp(-iga\boldsymbol{p})$. On the right-side, we have rewritten $\boldsymbol{U}$ in the eigenbasis of $\boldsymbol{A}$. The operator, $\boldsymbol{T}(ga)$ is known as a translation operator, and has the effect of shifting the pointer's position by ga, i.e., $\boldsymbol{T}(ga)\,|\psi(x)\rangle = |\psi(x-ga)\rangle$.

To understand the result of this evolution $\boldsymbol{U}$, we focus on a single term in the Eq. 2.3, $|a\rangle\langle a|\otimes\boldsymbol{T}(ga)$, with which we act on our initial total state:

$$\begin{aligned} |a\rangle\langle a|\otimes\boldsymbol{T}(ga)\,|\chi\rangle\,|\psi\rangle = |a\rangle\langle a|\chi\rangle\otimes\boldsymbol{T}(ga)\,|\psi(x)\rangle = \\ \langle a|\chi\rangle\,|a\rangle\otimes|\psi(x-ga)\rangle\,. \end{aligned} \tag{2.4}$$

The result of acting on the total initial state is to project the system state $|\chi\rangle$ onto the eigenstate $|a\rangle$. At the same time, the pointer state $|\psi(x)\rangle$ is shifted by an amount ga (i.e. the shift is proportional to the corresponding eigenvalue) to a new state $|\psi(x-ga)\rangle$. The mean of the position observable of the pointer, $\langle\boldsymbol{x}\rangle$, having initially pointed to zero, has now shifted to ga. Indeed, this is why the state is called the "pointer", as it can be thought of as the literal needle on the device that *points* to the result of the measurement that one observes on a gauge. The probability of this shift occurring is related to the overlap $\langle a|\chi\rangle$ of the initial system state and the eigenstate. The result of $\boldsymbol{U}$ acting on $|\chi\rangle\,|\psi\rangle$ is to make a superposition of all the individual possible results:

$$\boldsymbol{U}\,|\chi\rangle\,|\psi\rangle = \sum_n \langle a_n|\chi\rangle\,|a_n\rangle\otimes|\psi(x-ga_n)\rangle \tag{2.5}$$

The resulting bipartite state (i.e., a state which belongs to the tensor product space of two separate Hilbert spaces) has a very important form. Unlike the initial state $|\chi\rangle\,|\psi\rangle$, this state cannot be written as a product state $|\Phi\rangle\,|\Psi\rangle$, where $|\Phi\rangle\in\mathcal{S}$ and $|\Psi\rangle\in\mathcal{P}$. Because the states are separable, or independent, they lack any correlation. The lack of a product state for Eqn. 2.5 means that the system and pointer have become correlated; the total state can only be specified by knowing both the pointer and system states. This type of correlation is called *entanglement*, and is an important feature of the indirect measurement; Information is gained about a quantum system $|\chi\rangle$ not by observing it directly (like in direct measurements). Rather one obtains information indirectly, by looking at a pointer whose position has been entangled with the measurement eigenstates of the system (28).

Motivating Example Using Birefringent Crystals.

Here we consider an example that explicitly involves the von Neumann measurement model, and serves as an introduction to one of the main experimental tools that we will use. Consider a photon in a Gaussian spatial mode with polarization $\alpha |H\rangle + \beta |V\rangle$, passing through a birefringent crystal. Birefringence is a property of materials in which the crystal structure causes the refractive index of the material to be polarization dependent. That is, horizontally polarized light will see a different refractive index compared to vertically polarized light. When the k-vector of the incoming field is at an angle with respect to the crystal axis, an effect known as polarization dependent spatial walk-off, or simply just $walk\text{-}off$, occurs. Walk-off causes horizontally polarized light to be transmitted normally through the crystal, where as vertically polarized light will be displaced by g (more information about spatial walk-off can be found in the supplementary material at the end of chapter 4). The result is that the photon is now in a superposition of two parallel *paths*, one polarized as $|H\rangle$ and the other as $|V\rangle$. The probability to observe the photon in either one of these two paths will depend on the values α and β. This situation can be seen in Fig. 2.1.

We now treat this as a von Neumann measurement. We identify the system $\mathcal{S}$ as the polarization of the light, and it has a state $|\chi\rangle = \alpha |H\rangle + \beta |V\rangle$. The pointer is the Gaussian transverse spatial distribution of the photons, and is in an initial state $|\psi(x)\rangle$, with a spatial wavefunction

$$\langle x|\psi(x)\rangle = \psi(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{x^2}{4\sigma^2}\right).$$

We choose to measure the polarization of the light, which can be written as the observable $\boldsymbol{A} = 1 |V\rangle \langle V| + 0 |H\rangle \langle H| = 1\boldsymbol{\pi}_V + 0\boldsymbol{\pi}_H$. The interaction Hamiltonian, $\boldsymbol{\mathcal{H}} = g\boldsymbol{A} \otimes \boldsymbol{p}$, induces unitary evolution given by

$$\boldsymbol{U} = e^{i\boldsymbol{\mathcal{H}}} = \exp\left(-ig\boldsymbol{A} \otimes \boldsymbol{p}\right) = \boldsymbol{\pi}_V \otimes \boldsymbol{T}(g) + \boldsymbol{\pi}_H \otimes \boldsymbol{T}(0). \tag{2.6}$$

As the light propagates through the crystal it undergoes the evolution $\boldsymbol{U}$

$$\left(\boldsymbol{\pi}_V \otimes \boldsymbol{T}(g) + \boldsymbol{\pi}_H \otimes \boldsymbol{T}(0)\right) |\chi\rangle |\psi(x)\rangle = \beta |V\rangle |\psi(x-g)\rangle + \alpha |H\rangle |\psi(x)\rangle . \tag{2.7}$$

We see that the spatial distribution of the $|V\rangle$ polarized light has been shifted ($|\psi(x-g)\rangle$), and the $|H\rangle$ polarized light has not ($|\psi(x)\rangle$), just as in Fig. 2.1. We calculate the mean

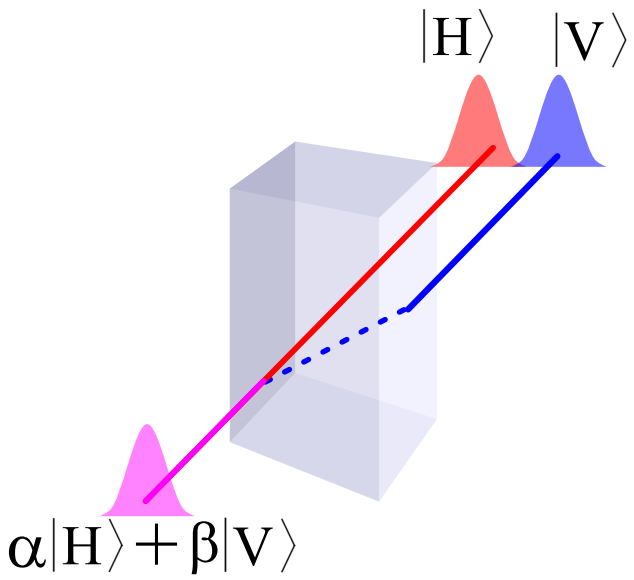

Figure 2.1: A birefringent crystal splits light according to its polarization, resulting in two parallel propagating beams. We assume here that the shift between the two parallel paths is much greater than the width of the transverse profiles. In this scenario, the result of measuring a single polarized photon can be unambiguously determined by observing which of the two lobes the photon was found in after the crystal.

shift of the pointer $\langle \boldsymbol{x} \rangle$ to retrieve the result of the measurement:

$$\langle \boldsymbol{x} \rangle = |\beta|^2 \langle \psi(x-g) | \boldsymbol{x} | \psi(x-g) \rangle + |\alpha|^2 \langle \psi(x) | \boldsymbol{x} | \psi(x) \rangle = |\beta|^2 g = \langle \boldsymbol{A} \rangle g. \tag{2.8}$$

The result shows that the mean position of the final pointer state is proportional to $|\beta|^2$, or equivalently, $\langle \boldsymbol{A} \rangle$. g is assumed to be known precisely, therefore the result of the measurement (i.e. whether or not the photon has been shifted) can be found by looking at the amount by which the pointer was shifted. If the initial system state was entirely vertically polarized ($|\chi\rangle = |V\rangle$), then $\beta = 1$, and the pointer is fully shifted by g. Conversely, if it was entirely horizontally polarized, ($|\chi\rangle = |H\rangle$), then $\beta = 0$, and the pointer remains un-shifted. This holds true for any value of $0 \leq \beta \leq 1$.

2.3 Strong and Weak Measurements

It is worth studying the overlap of the shifted pointer state $|\psi(x - ga)\rangle$ and the initial state. Given $\psi(x)$ is Gaussian with zero mean and width σ, then $\langle \psi(x - ga) | \psi(x) \rangle \propto \exp(-g^2 a^2 / 8\sigma^2)$. For a given eigenvalue a, we can study two regimes. In the first regime

$g^2a^2/8\sigma^2 \gg 1$. In this limit, the overlap $\langle\psi(x-ga)|\psi(x)\rangle$ vanishes. That is, the initial pointer state and the final pointer state are distinct (i.e. orthogonal) from one another. In the case where $\boldsymbol{A}$ is a projector, and thus $a \in \{0,1\}$, the overlap vanishes when $ga \gg \sigma$. Geometrically speaking, the shift imparted to the pointer is much greater than the width of the pointer itself. If one considers a single polarized photon passing through the crystal, it will land unambiguously in one of the two Gaussian lobes. This is called a *strong* measurement, and is the case found in conventional experiments (and in Fig 2.1 of our motivating example above). It is worth noticing that one drawback is that it has collapsed the state. While the state was originally in $\alpha\,|H\rangle + \beta\,|V\rangle$, when a photon is detected in in either of the two lobes after the crystal, it is either $|H\rangle$ or $|V\rangle$, but definitely not both. The result in Eqn. 2.8 is proportional to $|\beta|^2$, not β; we have lost phase information. From this single measurement, one cannot faithfully reconstruct $|\chi\rangle$.

As it turns out, there is some practical use of measurements that give ambiguous results. Looking at the other limiting case of the overlap, one can consider reducing the interaction strength so that $ga \ll \sigma$. In this case $\langle\psi(x-ga)|\psi(x)\rangle \approx 1$, meaning that the initial pointer and the shifted pointer cannot be distinguished. In the limit where $g \to 0$, the pointer now extends over multiple indicator marks, ga, and so the result of a single trial will be entirely ambiguous. This regime is known as *weak* measurement. While the result of a single trail does not give any results, the act of measurement does not greatly disturb the state of the system. This can be seen by calculating $\langle\chi|\,\langle\psi(x)|\,\boldsymbol{U}\,|\chi\rangle\,|\psi(x)\rangle$ and showing that in the weak limit this evaluates to unity. Moreover, averaged over many trials the mean measurement result is proportional to the average pointer position, $g\langle\boldsymbol{A}\rangle = \langle\boldsymbol{x}\rangle$, just as in Eqn. 2.8, where the measurement was considered strong. Even though the result of a single measurement is entirely ambiguous, we can still calculate the average result of each measurement.

Returning to the example of the birefringent crystal, we imagine the crystal has been cut to be thin enough such that the separation between the two Gaussian beams is much less than the width of the distribution. If one considers the measurement of a single polarized photon, one would not be able to distinguish which lobe it landed it, as shown in Fig. 2.2, and so the result is ambiguous. Regardless, the average result of the experiment can be obtained by averaging over many trials.

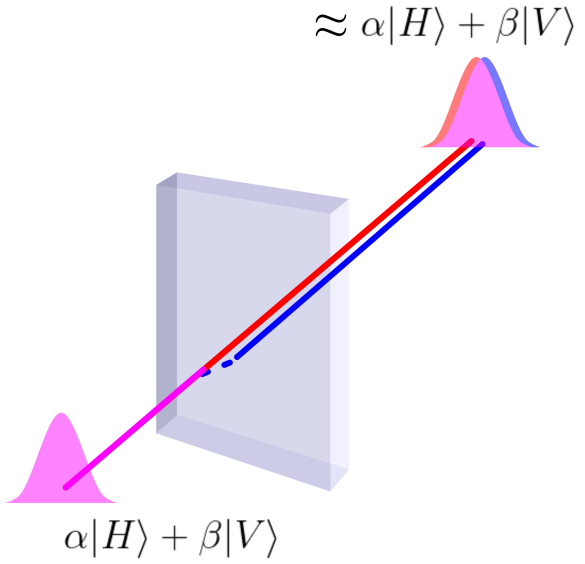

Figure 2.2: The birefringent crystal is cut such that the shift between the two parallel paths is much smaller than the width of the transverse profiles. In this scenario, the result of measuring a single polarized photon is entirely ambiguously as one cannot determine which of the two lobes the photon was found in after the crystal. While the result of a single measurement is undetermined, the advantage is that the state has not been collapsed. The average result of the measurement can be determined by measuring the (small) shift between the two paths.

2.4 Sequential Measurements

Weak measurements do not (significantly) disturb the initial state of the quantum system. This brings up the intriguing possibility of using weak measurements to perform sequences of measurements on the *same* quantum system (3). That is, because each measurement does not change the state, it appears possible to do sequences of non-commuting measurements on a single copy of the system with weak measurements. Formulating sequential measurements is a simple extension of the von Neumann model introduced above. Originally we introduced a pointer $|\psi(x)\rangle \in \mathcal{P}_x$ that points to the outcome of measuring $\boldsymbol{A}$, and the result of the measurement was $\langle x\rangle$. We introduce a second pointer $|\psi(y)\rangle \in \mathcal{P}_y$, that is independent of the first pointer (i.e. $[\boldsymbol{x}, \boldsymbol{y}] = 0$), and points to the outcome of measuring $\boldsymbol{B}$. If we measure first $\boldsymbol{A}$ then $\boldsymbol{B}$, the state evolves to

$$e^{ig_y \boldsymbol{B}\boldsymbol{p}_y} e^{ig_x \boldsymbol{A}\boldsymbol{p}_x} |\chi\rangle |\psi(x)\rangle |\psi(y)\rangle , \tag{2.9}$$

where $\boldsymbol{p}_{x(y)}$ is the momentum operator conjugate to $\boldsymbol{x}(\boldsymbol{y})$. The result of the sequential measurement is now the joint mean positions of the two pointers, $\langle \boldsymbol{x}^{\boldsymbol{A}} \boldsymbol{y}^{\boldsymbol{B}} \rangle$. The superscripts remind us which pointer corresponds to which operator. As well, they denote the order of the measurements as $\boldsymbol{x}$ and $\boldsymbol{y}$ commute, so their order in the expectation values does not necessarily reflect the order of measurements.

It is worth stopping here for a moment to think about some of the consequences of this sequential measurement result. First, how does this comply with the uncertainty principle? If $[\boldsymbol{A}, \boldsymbol{B}] \neq 0$ how can they be measured simultaneously? The uncertainty principle states that a single measurement cannot at the same time measure non-commuting observables with arbitrary certainty on both results. Here, the measurements of non-commuting observables are performed over many trials, and the results are averaged, which is not forbidden. The result of any single measurement obeys the uncertainty principle, but the weak measurement gives us the result averaged over many trials.

Secondly, and the question which concerns the first topic of my Masters work, is that of time ordering of weak measurements (6). Given that the measurement disturbance has been reduced so that it does not collapse the state, how does weakly measuring $\boldsymbol{A}$ followed by $\boldsymbol{B}$, compare with weakly measuring $\boldsymbol{B}$ followed by $\boldsymbol{A}$? Stated another way, are weak measurements time symmetric (22, 29)? In weak measurement, the disturbance has been reduced and so it does not collapse the state, which is ultimately the cause of ordering dependence in strong measurement, and so it should be the case that $\langle \boldsymbol{x}^{\boldsymbol{A}} \boldsymbol{y}^{\boldsymbol{B}} \rangle = \langle \boldsymbol{x}^{\boldsymbol{B}} \boldsymbol{y}^{\boldsymbol{A}} \rangle$. On the other hand, a weak measurement still nevertheless disturbs the state, even if less so, and as such the ordering should matter. According to this line of reasoning, it should be the case that $\langle \boldsymbol{x}^{\boldsymbol{A}} \boldsymbol{y}^{\boldsymbol{B}} \rangle \neq \langle \boldsymbol{x}^{\boldsymbol{B}} \boldsymbol{y}^{\boldsymbol{A}} \rangle$. As we discuss in the next chapter, the surprising answer is that both of these are true.

3

Time-Ordering of Weak Measurements

Preamble

This chapter is based on the following publication:

Experimental investigation of measurement-induced disturbance and time symmetry in quantum physics
D. Curic, M. C. Richardson, G. S. Thekkadath, J. Flrez, L. Giner, and J. S. Lundeen, Phys. Rev. A 97, 042128 (2018); doi: 10.1103/PhysRevA.97.042128

Unlike regular time evolution governed by the Schrödinger equation, standard quantum measurement appears to violate time-reversal symmetry. Measurement creates random disturbances (e.g., collapse) that prevents back-tracing the quantum state of the system. The effect of these disturbances is explicit in the results of subsequent measurements. In this way, the joint result of sequences of measurements depends on the order in time in which those measurements are performed. One might expect that if the disturbance could be eliminated this time-ordering dependence would vanish. Following a recent theoretical proposal (6), we experimentally investigate this dependence for a kind of measurement that creates an arbitrarily small disturbance, weak measurement. We perform various sequences of a set of polarization weak measurements on photons. We experimentally demonstrate that, although the weak measurements are minimally

disturbing, their time-ordering affects the outcome of the measurement sequence for quantum systems.

3.1 Introduction

A fundamental open question in physics is the role of time in quantum mechanics (19, 20, 30). While observables such as position and momentum are represented by operators, time in the Schrödinger equation appears only as ordinary number-parameter, just as in classical mechanics (20). In view of relativistic theories of physics which famously treat time and space on equal footing, this distinction is problematic. In fact, while Heisenberg's uncertainty principle between energy and time appears to suggest that a time operator conjugate to the total energy operator exists, attempts to create such an operator lead to contradictions (21). Similar issues confound attempts to create an observable for the time it takes for a particle to tunnel through a potential barrier, or even the time of arrival of a particle at a detector (7, 31, 32). Another issue is that it is widely believed that information is conserved in quantum physics. This follows from the unitary time-reversible evolution in the Schrödinger equation. Yet, it is possible that the passage and direction of time might be set and discerned only by sequences of 'events' (33). The only recordable 'events' in quantum mechanics are the results of measurements (29), which themselves are not unitary, time-reversible, or conserve information. In this work, we experimentally investigate sequential measurements on quantum systems in order to gain insight into the role that time plays in the theory.

Specifically, we test whether the results (i.e., 'events') of a sequence of measurements depends on the order in time in which they are performed. In both classical and quantum physics, measurements may induce a disturbance (6). Since the disturbance will affect subsequent measurements, the results of sequences of measurements may be dependent on the order in time (i.e., time-ordering) in which they are performed. In quantum physics, this is particularly apparent for sequences of incompatible observables, those that do not commute. However, this disturbance can be arbitrarily reduced, at the expense of information gain per measurement trial. Such minimally disturbing measurements are often referred to as weak measurement. Given that the disturbance

was the source of the time-ordering dependence, one might expect the time-ordering dependence to disappear for such minimally disturbing measurements.

Minimally-disturbing measurements have proven useful for probing quantum systems (4, 5, 15, 34, 35, 36). They have also recently attracted wide interest as a signal amplification technique and for studying paradoxes in quantum physics (15, 37) (for a review, please see (38)). Sequences of weak measurements and continuous weak measurements have also become important tools for exploring features of quantum mechanics that are impossible to study with conventional methods (39, 40, 41). It has been shown that the joint result of a sequence of weak measurements is invariant to time-orderings in classical physics (6, 22), as expected. Particularly relevant for this work, in quantum physics, sequences of weak measurements can be used for simultaneously measuring incompatible observables (5, 42). Although such weak measurements are minimally disturbing, the results are, surprisingly, predicted to vary with time-ordering, in contrast to the classical case.

We seek to experimentally demonstrate these time-ordering effects and investigate why they occur. We use the polarization of photons as our quantum system, and thus weakly measure incompatible observables, such as horizontal and diagonal polarization projectors. We show that for such observables the order in which weak measurements are made matters. We study this effect for sequences of two and three measurements. We also probe the role of quantum coherence in time-ordering by testing the case in which the state is incoherently polarized.

3.2 Theory

The notion of weak measurement naturally emerges from a general model for measurement known as the von Neumann or indirect measurement model (28). Almost all measurements, classical or quantum, fit within this model. In it, a measurement apparatus interacts with the measured system, therein disturbing it (6). Since the interaction creates the disturbance, an obvious method to reduce the latter is to weakly couple the measurement apparatus to the measured system. The measurement apparatus consists of a pointer $\mathcal{P}$ whose momentum $\boldsymbol{p}$ is coupled to measured observable $\boldsymbol{A}$

on the measured system $\mathcal{S}$ via unitary interaction

$$\boldsymbol{U} = \exp\left(-ig\boldsymbol{A}\otimes\boldsymbol{p}\right) = \sum_a |a\rangle\langle a| \otimes \boldsymbol{T}(ga). \tag{3.1}$$

Here, g denotes the interaction strength. On the right-side, we have rewritten $\boldsymbol{U}$ in the eigenbasis of $\boldsymbol{A}$, where $|a\rangle$ is an eigenstate of $\boldsymbol{A}$ with eigenvalue a. The translation operator, $\boldsymbol{T}(ga) = \exp(-iga\boldsymbol{p})$, shifts the pointer's position, i.e., $\boldsymbol{T}(ga)\,|x\rangle = |x - ga\rangle$, where $|x\rangle$ is a position eigenstate. If the initial pointer state $|\psi\rangle$ has a position width $\sigma < ga$ then the post-interaction pointer position $x = ga$ unambiguously indicates the measurement result a. This is the case commonly found in conventional measurements, such as a polarizing beam splitter (PBS), or measuring the spin of a silver atom with a Stern-Gerlach apparatus. With this explicit model, one can consider reducing the interaction strength so that $ga \ll \sigma$. Since the pointer now extends over multiple indicator marks, ga (where a is a particular value in the spectrum of $\boldsymbol{A}$), a single trial's measurement result will be ambiguous. However, averaged over many trials the mean measurement result is proportional to the average pointer position, $g\langle\boldsymbol{A}\rangle = \langle\boldsymbol{x}\rangle$, regardless of measurement strength. It is this average result, the expectation value of $\boldsymbol{A}$, that we will study.

To extend the above formalism to include a sequence of measurements $\boldsymbol{A}_N \cdots \boldsymbol{A}_2\boldsymbol{A}_1$, one composes a product of unitaries $\boldsymbol{U}_N \ldots \boldsymbol{U}_2\boldsymbol{U}_1$, where $\boldsymbol{U}_i$ takes the form of Eq. 3.1. Here, $\boldsymbol{A}_1$ is the first observable measured and $\boldsymbol{A}_N$ is the last. Thus, each observable $\boldsymbol{A}_i$ in the sequence of measurements is independently coupled to a distinct pointer $\mathcal{P}_i$ with state $|\psi(x_i)\rangle$. The final result of the measurement is the expectation value $\langle \boldsymbol{x}_1^{(\boldsymbol{A}_1)}\boldsymbol{x}_2^{(\boldsymbol{A}_2)}\cdots\boldsymbol{x}_N^{(\boldsymbol{A}_N)}\rangle/g^N$, where the superscript $(\boldsymbol{A}_i)$ is the observable to which the $\mathcal{P}_i$ pointer is coupled to (43). With this, we can consider the effect of reducing the interaction strength and testing different time-orderings of the measurements. For example, how does $\langle \boldsymbol{x}_1^{(\boldsymbol{A}_1)}\boldsymbol{x}_2^{(\boldsymbol{A}_2)}\rangle$ compare to $\langle \boldsymbol{x}_1^{(\boldsymbol{A}_2)}\boldsymbol{x}_2^{(\boldsymbol{A}_1)}\rangle$?

3.3 Experiment

The experimental setup is shown in Fig. 3.1. It is technically challenging to use spatially distinct systems as our pointer $\mathcal{P}$ and measured system $\mathcal{S}$. Instead, we use distinct degrees of freedom of a photon. The measured system $\mathcal{S}$ is the polarization degree of

freedom, whereas the pointer $\mathcal{P}$ is the photon's transverse position. Because the system is linear, the photons do not interact, instead propagating independently through the sequence of measurements, the spatial distribution for which satisfies Maxwell's equations (44). Therefore the results obtained from single photons would be the same as that obtained for a coherent state. As such we use a HeNe laser at 633 nm as a source of photons. This is followed by a polarizing beam splitter (PBS) and a rotatable half-wave plate (HWP), set at an angle θ, prepares an ensemble of identically polarized photons as our system input state, $|\theta\rangle = \cos(2\theta)|H\rangle + \sin(2\theta)|V\rangle$.

The coupling between $\mathcal{S}$ and $\mathcal{P}$ can be accomplished using polarization dependent walk-off in birefringent crystals. The walk-off transversely shifts the extraordinary polarized photons by g relative to ordinary polarized photons. A sequence of two weak measurements requires two independent pointers, which we take as the two transverse spatial degrees of freedom, x and y, of the photon, each with same initial wavefunction: $\langle x, y|\psi_x, \psi_y\rangle = \psi(x)\psi(y) = (2\pi\sigma^2)^{-\frac{1}{2}}\exp\left(-(x^2+y^2)/4\sigma^2\right)$. A beam expander magnifies the transverse Gaussian mode to set the width to $\sigma = 600\,\mu\text{m}$. With these two pointers, the measurement coupling is implemented with two walk-off crystals, one of which displaces the horizontal polarization along the x-axis, followed by an identical crystal rotated by 90° so that it shifts the vertical polarization along the y-axis. Both crystals impart a shift of $g = 160\,\mu\text{m}$ ensuring that we are in the weak measurement regime, $g/\sigma \approx 0.25$.

In order to change which observable each crystal implements we add waveplates that effectively rotate the basis of the measurement. With this, we use either the $\boldsymbol{x}_1 = \boldsymbol{x}$ or $\boldsymbol{x}_2 = \boldsymbol{y}$ positions to read out measurements of the $|I\rangle\langle I| = \boldsymbol{\pi}_I$, and $|J\rangle\langle J| = \boldsymbol{\pi}_J$ polarization projectors, depending on the ordering. We demonstrate the effect of time-ordering on weak measurements by comparing the result of a sequential measurement $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_I)}\boldsymbol{y}^{(\boldsymbol{\pi}_J)}\rangle$, where $\boldsymbol{A}_1 = \boldsymbol{\pi}_I$ and $\boldsymbol{A}_2 = \boldsymbol{\pi}_J$, with the reversed sequence $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_J)}\boldsymbol{y}^{(\boldsymbol{\pi}_I)}\rangle$, where $\boldsymbol{A}_1 = \boldsymbol{\pi}_J$ and $\boldsymbol{A}_2 = \boldsymbol{\pi}_I$. The expectation values are evaluated with respect to the evolved initial separable polarization and pointer state, $\boldsymbol{U}_2\boldsymbol{U}_1|\theta\rangle \otimes |\psi\rangle$, where the $\boldsymbol{U}_i$ take the form of Eqn. 1 for the respective operators $\boldsymbol{A}_i$. Experimentally the expectation values $\langle \boldsymbol{xy}\rangle$ of the photon transverse two-dimensional distribution are found by imaging onto a CMOS camera.

We begin the experiment by placing the first HWP in Fig. 3.1(b) at 0° and the second at 67.5° so that the two crystals implement a measurement of $\boldsymbol{\pi}_H$ followed by

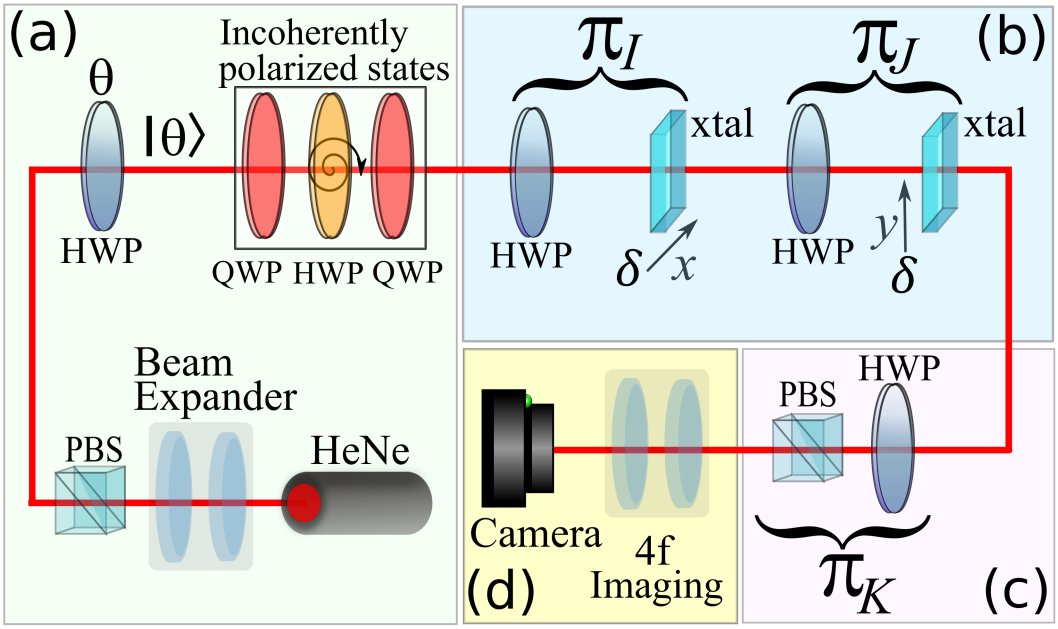

Figure 3.1: Setup to measure the effect of time-ordering of weak measurements. (a). State preparation: An attenuated HeNe laser at 633 nm with a coherence length of approximately 30 cm provides a source of photons. A beam expander is used to decrease the interaction strength by increasing the width σ of the pointer, i.e., the photon transverse distribution. The input system polarization state $|\theta\rangle$ is prepared by a polarizing beam splitter (PBS) and a half-wave plate (HWP). Using quarter-wave plates (QWP) and a HWP (QWP at 45° $\rightarrow$ HWP $\rightarrow$ QWP at 90°) incoherently polarized light can be generated. The HWP is attached to a motor that spins at a rate faster than the collection time of the camera. (b) Weak Measurements: Each weak measurement is implemented by a HWP followed by a walk-off crystal (xtal). The first effects the $\boldsymbol{\pi}_I$ projector by shifting $|I\rangle$ polarized light by $g < \sigma$ in the x direction. Likewise, the second implements the $\boldsymbol{\pi}_J$ projector by shifting $|J\rangle$ polarized light by g in the y direction. (c) Strong measurement: The combination of a HWP and a PBS realizes the third measurement of $\boldsymbol{\pi}_H$ when desired, and is taken out of the setup otherwise. (d) A $4f$ system images the shifted beam onto a CMOS camera.

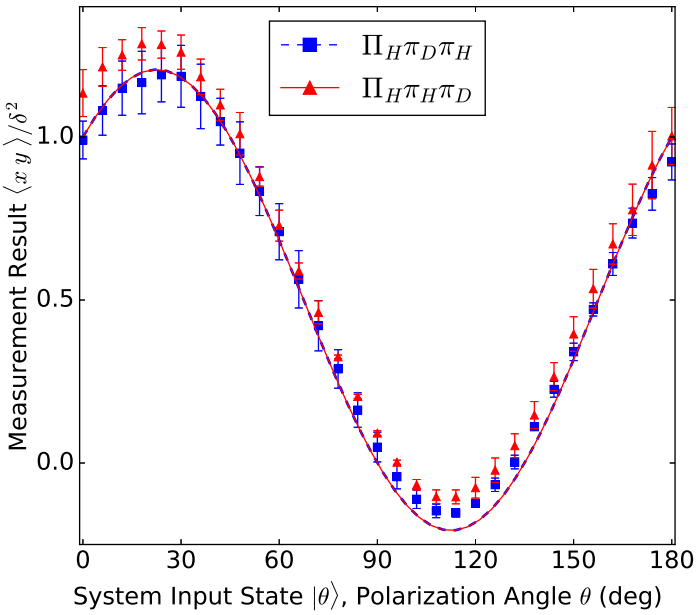

Figure 3.2: The measurement result $\langle \boldsymbol{xy} \rangle / g^2$ for a sequence of two measurements. In one, the sequence $\boldsymbol{\pi}_H \boldsymbol{\pi}_D$ is measured (red triangles), and in the other, $\boldsymbol{\pi}_D \boldsymbol{\pi}_H$ is measured (blue squares). Since the points agree within error, the plot shows that the results do not depend on the order in which the measurements are performed. The red solid and blue dashed line are the respective theoretical curves calculated by explicitly evaluating $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_D)} \boldsymbol{y}^{(\boldsymbol{\pi}_H)} \rangle$ and $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_H)} \boldsymbol{y}^{(\boldsymbol{\pi}_D)} \rangle$. Error bars are the standard error obtained by averaging over four experimental runs. Imperfections in the HWP birefringence likely introduces the differences between the experimental points and the theoretical curve, as they can create systematic errors not only when preparing the input polarization state $|\theta\rangle$, but also when aligning the walk-off crystals.

$\boldsymbol{\pi}_D$, two incompatible observables. By switching the first HWP to 22.5° and leaving the second as it is, the crystals now implement the reverse sequence, $\boldsymbol{\pi}_D$ followed by $\boldsymbol{\pi}_H$. In both cases, we record the joint result of the sequence $\langle \boldsymbol{xy} \rangle$, and compare the experimental results to the theoretical results calculated by evaluating the expectation values $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_D)} \boldsymbol{y}^{(\boldsymbol{\pi}_H)} \rangle$ and $\langle \boldsymbol{x}^{(\boldsymbol{\pi}_H)} \boldsymbol{y}^{(\boldsymbol{\pi}_D)} \rangle$. In Fig. 3.2 we plot this joint result as a function of the input system state angle, θ. The two orderings agree within errors. Thus, as expected since the measurement disturbance is now minimized, the joint result does not depend on the time-ordering of the measurements.

So far, nothing surprising has been revealed: when the measurement disturbance is

minimized, the result of a sequence of quantum measurements is indeed time-ordering invariant. However, there are fundamental phenomena in quantum physics that only appear in sequences of three or more measurements. An example is the violation of the Leggett-Garg inequality (45, 46, 47, 48, 49, 50). Hence, we extend the sequence to three measurements, where the third measurement is a conventional (i.e., 'strong') measurement of $\mathbf{\Pi}_K = |K\rangle\langle K|$ as implemented by a HWP and PBS (here the capital pi indicates a strong measurement). Our goal is again to test the role of measurement-order when the disturbance is not a factor. Since this added conventional measurement will substantially disturb the system and, thus, any subsequent measurements, we always perform it last.

The final joint-result of the three measurement sequence is

$$\begin{aligned}\langle \boldsymbol{\pi}_K \boldsymbol{\pi}_J \boldsymbol{\pi}_I \rangle = \frac{1}{g^2}\langle \mathbf{\Pi}_K \boldsymbol{x}^{(\boldsymbol{\pi}_J)} \boldsymbol{y}^{(\boldsymbol{\pi}_I)} \rangle = \\ \frac{1}{g^2}\int xy \mathrm{Prob}(x,y,K)\mathrm{d}x\mathrm{d}y,\end{aligned} \tag{3.2}$$

where $\mathrm{Prob}(x,y,K) = |\langle K,x,y|\boldsymbol{U}_2\boldsymbol{U}_1|\theta\rangle \otimes |\psi\rangle|^2$ is the probability that a given input photon with polarization $|\theta\rangle$ and spatial distribution $|\psi\rangle$ is transmitted through the PBS and is detected at transverse coordinate (x,y) on the camera. Since the last is a strong measurement, in Eq. 3.2 we directly evaluate the measurement outcome $\langle \mathbf{\Pi}_K \rangle$, rather than use the von Neumann formalism. We set $K = H$ for all of the following measurements. Figure 3.3(a) shows experimental results for two orderings of a sequence of three measurements, $\mathbf{\Pi}_H \boldsymbol{\pi}_D \boldsymbol{\pi}_H$ and $\mathbf{\Pi}_H \boldsymbol{\pi}_H \boldsymbol{\pi}_D$. The joint result of one ordering substantially disagrees with the other ordering. Theoretically, the difference is maximum at θ =45° and θ =135°. The nearest experimental points, 48° and 138°, differ by 0.4 ± 0.1 and 0.7 ± 0.1, respectively. Strikingly, quantum physics is not invariant to the time-ordering even though there is no obvious physical mechanism, such as measurement-induced disturbance, for this invariance.

Perhaps, while minimized, the residual disturbance still causes this time-ordering asymmetry. But, if this is the case, how could it possibly manifest in three measurements but not two? Specifically, the disturbance would necessarily need to propagate through the first two measurements to reach the third. Hence, one would expect two measurements would be time-asymmetric as well. A recent theoretical investigation

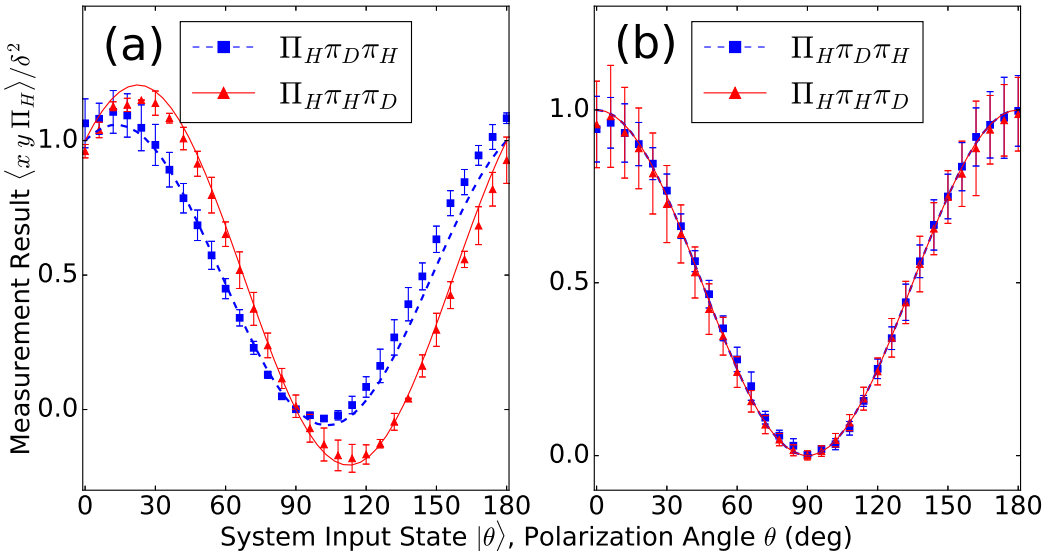

Figure 3.3: The measurement result $\langle \boldsymbol{xy\Pi}_H \rangle / g^2$ for a sequence of three measurements. In one, the sequence $\boldsymbol{\Pi}_H\boldsymbol{\pi}_D\boldsymbol{\pi}_H$ is performed (blue squares), and in the other, $\boldsymbol{\Pi}_H\boldsymbol{\pi}_H\boldsymbol{\pi}_D$ is performed (red triangles). The blue dashed and red solid lines are the respective theoretical curves calculated directly from Eq. 2. (a) The input system state is $|\theta\rangle$. The two distinct curves show that the order in which the measurements are made changes the measurement result. (b) An incoherently polarized system state $\rho(\theta)$ is used instead. Now the measurement result does not depend on the ordering of the measurements.

offers some mathematical insight (23). In it, our expectation value of N pointer positions $\langle \boldsymbol{x}_1^{(\boldsymbol{A}_1)} \boldsymbol{x}_2^{(\boldsymbol{A}_2)} \cdots \boldsymbol{x}_N^{(\boldsymbol{A}_N)} \rangle$ was found to be proportional to the recursively nested anti-commutator structure $\{\{\ldots\{\{\boldsymbol{A}_N, \boldsymbol{A}_{N-1}\}, \boldsymbol{A}_{N-2}\}, \ldots\}, \boldsymbol{A}_1\}$. While all the anti-commutators are symmetric under interchange of their two arguments, the only anti-commutator that is invariant under interchange of the measurements is the innermost one, which is non-nested. As such, it, and thus the expectation value, is invariant to the ordering of the last two measurements, $\boldsymbol{A}_N$ and $\boldsymbol{A}_{N-1}$. This provides a succinct mathematical description of why a sequence of two measurements always exhibits ordering invariance; the entire sequence is the last two measurements. The result of a sequence of three or more measurements will be invariant solely to the ordering of the last two measurements.

3.4 Discussion

While we have a mathematical prescription for when the time-ordering dependence of minimally invasive measurements matters, a physical intuition is still absent. That is, a mechanism (for example, back-action, decoherence, etc.) with which one can a-priori determine whether or not time-ordering will matter in a sequence of weak measurements is lacking. We attempt to experimentally shed light onto this problem. A Distinguishing feature of weak measurements is that they preserve the coherence of the measured system and the coherence of the pointer. This can allow a disturbance to propagate in unexpected ways, as in measurement back-action (51, 52). Since in a classical system this coherence would be absent, we test what happens to the time-ordering dependence as we decrease the coherence in the H-V basis of the initial system state.

To generate these reduced coherence (i.e., mixed) states, we send our polarized input state $|\theta\rangle$ through a rapidly spinning HWP that is sandwiched between two quarter-waveplates (QWP). Since this spinning is faster than the imaging camera acquisition time, the resulting state is effectively mixed: $\rho(\theta) = \sin^2(\theta) |H\rangle \langle H| + \cos^2(\theta) |V\rangle \langle V|$ (for more detail on how we generate mixed states, see the chapter 8 supplementary information, and mixed states themselves will be further discussed in the next chapter). We test the same pair of three-measurement sequences with these input states. The experiment results, shown in Fig. 3.3(b), show that the joint result of the sequence does

not depend on the order in which the observables were measured. This is strikingly different from the case found in Fig. 3.3(a); quantum coherence indeed appears to play an important role in time-ordering symmetry. However, it is not clear why this should be the case. As the theoretical curves calculated from Eqn. 2 suggest, the result is expected, but the underlying physical structure that brings about the ordering symmetry is not obvious, and warrants further study.

To summarize our experimental findings, in the case of two measurements, the order in which weak measurements are made does not impact the end result. But, in the case of three or more measurements, the order of the measurements matter. In short, we have shown that reducing the disturbance induced by measurements does not restore the time symmetry of quantum evolution, as exhibited by the Schrödinger equation. Our findings confirm a recent theoretical proposal on the time-ordering sequential weak measurements (6). While an intuitive mechanism for this time-ordering invariance is still not clear, we have shown that coherence plays a role in the time-ordering phenomena, but how it does so is not understood. We expect these results will motivate the development of closely related areas, such as whether different times of system can be considered separate Hilbert spaces (30), and how cause and effect can be identified in quantum systems (53).

3.5 Supplementary Material

Birefringence, Walk-off, and Displacement Crystal Alignment

When light enters a material, it travels at a speed, $v = c/n$, where n is the refractive index of the material. Certain materials display a polarization dependent refractive index, a property called birefringence. When a laser beam propagates through a birefringent crystals, it is not necessarily the case that the direction of the Poynting vector, $\vec{S}$, is co-linear to the propagation direction of the wave vector $\vec{k}$. That is, due to the birefringence, the intensity distribution may deviate from the direction defined by the wave vector. The angle between them, θ_w, is known as the walk-off and is determined by the index contrast and orientation of the crystal axis. This is called spatial walk-off, and is the property we exploit to separate our beams above.

As walk-off is a birefringence effect, it only affects one polarization. The polarization which will shift is determined by the optic axis, (labeled O in Fig. 3.4a)), which defines

a plane in the crystal. This optic axis is present due to the lattice arrangement of the constituent atoms, which leads to an anisotropy. Light whose polarization vector is along the plane (H in this case) experiences no walk off. It follows the ordinary laws of refraction, and is as such often called the *ordinary* polarization. Light not along the optical plane will undergo walk-off due to birefringence. This polarization would not obey the laws of refraction and is as such termed *extraordinary*. Another way to understand this phenomena is that each polarization component must independently satisfy Snell's law at the interface. However, as the refractive index is orientation dependent, both polarization components see a different polarization, and necessarily must have a different angle of refraction. The two split by an amount g, which is determined by θ_w and the length of the crystal, as shown in Fig. 3.4b). Assuming that the angle between the propagation direction and the crystal axis is θ, then the walk-off angle is

$$\theta_w = -\frac{1}{n_e}\frac{\partial n_e}{\partial \theta}, \tag{3.3}$$

where n_e is the extraordinary refractive index.

In aligning the crystal, first one must first find the optical axis. This can be done by first treating the crystal as a wave-plate. Waveplates in essence act in the same way as a birefringent crystal; by introducing a phase between orthogonal polarization, the state of the polarization is changed. However, if the polarization is aligned already to one of the optical axes, then it will (up to an overall phase) be unchanged. As such, we use a PBS to send $|H\rangle$ polarized light into the crystal. After the crystal we project onto $|V\rangle$ polarized light with a PBS. We rotate the crystal until we see that no light is passing through the PBS (one could of course project onto $|H\rangle$, but it is always easier to find a minimum in intensity rather than a maximum). When there is no light, we know that the optical axis is correctly aligned along the $|H\rangle$ polarization, as it is not transforming the state.

As the two polarizations have different path lengths, a phase offset is introduced between them. This phase difference needs to be compensated for in weak measurement. If one sends $|D\rangle$ polarized light into a thin birefringent crystal, the phase offset will result in a state $(|H\rangle + e^{i\phi}|V\rangle)/\sqrt{2} \neq |D\rangle$, which is not desired in a weak measurement. This can be compensated for by tilting the crystal forward by an angle Φ_c, as shown in Fig. 3.4c). The correct amount of tilt can be found by projecting onto $|A\rangle$ and imaging onto a camera. When the tilt $\Phi = \Phi_c$, the overlapping parts of the $|H\rangle$ polarized beam

and $|V\rangle$ polarized beam will be $|D\rangle$ polarized. Projecting onto $|A\rangle$ leaves the two $|H\rangle$ and $|V\rangle$ lobes with the $|D\rangle$ cut out, shown in Fig. 3.4d).

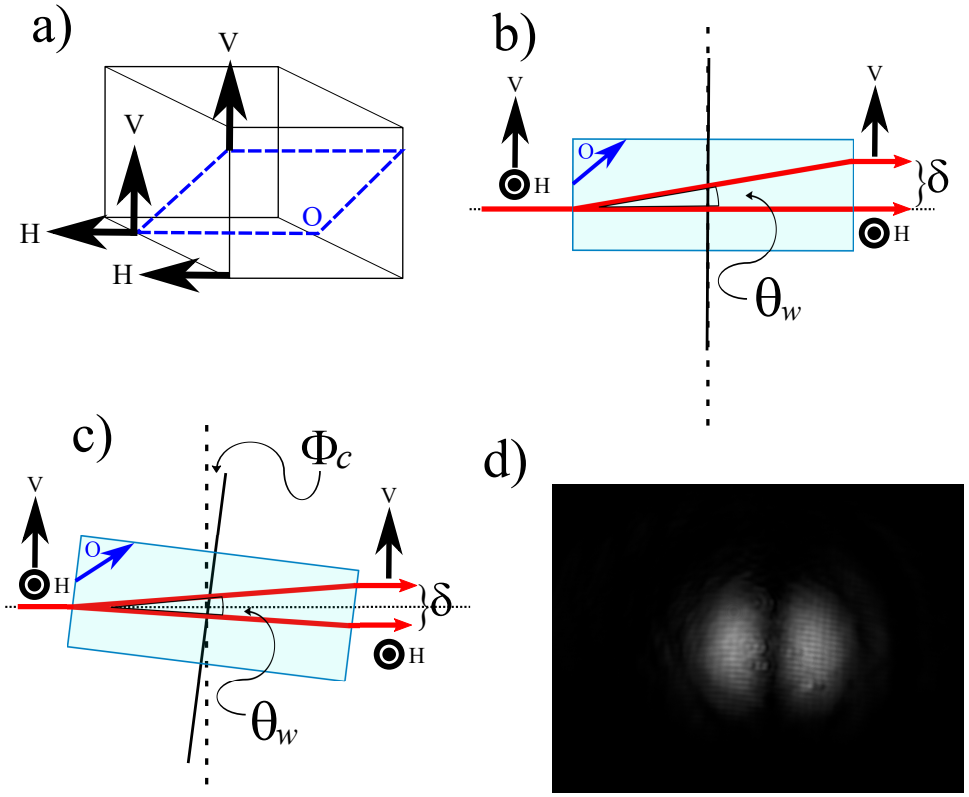

Figure 3.4: a) Walk-off in birefringent crystals. The polarization that is shifted is determined by the optical axis, labeled O. Light whose polarization is contained within the optical plane obeys the regular laws of refraction. On the other hand, polarizations with vectorial components outside the optical plane will undergo walk-off. **b)** Birefringent crystals spatially separate orthogonal polarization through spatial walk-off. The amount of separation, g, is determined by the walk-off angle θ_w, and the length of the crystal. This walk-off produces a phase offset between the two paths due to the difference in path length. **c)** This path length difference can be rectified by tilting the crystal forward by some compensation angle Φ_c. Regular refraction from Snell's law refracts the Horizontal path, making the path lengths equal. **d)** Φ_c can be found by projecting onto $|A\rangle$. When the tilt is correct, the state between the two lobes will be $|D\rangle$, therefore leaving a dark spot between the $|H\rangle$ and $|V\rangle$ lobes.

Auxiliary Results

Presented here are auxiliary results that were obtained during the course of the above experiment. These results concern weak measurements but are broader in scope. They concern the structure of weak measurement expectation values, and show a surprising relationship between operators in the system Hilbert space, and operators in the pointer Hilbert space. The three results are

$$\langle\psi|\prod_{i}^{N}\boldsymbol{x}_i^{\boldsymbol{A}_i}|\psi\rangle \propto \{\{\dots\{\{\boldsymbol{A}_1,\boldsymbol{A}_2\},\boldsymbol{A}_3\},\dots\},\boldsymbol{A}_N\}.$$

$$\langle\boldsymbol{x}_N^{\boldsymbol{A}_N}\prod_{i=1}^{N-1}\boldsymbol{p}_i^{\boldsymbol{A}_i}\rangle \propto [[\dots[[\boldsymbol{A}_N,\boldsymbol{A}_{N-1}],\boldsymbol{A}_{N-2}]\dots],\boldsymbol{A}_1]$$

$$\langle\boldsymbol{p}_N^{\boldsymbol{A}_N}\dots\rangle = 0$$

The first result is the one first shown in (23), albeit without proof. This relates the joint position expectation value to a nested anti-commutator sequence, as discussed in the main text. The second (and third) result is, as far as we know, not presented anywhere. It relates a similar joint expectation value, expect one looks at only momentum correlations for all but the last pointer. We see a similar expression, but this time replaced with anti-commutators. Notice that, unlike the previous expression, one cannot in general permute the order of operators as the anti-commutator is not symmetric to swapping of its argument. This is not true if $[\boldsymbol{A}_N,\boldsymbol{A}_{N-1}]=0$, as everything commutes with '0'. Note a more general result exists that combines the two. For a general sequence of both $\boldsymbol{x}$'s and $\boldsymbol{p}$'s, the result involves nested commutators and anti-commutators.

While these results are interesting in and of themselves, by far the most surprising one is the third one. This result suggests that if one measures the joint value where we look at the momentum of the last pointer, the result is zero, regardless of the sequence of $\boldsymbol{x}$'s and $\boldsymbol{p}$'s that come before. The significance of this is not clear, however it appears to be related to the fact that the last measurement in a sequence can always be made to be strong.

We conclude this section by proving the results above for the curious.

Proof for $\langle \prod_{i=1}^{N} \boldsymbol{x}_i^{\boldsymbol{A}_i} \rangle$

Show that

$$\langle \psi | \, \Pi_i^N \boldsymbol{x}_i^{\boldsymbol{A}_i} \, | \psi \rangle \propto \{\{\dots\{\{\boldsymbol{A}_1, \boldsymbol{A}_2\}, \boldsymbol{A}_3\}, \dots\}, \boldsymbol{A}_N\}.$$

proof

Proceed by induction. Suppose we wish to measure $\langle \boldsymbol{x}_1^{\boldsymbol{A}} \boldsymbol{x}_2^{\boldsymbol{B}} \rangle$. As in the main text, this notation indicates that the observable $\boldsymbol{B}$ is being measured second. We introduce a spectral decomposition for $\boldsymbol{A} = \sum_i a_i \boldsymbol{\pi}_i^{\boldsymbol{A}}$, and a similar one for $\boldsymbol{B}$. The expectation value is taken with respect to the state $\boldsymbol{U} \, |\Psi\rangle$ where $|\Psi\rangle = |\psi(x)\rangle \otimes |\psi(y)\rangle$, and $\boldsymbol{U} = \exp(ig\boldsymbol{B} \otimes \boldsymbol{p}_2) \exp(ig\boldsymbol{A} \otimes \boldsymbol{p}_1)$. Using the spectral expansion of $\boldsymbol{A}$ and $\boldsymbol{B}$ we get

$$\boldsymbol{U} = \sum_{ij} e^{iga_i \boldsymbol{p}_1} e^{igb_j \boldsymbol{p}_2} \boldsymbol{\pi}_j^{\boldsymbol{B}} \boldsymbol{\pi}_i^{\boldsymbol{A}} \tag{3.4}$$

The exponential operators are translation operators on the pointer states $|\Psi\rangle$. We apply $\boldsymbol{U}$ to our state to get

$$\boldsymbol{U} \, |\Psi\rangle = \sum_{ij} |\psi(x - ga_i)\rangle \, |\psi(y - gb_j)\rangle \, \boldsymbol{\pi}_j^{\boldsymbol{B}} \boldsymbol{\pi}_i^{\boldsymbol{A}}. \tag{3.5}$$

Evaluating the expectation value $\langle \boldsymbol{x}_1^{\boldsymbol{A}} \boldsymbol{x}_2^{\boldsymbol{B}} \rangle = \langle \Psi | \boldsymbol{U}^\dagger \boldsymbol{x}_1^{\boldsymbol{A}} \boldsymbol{x}_2^{\boldsymbol{B}} \boldsymbol{U} | \Psi \rangle$, to get the expression

$$\begin{aligned} \langle \boldsymbol{x}_1^{\boldsymbol{A}} \boldsymbol{x}_2^{\boldsymbol{B}} \rangle = \sum_{iji'j'} \langle \psi(x - ga_{i'}) | \, \boldsymbol{x}_1^{\boldsymbol{A}} \, |\psi(x - ga_i)\rangle \, \langle \psi(y - gb_{j'}) | \, \boldsymbol{x}_2^{\boldsymbol{B}} \, |\psi(y - gb_j)\rangle \times \\ \boldsymbol{\pi}_{i'}^{\boldsymbol{A}} \boldsymbol{\pi}_{j'}^{\boldsymbol{B}} \boldsymbol{\pi}_j^{\boldsymbol{B}} \boldsymbol{\pi}_i^{\boldsymbol{A}}. \end{aligned} \tag{3.6}$$

As $\boldsymbol{\pi}_{j'}^{\boldsymbol{B}}$ and $\boldsymbol{\pi}_j^{\boldsymbol{B}}$ are orthogonal projectors, their product is a Kronecker delta $\boldsymbol{\pi}_{j'}^{\boldsymbol{B}} \boldsymbol{\pi}_j^{\boldsymbol{B}} = \delta_{j'j} \boldsymbol{\pi}_j^{\boldsymbol{B}}$. As such we can eliminate one index in the above sum.

$$\begin{aligned} \langle \boldsymbol{x}_1^{\boldsymbol{A}} \boldsymbol{x}_2^{\boldsymbol{B}} \rangle = \sum_{iji'} \langle \psi(x - ga_{i'}) | \, \boldsymbol{x}_1^{\boldsymbol{A}} \, |\psi(x - ga_i)\rangle \, \langle \psi(y - gb_j) | \, \boldsymbol{x}_2^{\boldsymbol{B}} \, |\psi(y - gb_j)\rangle \times \\ \boldsymbol{\pi}_{i'}^{\boldsymbol{A}} \boldsymbol{\pi}_j^{\boldsymbol{B}} \boldsymbol{\pi}_i^{\boldsymbol{A}} \end{aligned} \tag{3.7}$$

Now we want to evaluate the two expectation values in the above sum. For Gaussian pointers centered initially at the origin, the expectation values evaluate as

$$\langle \psi(y - gb_j) | \, \boldsymbol{x}_2^{\boldsymbol{B}} \, |\psi(y - gb_j)\rangle = gb_j$$

and

$$\langle\psi(x-ga_{i'})|\,\boldsymbol{x}_1^{\boldsymbol{A}}\,|\psi(x-ga_i)\rangle = \frac{g(a_{i'}+a_i)}{2}e^{\frac{g^2(a_{i'}-b_i)^2}{8\sigma^2}} \to \frac{g(a_{i'}+a_i)}{2} \quad \text{as } g\to 0.$$

Notice that the first expectation value implies that the strength of the last measurement ($\boldsymbol{B}$) does not matter, as the expectation value will always be gb_j. This result is true for the last measurement in any sequence. From these expectation values we can write the sum as

$$\frac{g^2}{2}\sum_{ii'j}(a_{i'}+a_i)b_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}}. \tag{3.8}$$

We can split the above sum into two

$$\sum_{ii'j}(a_{i'}+a_i)b_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}} = \sum_{ii'j}a'_ib_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}} + \sum_{ii'j}a_ib_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}}. \tag{3.9}$$

Focusing on the first term and using the spectral form of $\boldsymbol{A}$ and $\boldsymbol{B}$, and $\sum_i\boldsymbol{\pi}_i=\mathbf{1}$, we can simplify this further.

$$\sum_{ii'j}a'_ib_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}} = \sum_i a'_i\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\sum_j b_j\boldsymbol{\pi}_j^{\boldsymbol{B}}\sum_i\boldsymbol{\pi}_i^{\boldsymbol{A}} = \boldsymbol{AB}\mathbf{1} = \boldsymbol{AB}.$$

Using the same algebra, we can show that the second term in Eqn. 3.9 equates to $\boldsymbol{BA}$, and so we get the result that

$$\frac{g^2}{2}\sum_{ii'j}(a_{i'}+a_i)b_j\boldsymbol{\pi}_{i'}^{\boldsymbol{A}}\boldsymbol{\pi}_j^{\boldsymbol{B}}\boldsymbol{\pi}_i^{\boldsymbol{A}} = \boldsymbol{AB}+\boldsymbol{BA} = \{\boldsymbol{A},\boldsymbol{B}\}, \tag{3.10}$$

This proves the base case that

$$\langle\Psi|\,\boldsymbol{x}_1^{\boldsymbol{A}}\boldsymbol{x}_2^{\boldsymbol{B}}\,|\Psi\rangle = \frac{g^2}{2}\{\boldsymbol{A},\boldsymbol{B}\}.$$

For the inductive case, we suppose that for $N=K-1$ operators, the following holds

$$\langle\Psi|\,\Pi_{i=2}^K\boldsymbol{x}_i^{\boldsymbol{A}_i}\,|\Psi\rangle \propto \{\{\dots\{\{\boldsymbol{A}_K,\boldsymbol{A}_{K-1}\},\boldsymbol{A}_{K-2}\},\dots\},\boldsymbol{A}_2\} = \boldsymbol{C} \tag{3.11}$$

where again, the subscript denotes the observables position in the sequence. Then, for $N=K$ operators (note we are adding a new operator at the *begining* of the sequence)

$$\langle\Psi|\,\Pi_{i=1}^K\boldsymbol{x}_i^{\boldsymbol{A}_i}\,|\Psi\rangle \propto$$

$$\sum_{i_1 i_2 \ldots i_K} \sum_{i'_1 i'_2 \ldots i'_{K-1}} a^K_{i_K}(a^{K-1}_{i_{K-1}} + a^{K-1}_{i'_{K-1}}) \ldots (a^1_{i_2} + a^1_{i'_2})(a^1_{i_1} + a^1_{i'_1}) \times \\ \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i'_1} \boldsymbol{\pi}^{\boldsymbol{A}_2}_{i'_2} \ldots \boldsymbol{\pi}^{\boldsymbol{A}_{K-1}}_{i'_{K-1}} \boldsymbol{\pi}^{\boldsymbol{A}_K}_{i_K} \boldsymbol{\pi}^{\boldsymbol{A}_{K-1}}_{i_{K-1}} \ldots \boldsymbol{\pi}^{\boldsymbol{A}_2}_{i_2} \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i_1} = \tag{3.12}$$

$$\sum_{i_1 i'_1} (a^1_{i_1} + a^1_{i'_1}) \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i'_1} \left(\sum_{i_1 i_2 \ldots i_K} \sum_{i'_1 i'_2 \ldots i'_{K-1}} a^K_{i_K}(a^{K-1}_{i_{K-1}} + a^{K-1}_{i'_{K-1}}) \ldots (a^1_{i_2} + a^1_{i'_2}) \boldsymbol{\pi}^{\boldsymbol{A}_2}_{i'_2} \ldots \boldsymbol{\pi}^{\boldsymbol{A}_{K-1}}_{i'_{K-1}} \boldsymbol{\pi}^{\boldsymbol{A}_K}_{i_K} \boldsymbol{\pi}^{\boldsymbol{A}_{K-1}}_{i_{K-1}} \ldots \boldsymbol{\pi}^{\boldsymbol{A}_2}_{i_2} \right) \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i_1}. \tag{3.13}$$

By inductive hypothesis, the expression inside the brackets is $\boldsymbol{C}$, so we get

$$\sum_{i_1 i'_1} (a^1_{i_1} + a^1_{i'_1}) \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i'_1} \boldsymbol{C} \boldsymbol{\pi}^{\boldsymbol{A}_1}_{i_1} = \boldsymbol{A}_1 \boldsymbol{C} + \boldsymbol{C} \boldsymbol{A}_1 = \{\boldsymbol{C}, \boldsymbol{A}_1\} \tag{3.14}$$

This completes the proof. The proof for

$$\langle \boldsymbol{x}^{\boldsymbol{A}_N}_N \prod_{i=1}^{N-1} \boldsymbol{p}^{\boldsymbol{A}_i}_i \rangle \propto [[\ldots[[\boldsymbol{A}_N, \boldsymbol{A}_{N-1}], \boldsymbol{A}_{N-2}]\ldots], \boldsymbol{A}_1]$$

is nearly identical.

Proof that $\langle p^{\boldsymbol{A}_N}_N \ldots \rangle = 0$

The joint expectation value with the momentum of the last pointer is zero i.e. $\langle p^{\boldsymbol{A}_N}_N \ldots \rangle = 0$, where any combination of x's and p's may be before $p^{\boldsymbol{A}_N}_N$.

proof

$$\langle \boldsymbol{p}^{\boldsymbol{A}_N}_N \ldots \rangle = \sum_{ii' \ldots} \langle \psi(x - g a^N_{i'}) | \boldsymbol{p}^{\boldsymbol{A}_N}_N | \psi(x - g a^N_i) \rangle \ldots\ldots \boldsymbol{\pi}^{\boldsymbol{A}_N}_{i'} \boldsymbol{\pi}^{\boldsymbol{A}_N}_i \cdots = \\ \sum_{i \ldots} \langle \psi(x - g a^N_i) | \boldsymbol{p}^{\boldsymbol{A}_N}_N | \psi(x - g a^N_i) \rangle \ldots\ldots \boldsymbol{\pi}^{\boldsymbol{A}_N}_i \ldots \tag{3.15}$$

But $\langle \psi(x - g a^N_i) | \boldsymbol{p}^{\boldsymbol{A}_N}_N | \psi(x - g a^N_i) \rangle = 0$ because a shift in x is a phase in Fourier space. These phases cancel out (because you get the conjugate from the bra), and so you have the expectation value of an un-shifted Gaussian. This completes the proof.

Notice that if we had post-selected, say on the state $|v\rangle$, this would change the above equation to

$$\langle \boldsymbol{p}^{\boldsymbol{A}_N}_N \ldots \rangle \propto \sum_{ii' \ldots} \langle \psi(x - g a^N_{i'}) | \boldsymbol{p}^{\boldsymbol{A}_N}_N | \psi(x - g a^N_i) \rangle \ldots\ldots \boldsymbol{\pi}^{\boldsymbol{A}_N}_{i'} \boldsymbol{\pi}_v \boldsymbol{\pi}^{\boldsymbol{A}_N}_i \cdots \neq 0 \tag{3.16}$$

since the phases no longer cancel in the integral because the $\boldsymbol{\pi}_{i'}^{\boldsymbol{A}_N}$ and $\boldsymbol{\pi}_{i}^{\boldsymbol{A}_N}$ no longer cancel out and get rid of one of the indicies.

4

Coherence in Sequential Measurements, Part I

Preamble

This chapter introduces the background for the next topic of the thesis which discusses how coherence manifests in sequential measurement. We introduce a generalized quantum state called the density matrix, allowing us to treat statistical mixtures of states. We discuss entanglement and its connection to decoherence in open quantum systems, which leads into a discussion of the propagation of coherence in strong sequential measurements. We find that coherence between pointer states remains even after sequences of non-commuting measurements, a result which is not consistent with the collapse formalism of measurement.

4.1 The Density Matrix

When discussing quantum states, we have so far made the implicit assumption that we *know* that the system is in state $|\psi\rangle$. However, this is not always possible. For example, in an open system (i.e. a system that interacts freely with the surrounding environment), collisions between a quantum system and its surroundings may cause thermal excitation that the observer is unaware of. As another example, consider a single photon source that non-deterministically has probability $1/2$ of being $|H\rangle$ polarized, and probability $1/2$ of being $|V\rangle$ polarized. How does one treat such a state. In what way is this different to writing $|D\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$?

One approach might be to perform calculations via a table. Suppose the goal of an experiment is to measure the expectation value of an observable $\boldsymbol{A}$. The system is known to be in one of the states $\{|\psi_i\rangle\}_{i=1}^{N}$, each with an associated probability of p_i. If the system is in state $|\psi_j\rangle$, then the expectation value is $\langle \boldsymbol{A} \rangle_j = \langle \psi_j | \boldsymbol{A} | \psi_j \rangle$. We can calculate $\langle \boldsymbol{A} \rangle_j$ for all N of the possible states $|\psi_j\rangle \in \{|\psi_i\rangle\}$ and list the results in a table.

State	Expectation Value	Probability of State
$\|\psi_1\rangle$	$\langle \psi_1 \| \boldsymbol{A} \| \psi_1 \rangle$	p_1
$\|\psi_2\rangle$	$\langle \psi_2 \| \boldsymbol{A} \| \psi_2 \rangle$	p_2
$\vdots$	$\vdots$	$\vdots$
$\|\psi_N\rangle$	$\langle \psi_N \| \boldsymbol{A} \| \psi_N \rangle$	p_N

The system is in one of these $|\psi_i\rangle$ states with a particular probability. Therefore, the expectation value is the sum of all the possible results, weighted by the associated probability of that result occurring:

$$\langle \boldsymbol{A} \rangle = \sum_i p_i \langle \psi_i | \boldsymbol{A} | \psi_i \rangle . \tag{4.1}$$

Deriving the Density Matrix

In 1927 John von Neumann and Landau introduced a compact method of performing the above calculation by generalizing the quantum state to allow for statistical mixtures. This generalized state is called the density matrix (27, 54). We will introduce the density matrix in a constructive manner. Starting from Eqn. 4.1, we expand $\boldsymbol{A}$ into an arbitrary basis, $\boldsymbol{A} = \sum_{nm} c_{nm} |n\rangle \langle m|$. As this is not necessarily the eigenbasis of $\boldsymbol{A}$, the expression has non-diagonal terms and the coefficients $c_{nm} = \langle n | \boldsymbol{A} | m \rangle$ may be complex. In this basis Eqn. 4.1 becomes

$$\langle \boldsymbol{A} \rangle = \sum_{nm} \sum_i p_i c_{nm} \langle \psi_i | n \rangle \langle m | \psi_i \rangle ,$$

which can be re-written as

$$\langle \boldsymbol{A} \rangle = \sum_{nm} c_{nm} \langle m | \left(\sum_i p_i |\psi_i\rangle \langle \psi_i | \right) | n \rangle . \tag{4.2}$$

Notice that all of the statistical properties associated with the system been condensed into the operator in the brackets. We write a shorthand for this operator

$$\boldsymbol{\rho} = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| . \tag{4.3}$$

$\boldsymbol{\rho}$ is the *density matrix*[1] and generalizes the concept of the quantum state. Rather than the system being in a single state $|\psi\rangle$, it is now in a mixture (*not* superposition) of possible states, weighted by the probability to be in that state. The density matrix has the following properties:

1. $\boldsymbol{\rho}$ is Hermitian, $\boldsymbol{\rho} = \boldsymbol{\rho}^\dagger$
2. $\boldsymbol{\rho}$ is positive semi-definite, meaning all its eigenvalues are positive or zero
3. $\boldsymbol{\rho}$ is normalized, $\mathrm{Tr}(\boldsymbol{\rho}) = 1$.

The notation $\mathrm{Tr}(\cdot)$ above indicates the matrix trace, and is equal to the sum of the diagonal elements of $\boldsymbol{\rho}$. The trace notation is useful in emphasizing the basis independence of a particular property of the density matrix. Because of this basis independence, the trace can be written as $\mathrm{Tr}(\cdot) = \sum_n \langle n| \cdot |n\rangle$ for any arbitrary basis $\{|n\rangle\}$.

The density matrix formalism introduces the concept of 'mixedness', or *purity* of the state. So far we have largely worked with *pure* states. These are subset of states for which $\boldsymbol{\rho} = |\psi\rangle \langle\psi|$ (i.e. there exists a basis in which the probability of one state is 1, and the rest are zero). Pure states are idempotent, meaning that $\boldsymbol{\rho}^2 = \boldsymbol{\rho}$. The most common test for purity is to calculate $\mathrm{Tr}(\boldsymbol{\rho}^2)$. Because $\mathrm{Tr}(\boldsymbol{\rho})$=1, $\mathrm{Tr}(\boldsymbol{\rho}^2) = 1$ if and only if $\boldsymbol{\rho}$ is pure. In most bases, it will not be clear if $\boldsymbol{\rho}$ is pure or not. The basis independence of Tr provides a simple way of checking the purity. The purity ranges from 1 (pure) to $1/d$, where d is the dimension of the Hilbert space. States for which $\mathrm{Tr}(\boldsymbol{\rho}^2) = 1/d$ are called *totally mixed* states, and they always take the form $\boldsymbol{\rho} = \mathbf{1}/d$, where the bold $\mathbf{1}$ is the identity operator. These represent states with a totally uniform statistical distribution. For this reason totally mixed states are also called the maximum entropy states, as they are the states with the most uncertainty.

[1] Often this is also called the density *operator*. However, it should be noted that $\boldsymbol{\rho}$ is time dependent in the Schrödinger picture, and stationary in the Heisenberg picture. This is how *states* evolve, whereas operators evolve in the opposite way.

4.2 The Partial Trace

Most applications of quantum mechanics deal with states that are not localized to one Hilbert space, but are rather found in a bipartite state belonging to two different Hilbert spaces. Recall that this is how the von Neumann measurement works. Prior to the measurement process, the system and pointer states are independent. The total state can be written as simply the product of the two separate states $|\Psi\rangle_{\mathcal{S}\otimes\mathcal{P}} = |\chi\rangle_{\mathcal{S}} \otimes |\psi\rangle_{\mathcal{P}}$, where $|\chi\rangle_{\mathcal{S}} \in \mathcal{S}$ and $|\psi\rangle_{\mathcal{P}} \in \mathcal{P}$. The fact that $|\Psi\rangle_{\mathcal{S}\otimes\mathcal{P}}$ is a product means there is no correlation between the two states, and they may be treated independently. After the von Neumann interaction the state evolves to $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}} = \boldsymbol{U}\,|\Psi\rangle_{\mathcal{S}\otimes\mathcal{P}}$. In general $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}}$ *cannot* be written as a product state, i.e. $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}} \neq |\chi'\rangle_{\mathcal{S}} \otimes |\psi'\rangle_{\mathcal{P}}$, where $|\chi'\rangle_{\mathcal{S}} \in \mathcal{S}$ and $|\psi'\rangle_{\mathcal{P}} \in \mathcal{P}$. These *non-seperable* states have correlations that cannot be obtained from the subsystems separately (i.e. they are entangled).

Some classical intuition can be had by considering classical probability theory. Given the marginal distributions $P(x)$ and $P(y)$ shown in Fig. 4.1a, the best guess (assuming no additional information), for the joint probability distribution is $P(x, y) = P(x)P(y)$, shown in Fig. 4.1b. The actual joint probability distribution is shown in Fig. 4.1c. The distinguishing feature between Fig. 4.1b and Fig. 4.1c is the lack of *correlation* between x and y in the former. If only the marginal distributions are known, the correlations between them cannot be recovered. Information has been lost in the process.

The correlation between entangled states is fragile, and is lost naturally in bipartite systems when a subsystem interacts with an external environment in a manner which we cannot control. Consider for example a pair (call them A and B) of polarization entangled photons generated from a nonlinear crystal. The total bipartite state of the two photons is written as

$$|\Psi\rangle = (|H\rangle_A\,|V\rangle_B + |V\rangle_A\,|H\rangle_B)/\sqrt{2}.$$

If we measure the B photon (say with a PBS) and learn that it is $|H\rangle_B$, we *know*, without having to perform any additional measurements, that photon A is $|V\rangle_A$, and vice versa. The probability that the B photon is $|H\rangle_B$ is $1/2$, but given that information, we know with probability 1 that the A photon is $|V\rangle_A$. That is, given we know the state of the B photon, the A photon is in a pure state.

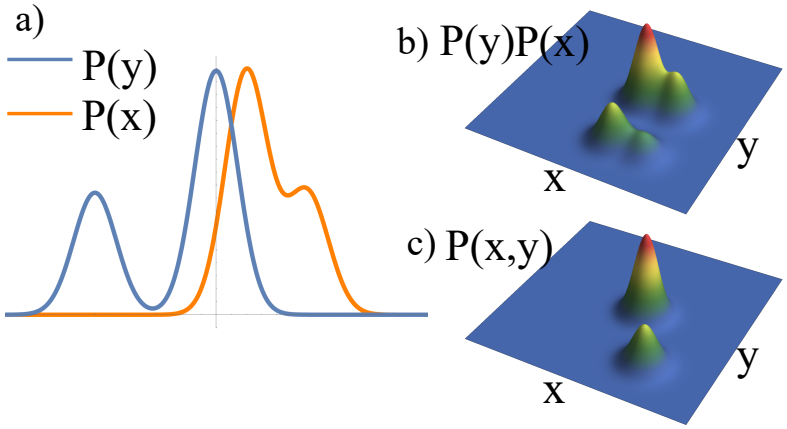

Figure 4.1: ***a***) Given the marginal distributions $P(x)$ and $P(y)$, it is impossible to reconstruct the original joint distributions. The best guess one might have is to assume they are uncorrelated (***b***), but this does not reproduce the original distribution (***c***). This is analogous to the non-separability of entangled states, $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}} \neq |\chi'\rangle_{\mathcal{S}} \otimes |\psi'\rangle_{\mathcal{P}}$.

Suppose now that on the way to the PBS, the B photon hits a dust particle and is absorbed. We no longer know what the state of the A photon with certainty. The state of the A photon might be $|H\rangle$ or $|V\rangle$, and it has equal probability to be either. In other words, the state of A is now a mixed state given by $\boldsymbol{\rho}_A = \mathbf{1}/2$. More specifically, it is a totally mixed state. By symmetry, we can also say that $\boldsymbol{\rho}_B = \mathbf{1}/2$. It is straight forward to check that $\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B \neq |\Psi\rangle\langle\Psi|$. That is, even if we know the subsystems separately, we cannot recover the original bipartite state, just as in the classical example above. The mathematical formalism for this process is known as the *partial trace*. It is identical to the trace, but carried over only one subsystem, as opposed to every subsystem. In statistics, this would also be referred to as "integrating out" a variable. We denote this by $\mathrm{Tr}_{\mathcal{X}}$ where $\mathcal{X}$ is the subsystem (e.g. $\mathcal{S}$ or $\mathcal{P}$) that is being traced over.

We see now that mixed states correspond to bipartite pure states where one of the parties has been lost. Indeed, we can use purity to test for entanglement. Given the task to determine if a state $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}}$ is entangled or not, one calculates the purity of the reduced density matrix belonging to the subsystems $\mathcal{S}$ or $\mathcal{P}$, for example, $\boldsymbol{\rho}_{\mathcal{P}} = \mathrm{Tr}_{\mathcal{S}}(|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}}\langle\Psi'|)$. If the purity of $\boldsymbol{\rho}_{\mathcal{P}}$ is less than 1 (i.e. not a pure state), then $|\Psi'\rangle_{\mathcal{S}\otimes\mathcal{P}}$ is entangled. The purity of the subsystem is unity only in the event where the two subsystems are totally uncorrelated (i.e. not entangled). This correspondence between mixed states and pure bipartite states provides a useful tool. We introduce a new

subsystem $\mathcal{E}$, often called the *environment*, which is spanned by the states $|e\rangle$. We then define a new state, $|\rho\rangle$, called the *purification* of $\boldsymbol{\rho}_A$ such that

$$|\rho\rangle = \frac{1}{\sqrt{2}}(|e_1\rangle_{\mathcal{E}} |H\rangle_A + |e_2\rangle_{\mathcal{E}} |V\rangle_A). \tag{4.4}$$

We can see that $|\rho\rangle$ is of the same form as $|\Psi\rangle$ above (i.e. it is entangled). Using the partial trace, one recovers $\boldsymbol{\rho}_A = \mathrm{Tr}_{\mathcal{E}}(|\rho\rangle \langle\rho|)$. Purification is often a useful tool when dealing with mixed states, as it allows one to treat them mathematically as pure states, which are often easier to handle. We also see that purification induces coherence. If one can gain information about the environment, then a totally decohered state can be made to have coherence.

Mixtures vs. Superposition

At this point it may still be unclear why the density matrix formalism is needed. To see this, we continue with the example of the photon pair source. We now place a HWP in the path of photon A that performs the transformation $|H\rangle$ to $|D\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)$ and $|V\rangle$ to $|A\rangle = 1/\sqrt{2}(|H\rangle - |V\rangle)$. We also place a polarizer that projects onto $|D\rangle$ after the waveplate. There are several cases that can occur:

1. Measuring B returns $|V\rangle \implies$ Photon A is $|H\rangle \implies |H\rangle \xrightarrow{\text{HWP}} |D\rangle \implies$ **photon makes it through polarizer with probability 1**.
2. Measuring B returns $|H\rangle \implies$ Photon A is $|V\rangle \implies |V\rangle \xrightarrow{\text{HWP}} |A\rangle \implies$ **photon makes it through polarizer with probability 0**.
3. Photon B is lost $\implies$ Photon A is $\mathbf{1}/2 \implies \mathbf{1}/2 \xrightarrow{\text{HWP}} \mathbf{1}/2 \implies$ **photon makes it through polarizer with probability 0.5**.

Now we rotate the polarizer to project onto $|A\rangle$.

4. Measuring B returns $|V\rangle \implies$ Photon A is $|H\rangle \implies |H\rangle \xrightarrow{\text{HWP}} |D\rangle \implies$ **photon makes it through polarizer with probability 0**.
5. Measuring B returns $|H\rangle \implies$ Photon A is $|V\rangle \implies |V\rangle \xrightarrow{\text{HWP}} |A\rangle \implies$ **photon makes it through polarizer with probability 1**.

6. Photon B is lost $\implies$ Photon A is $\mathbf{1}/2 \implies \mathbf{1}/2 \xrightarrow{\text{HWP}} \mathbf{1}/2 \implies$ **photon makes it through polarizer with probability 0.5**.

We can see that there is a fundamental difference between the pure and mixed state. We can see an illustration of this scenario in Fig. 4.2. From a purely mathematical

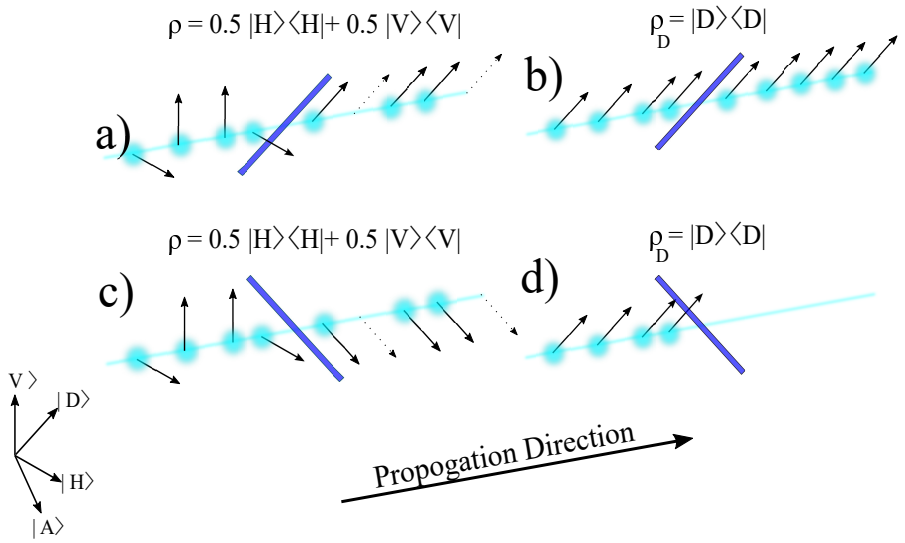

Figure 4.2: Illustrating the difference between the dynamics of the state $\boldsymbol{\rho}$, which is a statistical mixture of $|H\rangle$ and $|V\rangle$, and the state $|D\rangle = 1/\sqrt{2}(|H\rangle+|V\rangle)$, which is a quantum superposition. Here the blue bars represent the orientation of a polarizer, projecting onto $|D\rangle$ in panels a) and b), and $|A\rangle$ in panels c) and d). **a**) Half the light is lost because the projection onto $|D\rangle$ may fail, with probability 0.5. **b**) No light is lost because of the phase relationship between $|H\rangle$ and $|V\rangle$. **c**) Half the light is lost because the projection onto $|A\rangle$ may fail, with probability 0.5. **d**) All the light is lost because of the phase relationship between $|H\rangle$ and $|V\rangle$.

perspective, lets compare the density matrices in the above examples:

$$\boldsymbol{\rho}_D = |D\rangle\langle D| = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \quad \text{and} \quad \boldsymbol{\rho} = 0.5\,|H\rangle\langle H| + 0.5\,|V\rangle\langle V| = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \quad (4.5)$$

The distinguishing feature is a lack of *coherence* in the mixed state, which is indicated by the lack of off-diagonal terms in the density matrix. Physically speaking, coherence is defined a definite phase between $|H\rangle$ and $|V\rangle$. In $\boldsymbol{\rho}$ there is only random phase between the two. More fundamentally, the distinction is that a mixture is either $|H\rangle$ or $|V\rangle$, we simply do not know which do not know which. On the other hand, $|D\rangle =$

$1/\sqrt{2}(|H\rangle + |V\rangle)$ is *neither* $|H\rangle$ nor $|V\rangle$, rather it is something totally different. This is why the state $|D\rangle$ is always accepted by the polarizer in Fig. 4.2b, but always rejected in Fig. 4.2d. Put another way, one can view coherence as a measure of how "quantum" a state is. Statistical mixtures can be viewed as "classical", where as superposition is "quantum". Coherence is key in the study of the quantum to classical transition (55).

4.3 Decoherence in Measurement

As we've seen above, when a subsystem is lost, or cannot be seen, it must be traced out. By this logic, the quantum state should be traced out after having interacted with the measurement apparatus. After all, we never *see* the quantum system. Our interaction with it is entirely mediated by the measurement device (8).

Assuming that $\boldsymbol{U} = \exp(i\delta \boldsymbol{A}\boldsymbol{p})$, where $\boldsymbol{A} = \sum_n a_n |a_n\rangle \langle a_n|$, the state of the system and pointer post interaction is

$$\boldsymbol{U} |\chi\rangle |\psi\rangle = \sum_n \langle a_n|\chi\rangle |a_n\rangle \otimes |\psi(x - \delta a_n)\rangle . \tag{4.6}$$

Because we do not see the quantum system, we must trace it out, leaving us with the state

$$\boldsymbol{\rho}_{\mathcal{P}} = \mathrm{Tr}_{\mathcal{S}} \left(\boldsymbol{U} |\chi\rangle |\psi\rangle \langle\chi| \langle\psi| \boldsymbol{U}^{\dagger} \right) = \sum_n |\langle a_n|\chi\rangle|^2 |\psi(x - \delta a_n)\rangle \langle\psi(x - \delta a_n)| . \tag{4.7}$$

The pointer is now in a statistical mixture (i.e. not "quantum") of shifted position states. The probability for us to observe the photon at any single position δa_n is $|\langle a_n|\chi\rangle|^2$ which is exactly Born's rule. This is the (more) rigorous formulation of the collapse of the wave function.

4.4 Decoherence in Sequential Measurement

In the previous chapter we discussed how weak sequential measurements have an interesting time ordering structure. We will now focus on a related topic of the coherence properties of pointers after sequential measurements of non-commuting observables. Specifically, we want to see how coherence propagates in sequential measurement (if it does). We consider a sequence of non-commutative measurements, $\boldsymbol{A}$ then $\boldsymbol{B}$ then $\boldsymbol{C}$, performed by devices that point to two possible outcomes $(a_{1,2}, b_{1,2}, c_{1,2})$, denoting the

result of each measurement. We will perform the measurements in two different ways, one in which we use an intuitive (and as we will see later, incorrect) approach, and then using the formal von Neumann model of measurement to calculate the pointer density matrix, and then we compare to two results. We saw that for one measurement the pointer density matrix is totally decohered. Does this remain true for sequential measurements?

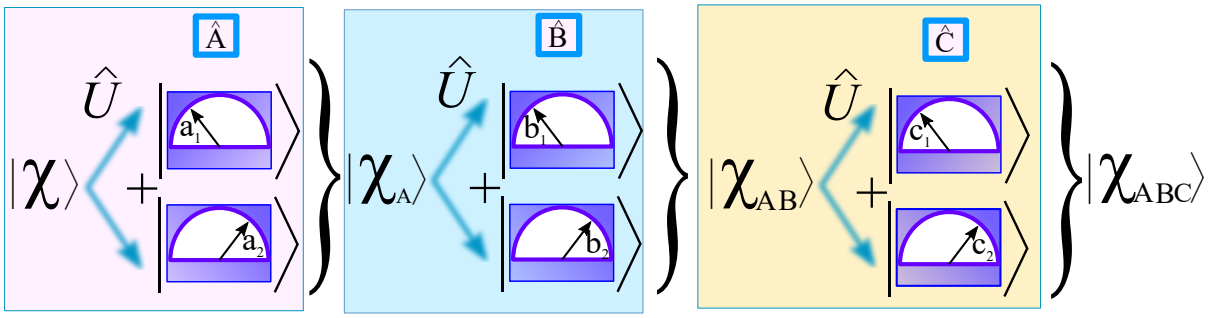

Figure 4.3: A graphical representation of the von Neumann measurement model for three sequential measurements. The quantum system, $|\chi\rangle$, interacts with the measurement device, moving the dial to a position based off the outcome of the measurement. The total state of both the measurement device and quantum system are now in an entangled superposition of possible results, $|\chi_A\rangle$ (denoted by the '+' signs). The process repeated for each of the measurements.

The Intuitive Approach

Experimentally, it is difficult to use external degrees of freedom to act as a measurement device (for example, the spin of some atom, or a separate photon that is emitted based off the result measurement). As such, we can imagine using the internal degrees of freedom of the photon, such as the path degree of freedom. In this 'path encoding', photons are in a superpositions of paths, like for example the shifted and un-shifted paths of a birefringent crystal.

We perform the following sequence of (strong) measurements of polarization observables: $\boldsymbol{\pi}_H$ then $\boldsymbol{\pi}_D$ then $\boldsymbol{\pi}_H$. We assume that no measurements have been made in the system's past prior to the first measurement. Because of this, we cannot have any information about it, and so the initial state is totally mixed ($\boldsymbol{\rho} = \mathbf{1}/2$). A conceptual experiment that encodes onto the path degree of freedom can be constructed from an array of PBSs and HWPs at 22.5°. The pointer states are the transmitted ($|T\rangle$) or

reflected ($|R\rangle$) paths. The corresponding experiment would look something like what is shown in 4.4. After the first measurement of $\boldsymbol{\pi}_H$ the pointer density matrix is

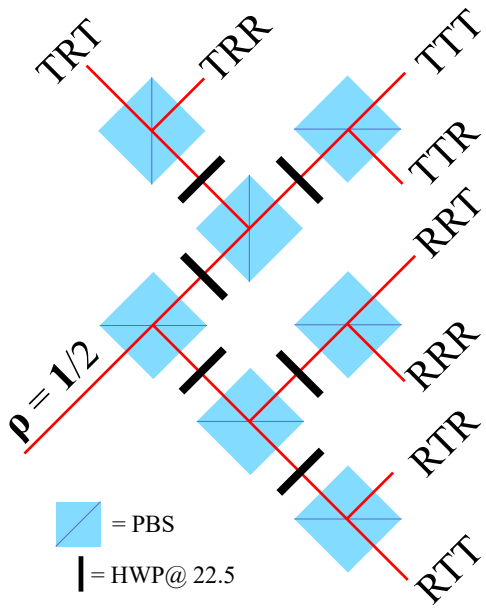

Figure 4.4: A possible experimental arrangement to test how coherence propagates through sequential measurements. The sequence $\boldsymbol{\pi}_H\boldsymbol{\pi}_D\boldsymbol{\pi}_H$ is implemented using polarizing beam splitters (PBS) and half waveplates (HWP). The pointer states are the different paths that the photon might take as it is transmitted (T) or reflected (R) through the beam splitter network.

$$\boldsymbol{\rho}_{\mathcal{P}_1} = 0.5\,|T\rangle\langle T| + 0.5\,|R\rangle\langle R|\,. \tag{4.8}$$

The second measurement is that of $\boldsymbol{\pi}_D$. There are now four pointers, those being the twice transmitted path (TT), reflected twice (RR), and of the combined cases (TR and RT). The possible results are all equi-probable, as $|\langle H|D\rangle|^2 = 1/2$, resulting in the pointer density matrix

$$\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2} = 0.25\,|TT\rangle\langle TT| + 0.25\,|TR\rangle\langle TR| + 0.25\,|RT\rangle\langle RT| + 0.25\,|RR\rangle\langle RR|\,. \tag{4.9}$$

The final measurement is $\boldsymbol{\pi}_H$. Again the pointer states are in an equi-probable mix of the possible pointer states

$$\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3} = \frac{1}{8}(|TTT\rangle\langle TTT| + |TTR\rangle\langle TTR| + |TRT\rangle\langle TRT| + |RTT\rangle\langle RTT|$$

$$+ |TRR\rangle\langle TRR| + |RTR\rangle\langle RTR| + |RRT\rangle\langle RRT| + |RRR\rangle\langle RRR|),$$

or, in matrix form,

$$\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3} =$$

$$\frac{1}{8}\begin{array}{c} \begin{array}{cccccccc} \text{TTT} & \text{TTR} & \text{TRT} & \text{RTT} & \text{TRR} & \text{RTR} & \text{RRT} & \text{RRR} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} \text{TTT} \\ \text{TTR} \\ \text{TRT} \\ \text{RTT} \\ \text{TRR} \\ \text{RTR} \\ \text{RRT} \\ \text{RRR} \end{array} \tag{4.10}$$

The lack of any off-diagonals indicates there is no coherence, or in other words, no superposition is taking place. This is perhaps not a surprising, and can be traced back to the collapse of the wavefunction when measuring incompatible observables.

Von Neumann Model Approach

Now that we have a intuitive feeling for what the result of sequential measurements should yield, we check this using the von Neumann model. To simplify the math we introduce an auxiliary environment Hilbert space to purify the density matrix:

$$\mathbf{1}_{\mathbb{S}}/2 \xrightarrow{\text{purification}} |\rho\rangle = \frac{1}{\sqrt{2}} \sum_n |\phi_n\rangle |n\rangle , \tag{4.11}$$

and

$$\boldsymbol{\rho} \xrightarrow{\text{purification}} |\rho\rangle |\psi(x)\rangle , \tag{4.12}$$

where $|\phi_n\rangle$ are the environment basis states that will be traced over, and $|n=0\rangle = |H\rangle$ and $|n=1\rangle = |V\rangle$. In this context the environment can represent all the past interactions that the system has had, which we are not privy to (8). The evolution is given by $\boldsymbol{U} = \exp(ig\boldsymbol{\pi}_H\boldsymbol{p})$. Tracing out the quantum state post interaction gives the pointer density matrix

$$\boldsymbol{\rho}_{\mathcal{P}} = 0.5\,|\psi(x-g)\rangle\langle\psi(x-g)| + 0.5\,|\psi(x)\rangle\langle\psi(x)|).$$

This is of the same form as Eq. 4.8 if we take $|\psi(x-g)\rangle = |T\rangle$ and $|\psi(x)\rangle = |R\rangle$. The evolution operator for a sequence of two measurements is $\boldsymbol{U} = \boldsymbol{U}_2\boldsymbol{U}_1 = \exp(ig\boldsymbol{\pi}_D\boldsymbol{p}_2)\exp(ig\boldsymbol{\pi}_H\boldsymbol{p}_1)$. Here $\boldsymbol{p}_1$ is the momentum operator of the first pointer, $|\psi(x_1)\rangle$, and $\boldsymbol{p}_2$ is the momentum operator of the second pointer $|\psi(x_2)\rangle$. The initial state is

$$\boldsymbol{\rho} \xrightarrow{\text{purification}} |\rho\rangle\,|\psi(x_1)\rangle\,|\psi(x_2)\rangle\,, \tag{4.13}$$

Tracing out the quantum state post interaction gives the two pointer density matrix

$$\boldsymbol{\rho}_{\mathcal{P}_1,\mathcal{P}_2} = 0.25(|\psi(x_1-g,x_2-g)\rangle\langle\psi(x_1-g,x_2-g)| + |\psi(x_1-g,x_2)\rangle\langle\psi(x_1-g,x_2)| +$$

$$|\psi(x_1,x_2-g)\rangle\langle\psi(x_1,x_2-g)| + |\psi(x_1,x_2)\rangle\langle\psi(x_1,x_2)|)$$

which is, again, the same form as Eqn. 4.9, if we again take $|\psi(x-g)\rangle = |T\rangle$ and $|\psi(x)\rangle = |R\rangle$.

It appears at this point that we have a good understanding of how coherence propagates (or does not) through sequential measurement. For completeness sake we compute the pointer density matrix for three measurements. The new evolution operator is $\boldsymbol{U} = \boldsymbol{U}_3\boldsymbol{U}_2\boldsymbol{U}_1 = \exp(ig\boldsymbol{\pi}_H\boldsymbol{p}_3)\exp(ig\boldsymbol{\pi}_D\boldsymbol{p}_2)\exp(ig\boldsymbol{\pi}_H\boldsymbol{p}_1)$, and the initial state is $|\rho\rangle\,|\psi(x_1)\rangle\,|\psi(x_2)\rangle\,|\psi(x_3)\rangle$. Tracing out the quantum state post interaction gives the three pointer density matrix

$$\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3} =$$

$$\frac{1}{8}\begin{array}{c} \begin{array}{cccccccc} \text{TTT} & \text{TTR} & \text{TRT} & \text{RTT} & \text{TRR} & \text{RTR} & \text{RRT} & \text{RRR} \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \text{TTT} \\ \text{TTR} \\ \text{TRT} \\ \text{RTT} \\ \text{TRR} \\ \text{RTR} \\ \text{RRT} \\ \text{RRR} \end{array} . \tag{4.14}$$

Surprisingly, this is *not* what we had in Eqn. 4.10. We now have coherences between pointers as indicated by the off-diagonal terms.

Interpretation

Why did our initial approach, which appears intuitively correct, differ from the von Neumann model? The difference is an implicit assumption made in the former approach. We had assumed that after each successive measurement the state collapses. We made this assumption when we calculated $\boldsymbol{\rho}_{\mathcal{P}_1}$ and then used it to calculate $\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2}$, and then again used that result to get $\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3}$. This is tracing out the system between each PBS. However, information is only ever obtained at the end of the sequence of three measurements. The system should not be traced out before then. Stated in yet another way; *it is wrong to treat sequential measurements as a sequence of individual measurements.*

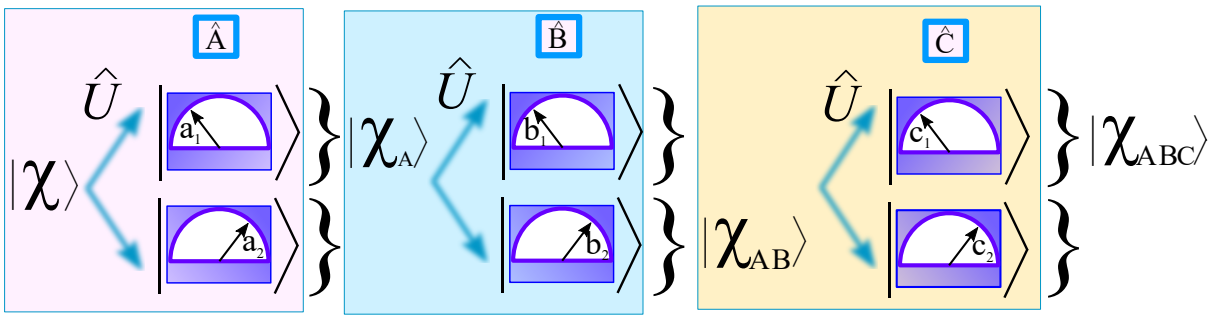

Figure 4.5: The graphical representation of the intuitive approach, which is different from Fig. 4.3. Because we assumed that the state collapses between each measurement, the pointer density matrix becomes an incoherent mixture of all possible results, rather than a superposition (denoted by the lack of '+' signs).

This is a rather bold claim; the projection done by the PBS does not on its own collapse the state. This, of course, requires verification. The most obvious way to do so is to measure, or reconstruct, the state $\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3}$ that is encoded by the different possible paths a photon may take, and show that it has coherences where one would not expect to see if the state was collapsing between each PBS. Unfortunately state reconstruction of many path modes is generally difficult to do. Generally speaking, one would have to interfere each of the paths pairwise. This would not only be experimentally difficult to perform in the lab, but would also require $8^2 = 64$ measurements (or more generally, on the order of d^2 measurements, where d is the dimension of the Hilbert space). As such, a method of reconstructing large path-encoded density matrices is required. The

next chapter discusses the method I developed in the course of my masters to easily perform this state reconstruction for many path modes.

5

High-Dimensional Multi-Rail Quantum State Estimation

Preamble

Quantum information protocols often rely on tomographic techniques to determine the state of the system. A popular method of encoding information is on the different path modes photons may take. However reconstruction of states encoded in a large number of paths is often prohibitively resource intensive. Addressing this, we present a simple method of determining the state of photon in a superposition of d rails (i.e., beam paths) using Fourier transforms. We demonstrate the technique by measuring several 6-dimensional density matrices and computing metrics such as state fidelity and state purity.

5.1 Introduction

Determining the quantum state of a physical system is a key task in quantum physics, in particular in quantum metrology, quantum information and quantum cryptography. As the quantum state $\boldsymbol{\rho}$ mathematically defines all properties of the system, determining it is often a priority. Often then the experiment is judged based on how it preserved certain quantum features such as entanglement, coherence, and so on. As such, many techniques of reconstructing a variety of different types of physical quantum systems exist (56, 57, 58, 59). Considerable effort has been put towards the development of

these techniques as well as supplementary likelihood methods that post process the result. These have even found application in classical image processing (60).

Despite its importance, relatively few existing methods can easily accommodate reconstructing the path states with more dimensions than a qubit. Generally, such reconstruction methods require $\mathcal{O}(d^2)$ measurement settings for a d-dimensional system. Common techniques reconstruct the state by projecting onto an over-complete set of projectors. Despite considerable effort dedicated to optimizing these schemes (61, 62), resource considerations typically make states with high dimensions difficult to reconstruct. This problem becomes significant as information processing moves towards higher dimensions in order to reach higher bit-rates and better security (63, 64).

A popular method of increasing the dimension of a system is by using the different paths that photons may take. That is, rather than encoding information in say, the horizontal or vertical polarizations of the photon, the information is encoded as a superposition of all the possible paths the photon may take through the experiment (say for example, the transmitted or reflected path of a beam splitter). One advantage is that the dimension is not constrained just two (like in polarization), and can be increased just by adding in more possible paths (e.g. from a beam splitter network). As an important example, any discrete $N \times N$ unitary can be realized in the lab using a universal multi-port interferometer, which produces a coherent superposition of N path modes (65, 66). The use of path modes is especially popular in on-chip photonic devices due to the difficultly of fabricating devices that maintain the polarization state (67) (for an example where it was managed, see (68)). Indeed, implementations of on-chip multi-port interferometers are an especially popular topic (69, 70). In free-space however, path modes are not a commonly used method of state encoding. This is in part due to the difficulty of reconstructing the state of a large number of path modes. Suffice it to say, the ability to simply reconstruct the state of a large number of path modes is of great use and importance in quantum communication.

Unfortunately, a simple way of performing reconstructing large states is lacking. One method is to interfere each path pairwise, which requires d^2 operations (71, 72). We propose and implement a simple method of performing tomography of path modes. This method exploits the physical manifestation of state coherence, i.e. interference. We show how the state, encoded into transverse path modes, may be reconstructed entirely by using both real and k-space simultaneously with a one dimensional Fourier transform

performed by a cylindrical lens. This, along with a second discrete Fourier transform in post-analysis, allows us to easily reconstruct high dimensional path-encoded states. This method is conceptually similar to the reconstruction of the optical coherency matrix or two particle interferometry (73, 74), however we base our discussion in the context of density matrices as this method is not inherently optical, and works only on the premise of interference brought about the wave-like nature of quantum objects (75).

5.2 Theory

The mutual coherences describes the relative phase between states that physically manifest as interference, and serves as a measure of the indistinguishably of the degrees of freedom of the quantum system (76). The greater the indistinguishablility, the larger the magnitude of the coherence will be and the more the two states will interfere. The magnitude of the coherence is directly related to the interference visibility. The phase of the interference pattern is set by the phase of the coherence. Given that information about the coherences is encoded into the phase and amplitude of the oscillating interference pattern, it seems natural to consider Fourier transforms (FT) as a natural tool for reconstructing the state $\boldsymbol{\rho}$.

As an illustrative example, we reconstruct the state of two path modes $|\psi_0\rangle$ and $|\psi_1\rangle$, centered at $x = \delta_0$ and $x = \delta_1$ respectively. The state is described by a 2×2 density matrix $\boldsymbol{\rho} = \sum_{ij} a_{ij} |\psi_i\rangle \langle\psi_j|$. In position space, the path modes are assumed to be Gaussian wavefunctions,

$$\psi(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-\frac{x^2}{4\sigma^2}},$$

with width σ. The path modes can be interfered by looking in momentum (Fourier) space. In an optical setting, this is easily accomplished by using a lens. A lens maps the field present one focal length (f) before it, $E(z = -f, \vec{r})$, to $E(z = f, \vec{r} = f\vec{k}/k_T)$. Here $\vec{r}$ is the transverse position and $\vec{k}$ is the transverse component of the total momentum vector, and k_T is the magnitude of the total momentum vector (i.e. the wavenumber). Therefore the field located at $z = f$ is the Fourier transform of the field at $z = -f$, up to some scaling factor f/k_T. In this sense, a lens is an analogue Fourier transform computer. We call the step in which we interfere the paths using a lens the Optical Fourier transform (OFT). In momentum space, the density matrix is $\rho(k, k') = \langle k| \boldsymbol{\rho} |k'\rangle$.

However, when the interference pattern is observed (for example, on a camera), only see the diagonal elements remain, i.e. $|k'\rangle = |k\rangle$. As such we get $\rho(k) = \langle k|\,\boldsymbol{\rho}\,|k\rangle = |\tilde{\psi}(k)|^2 \sum_{i,j} a_{ij} \exp(i\Delta_{ij}k)$, where $\Delta_{ij} = \delta_i - \delta_j$ is the pairwise distance between the paths i and j. Writing $a_{ij} = |a_{ij}|e^{i\phi_{ij}}$, this can be simplified to

$$\rho(k) = |\tilde{\psi}(k)|^2(a_{00} + a_{11} + 2|a_{12}|\cos(\Delta_{12}k + \phi_{12})), \tag{5.1}$$

where ϕ_{12} is the phase of a_{12}. The first two terms are diagonal elements of $\boldsymbol{\rho}$, and necessarily sum to unity. We see the visibility of the interference term (i.e. the cosine term), $2|a_{ij}|$, is directly proportional to the magnitude of a_{ij}, and the phase of the interference, ϕ_{ij}, is the same as the phase of a_{ij}. The entire interference pattern is multiplied by an overall envelope $|\tilde{\psi}(k)|^2$.

We can extract the amplitude and phase information from the interference pattern in Eq. 5.1 by using a second FT. We call this the Discrete FT (or DFT), as it is performed on a computer. Analytically the DFT step gives

$$\begin{gathered} F\{\rho(k)\}(x) = \\ (a_{00} + a_{11})\delta(x) + a_{12}\delta(x - \Delta_{12}) + a_{21}\delta(x + \Delta_{12}). \end{gathered} \tag{5.2}$$

From the above we can identify that $a_{12} = F\{\rho(k)\}(x = \Delta_{12})$. While interfering paths using a lens works well for two paths, extending to higher dimensions introduces the possibility of overlapping frequencies due to equal spacing. This obscures the coherence as any paths with the same spacing will appear at the same frequency. We now discuss this problem and present our solution.

5.2.1 Extending to Higher Dimensions

We now extend the method to higher dimensions by introducing more path modes. The states $\{|\psi_i\rangle\}$ are now spatial modes in the transverse xy-plane. Path modes are assumed to have rotationally symmetric Gaussian wavefunctions, with equal width σ in both the x and y directions. The i-th path mode is at position $\vec{\delta_i} = (d_{x_i}, d_{y_i})$, and the subscripts $\psi_i(x, y)$ denote the state centered around $\vec{\delta_i}$; $\psi_i(x, y) = \psi(x - \delta_{x_i})\psi(y - \delta_{y_i})$. The set of positions $\{\vec{\delta_i}\}$ is called the *geometry*, and the spacing between any two paths i and j is $\vec{\Delta}_{ij} = \vec{\delta_i} - \vec{\delta_j}$. We assume that the paths are spatially distinct ($|\vec{\Delta}_{ij}| \gg \sigma$ for all i, j), so that the states form an approximately orthonormal basis for the Hilbert space.

Typical lenses perform two dimensional (2D) OFTs along both transverse axes xy. If $\vec{\Delta}_{ij}$ and $\vec{\Delta}_{i'j'}$ are collinear, the magnitude of each pairwise distance must be unique for the reconstruction method to work. Otherwise, the interferences will overlap on the DFT and cannot be distinguished. However, we allow the spacing to be not unique (i.e. $|\vec{\Delta}_{ij}| = |\vec{\Delta}_{i'j'}|$, for $(i,j) \neq (i',j')$), so long as $\vec{\Delta}_{ij}$ and $\vec{\Delta}_{i'j'}$ are not collinear. In this situation the interference pattern can be distinguished by the direction of the oscillations of the interference, as they will be in the direction of $\vec{\Delta}_{ij}$. This means the ideal geometry is one where every pairwise distance is unique for each direction. However, it is experimentally challenging to produce such a geometry, especially for a large number of path modes. It is often simpler to produce a geometry where each pairwise distance is not unique, and they have the same k-vector (for example, the geometry in Fig. 5.1a). In this case, implementing the OFT using a normal lens will fail as the coherences cannot be distinguished by either the frequency or direction of their k-vector, and may even cancel if out of phase.

We overcome this problem by spatially separating the interference patterns that have the same k-vector with a cylindrical lens. The cylindrical lens performs a one dimensional (1D) OFT along a single axis, which we refer to as the OFT-axis, while leaving the orthogonal axis unchanged, as shown in Fig. 5.1b. This means one does not have to consider interference patterns with the same k-vector having the same frequency, as they will be spatially separated in the axis perpendicular to the OFT-axis. By rotating the OFT-axis to every angle defined by the path modes i and j, given by $\theta_{ij} = \tan^{-1}((\delta_{y_i} - \delta_{y_j})/((\delta_{x_i} - \delta_{x_j})))$ (shown in Fig. 5.1c), and taking the DFT for all interference patterns obtained, one can recover the coherences of $\boldsymbol{\rho}$, an example of which is shown in Fig. 5.1d.

We present the formal theory by starting with a density matrix of path modes given by $\boldsymbol{\rho} = \sum_{ij}^{N} a_{ij} \left|\psi_i(x,y)\right\rangle \left\langle\psi_j(x,y)\right|$. Here the spatial state $\left|\psi_i(x,y)\right\rangle$ is separable in x and y space, so that $\left|\psi_i(x,y)\right\rangle = \left|\psi_i(x)\right\rangle \left|\psi_i(y)\right\rangle$. There are two approaches that one may take to introduce the OFT mathematically, of which we choose the latter. Either the OFT-axis is rotated in the $xy-$plane or, leaving the OFT-axis along, say, the y-axis, we rotate the wavefunctions in the $xy-$plane instead. This latter method is simpler as Gaussians are rotationally invariant, reducing the rotation to finding positions $\{\vec{\delta}_i\}_i^N$ in the rotated basis, then calculating the FT along the y-axis. If $R(\theta)$ is the usual rotation by θ matrix, then the positions in the rotated basis are $\vec{\delta}' = \{R(\theta)\vec{\delta}_i\}_i^N = \{\vec{\delta}'_i\}_i^N$. In

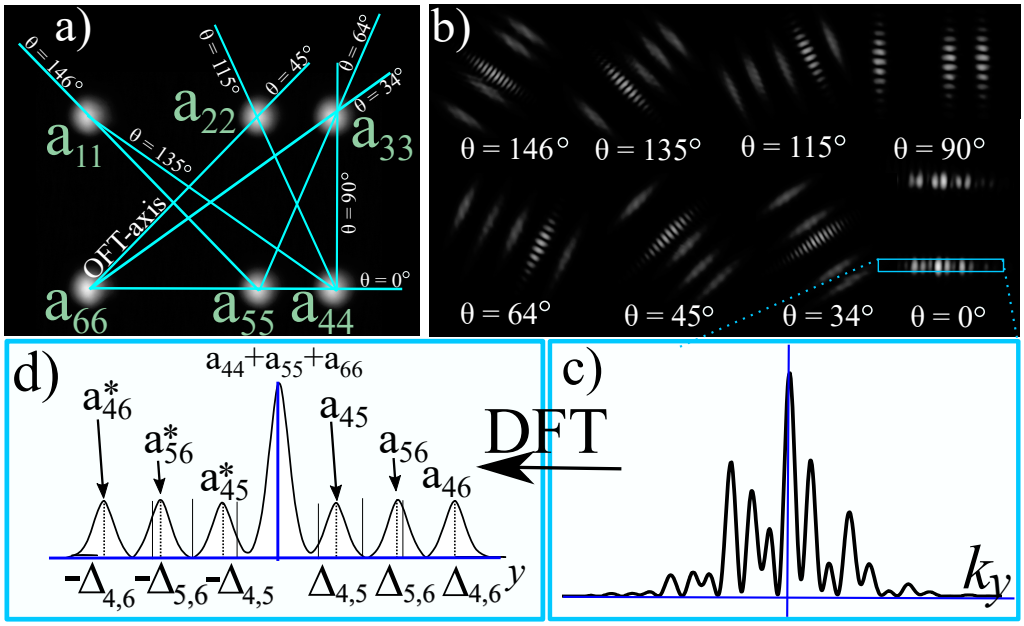

Figure 5.1: a) The six path modes, labeled by their density matrix elements a_{11} through a_{66}, are shown in the figure and encode the state $\boldsymbol{\rho}$. The optical Fourier transform axis (OFT-axis) rotates to interfere each pair of paths, as indicated by the lines. **b)** The corresponding OFT for each angle. The diagonal elements can be recovered from the non-interfering Gaussians. **c)** Each pattern is recorded and analyzed one at a time via discrete Fourier transform (DFT) by taking a one pixel wide 'slice' through the interference pattern. **d)** The Fourier transform of the interference pattern (For illustrative purposes we plot the magnitude). The normalization is obtained by summing the zero frequency peaks of each interference pattern present in the panel. All panels contain real data.

this new basis the wavefunctions change to $\psi_i'(x,y) = \psi(x-\delta_{x_i}')\psi(y-\delta_{y_i}')$. In the following discussion we drop the primed notation for notational clarity, assuming that states have been rotated to the appropriate angles.

We perform the 1D OFT along the y-axis, while leaving the $x-$axis untouched. Physically this is the step performed by the cylindrical lens, except we are in a frame of reference of the lens, (i.e. the lens is stationary and the path geometry is rotating). The interference pattern displayed on a screen is the probability distribution of the momentum representation. This leaves only the diagonal elements of the density matrix for the y direction:

$$\begin{gathered}\boldsymbol{\rho}(k_y) = \langle k_y| \boldsymbol{\rho} |k_y\rangle = \\ \sum_{ij}^{N} a_{ij} \langle k_y|\psi_i(y)\rangle \langle \psi_j(y)|k_y\rangle |\psi_i(x)\rangle \langle \psi_j(x)| = \\ |\widetilde{\psi}(k_y)|^2 \sum_{ij}^{N} a_{ij} e^{i\Delta_y^{(i,j)}k_y} |\psi_i(x)\rangle \langle \psi_j(x)| .\end{gathered} \tag{5.3}$$

Here $\Delta_y^{(i,j)} = \delta_{y_i} - \delta_{y_j}$ is the vertical separation between $\vec{\delta}_i$ and $\vec{\delta}_j$. Note that Eq. 5.3 is still an operator in x-space as we are only looking at the momentum distribution in the y-space (hence why $\boldsymbol{\rho}(k_y)$ is bolded). We will drop the envelope $|\widetilde{\psi}(k_y)|^2$ for convenience. In practice, this simply widens the FT signal at each frequency, but the relative heights are unchanged.

To analyze this sum, we split the expression into isolated terms. We separate the terms where $i = j$, and consider only terms with the same δ_{x_m}, as these are the terms that interfere. This is equivalent to calculating $\langle x = \delta_{x_m}| \boldsymbol{\rho}(k_y) |x = \delta_{x_m}\rangle$. Note that $\langle \psi_i(x,y)|\psi_j(x,y)\rangle = 0$ does not imply that $\langle \psi_i(x)|\psi_j(x)\rangle = 0$, as δ_{x_i} may be equal to δ_{x_j}, so long as $\delta_{y_i} \neq \delta_{y_j}$. Because the path modes were assumed to not overlap significantly ($|\vec{\Delta}_{ij}| \gg \sigma$), and path whose x-space wavefunction is not centered at δ_m will be zero at δ_m (i.e. $\psi(\delta_{x_m} - \delta_{x_n}) = 0$ if $\delta_{x_m} \neq \delta_{x_n}$). As such, we keep only terms where $\delta_{x_i} = \delta_{x_j} = \delta_{x_m}$, which we express through delta functions (these are bold for

clarity):

$$\begin{aligned}\rho(k_y, \delta_{x_m}) = \langle x = \delta_{x_m} | \boldsymbol{\rho}(k_y) | x = \delta_{x_m} \rangle = \\ \sum_{i}^{N} a_{ii} |\psi_i(\delta_{x_i})|^2 \boldsymbol{\delta}(\delta_{x_i}, \delta_{x_m}) + \\ \sum_{i \neq j}^{N} a_{ij} \psi_i(\delta_{x_i}) \psi_j^*(\delta_{x_j}) e^{i\Delta_y^{(i,j)} k_y} \boldsymbol{\delta}(\delta_{x_i}, \delta_{x_m}) \boldsymbol{\delta}(\delta_{x_j}, \delta_{x_m}).\end{aligned} \tag{5.4}$$

The remaining $\psi_i(x)$ have the same x-displacement meaning that we can equate $\psi_i(x) = \psi_j(x)$. Consequently, $\psi_i(x)\psi_j^*(x) = \psi(x - \delta_{x_i})\psi^*(x - \delta_{x_j}) = |\psi(x - \delta_{x_m})|^2$. As such, we factor out $|\psi(\delta_{x_m})|^2$ and drop it in the analysis.

Finally, we perform the step done by the DFT by taking the FT in the k_y direction, returning us back to real space. This gives the final result:

$$\begin{aligned}\int \mathrm{d}k_y \rho(k_y, \delta_{x_m}) e^{ik_y y} = \sum_{i}^{N} a_{ii} \boldsymbol{\delta}(\delta_{x_i}, \delta_{x_m}) \delta(y) + \\ \sum_{i \neq j}^{N} a_{ij} \delta(y - \Delta_y^{(i,j)}) \boldsymbol{\delta}(\delta_{x_i}, \delta_{x_m}) \boldsymbol{\delta}(\delta_{x_j}, \delta_{x_m}).\end{aligned} \tag{5.5}$$

The first term is the sum of the intensities of all path modes with the same x displacement δ_{x_m}, and appears at the 'frequency' $y = 0$. In the second term, each frequency $y = \Delta_y^{(i,j)}$ has a height of a_{ij}, giving the desired result. The coherence $a_{i,j}$ can be resolved so long as $\Delta_y^{(i,j)} = \Delta_y^{(n,m)}$ only if $(i, j) = (n, m)$ (i.e. each pairwise distance is unique). Otherwise, if $\Delta_y^{(i,j)} = \Delta_y^{(n,m)}$ but $(i, j) \neq (n, m)$, the height of the peak will be $a_{i,j} + a_{n,m}$, meaning we cannot determine the coherence.

It is possible to obtain the diagonal elements of $\boldsymbol{\rho}$ at the same time. This is convenient as the experimental setup does not need to be changed, and light is not lost. For a given θ_{ij}, only those path modes whose angle between them is θ_{ij} will interfere. The other modes will still be transformed along the OFT-axis, but their intensities will remain unchanged. This can be seen in Fig. 5.1b. Because the intensities remain unchanged, one directly obtains the diagonal values from the intensity of these lobes, divided by the total intensity. The total intensity is also used in the DFT to obtain the correctly normalized coherence values (the coherence is equal to the peak height divided by the total intensity).

The exact number of angles θ_{ij} (or experiment configurations) required depends on the specific geometry. The Hermiticity of $\boldsymbol{\rho}$ reduces the number of angles needed

to obtain the coherences is no greater than $N(N-1)/2$ (we neglect any extra angles needed to calculate the diagonals). However, judicious choice of geometry can reduce this number significantly. As an example we construct a geometry explicitly in the following way: First we generate the following 1D geometry: $\delta = \{\delta_j = \sum_{i=0}^{j} kr^j | 0 < r < 1, k > 0\}_{j=1}^{N}$, i.e. the position of the j-th path is the j-th partial sum of a geometric series. Next we generate the 2D geometry from the Cartesian product: $\vec{\delta} = \delta \times \delta$. For this geometry, the number of angles needed, $\eta(N)$, is $\mathcal{O}(N^4) = \mathcal{O}(d^2)$, the same as traditional tomography. More generally, if $d = n \times m$ $(n < m)$ paths are arranged in an $n \times m$-lattice, constructed similarly to the above but truncated in, say, the y-direction, the number of angles can be reduced. For example, in the limiting case where $n = 1$, only one angle is needed to obtain all the coherences.

This suggests more rectangular geometries result in the least number of angles. However, this makes the maximum pairwise distance, $\Delta_{\max} = \max_{ij}\{\Delta_{i,j}\}$, becomes large compared to the lens and camera screen area. This which can lead to aberration, or even aperturing. Moreover, for lasers with large spectral bandwidths, chromatic aberration may wash away the visibility of the interference pattern. The sampling frequency of the camera, $1/\gamma$, where γ is the pixel length, must also be taken into account. The interference pattern on the screen oscillates with a frequency of $2\pi\Delta_{i,j}/\lambda f$, where f is the focal length of the lens. The Nyquist condition requires that $\Delta_{\max} < \lambda f/\pi\gamma$.

While the above discussion is concerned exclusively with path modes, one can always convert other degrees of freedom into path degrees of freedom (77, 78, 79, 80). In the laboratory, polarization states are the standard degree of freedom for information encoding because of the wide body of robust and diverse computational protocols available, and ease of use (81, 82, 83). However, when communication over large distances is required, the polarization states are often transformed into path states (84, 85). The method presented would allow for practical state estimation, without having to first transform into another encoding. Further more, reconstructing bipartite states, for example path-polarization states, can be accomplished simply by converting the polarization degree into a path degree.

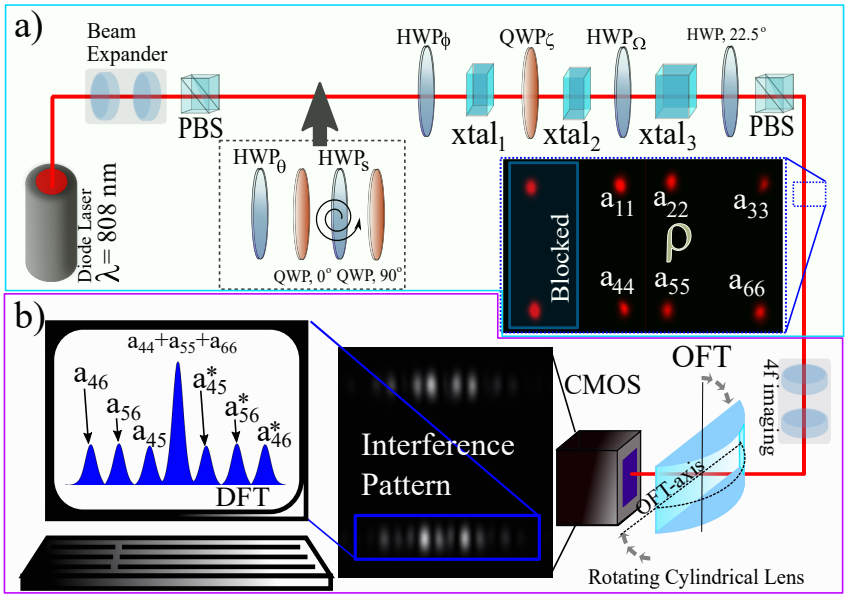

Figure 5.2: Demonstrating the tomography method experimentally. **(a)**. In the blue box is the state preparation: A beam expander sets the Rayleigh length of a 808 nm diode laser so that all the optics are within it. A series of displacement crystals ($\text{xtal}_{1,2,3}$) and both half and quarter waveplates (HWP_ϕ, QWP_ζ, and HWP_Ω), generate the state $\boldsymbol{\rho}$. The eight spatial paths produced by the three crystals are not compatible with the tomography. Therefore two of the eight paths are blocked to produce a six dimensional state. A set of HWP and QWP may optionally be inserted to form a mixed state by rapidly spinning HWP_s, whose purity is a function of the angle of HWP_θ. **(b)** The analysis is presented in the purple box: A $4f$ system ($f_1 = 1000$ mm and $f_2 = 400$ mm) images the six path modes onto a CMOS camera. A rotating cylindrical lens ($f = 250$ mm) performs the optical Fourier transform (OFT) along the OFT-axis. The coherences are obtained by post-analysis on computer. A slice of each interference pattern is taken and analyzed with a discrete FT on a computer. No window functions or filtering are used in this paper. Image contains raw data of three "slit" interference and the corresponding DFT.

5.3 Experiment

We demonstrate our technique by reconstructing density matrices produced by the optical setup shown in Fig. 5.2. Using polarized light from a diode laser at $\lambda = 808$ nm, we use a series of waveplates and calcite beam displacers, with the crystal axis cut at 45°, to generate $N = 6$ paths which encode $\boldsymbol{\rho}$. A $4f$ system ensures that all optics are within the Rayleigh length of the beam. Beam displacers provide a compact method of producing a geometry of path modes, and offer robust phase stability between paths. The first beam displacer (xtal_1 in Fig. 5.2a) is 2.8 cm in length, resulting in a transverse shift by $\delta_x = 2.7$ mm. The crystal axis is orientated such that $|H\rangle$ polarized light is shifted horizontally. Xtal_2 shifts $|V\rangle$ polarized light vertically by $\delta_y = 2.7$ mm as well. Lastly, xtal_3 is 4.1 cm in length and shifts $|H\rangle$ polarized light by $\delta_X = 4$ mm horizontally. In total, the sequence of beam displacers results in eight paths. However, this geometry is incompatible with the tomography method as the horizontal spacing between the right-most two paths and left-most two paths is the same, making the coherences indiscernible on the DFT. Therefore, we block two of the eight resulting paths to obtain a geometry of two rows and three columns. The resulting six dimensional state $\boldsymbol{\rho}(\phi, \zeta, \Omega)$ is determined by the angles (ϕ, ζ, Ω) of the corresponding waveplates (labeled HWP_ϕ, HWP_Ω and QWP_ζ in Fig. 5.2a), preceding each crystal. We note that in this situation a 2D OFT would not reconstruct the state properly as the geometry contains repeating spacings along multiple transverse directions.

The spatial modes are imaged onto a CMOS camera using a $4f$ system where $f_1 = 1000$ mm and $f_2 = 400$ mm, and background subtracted to remove noise. The OFT is performed using a cylindrical lens ($f = 250$ mm), placed one focal length before the camera. This lens rotates in an automated mount to each of the angles θ_{ij}, of which there are eight. The 1D OFT produces a set of interference patterns orientated at θ_{ij} with respect to the horizontal, spatially separated in the axis perpendicular to the OFT-axis. To obtain the coherence, first we rotated the image back by $\theta = -\theta_{ij}$ so to orient the interference patterns along the horizontal. A line is then taken through the center of each interference pattern present (this is the projection onto $|x = \delta_{x_m}\rangle$), resulting in a one dimensional list of the interference pattern. This list is then discrete

Fourier transformed to obtain the coherence values. An example of this procedure is shown pictorially in Fig. 5.3

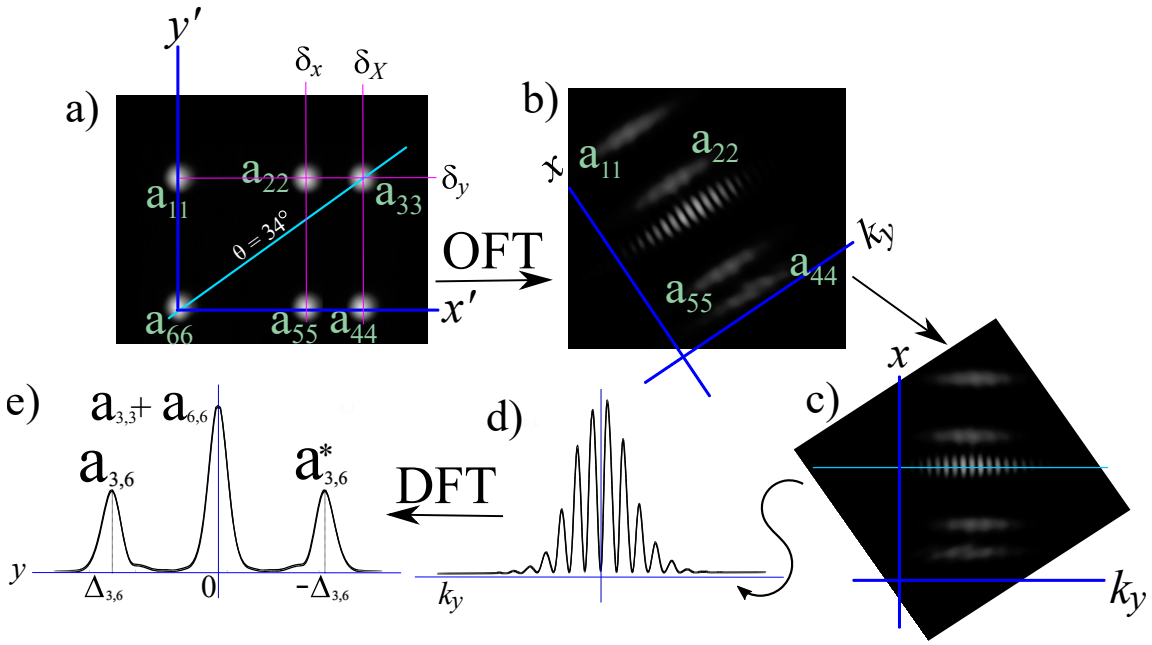

Figure 5.3: A step by step example of the reconstruction procedure. **a**) The OFT-axis is set at $\theta_{3,6} = 34°$ so as to interfere the paths labeled $a_{3,3}$ and $a_{6,6}$. **b**) The interference produced via optical Fourier transform is performed by the cylindrical lens is imaged onto the CCD camera. **c**) In post analysis, the image is rotated by $\theta = -34°$ so as to align the interference pattern with the horizontal. This allows us to easily take a one pixel wide line through the center of the interference pattern (i.e. project onto $|x = \delta_X \cos(34°) - \delta_y \sin(34°)\rangle$). **d**) The resulting 1D data set. **e**) We take the DFT of the resulting data. The coherences show up as peaks (only magnitude is shown in picture) at a 'frequency' determined by the spacing between paths $a_{3,3}$ and $a_{6,6}$. The peaks are not normalized, and must be normalized by the total intensity, which is the height of the zero-frequency peak, plus the intensity of all the other 'lobes' in **b**. All images are background subtracted.

The method thus stated correctly gives the magnitude of each of the coherence. However, as is, the phases will be off by a constant that depends on the interference patterns position on the camera. To get the phase, two reference states are needed. The first reference sets the phase up to conjugation. This is because a_{ij} cannot be uniquely identified as the peak on the left or the peak on the right. Thus the second reference is needed to determine if the DFT peak at Δ_{ij} is the coherence a_{ij} or its

conjugate a_{ij}^*. Any two different states whose coherences have a defined phase may be used as a reference.

5.4 Results

The tomography method is demonstrated with $N = 6$ dimensional states of the form $\boldsymbol{\rho}(\phi = 22.5°, \zeta, \Omega = 22.5°) = \boldsymbol{\rho}(\zeta)$ (i.e., all HWP are held fixed), whose coherences are a function of the QWP angle ζ. Fig. 5.4a shows the real and imaginary parts of the fifteen coherences (not including the conjugates). As the experimental setup is constrained in terms of the coherences it can produce, only several values for the real and imaginary parts are possible. Nevertheless, Fig. 5.4 shows good agreement between the experimental values (dots) and the theoretical values (dotted lines). Fig. 5.4b shows the experimental and theoretical density matrix for $\zeta = 11.25°$ in a side by side comparison.

Next, we calculate the state fidelity, $F(\boldsymbol{\rho}, \boldsymbol{\sigma}) = \mathrm{Tr}[\sqrt{\sqrt{\boldsymbol{\rho}}\boldsymbol{\sigma}\sqrt{\boldsymbol{\rho}}}]$, which is a measure of how close two density matrices $\boldsymbol{\rho}$, and $\boldsymbol{\sigma}$ are. The fidelity is bound between zero and unity, and $F(\boldsymbol{\rho}, \boldsymbol{\sigma}) = 1$ only if $\boldsymbol{\rho} = \boldsymbol{\sigma}$. We calculate the fidelity between the experimentally reconstructed state $\boldsymbol{\rho}(\phi, \zeta, \Omega)$ and the theoretically predicted one $\boldsymbol{\rho}_{\mathrm{th}}$. This is done for states where either HWP_ϕ, HWP_Ω, or the QWP_ζ are varied individually. If a HWP is held fixed it is held at 22.5°. If either HWP_ϕ or HWP_Ω are rotated then $\zeta = 45°$. Fig. 5.4 displays the results. The average fidelity for HWP_ϕ is 0.987 ± 0.001, 0.9893 ± 0.0008 for HWP_Ω, and 0.979 ± 0.001 for QWP_ζ, and the average over all data points is 0.9852 ± 0.0008.

Finally we measure the state purity $\mathrm{Tr}[\boldsymbol{\rho}^2]$. We introduce a series of HWP_θ-QWP-HWP_s-QWP. Here the QWP are 90° apart. The subscript s on HWP_s indicates that it is rapidly spinning in a mount. HWP_s spins faster than the collection time of the camera. This mixes the polarizations, which in turn mixes the spatial mode density matrix. HWP_θ controls the degree of mixing, where $\theta = 0°$ produces a pure state and $\theta = 22.5°$ is totally mixed. The purity of the resulting spatial state $\boldsymbol{\rho}(\theta)$ is $\mathrm{Tr}[\boldsymbol{\rho}^2] = 1/9(5{+}4\cos(4\theta))$. Note that spatial state is never entirely mixed, as the purity reaches a theoretical minimum at 5/9. Fig. 5.4 shows that the measured purity is agrees strongly with the theory.

We also note that the reconstruction method satisfies the unity trace and hermiticity conditions, but, like in most reconstruction methods, the state may not be positive

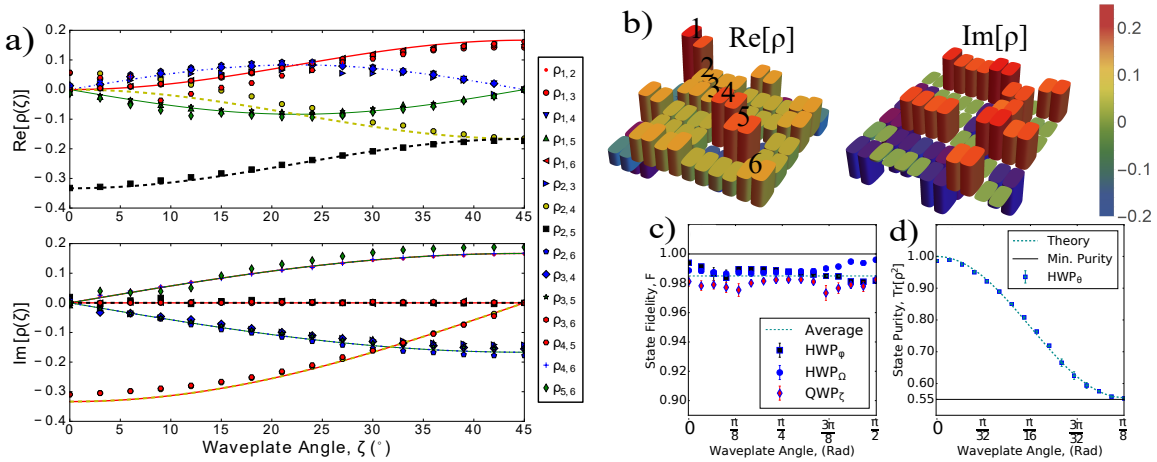

Figure 5.4: a) The experimental (dots) and theoretical (curve) coherences of the density matrix $\boldsymbol{\rho}$ (not including conjugate values), produced by rotating QWP_ζ in Fig. 5.2. As the coherences are constrained by the experimental setup, only a few unique values appear. As such, data points overlap. Note, error bars were omitted for clarity but were calculated to be $\mathcal{O}(10^{-2})$ for each data point. **b)** The experimentally reconstructed $d = 6$ state $\boldsymbol{\rho}(\zeta = 11.25°)$, with the Re[$\boldsymbol{\rho}$] on the left and Im[$\boldsymbol{\rho}$] on the right. Each diagram represents a 6×6 matrix, with theoretical elements to the right of experimental elements. The diagonal elements are labeled by their corresponding path label. The measured fidelity is 0.9911. **c)** The fidelity $F(\boldsymbol{\rho}, \boldsymbol{\rho}_{\text{th.}})$, as a function of angles for HWP_ϕ, HWP_Ω, or QWP_ζ shown in Fig. 5.2. The fidelity is close to unity, meaning $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\text{th.}}$ are nearly equal. Averaged over all points the fidelity is 0.9852 ± 0.0008 (dashed line). **d)** The purity of $\boldsymbol{\rho}(\theta)$ as a function of the angle of HWP_θ in Fig. 5.2. The dashed line is the theoretical purity and the solid line indicates the smallest purity attainable by mixing the state using this method.

semi-definite. This can be seen by the fact that the purity of our pure states is not one, and may in fact even exceed one. This is due to measurement uncertainties, which may arise from smearing of the visibility due to the lens, as well as movement of the beam or interference fringes due to phase instability. If there is considerable movement, averaging the images will also harm the result as the interference pattern can become washed out, resulting in a lower calculated visibility. Imposing positive semi-definitivity would require maximum-likelihood algorithms which perform parameter estimation to 'guess' the most likely physical state. Nevertheless, we see that, even without these likelihood algorithms, our calculated metrics, such as fidelity and purity, are still very high.

5.5 Summary

The above results show the tomography method is able to faithfully reconstruct several different metrics for a six dimensional density matrix. The reconstruction method works by extracting the coherences from interference pattern. To avoid having multiple signals with the same spatial frequency overlap, we utilize a cylindrical lens to separate the interference patterns spatially. In this sense the method uses both position space and k-space simultaneously to reconstruct the state. We emphasize that the above results reflect a proof of principle; the tomography method presented above can, in theory, accommodate much larger dimensions, given they conform to the physical constraints (e.g. pixel size, camera size, aberration etc.) discussed above.

In summary, we present a simple method to measure of large dimensional density matrices of path modes using a 1D-FT. We demonstrate the method experimentally by calculating several metrics for six dimensional matrices. While the above discussion is focused on spatial modes, it is simple to extend it to other degrees of freedom by first converting the desired degree of freedom to a spatial one (e.g. going from polarization/time/angular momentum encoding to path encoding). The method directly reconstructs the state by exploiting the physical manifestation of quantum coherence (i.e. interference).

6

Coherence in Sequential Measurements, Part II

Preamble

In this chapter, we conclude our investigation into the propagation of coherence through sequential measurement. In chapter 4, we saw that the pointer state density matrix of three measurements differs from the one predicted by the von Neumann model if we assume that each projective measurement collapses the wavefunction.To properly test this, we needed to devise a technique with which to measure a density matrix encoded in eight possible path modes that the photons may take. In chapter 5, we showed how this can be done using a rotating cylindrical lens. We now utilize the method and discuss the results.

6.1 Introduction

In chapter 4 we saw that a sequence of individual measurements is in some way fundamentally different from sequential measurements. In the former, the wavefunction collapses due to each projective measurement in the sequence. In the latter case the wavefunction does not collapse until the very end of the experiment, when the result is read off from the pointers. This allows coherence between pointer states to remain when there otherwise would have been none.

The difference between these two models, which we now refer to as the "collapse" model and the "von Neumann" models of sequential measurement, is present in the

three pointer density matrices obtained under each assumption, $\boldsymbol{\rho}_{\text{col}}$ and $\boldsymbol{\rho}_{\text{VN}}$, respectively.

$$\boldsymbol{\rho}_{\text{col}} = \frac{1}{8}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$\boldsymbol{\rho}_{\text{VN}} = \frac{1}{8}\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Recall that only the diagonals constitute probabilities, which is why $\boldsymbol{\rho}_{\text{VN}}$ is normalized by 1/8. We can also represent this difference pictorially as shown in Fig. 6.1. If the wavefunction collapses between each measurement then sequentially each subsequent measurement depends only on the result of the measurement prior to it. That is, no information from the measurement of $\boldsymbol{A}$ influences the result of measurement $\boldsymbol{C}$. Stated in yet another way, the sequential measurement process is Markovian. On the other hand, if the von Neumann model is correct, collapse only occurs at the end of the three measurements. Because of this, information from the first measurement may remain, influencing the final measurement result, which manifests itself as coherence between pointers.

The collapse versus non-collapse feature manifests itself in the lack of coherence between pointer states in $\boldsymbol{\rho}_{\text{col}}$. This difference is still somewhat abstract. Reconstructing the state matrix and showing it is equal to one of the two matrices can obscure a deeper meaning. We can rephrase the difference in another way. The von Neumann model predicts that, in a sequence of one or two measurements, if the system state is totally mixed (or in some sense classical), the pointer state is also totally mixed. The purity of the pointer density matrix is minimal when the purity of the system density matrix is minimal. However, when three (or more) measurements are present, the pointer

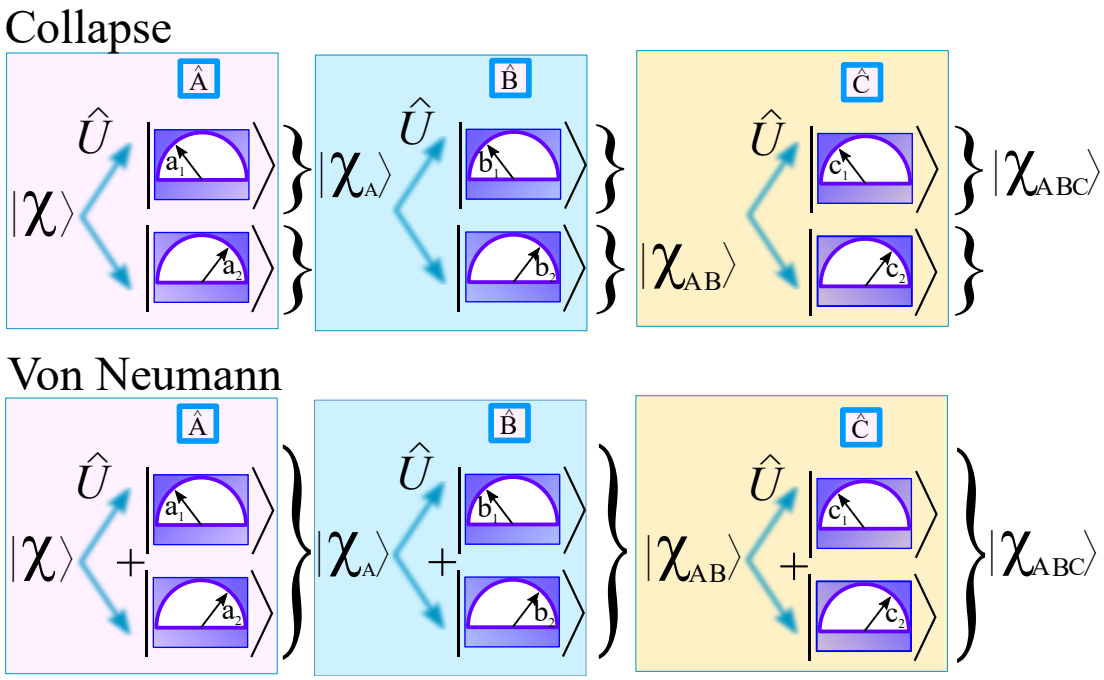

Figure 6.1: A pictorial representation of the difference between the collapse model and the prediction made by the von Neumann model. In the collapse model, the wavefunction collapses following each of the measurements in the sequence, producing an incoherent mixture of pointers, given by $\boldsymbol{\rho}_{\text{col}}$. The von Neumann model on the other hand predicts that superposition remains after each measurement. This allows information from the first measurement to persist to the last measurement.

states are not totally mixed, meaning that superposition is taking place somehow. In this case, the purity of the pointer density matrix is not minimal when the purity of the system density matrix is. We therefore reconstruct the pointer density matrix and calculate the purity, and check if the purity is minimal when the system state purity is minimal. This simplifies the experimental method as the purity depends only on the magnitude of the density matrix elements, and not the phase.

6.2 Experiment

The experimental setup is conceptually similar to the setup presented in chapter 5. Instead of a network of polarizing beam splitters and waveplates, we use beam-displacers to shift a polarized beam transversely based on its polarization. This is more convenient as the beams propagate parallel to one another. The first beam displacer (xtal_1)

shifts horizontally polarized light horizontally by $\delta_x = 2.7$ mm. This performs the measurement of $\boldsymbol{\pi}_H$. The second beam displacer (xtal_2) shifts vertically polarized light vertically by $\delta_y = 2.7$ mm as well. This crystal is preceded and followed by HWPs at 22.5° so that together they constitute a measurement of $\boldsymbol{\pi}_D$. Lastly, the final beam displacer (xtal_3) shifts horizontally polarized light by $\delta_X = 4$ mm horizontally, which is the final measurement of $\boldsymbol{\pi}_H$. The geometry of the path modes produced by this arrangement is $\{(0,0),(\delta_x,0),(0,\delta_y),(\delta_x,\delta_y),(\delta_X,0),(\delta_X,\delta_y),(\delta_X+\delta_x,0),(\delta_X+\delta_x,\delta_y)\}$. The crystals and geometry are shown in Fig. 6.2, as well as the coherences predicted by the von Neumann model, which are denoted by dashed blue lines between the coherent paths.

The diagonal elements can be obtained using the method presented in chapter 5. Either retrieve the diagonals from non-interfering paths after the 1D OFT, or divide the intensity of each path by the total intensity. Because the horizontal direction has repeated spacings (the left most two and right most two have a spacing of 2.7 mm as shown in Fig. 6.2), the Fourier method will not work as presented in chapter 5. The blue lines in Fig. 6.2 represent the expected coherences, and show that we should not expect to see interference. However, because the spacing is the same, we cannot distinguish between a lack of interference, and a combination of two interference patterns which are perfectly out of phase. As such we need to modify the method presented in chapter 5.

The coherences between vertically separated paths $\{a_{18}, a_{27}, a_{36}, a_{45}\}$ are easily recoverable by aligning the OFT-axis at 90° with respect to the horizontal. Coherences between horizontally separated path modes, such as a_{87} and a_{56} have the aforementioned problem of equal spacing. One solution is to block certain paths. For example, if the left-most two paths paths a_{11} and a_{88} are blocked, the issue of equal spacing is no longer present. This would require additional normalization by the probability for the photon to end up in the blocked paths. However, we can exploit the fact that the coherences between elements along the horizontal are expected to be zero. When blocking a_{11} and a_{88}, the remaining beams along the horizontal have unequal spacing. The only way that no interference occurs is if all the coherences are zero, at which point normalization does not matter. Even if the coherences are out of phase, they

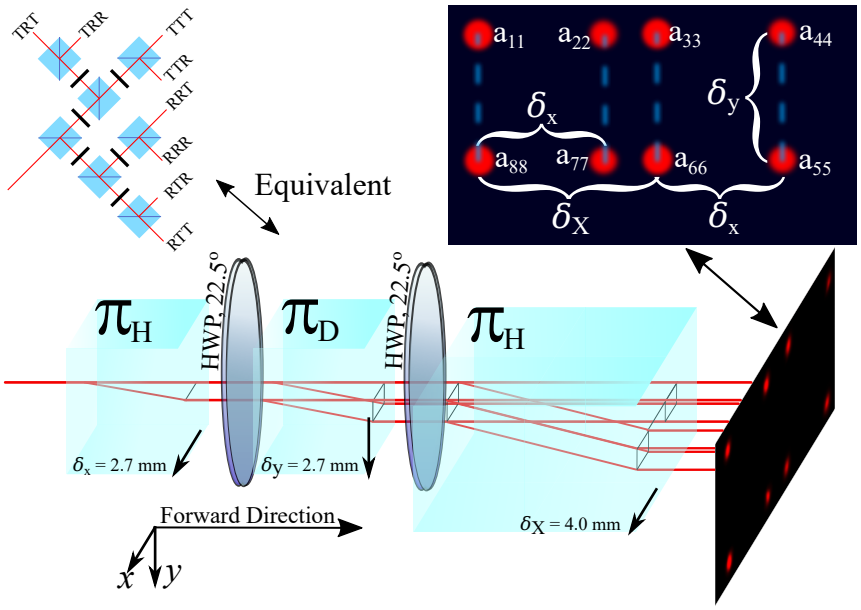

Figure 6.2: Implementation of a sequence of non-commuting measurements ($\boldsymbol{\pi}_H\boldsymbol{\pi}_D\boldsymbol{\pi}_H$) using a sequence of displacement crystals and waveplates. This is equivalent to the beam splitter network shown in chapter 3, but has the advantage of requiring fewer resources, and the resulting beams co-propagate. The co-propagating beams allow us to easily to perform the 'double-slit' interference needed for the reconstruction technique, but also provides robust phase stability between each path mode due to their proximity as well as sharing all the same optics, so vibration is not an issue. The first crystal performs $\boldsymbol{\pi}_H$ by shifting $|H\rangle$ polarized light along the x direction by $\delta_x = 2.7$ mm. The second crystal performs $\boldsymbol{\pi}_D$ by shifting $|V\rangle$ polarized light along the y direction by $\delta_y = 2.7$ mm. The last crystal performs $\boldsymbol{\pi}_H$ by shifting $|H\rangle$ polarized light along the x direction by $\delta_X = 4.0$ mm. The resulting geometry for the eight beam paths is shown. If the von Neumann model is correct and the measurements do not fully collapse the wavefunction, then one would expect to have coherence between the vertically displaced paths, as indicated by the blue dashed lines.

will not cancel out because they have different frequencies. Of course, if significant[1] interference is seen, we would need to take into account the probability for the photon to end up in the blocked path.

We use the technique presented in the chapter 8 supplementary information to generate mixed states. A polarized input state $|\theta\rangle$ is passed through a rapidly spinning HWP sandwiched between two quarter-waveplates (QWP). This produces the state

$$\boldsymbol{\rho}_{\mathcal{S}} = \begin{pmatrix} \sin\theta^2 & C\Omega/\omega \\ C^*\Omega/\omega & \cos\theta^2 \end{pmatrix},$$

where $|C| \leq \sin\theta\cos\theta$, ω is angular velocity of the spinning waveplate, and Ω is the camera integration time. If $\omega \gg \Omega$, the resulting state is mixed: $\boldsymbol{\rho}_{\mathcal{S}}(\theta) = \sin^2(\theta)\,|H\rangle\langle H| + \cos^2(\theta)\,|V\rangle\langle V|$. This state produces path modes whose intensities depend on the angle θ. Because the analysis code works more easily for non-zero intensities, we transform the state to

$$\boldsymbol{\rho}_{\mathcal{S}} = \frac{1}{2}\begin{pmatrix} 1 & -\cos 2\theta \\ -\cos 2\theta & 1 \end{pmatrix},$$

using a HWP at 22.5°. The resulting paths have equal and constant intensities, making the reconstruction method easier to perform. The full experimental setup is shown in Fig. 6.3, as well as a measurement of the state $\boldsymbol{\rho}_{\mathcal{S}}$ showing the mixed state generation works.

[1]Even if there is zero coherence the reconstruction technique will return some small value due to error, this is What is meant by 'significant'. These errors occur due to imperfections in beams splitters and displacement crystals. As well, faint (i.e. low visibility) high frequency interference can be discerned in the images and is believed to be due to the films on the surface of the CMOS camera.

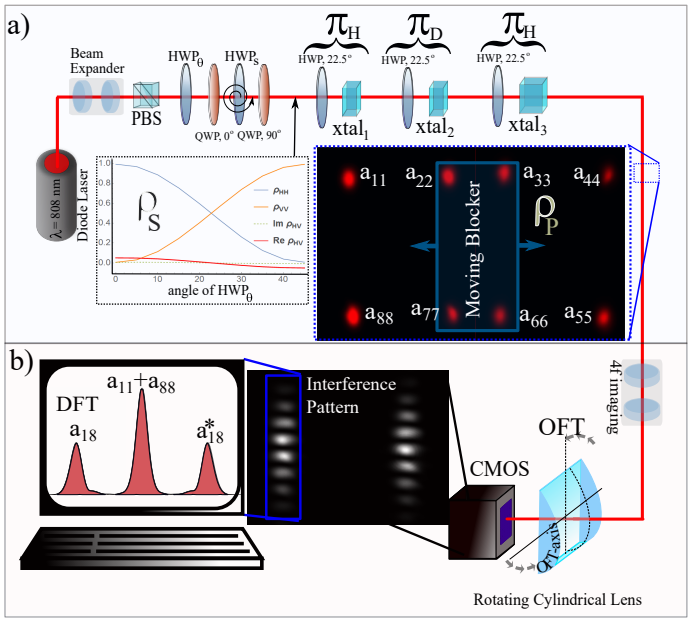

Figure 6.3: Experimental setup to test the differences predicted by the collapse model and the von Neumann model. **a)**. State preparation and sequential measurements: A beam expander sets the Rayleigh length of a 808 nm diode laser so that all the optics are within it. A set of HWP and QWP generates a mixed state by rapidly spinning HWP$_s$ faster than the collection time of the camera. The purity of the state is a function of the angle of HWP$_\theta$, and the measured mixed state is plotted in a graph. A series of displacement crystals (xtal$_{1,2,3}$) and half waveplates at 22.5°, perform a sequence of non-commuting measurements $\boldsymbol{\pi}_H\boldsymbol{\pi}_D\boldsymbol{\pi}_H$. The resulting eight dimensional path encoded density matrix is reconstructed. Repeated spacings between paths produce oscillation patterns with equal frequency, which therefore overlap on the Fourier transform. To avoid this problem we place a move-able beam block and reconstruct the state in chunks. **b)** State Reconstruction: A $4f$ system ($f_1 = 1000$ mm and $f_2 = 400$ mm) images the six path modes onto a CMOS camera. A rotating cylindrical lens ($f = 250$ mm) performs the optical Fourier transform (OFT) along the OFT-axis, and the coherences are obtained by discrete FT (DFT) the interference patterns on a computer.

6.3 Results

We define a normalized purity of a state $\boldsymbol{\rho}$ to be $\mathrm{P}_N(\boldsymbol{\rho}) = \mathrm{Purity}(\boldsymbol{\rho})/\mathrm{Purity}(\boldsymbol{\rho})_{\min}$, such that P_N is equal to 1 for a totally mixed state. In Fig. 6.4 we plot $\mathrm{P}_N(\boldsymbol{\rho}_\mathcal{P})$ against $\mathrm{P}_N(\boldsymbol{\rho}_\mathcal{S})$ for the case of one, two and three measurements in a sequence. We see that for a 'sequence' of one measurement, $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{P}_1})$ is always equal to one, as expected. Similarly for two measurements, we see that $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2}) = 1$ when $\mathrm{P}_N(\boldsymbol{\rho}_\mathcal{S}) = 1$. In this

case the pointer states may also have coherence if the system has coherence. Finally, for three measurements, we see that $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3}) \neq 1$ when $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{S}}) = 1$, meaning that the pointers have coherence when the system does not. This result supports the prediction made by the von Neumann model and goes against the idea that the wavefunction collapses between each measurement.

6.4 Discussion

The result of the experiment shows that in a sequence of three measurement, coherence is retained between the pointer states. This result is incompatible with the notion that the wavefunction collapses after each measurement. This implies that the result of the first measurement may influence the result of the last measurement, even though the non-commuting nature of the measurements implies that all information of the previous result should be erased.

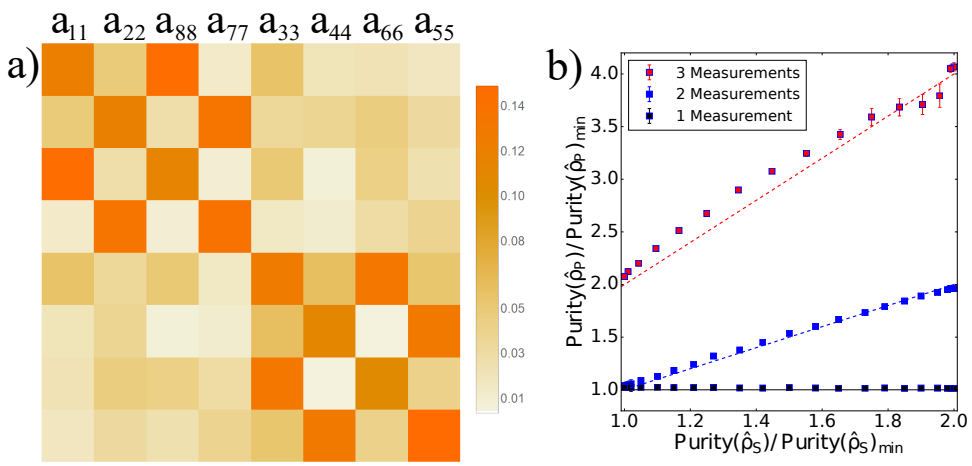

Figure 6.4: a) Reconstruction of the absolute value of the three pointer density matrix obtained with the input state $\boldsymbol{\rho}_{\mathcal{S}} = \mathbf{1}/2$. The reconstruction clearly shows coherences in the off-diagonals which is not predicted by the collapse model of sequential measurement. The labels on the top indicate which path mode the matrix element belongs to. Off-diagonals are the coherences between said path modes. **b)** The normalized purity, $\mathrm{P}_N(\boldsymbol{\rho})$ of the pointer density matrix $\boldsymbol{\rho}_{\mathcal{P}}$ for one, two, and three measurements as a function of the normalized purity of the quantum system density matrix. For two and three measurements we see that when $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{S}}) = 1$ (i.e. $\boldsymbol{\rho}_{\mathcal{S}} = \mathbf{1}/2$), then $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{P}}) = 1$, which is consistent with the collapse model. But for three measurements, $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{P}}) \neq 1$ when $\mathrm{P}_N(\boldsymbol{\rho}_{\mathcal{S}}) = 1$, which is inconsistent with the collapse model.

The result also suggests that it is wrong to treat sequential measurement on a single quantum particle through continuous application of the Born rule. Indeed, the questionable application of the Born rule is starting to gain more attention as sequential measurements become more popular (86). One might wonder how does a sequence of three beamsplitters, separated by waveplates, fit into this discussion? Is this not sequential measurement? In fact, it is more appropriate to say that this is a sequence of three individual measurements. This is because one only considers one output port of each beamsplitter. Measurements are performed conditionally on the outcome of the previous measurement. As such, the experimenter has access to the result of each individual measurement, in turn collapsing the wavefunction.

An interesting point is that the discrepancy between the collapse model and the von Neumann model only manifests in sequences of three (or more) measurements. For the case of one and two measurements, the two models agree (i.e. the separate pointer states are mutually incoherent and follow the Born rule). This is surprising given that we saw a similar discrepancy between two and three measurements in the time ordering of weak measurements experiment. The true significance of this is not entirely clear. Some possible interpretations may be that the first and last measurements act as a kind of boundary conditions. The middle measurement is then somehow determined by both these 'past' and 'future' conditions. This is in some sense not too dissimilar from the two state vector formalism popularized by Aharonov and Vaidman, which introduces a time-symmetrized approach to standard quantum mechanics (87).

At the moment the true implications of the result are speculative. Perhaps these results hold some insight on the ontological nature of the wavefunction, as it is not the beamsplitters which collapse the wavefunction, but rather our access to the result of the measurements. To me, this suggests the wavefunction is a bookkeeping tool, however the authors of the original theory paper take it as evidence for the opposite interpretation (i.e. the wavefunction is real) (8). Perhaps the results hold some significance as a microcosm of quantum measurements, with the first measurement acting as the state preparation, middle acting as the process in the lab, and the last measurement being the readout. Despite all the different possible speculative routes one may pursue, what we can claim for certain is that the coherences found in the pointer density matrix are not compatible with the traditional collapse model of measurement, which is a foundational pillar for quantum mechanics.

7

Conclusion

7.1 Summary

Despite the theoretical and technological triumphs that researchers have had in the field of quantum mechanics, shockingly little is still understood about some of the most foundational aspects of the theory. This thesis was aimed at specifically studying foundational problems, particularly in the area of quantum measurement. We did this through the metaphorical lens of sequential measurements, with hopes to understand the properties of the collapse of the wavefunction. We did this in two regimes; In the first, we use sequences of weak measurements to study the related concept of time reversibility of quantum measurements. Quantum measurements are famously time asymmetric, the dynamics going forward and backward are not the same, which runs counter to the time symmetry found in the rest of physics. Weak measurements reveal almost no information about the outcome of a single measurement, but as a consequence they do not disturb the wavefunction either. Therefore, we test to see if in the weak measurement limit whether or not time asymmetry persists. We find that making the measurements does not in fact recover time symmetry. This makes time asymmetry a fundamental aspect of the quantum measurement. However, we also see that this asymmetry manifests in interesting ways. In particular, we saw that time symmetry *is* present for a sequence of two non-commuting measurements, but not for a sequence of three (or potentially more) measurements.

In the second project we looked at sequences of strong measurements. Here we test a prediction made by the von Neumann model, as pointed out in (8). This prediction

claims that, because you do not gain the experimental results until the end of the sequence, the wavefunction does not need to collapse between successive non-commuting measurements. This allows for correlations to exist between pointer states that would not be there if one assumed that the wavefunction collapsed between each successive measurement. We develop a new way of performing state reconstruction based off using a one-dimensional Fourier transform performed by a cylindrical lens, and show that the prediction made in (8) is indeed correct; coherences do in fact persist between the sequence of strong measurements.

The measurement problem, and its related problems such as that of the time-asymmetry of measurement, remains one of biggest conceptual hurdles of quantum mechanics. Researchers are able to make great strides using the Copenhagen interpretation of measurement, which was famously paraphrased by Nathaniel David Mermin as "shut up and calculate". However, the fact that such a fundamental concept is still so poorly understood is a black mark on the otherwise excellent predictive power that the theory holds. Perhaps one of the most interesting questions is, how is it possible that the mathematical structure is so stable given its foundation is so shaky? While there is still much work left to be done, we are hopeful that combining theoretical and experimental research into the topic may open new doors and finally allow us to resolve this problem.

7.2 Future Work

Here I compile a small list of possible avenues for future work with respect to studying sequential measurements and discuss them.

1. The content presented in the collapse vs. Von Neumann model of sequential measurement is only half of the topic. The work shown in (8) shows how the collapse model is retrieved when considering amplified measurements. These can be thought of measurement witnesses: pointers that confirm the result of a measurement. In some sense this is how a real measurement apparatus works. All the particles that constitute the measurement apparatus must all confirm the result of the measurement. It cannot be the case that half the particles disagree with the other half. The collapse model is recovered if one loses track of these measurement witnesses. One can imagine when an apparatus is composed of $O(10^{23})$

measurement witnesses (i.e. particles), it is easy to lose track of them. As we saw in chapter 3, losing track of a particle corresponds to tracing over subsystems of the total Hilbert space, which in turn decoheres the system.

The preliminary part of this part of the experiment has been started. The difficult part has been to introduce a measurement witness, which requires additional degrees of freedom. However, we cannot simply introduce new paths as we also need to be able to easily 'ignore' of this degree of freedom, which is difficult to do with paths. One obvious method is to use the time degree of freedom, the pointer is the temporal distribution, and the shifts are implemented with delay lines. Losing track of this degree of freedom is accomplished because the camera is slow, and cannot resolve these differences in time.

2. We would also like to implement the experiment using single photons instead of a laser beam. While the results are not expected to change, the experimental data will be perhaps more convincing as there is no avenue for coherence to potentially propagate using some unknown means.

3. To make the states reconstructed from the method introduced in chapter 5 physical, a likelihood algorithm is needed. Popular likelihood algorithms are available for public use. However, these are written in terms of coincidence counts of single photon detectors. As such, a maximum likelihood parameter estimation program would need to be written from the ground up to accommodate the presented method.

4. Working with the von Neumann model has made it clear that photons have frustratingly few degrees of freedom that are easy to work with. Spatial, polarization and time are common degrees of freedom, however, if one wants to do longer sequences of measurements, more degrees of freedom are needed. Energy is another degree of freedom, but it is notoriously difficult to work with. As such, I am always on the look out for new methods of utilizing the different internal or external degrees of freedom of photons. For example, a few proposals have looked at utilizing the number statistics of the photons are a possible pointer (e.g. (88)).

8

Supplementary Information

8.1 Mixed State Generation

Each of the above projects required the generation of mixed states. Here I discuss how we produce the statistical mixture states. In order to do this, we place a rapidly spinning half wave plate, sandwiched between two quarter waveplates set 90 degrees apart from one another. The first quarter waveplate is at 45 degrees. This can, and the second is 90 degrees away. This is placed after the initial state preparation, which produces the state ψ. The rate at which the half-wave plate spins (ω), is made to be much much greater than the rate of collection of the camera (Ω).

A waveplate adds a phase ϕ to a state perpendicular to its fast axis. Representing the waveplate as an operator $\boldsymbol{U}_\phi$, the operator leaves states aligned with its fast axis, denoted as $|\theta\rangle = \cos\theta\,|H\rangle + \sin\theta\,|V\rangle$ unchanged, and adds a phase $e^{i\phi}$ to the perpendicular state $\left|\theta^\perp\right\rangle = |\theta + \pi/2\rangle$. As such the waveplate operator can be written as

$$\boldsymbol{U}_\phi(\theta) = |\theta\rangle\langle\theta| + e^{i\phi}\left|\theta^\perp\right\rangle\left\langle\theta^\perp\right|. \tag{8.1}$$

Therefore, a half waveplate, with its fast axis orientated at θ with respect to the horizontal, $\boldsymbol{U}_{\lambda/2}(\theta) = \boldsymbol{U}_{\phi=\pi}(\theta)$ can be written as

$$\boldsymbol{U}_{\lambda/2}(\theta) = |\theta\rangle\langle\theta| - \left|\theta^\perp\right\rangle\left\langle\theta^\perp\right|.$$

Similarly, a quarter waveplate, with its fast axis orientated at θ with respect to the horizontal, $\boldsymbol{U}_{\lambda/4}(\theta) = \boldsymbol{U}_{\phi=\pi/2}(\theta)$ can be written as

$$\boldsymbol{U}_{\lambda/4}(\theta) = |\theta\rangle\langle\theta| + i\left|\theta^\perp\right\rangle\left\langle\theta^\perp\right|.$$

Denoting $\boldsymbol{U}^{(1)}_{\lambda/4}$ and $\boldsymbol{U}^{(2)}_{\lambda/4}$ as the first and second quarter waveplates, and $\boldsymbol{U}_{\lambda/2}(\omega t)$ as the spinning waveplate, we get that the final state is related to the initial state $\boldsymbol{\rho} = |\theta\rangle \langle\theta|$ by

$$\boldsymbol{\rho}_f = \frac{\Omega}{2\pi} \int_0^{\frac{2\pi}{\Omega}} \boldsymbol{U}^{(2)}_{\frac{\lambda}{4}}(5\pi/4)\boldsymbol{U}_{\frac{\lambda}{2}}(\omega t)\boldsymbol{U}^{(1)}_{\frac{\lambda}{4}}(\pi/4)\hat{\rho}\boldsymbol{U}^{(1)}_{\frac{\lambda}{4}}(\pi/4)^{\dagger}\boldsymbol{U}^{\dagger}_{\frac{\lambda}{2}}(\omega t)\boldsymbol{U}^{(2)}_{\frac{\lambda}{4}}(5\pi/4)^{\dagger}\mathrm{d}t.$$

Written in the H, V basis, the result is

$$\boldsymbol{\rho}_f = \begin{pmatrix} \sin\theta^2 & C\Omega/\omega \\ C^*\Omega/\omega & \cos\theta^2 \end{pmatrix},$$

where C is some complex number such that $|C|\Omega/\omega \leq \sin\theta\cos\theta$, by positive-definitivity of $\boldsymbol{\rho}$. We see that the faster the waveplate spins, the smaller the off-diagonal terms become, while the diagonal terms are only effected by the angle of the first waveplate.

References

[1] The Economist. **Quantum technology is beginning to come into its own**, 2017. 1

[2] DN Basov, RD Averitt, and D Hsieh. **Towards properties on demand in quantum materials**. *Nature materials*, **16**(11):1077, 2017. 1

[3] Graeme Mitchison, Richard Jozsa, and Sandu Popescu. **Sequential weak measurement**. *Physical Review A*, **76**(6):062105, 2007. 2, 15

[4] Jeff S Lundeen, Brandon Sutherland, Aabid Patel, Corey Stewart, and Charles Bamber. **Direct measurement of the quantum wavefunction**. *Nature (London)*, **474**(7350):188–191, 2011. 2, 19

[5] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen. **Direct Measurement of the Density Matrix of a Quantum System**. *Phys. Rev. Lett.*, **117**:120401, Sep 2016. 2, 5, 19

[6] Adam Bednorz, Kurt Franke, and Wolfgang Belzig. **Noninvasiveness and time symmetry of weak measurements**. *New J. Phys.*, **15**(2):023043, 2013. 2, 5, 16, 17, 18, 19, 27

[7] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao. **Measurement of the single-photon tunneling time**. *Phys. Rev. Lett.*, **71**:708–711, Aug 1993. 2, 18

[8] Jennifer R Glick and Christoph Adami. **Markovian and non-Markovian quantum measurements**. *arXiv preprint arXiv:1701.05636*, 2017. 2, 5, 43, 46, 73, 74, 75

[9] Werner Heisenberg. **The physical principles of quantum mechanics**. *U. Chicago Press, Chicago*, **21**, 1930. 3

[10] PAM Dirac. **The Principles of Quantum Mechanics (Clarendon, Oxford)**, 1947. 3

[11] John Von Neumann. *Mathematical foundations of quantum mechanics*. Number 2. Princeton university press, 1955. 3

[12] Gerhart Lüders. **Concerning the state-change due to the measurement process**. *Annalen der Physik*, **15**(9):663–670, 2006. 4

[13] Karl Kraus. *States, effects and operations: fundamental notions of quantum theory*. Springer, 1983. 4

[14] Roman Schnabel. **Squeezed states of light and their applications in laser interferometers**. *Physics Reports*, **684**:1–51, 2017. 4

[15] Yakir Aharonov, David Z. Albert, and Lev Vaidman. **How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100**. *Phys. Rev. Lett.*, **60**:1351–1354, Apr 1988. 4, 9, 19

[16] J Abadie, BP Abbott, R Abbott, TD Abbott, M Abernathy, C Adams, R Adhikari, C Affeldt, B Allen, GS Allen, et al. **A gravitational wave observatory operating beyond the quantum shot-noise limit**. *Nature Physics*, **7**(12):962, 2011. 4

[17] Levon Chakhmakhchyan, Stéphane Guérin, Joshua Nunn, and Animesh Datta. **Compact entanglement distillery using realistic quantum memories**. *Phys. Rev. A*, **88**:042312, Oct 2013. 4

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. **Evidence for the 2π Decay of the K_2^0 Meson**. *Phys. Rev. Lett.*, **13**:138–140, Jul 1964. 4

[19] Gonzalo Muga, R Sala Mayato, and Inigo Egusquiza. *Time in quantum mechanics*, **734**. Springer Science & Business Media, 2007. 5, 18

[20] Jan Hilgevoord. **Time in quantum mechanics: a story of confusion**. *Stud. Hist. Philos. Mod. Phys.*, **36**(1):29–60, 2005. 5, 18

[21] W Pauli. **Encyclopedia of Physics, vol. 5**. *Encyclopedia of Physics*, **5**, 1958. 5, 18

[22] Kurt Franke, Adam Bednorz, and Wolfgang Belzig. **Time asymmetry in weak measurements**. *Phys. Scripta*, **2012**(T151):014013, 2012. 5, 16, 19

[23] Lajos Diósi. **Structural features of sequential weak measurements**. *Phys. Rev. A.*, **94**(1):010103, 2016. 5, 26, 31

[24] Michel H Devoret, Andreas Wallraff, and John M Martinis. **Superconducting qubits: A short review**. *arXiv preprint cond-mat/0411174*, 2004. 7

[25] Barry C. Sanders. **Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement**. *Phys. Rev. A*, **40**:2417–2427, Sep 1989. 9

[26] D. T. Smithey, M. Beck, J. Cooper, and M. G. Raymer. **Measurement of number-phase uncertainty relations of optical fields**. *Phys. Rev. A*, **48**:3159–3167, Oct 1993. 9

[27] John Von Neumann. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton university press, 2018. 10, 37

[28] Howard M Wiseman and Gerard J Milburn. *Quantum measurement and control*. Cambridge university press, 2009. 11, 19

[29] Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz. **Time Symmetry in the Quantum Process of Measurement**. *Phys. Rev.*, **134**:B1410–B1416, Jun 1964. 16, 18

[30] Dominic Horsman, Chris Heunen, Matthew F. Pusey, Jonathan Barrett, and Robert W. Spekkens. **Can a quantum state over time resemble a quantum state at a single time?** *Proc. R. Soc. A*, **473**(2205), 2017. 18, 27

[31] EH Hauge and JA Støvneng. **Tunneling times: a critical review**. *Rev. Mod. Phys.*, **61**(4):917, 1989. 18

[32] Juan Gonzalo Muga and C Richard Leavens. **Arrival time in quantum mechanics**. *Phys. Rep.*, **338**(4):353–438, 2000. 18

[33] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. **Quantum time**. *Phys. Rev. D*, **92**:045033, Aug 2015. 18

[34] Jeff Z Salvail, Megan Agnew, Allan S Johnson, Eliot Bolduc, Jonathan Leach, and Robert W Boyd. **Full characterization of polarization states of light via direct measurement**. *Nat. Photonics*, **7**(4):316–321, 2013. 19

[35] Yong-Su Kim, Jong-Chan Lee, Osung Kwon, and Yoon-Ho Kim. **Protecting entanglement from decoherence using weak measurement and quantum measurement reversal**. *Nature Physics*, **8**(2):117, 2012. 19

[36] S. E. Ahnert and M. C. Payne. **Weak measurement of the arrival times of single photons and pairs of entangled photons**. *Phys. Rev. A*, **69**:042103, Apr 2004. 19

[37] J. S. Lundeen and A. M. Steinberg. **Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox**. *Phys. Rev. Lett.*, **102**:020404, Jan 2009. 19

[38] Justin Dressel, Mehul Malik, Filippo M. Miatto, Andrew N. Jordan, and Robert W. Boyd. **Colloquium**. *Rev. Mod. Phys.*, **86**:307–316, Mar 2014. 19

[39] Graeme Mitchison, Richard Jozsa, and Sandu Popescu. **Sequential weak measurement**. *Phys. Rev. A*, **76**:062105, Dec 2007. 19

[40] Greg A. Smith, Souma Chaudhury, Andrew Silberfarb, Ivan H. Deutsch, and Poul S. Jessen. **Continuous Weak Measurement and Nonlinear Dynamics in a Cold Spin Ensemble**. *Phys. Rev. Lett.*, **93**:163602, Oct 2004. 19

[41] Justin Dressel, Areeya Chantasri, Andrew N. Jordan, and Alexander N. Korotkov. **Arrow of Time for Continuous Quantum Measurement**. *Phys. Rev. Lett.*, **119**:220507, Dec 2017. 19

[42] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese. **Measuring Incompatible Observables by Exploiting Sequential Weak Values**. *Phys. Rev. Lett.*, **117**:170402, Oct 2016. 19

[43] J.S. Lundeen and K.J. Resch. **Practical measurement of joint weak values and their connection to the annihilation operator**. *Phys. Lett.*, **334**(5):337 – 344, 2005. 20

[44] Iwo Białynicki-Birula. **On the wave function of the photon**. *Acta Physica Polonica A*, **1**(86):97–116, 1994. 21

[45] A. J. Leggett and Anupam Garg. **Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?** *Phys. Rev. Lett.*, **54**:857–860, Mar 1985. 24

[46] ME Goggin, MP Almeida, Marco Barbieri, BP Lanyon, JL OBrien, AG White, and GJ Pryde. **Violation of the Leggett–Garg inequality with weak measurements of photons**. *Proc. Nat. Acad. Sci. U.S.A*, **108**(4):1256–1261, 2011. 24

[47] Nathan S. Williams and Andrew N. Jordan. **Weak Values and the Leggett-Garg Inequality in Solid-State Qubits**. *Phys. Rev. Lett.*, **100**:026804, Jan 2008. 24

[48] Alessio Avella, Fabrizio Piacentini, Michelangelo Borsarelli, Marco Barbieri, Marco Gramegna, Rudi Lussana, Federica Villa, Alberto Tosi, Ivo Pietro Degiovanni, and Marco Genovese. **Anomalous weak values and the violation of a multiple-measurement Leggett-Garg inequality**. *Phys. Rev. A*, **96**:052123, Nov 2017. 24

[49] Eleonora Nagali, Simone Felicetti, Pierre-Louis de Assis, Vincenzo D'Ambrosio, Radim Filip, and Fabio Sciarrino. **Testing sequential quantum measurements: how can maximal knowledge be extracted?** *Scientific reports*, **2**:443, 2012. 24

[50] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan. **Experimental Violation of Two-Party Leggett-Garg Inequalities with Semiweak Measurements**. *Phys. Rev. Lett.*, **106**:040402, Jan 2011. 24

[51] Aephraim M. Steinberg. **How Much Time Does a Tunneling Particle Spend in the Barrier Region?** *Phys. Rev. Lett.*, **74**:2405–2409, Mar 1995. 26

[52] J. Dressel and A. N. Jordan. **Significance of the imaginary part of the weak value**. *Phys. Rev. A*, **85**:012107, Jan 2012. 26

[53] Katja Ried, Megan Agnew, Lydia Vermeyden, Dominik Janzing, Robert W Spekkens, and Kevin J Resch. **A quantum advantage for inferring causal structure**. *Nature Physics*, **11**(5):414–420, 2015. 27

[54] L. Landau. **Das Dämpfungsproblem in der Wellenmechanik**. *Zeitschrift für Physik*, **45**(5):430–441, May 1927. 37

[55] Kristian Camilleri. **A history of entanglement: Decoherence and the interpretation problem**. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **40**(4):290–302, 2009. 43

[56] Matteo GA Paris and Massimiliano F Sacchi. **Quantum tomography**. *Advances in Imaging and Electron Physics*, **128**:206–309, 2003. 50

[57] JS Lundeen, A Feito, H Coldenstrodt-Ronge, KL Pregnell, Ch Silberhorn, TC Ralph, J Eisert, MB Plenio, and IA Walmsley. **Tomography of quantum detectors**. *Nature Physics*, **5**(1):27, 2009. 50

[58] Daniel FV James, Paul G Kwiat, William J Munro, and Andrew G White. **On the measurement of qubits**. In *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, pages 509–538. World Scientific, 2005. 50

[59] L. Rippe, B. Julsgaard, A. Walther, Yan Ying, and S. Kröll. **Experimental quantum-state tomography of a solid-state qubit**. *Phys. Rev. A*, **77**:022307, Feb 2008. 50

[60] Giacomo M D'Ariano, Chiara Macchiavello, and Matteo GA Paris. **A fictitious photons method for tomographic imaging**. *Optics communications*, **129**(1-2):6–12, 1996. 51

[61] William K Wootters and Brian D Fields. **Optimal state-determination by mutually unbiased measurements**. *Annals of Physics*, **191**(2):363–381, 1989. 51

[62] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd. **Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures**. *Phys. Rev. X*, **5**:041006, Oct 2015. 51

[63] H Bechmann-Pasquinucci and Wolfgang Tittel. **Quantum cryptography using larger alphabets**. *Physical Review A*, **61**(6):062308, 2000. 51

[64] Nicolas J Cerf, Mohamed Bourennane, Anders Karlsson, and Nicolas Gisin. **Security of quantum key distribution using d-level systems**. *Physical Review Letters*, **88**(12):127902, 2002. 51

[65] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. **Experimental realization of any discrete unitary operator**. *Phys. Rev. Lett.*, **73**:58–61, Jul 1994. 51

[66] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer, and Ian A. Walmsley. **Optimal design for universal multiport interferometers**. *Optica*, **3**(12):1460–1465, Dec 2016. 51

[67] Jonathan CF Matthews, Alberto Politi, André Stefanov, and Jeremy L O'brien. **Manipulation of multiphoton entanglement in waveguide quantum circuits**. *Nature Photonics*, **3**(6):346, 2009. 51

[68] Nobuyuki Matsuda, Hanna Le Jeannic, Hiroshi Fukuda, Tai Tsuchizawa, William John Munro, Kaoru Shimizu, Koji Yamada, Yasuhiro Tokura, and Hiroki Takesue. **A monolithically integrated polarization entangled photon pair source on a silicon chip**. *Scientific reports*, **2**:817, 2012. 51

[69] Benjamin J Metcalf, Nicholas Thomas-Peter, Justin B Spring, Dmytro Kundys, Matthew A Broome, Peter C Humphreys, Xian-Min Jin, Marco Barbieri, W Steven Kolthammer, James C Gates, et al. **Multiphoton quantum interference in a multiport integrated photonic device**. *Nature communications*, **4**:1356, 2013. 51

[70] Alberto Peruzzo, Anthony Laing, Alberto Politi, Terry Rudolph, and Jeremy L O'brien. **Multimode quantum interference of photons in multiport integrated devices**. *Nature communications*, **2**:224, 2011. 51

[71] Nicolò Spagnolo, Enrico Maiorino, Chiara Vitelli, Marco Bentivegna, Andrea Crespi, Roberta Ramponi, Paolo Mataloni, Roberto Osellame, and Fabio Sciarrino. **Learning an unknown transformation via a genetic approach**. *Scientific Reports*, **7**(1):14316, 2017. 51

[72] Ayman F. Abouraddy, Timothy M. Yarnall, and Bahaa E. A. Saleh. **Generalized optical interferometry for modal analysis in arbitrary degrees of freedom**. *Opt. Lett.*, **37**(14):2889–2891, Jul 2012. 51

[73] Kumel H Kagalwala, H Esat Kondakci, Ayman F Abouraddy, and Bahaa EA Saleh. **Optical coherency matrix tomography**. *Scientific reports*, **5**:15333, 2015. 52

[74] Michael A. Horne, Abner Shimony, and Anton Zeilinger. **Two-particle interferometry**. *Phys. Rev. Lett.*, **62**:2209–2212, May 1989. 52

[75] Marian O Scully, Berthold-Georg Englert, and Herbert Walther. **Quantum optical tests of complementarity**. *Nature*, **351**(6322):111–116, 1991. 52

[76] L. Mandel. **Coherence and indistinguishability**. *Opt. Lett.*, **16**(23):1882–1883, Dec 1991. 52

[77] Lan-Tian Feng, Ming Zhang, Zhi-Yuan Zhou, Ming Li, Xiao Xiong, Le Yu, Bao-Sen Shi, Guo-Ping Guo, Dao-Xin Dai, Xi-Feng Ren, et al. **On-chip coherent conversion of photonic quantum entanglement between different degrees of freedom**. *Nature communications*, **7**:11985, 2016. 58

[78] J. Brendel, N. Gisin, W. Tittel, and H. Zbinden. **Pulsed Energy-Time Entangled Twin-Photon Source for Quantum Communication**. *Phys. Rev. Lett.*, **82**:2594–2597, Mar 1999. 58

[79] Connor Kupchak, Philip J. Bustard, Khabat Heshami, Jennifer Erskine, Michael Spanner, Duncan G. England, and Benjamin J. Sussman. **Time-bin-to-polarization conversion of ultrafast photonic qubits**. *Phys. Rev. A*, **96**:053812, Nov 2017. 58

[80] Wuhong Zhang, Qianqian Qi, Jie Zhou, and Lixiang Chen. **Mimicking Faraday Rotation to Sort the Orbital Angular Momentum of Light**. *Phys. Rev. Lett.*, **112**:153601, Apr 2014. 58

[81] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. **Linear optical quantum computing with photonic qubits**. *Rev. Mod. Phys.*, **79**:135–174, Jan 2007. 58

[82] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. **A scheme for efficient quantum computation with linear optics**. *nature*, **409**(6816):46, 2001. 58

[83] Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Żukowski. **Multiphoton entanglement and interferometry**. *Reviews of Modern Physics*, **84**(2):777, 2012. 58

[84] N. Bergamasco, M. Menotti, J. E. Sipe, and M. Liscidini. **Generation of Path-Encoded Greenberger-Horne-Zeilinger States**. *Phys. Rev. Applied*, **8**:054014, Nov 2017. 58

[85] Jeremy L O'brien, Akira Furusawa, and Jelena Vučković. **Photonic quantum technologies**. *Nature Photonics*, **3**(12):687, 2009. 58

[86] Sally Shrapnel, Fabio Costa, and Gerard Milburn. **Updating the Born rule**. *New Journal of Physics*, **20**(5):053010, 2018. 73

[87] Yakir Aharonov and Lev Vaidman. **The two-state vector formalism: an updated review**. In *Time in quantum mechanics*, pages 399–447. Springer, 2008. 73

[88] Bertúlio de Lima Bernardo. **Photon statistics as a probe for weak measurements**. *JOSA B*, **31**(7):1494–1498, 2014. 76