

Generalized Measurement and Post-selection  
in Optical Quantum Information

by

Jeffrey Stephen Lundeen

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Physics  
University of Toronto

© Copyright by Jeffrey S. Lundeen (2006)

# Abstract

Generalized measurement and post-selection  
in optical quantum information

Jeffrey S. Lundeen

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2006

Post-selection is the selection of a subset of an ensemble based on a measurement on that ensemble. In the last few years, post-selection has taken on new importance as a mechanism to induce the optical nonlinearities required for quantum logic gates. I present experiments and theory that explore post-selection in photonic systems.

First, I experimentally demonstrate effective nonlinear absorption between two single photons. I show that this photon-exchange effect can be understood as a post-selection induced nonlinearity similar to the one that occurs in Hong-Ou-Mandel interference.

I then go on to describe an experimental implementation of Hardy's Paradox, which is one of the three great thought-experiments about entanglement and correlations in quantum mechanics (the others being the EPR and GHZ "paradoxes"). In the place of the electron-positron annihilation in the original proposal, I apply, for the first time, a post-selection-based absorptive photon switch that coworkers and I had previously developed.

Weak measurement promises to be an important technique for characterizing post-selected quantum systems during their evolution. The technique minimizes disturbance to the system and therefore can be used in situ. In a third experiment, I attempt to resolve a recent debate: Which is more fundamental – complementarity

or the uncertainty principle? Specifically, can interference be destroyed by a which-path measurement without transferring momentum to the particle? This question is difficult to investigate as any measurement of the momentum transfer causes collapse, severely disturbing the system. I circumvent this problem with the technique of weak measurement and show that which-path measurements do indeed induce a momentum transfer as required by the uncertainty principle.

I then proceed to develop theoretically a new method for performing multiparticle weak measurements that is unique because it does not require single-particle-level interactions. This theory provides new insight into the importance of the imaginary component of the weak measurement result.

In a fourth experiment, I use the latter theory and return to Hardy's Paradox to perform the first weak measurement of an entangled system. I post-select on the paradoxical measurement result while weakly measuring a set of observables. The results are surprising in that they no longer conflict; they now satisfy a classical-logic truth-table.

This work is dedicated to my family. To my parents, for enduring my childhood experiments investigating the breeding of gerbil colonies, the chemical reactions between household liquids, the optical properties of powerful magnifying glasses (and, in the process, combustion), the cryogenic preservation of stinging insects in liquid nitrogen, and so on. And to my partner, Laura, for encouraging me in my adult experiments in quantum optics.

# Contents

<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Context . . . . .	1
1.2 Spontaneous Parametric Downconversion . . . . .	6
1.3 Post-selection . . . . .	12
1.3.1 Photodetector post-selection . . . . .	13
1.3.2 Hong-Ou-Mandel plus photodetector post-selection . . . . .	14
1.3.3 Hong-Ou-Mandel plus POVM . . . . .	16
1.3.4 Nonlinear unitary transformations with post-selection . . . . .	19
1.3.5 Conclusion . . . . .	20
1.4 Weak Measurement . . . . .	21
1.4.1 Introduction . . . . .	21
1.4.2 The weak value . . . . .	23
1.4.3 Properties of weak values . . . . .	30
1.4.4 Experimental weak measurement . . . . .	32
1.4.5 The meaning of the weak value . . . . .	36
1.4.6 Conclusion . . . . .	45
<b>2 The photon-exchange effect</b>	<b>46</b>
2.1 Introduction . . . . .	46
2.2 Theory . . . . .	48
2.3 Experiment . . . . .	51
2.4 Results . . . . .	54
2.5 Is this a quantum effect? . . . . .	57
2.6 Conclusion . . . . .	62
<b>3 Hardy's Paradox</b>	<b>65</b>
3.1 Introduction . . . . .	65

3.2	Interaction-free measurements. . . . .	66
3.3	Intuitive version of Hardy's Paradox . . . . .	69
3.4	Quantum mechanical model . . . . .	69
3.5	Rigorous version of Hardy's Paradox . . . . .	72
3.6	The switch . . . . .	74
3.7	Experimental implementation . . . . .	76
3.8	Results . . . . .	80
<b>4</b>	<b>Momentum transfer and Which-Way Measurement</b>	<b>83</b>
4.1	Introduction . . . . .	83
4.2	The Idea . . . . .	86
4.3	Experiment . . . . .	89
4.4	Theoretical modelling . . . . .	92
4.5	Results . . . . .	97
4.6	Quantum eraser . . . . .	100
4.7	Conclusion and discussion . . . . .	101
<b>5</b>	<b>Measurement of Joint Weak Values</b>	<b>104</b>
5.1	Introduction . . . . .	104
5.2	Theory . . . . .	107
5.2.1	Background . . . . .	107
5.2.2	The annihilation operator and weak measurement . . . . .	108
5.2.3	Joint weak values . . . . .	109
5.2.4	The $N$ th-order joint weak value . . . . .	112
5.2.5	An example . . . . .	114
5.3	Spin pointers . . . . .	114
5.4	Conclusion . . . . .	118
<b>6</b>	<b>Weak measurement and Hardy's Paradox</b>	<b>119</b>
6.1	Introduction . . . . .	119
6.2	Theory . . . . .	121
6.3	Experiment . . . . .	126
6.4	Results . . . . .	131
6.5	Conclusion . . . . .	134
<b>7</b>	<b>Conclusion</b>	<b>135</b>
	<b>Bibliography</b>	<b>137</b>
<b>A</b>	<b>Conditions necessary for weak measurement</b>	<b>156</b>
A.1	Pointer width . . . . .	156
A.2	Pointer shape . . . . .	157
A.3	Pointer coherence . . . . .	157

**B The  $N$ th-order joint weak value with  $N$  spin pointers****160**

# List of Figures

1.1	Spontaneous parametric downconversion . . . . .	8
1.2	Kwiat’s double-crystal entangled state creation . . . . .	11
1.3	The Hong-Ou-Mandel Effect . . . . .	15
1.4	Entanglement of photons with QND measurements . . . . .	18
1.5	The von Neumann measurement model . . . . .	25
1.6	Generalized measurement . . . . .	27
1.7	The Hulet weak measurement experiment . . . . .	34
1.8	The weak measurement of an arbitrary wavefunction . . . . .	44
2.1	The absorptive version of the Franson proposal . . . . .	48
2.2	General schematic of the experiment . . . . .	51
2.3	Experimental setup of photon-exchange effect experiment . . . . .	53
2.4	Data for photon-exchange effect . . . . .	54
2.5	The anomolous two-photon transmission curve. . . . .	56
2.6	Coincidence rate for classical pulses . . . . .	60
2.7	Classical analogue of the photon exchange phenomena . . . . .	61
2.8	Classical and quantum predictions for anomalous transmission of photon pairs . . . . .	61
2.9	HOM-based description of the photon-exchange effect . . . . .	63
3.1	An interaction-free measurement . . . . .	68
3.2	Hardy’s Paradox . . . . .	70
3.3	The low-light level switch . . . . .	75
3.4	The experimental setup for Hardy’s Paradox. . . . .	77
3.5	The phase-stabilization apparatus . . . . .	79
3.6	Photon “annihilation” in Hardy’s Paradox . . . . .	80
3.7	Mach-Zehnders in Hardy’s Paradox . . . . .	82
4.1	Weak measurement of momentum transfer from a which-way measurement . . . . .	90
4.2	Initial photon probability distribution . . . . .	93
4.3	Optical model predictions for the weak value . . . . .	96



4.4	Example weak measurement of the initial momentum . . . . .	98
4.5	The weakly measured momentum transfer . . . . .	99
4.6	Weakly measured momentum transfer with a Quantum Eraser . . . . .	100
5.1	Pointer correlations from joint weak measurements. . . . .	117
6.1	Schematic for practical weak measurements in Hardy's Paradox . . . . .	122
6.2	Experimental Setup . . . . .	128
6.3	Apparatus for joint weak measurement in Hardy's Paradox . . . . .	130
6.4	Raw data for the weak value of the occupation of the inner arms. . . . .	132

# List of Tables

6.1	Truth table for weak values in Hardy's Paradox . . . . .	125
6.2	Raw data for the weak measurement of the occupation of the inner arms	131
6.3	Raw data for the weak measurement of the occupation of the outer arms	132
6.4	A truth table containing the joint and individual experimental weak values in Hardy's Paradox . . . . .	133

## Acknowledgements

First and foremost, I thank my M.Sc. and Ph.D. advisor Aephraim Steinberg. Aephraim created an atmosphere that fostered scientific curiosity, provided the support for the hard work research requires, and encouraged engagement in hot developments in the field. I have endeavoured to emulate his dazzling ability as a presenter and writer to confidently lead his audience through complicated and/or subtle scientific arguments.

I was privileged to begin my graduate career with Kevin Resch as a coworker. Kevin made my transition into experimental research much easier by patiently teaching me the basic tools of the trade. He was a fun person to work with and was always game to debate new and controversial ideas in quantum optics. I aspire to have his work ethic and ability to persevere with hard problems.

I thank my family for giving me financial and emotional support during my graduate career. Family Sunday dinners were always a welcome breather from the academic world. As well, they generously provided me with a laptop to write this thesis on and a home to live in while finishing it.

I especially thank my partner of ten years, Laura McCurdy. She has supported me through some rough times. I appreciate her efforts to keep me from becoming socially inept, to distract me from the scientific problems on my mind, and to ensure I stay healthy.

Later in my Ph.D, I was fortunate to work with Morgan Mitchell, Rob Adamson, Krister Shalm, and Reza Mir. Not only have they helped me build my skills as scientist but they were really fun people to hang out with.

A number of people in physics have helped make my Ph.D. a mostly enjoyable experience. I thank Laure Wawrezinieck, Lauren Seagel, Naomi Ginsberg, Sal Maone, Stefan Myrskog, Mirco Sierke, Terry Rudolph, Chris Ellenor, Brock Wilson, Jalani (Fox) Kaneem, Karen Sauke, Fan Wang, Ana Jofre, Masoud Mohseni, Alan Stummer, and Matt Partlow. I also thank my Toronto friends Jonn Barton, Jeremy Paczuski,

Colin Oakes, and Brian Oakes for excusing my long unexplained absences from their lives during my Ph.D.

I thank the American Physical Society, Elsevier, and my coauthors Kevin Resch, Aephraim Steinberg, Geoff Lapaire, John Sipe, Reza Mir, Morgan Mitchell, Josh Garretson and Howard Wiseman for allowing the reprinting of copyrighted materials. I appreciate the personal financial support provided to me by the University of Toronto Open Fellowship, NSERC, and the Walter C. Sumner Foundation.

## Scientific Acknowledgements

Kevin Resch had the idea for the single-photon level switch described in Chapter 2 that we used to implement Hardy's Paradox. John Sipe conceived of the idea for the photon-exchange effect in Chapter 2 and Geoff Lapaire developed a theoretical description of the experiment in the photon-wavefunction picture. Howard Wiseman came up with the idea for a weak measurement of momentum transfer in the double-slit experiment described in Chapter 4. Reza Mir took the data for that experiment and, along with Morgan Mitchell, performed most of the analysis. Josh Garretson made numerical predictions for the experiment based on weak measurement theory. The theory in Chapter 5 grew out of discussions with Kevin Resch and Aephraim Steinberg.

## Publication List

The following papers were published during the completion of this Ph.D. Chapters 2–6 are based on some of these works.

“A double-slit ‘which-way’ experiment addressing the complementarity-uncertainty debate,” R. Mir, J.S. Lundeen, M.W. Mitchell, A.M. Steinberg, H. M. Wiseman, and J. L. Garretson, submitted (May 2005).

“Comment on ‘Linear optics implementation of weak values in Hardy’s Paradox’,” J.S. Lundeen, K.J. Resch, and A.M. Steinberg, *Physical Review A*, **72**, 016101 (2005).

“Practical measurement of joint weak values and their connection to the annihilation operator,” J.S. Lundeen and K.J. Resch, *Physics Letters A*, **334**, 337 (2005).

“Super-resolving phase measurements with a multi-photon entangled state,” M.W. Mitchell, J.S. Lundeen, and A.M. Steinberg, *Nature*, **429**, 161 (2004).

“Photon-exchange effects on photon-pair transmission,” K.J. Resch, G.G. Lapaire, J.S. Lundeen, J.E. Sipe, A.M. Steinberg, *Physical Review A*, **69**, 063814 (2004).

“Experimental Realization of the Quantum Box Problem,” K.J. Resch, J.S. Lundeen, A.M. Steinberg, *Physics Letters A*, **324**, 125 (2004).

“Experimental application of decoherence-free subspaces in a optical quantum computing algorithm,” M. Mohseni, J.S. Lundeen, K.J. Resch, and A.M. Steinberg, *Physical Review Letters*, **91**, 187903 (2003).

“Practical creation and detection of polarization Bell states using parametric down-conversion,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *The Physics of Communication*, Antoniou, Sadovnichy, and Walther, eds., World Scientific, pp 437-451 (2003).

“Conditional-phase switch at the single-photon level,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *Physical Review Letters*, **89**, 037904 (2002).

“Quantum state preparation and conditional coherence,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *Physical Review Letters*, **88**, 113601 (2002).

“Electromagnetically induced opacity for photon pairs,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *Journal of Modern Optics*, **49**, 487 (2002).

“Nonlinear optics with less than one photon,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *Physical Review Letters*, **87**, 123603 (2001).

“Total reflection cannot occur with a negative time delay,” K.J. Resch, J.S. Lundeen, A.M. Steinberg, *IEEE Journal of Quantum Electronics*, **37**, 794 (2001).

“Comment on ‘Manipulating the frequency-entangled states by an acoustic-optical modulator’,” K.J. Resch, S.H. Myrskog, J.S. Lundeen, and A.M. Steinberg, *Physical Review A*, **64**, 056101 (2001).

“Experimental observation of nonclassical effects on single-photon detection rates,” K.J. Resch, J.S. Lundeen, and A.M. Steinberg, *Physical Review A*, **63**, 020102(R) (2001).

# Chapter 1

## Introduction

### 1.1 Context

Over the past 100 years, the strange features of quantum theory, such as collapse and entanglement, have grown from curiosities to the basis for a whole new field – quantum information. This transformation has been particularly marked in the field of quantum optics, which initially was largely focused on the theory and application of the laser. From the late seventies to the early eighties, this focus began to change when the first tests of Bell’s inequalities were conducted with photons [1][2], making optics the first system in which entanglement was sought out for its own value. At that time few guessed that this line of investigation would lead to anything but odd gedanken experiments [3]. In the 80s, however, two remarkable papers not only broached the possibility that these strange features of quantum mechanics were actually useful, but also demonstrated this. Specifically, in 1982 Bennett and Brassard demonstrated that the existence of quantum collapse ensured that elements of a secret key could be distributed securely by single particles [4]. And in 1985, David Deutsch introduced the quantum Turing machine and demonstrated that entangled quantum states could be manipulated to solve a particular mathematical problem faster than any known classical algorithm [5][6][7]. These two papers gave birth to quantum com-



munication and quantum computation, the two main branches of the field of quantum information. Subsequently, many key experiments were first performed with photons, including quantum teleportation [8][9]?? and dense coding[10].

Despite the predominance of quantum optics in experimental quantum information, the search for a suitable physical system as a platform for quantum computing largely excluded optics. This was because the manipulation of entanglement, including the creation of it, was thought to be necessary for quantum computing [11]. However, schemes for quantum computing with optics relied on large nonlinear interactions between single photons [12] to build the quantum gates, such as the C-PHASE, necessary to manipulate entanglement. The problem was that these nonlinear interactions were, and largely still are, unavailable to us [13].

As I began my graduate work at the University of Toronto, research groups were becoming adept at producing increasingly complicated entangled states with larger numbers of photons [14][15]. In addition, we and other groups demonstrated a number of schemes with enhanced nonlinearities [16][17][18][19][20]. Unfortunately, these schemes were either technologically prohibitive or had severe restrictions on the conditions in which they could be used. There were still no feasible prospects for quantum computing with photons. This changed with a surprising paper by Knill, Laflamme and Milburn (KLM), which showed that a C-PHASE gate could be built with only linear-optical elements, such as beamsplitters and waveplates [21]. These elements coupled extra optical modes (known as ancillas) to those containing the qubit photons. A unique characteristic of KLM's device was that it only functioned upon successful post-selection (based on the detection of photons in the ancilla optical modes). Combined with a couple of other tricks, they showed that quantum computing with photons was, if not immediately feasible, at least conceivable. The concept of measurement-induced nonlinearities was later shown to be closely related to the concept of cluster-state computing, in which quantum information processing is driven by single-particle measurements on a lattice of entangled particles [22][23].

Nonetheless, our understanding of these post-selected or measurement-driven systems is still in its infancy [23]. This is compounded by our lack of understanding of entanglement; the field is still not clear on its definition, the classification of its types, or its role in quantum computing [24]. This thesis presents a series of experiments that explore post-selection in quantum mechanics.

First, we describe an experimental demonstration of an effect proposed by John Sipe, in which there is an effective nonlinear absorption between two single photons. The two photons, separated by a time interval, pass through an atomic cloud with a narrow absorption feature. At a particular time interval, we observe an anomalous absorption that differs from the product of the absorption of the individual photons. This photon-exchange effect can be understood as a post-selection induced nonlinearity similar to that which occurs in Hong-Ou-Mandel interference.

Hardy's Paradox is one of the three great thought-experiments about entanglement and correlations in quantum mechanics (the others being the EPR [25] and GHZ [26] "paradoxes"). John Bell, following EPR, proved what is now known as Bell's theorem [27]. This theorem states that any local realistic description of the world is incompatible with quantum mechanics. A local theory is one in which distant events and objects can not influence one another. A realistic theory is one in which a measurement of any property of an object can be predicted with certainty. Or, more simplistically, properties of objects must be set before they are observed. In contrast, in the standard formulation of quantum mechanics one can only predict the probability of a particular measurement result. Closely related are Bell inequalities, which place restrictions on measurement results solely using standard probability theory (i.e. no physical theory is assumed). Implicit in this probability theory are the assumptions of locality and reality. Measurements on entangled systems have violated these inequalities [1][2] and thus shown that the world is either nonlocal or not realistic, a rather unpalatable choice for many physicists.

In a similar spirit to John Bell's work, Lucien Hardy proposed a simple entan-

gled system in which quantum mechanics predicts measurement results that not only contradict standard probability theory but also classical logic [28]. In the second experiment, we implement Hardy’s Paradox using an absorptive switch we had previously developed [19]. The switch operates at the level of single photons and functions through a combination of quantum interference and post-selection. This experiment is the first demonstration of an application using this switch.

Weak measurement promises to be an important technique for characterizing post-selected quantum systems during their evolution [29]. The technique minimizes disturbance to the system and therefore can be used in situ. In a third experiment, we attempt to resolve a debate that occurred within the pages of *Nature* [30][31]: Which is more fundamental – complementarity or the uncertainty principle? Specifically, can interference be destroyed by a which-path measurement without transferring momentum to the particle? This question is difficult to investigate as any measurement of the momentum transfer causes collapse, severely disturbing the system. We solve this difficulty by weakly measuring the initial momentum before the which-path measurement and then post-selecting on a final momentum [32]. The difference in the two momenta is the transfer. This experiment, the first measurement of momentum transfer in an interferometer, showed that which-path measurements do indeed induce a momentum transfer, as required by the uncertainty principle.

We then proceed to theoretically develop a new method for performing multiparticle weak measurements. As described above, post-selection has become an important tool in driving quantum logic gates in linear-optics circuits. However, these gates involve multiparticle entanglement, and therefore weak measuring typically requires single-particle-level nonlinearities. Our method is unique because it circumvents this problem. In addition, the theory we develop reveals the importance of the imaginary component of the weak measurement result.

In a fourth experiment, we return to Hardy’s Paradox to perform the first weak measurement of an entangled system. We ask what becomes of the classical-logic

contradiction when the measurements involved are non-disturbing and thus can be measured simultaneously. We post-select on the paradoxical measurement result while weakly measuring a set of observables [33]. The results are surprising in that they no longer conflict; they now satisfy a classical-logic truth-table. However, this resolution comes at an expense: One of the entries in the table is negative – an impossible result for normal measurements, but for weak measurements this is a characteristic of the entangled Hardy state.

Before describing these experiments in detail, I will first introduce some theoretical and experimental background.

## 1.2 Spontaneous Parametric Downconversion

Spontaneous parametric downconversion (SPDC) is currently the most widely used source of photons for studies of quantum optics, and particularly quantum information. Since most of the experiments in this thesis involve SPDC (with the exception of the double-slit experiment in (chapter 4), a short review of its properties is useful. In SPDC, an intense pump beam passes through a  $\chi^{(2)}$  nonlinear crystal [34][35]. Occasionally, one of the pump photons splits into two *downconverted* (DC) photons, which are typically labelled “signal” and “idler”. Since this is a parametric process (the crystal is returned to its initial state), energy is conserved:

$$\hbar\omega_p = \hbar\omega_i + \hbar\omega_s. \quad (1.1)$$

For a crystal which is infinite in all three dimensions, momentum is also perfectly conserved between the three beams:

$$\mathbf{k}_p = \mathbf{k}_i + \mathbf{k}_s. \quad (1.2)$$

In the field of nonlinear optics, creating the conditions under which this conservation is possible is called *phase matching* [36].

In a crystal with finite dimensions, the momentum conservation (and hence phase matching) need not be perfect. Typically, the crystal is considered infinite in the directions parallel to the crystal interface,  $\hat{x}$  and  $\hat{y}$ , which is correct as long as the crystal extends beyond the pump beam diameter. Much like in specular reflection or refraction, spatial invariance along these directions results in perfect momentum conservation:

$$\hbar k_p^x = \hbar k_i^x + \hbar k_s^x \quad (1.3)$$

$$\hbar k_p^y = \hbar k_i^y + \hbar k_s^y. \quad (1.4)$$

However, in the  $\hat{z}$  direction, which is usually taken to be parallel to the pump propagation direction, a realistic crystal has a finite length of  $L$ . This has the consequence that since the three waves (signal, idler and pump) interact over a finite length, the  $\hat{z}$ -component of the momentum only needs to be conserved to within  $\sim \hbar/L$  for down-conversion to occur. For a specific pump frequency and direction, the probability that a pair of downconverted photons will be created depends on the wavevectors of the photons [37],

$$P(\mathbf{k}_p, \mathbf{k}_i, \mathbf{k}_s, \omega_p, \omega_i, \omega_s) \propto \left| \int_{-L/2}^{L/2} e^{i\Delta k_z z} dz \right|^2 = \text{sinc}^2(\Delta k_z L/2), \quad (1.5)$$

where,

$$\Delta k_z = k_p^z - k_i^z - k_s^z. \quad (1.6)$$

The finite width of the sinc function implies that there is some probability for down-conversion even with imperfect momentum conservation. As the crystal gets longer the sinc function gets narrower and the requirement for perfect phase matching gets stricter.

Although frequency does not appear explicitly in Eq. 1.5, the downconversion probability depends on the frequency of the idler, signal and pump waves through the frequency dependence of the wavevectors, which results from dispersion in the crystal. For typical crystal lengths ( $L = 0.1 \text{ cm} - 10 \text{ cm}$ ),  $P$  drops off quickly with  $\Delta k_z$ . Thus, the crystal is usually chosen for perfect phasematching. The search for materials in which the three beams can be phasematched has been and is still an area of great research effort. The difficulty arises from the fact that in the type of dispersion common to most transparent materials, *normal dispersion*, the refractive index increases with frequency. Consequently, the pump wave will typically experience a larger index of refraction than the downconverted photons, which makes it impossible to satisfy Eq. 1.1 and Eq. 1.2 simultaneously.

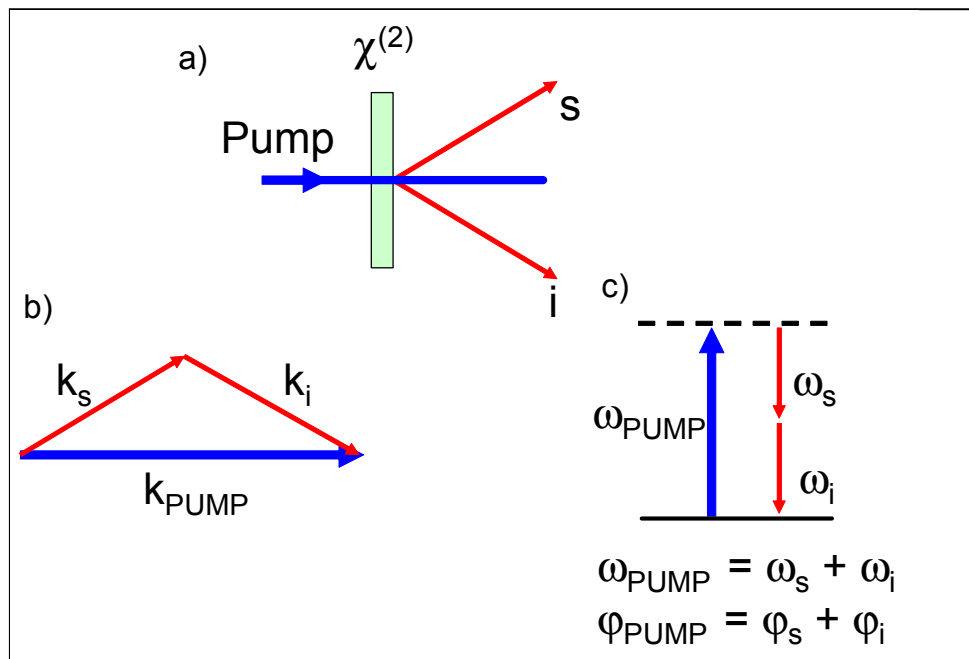


Figure 1.1: a) In spontaneous parametric downconversion a pump photon splits into a signal (s) and idler (i) photon pair inside a  $\chi^{(2)}$  nonlinear crystal. b) Inside the crystal, the wavevectors of the idler and signal photons must add to be equal to the pump wavevector, conserving momentum. c) Energy is conserved, forcing the idler and signal frequencies to sum to the pump frequency. The phase of the outgoing beams must also sum to equal the pump phase.

One strategy to overcome this problem is to use a  $\chi^{(2)}$  nonlinear crystal that is birefringent. The crystal has two different refractive indices,  $n_e$  for the extraordinary polarization (along the crystal axis), and  $n_o$  for the ordinary polarization (perpendicular to the crystal axis). Both of these indices depend on wavelength due to dispersion, which we model with Sellmeyer equations. Consider light with wavelength  $\lambda$  (in vacuum) that travels through the crystal at some angle  $\theta$  relative to the crystal axis. The extraordinary polarization is now defined as the polarization component that lies in a plane formed by the crystal axis and the beam direction. It will experience a refractive index that is simply a weighted quadratic sum of the two fundamental indices,

$$n_E(\theta, \lambda) = \frac{1}{\sqrt{\left(\frac{\sin(\theta)}{n_e(\lambda)}\right)^2 + \left(\frac{\cos(\theta)}{n_o(\lambda)}\right)^2}}. \quad (1.7)$$

Consequently, by changing the angle of propagation one can tune the refractive index of any of the beams involved until perfect phasematching is achieved. Conversely, for a set pump frequency and angle, perfect phasematching occurs for signal and idler photons with a range of frequencies. Each frequency pair in this range,  $\omega_s$  and  $\omega_i = \omega_s - \omega_p$ , occurs at a pair of angles  $\theta_i$  and  $\theta_s$  such that momentum is conserved.

The experiments in this thesis use Beta Barium Borate (BBO) crystal. This crystal does not have the highest nonlinear coefficient of any  $\chi^{(2)}$  nonlinear crystal. However, the  $\chi^{(2)}$  nonlinearity is generally a function of the frequency of the fields as well as the crystal angles. BBO has one of the largest  $\chi^{(2)}$  nonlinearities when it is angle-tuned for SPDC from a 400nm pump to 800nm photon pairs. As well, it is hard, and has a relatively small susceptibility to moisture (compared to other nonlinear crystals such as Potassium Dihydrogen Phosphate, KDP). Perfect phasematching can be achieved in two different ways for BBO. Type I occurs when the pump is e polarized and the two downconverted photons have o polarizations. Type II, which is used in this thesis, occurs when the pump is e polarized, and one downconverted photon is o and the other e polarized.



Downconversion has been used for a variety of purposes in quantum optics and quantum information. Perhaps the easiest application to understand is the creation of single-photon states. For an intense pump beam (that we approximate with a classical field), the nonlinear interaction Hamiltonian of the crystal can be modelled by

$$\mathcal{H} = g\hat{a}_i^\dagger\hat{a}_s^\dagger + g^*\hat{a}_i\hat{a}_s, \quad (1.8)$$

where  $\hat{a}_n^\dagger$  is the creation operator of mode  $n$ , and  $g$  is a constant, assumed to be small, that combines the amplitude of pump field and the nonlinear coupling. In this highly simplified model, we neglect frequency, transverse wavevector and polarization degrees of freedom. Since there is initially vacuum in the idler and signal modes we neglect the annihilation operators in the interaction Hamiltonian. (This would not be valid if  $g$  was large, in which case the process is called parametric generation.) Subsequently, the following state is produced by an interaction lasting time  $t$ ,

$$|\psi\rangle = |0\rangle - \frac{igt}{\hbar} |1_i 1_s\rangle - \frac{1}{2} \left(\frac{gt}{\hbar}\right)^2 |2_i 2_s\rangle + \frac{i}{6} \left(\frac{gt}{\hbar}\right)^3 + \dots \quad (1.9)$$

A photon pair will be produced with a small probability equal to  $(gt/\hbar)^2$ . Typically, the pump intensity is low enough that we are able to neglect terms higher than first-order. In this regime, the detection of a photon in the signal mode ideally indicates the presence of one photon in the idler mode with certainty [38]. This process is termed *heralded* single-photon production. The spontaneous nature of the process, as evidenced by the predominance of the vacuum term in the state, precludes the preferable process of *triggered* single-photon production, where a photon is produced with a high probability in a specific time interval. In summary, photon pairs are produced at random but the detection of one signals the presence of a single-photon in the other mode.

The spontaneous nature of SPDC is not always a drawback as it enables the creation of entangled photon states. This idea was introduced by Kwiat *et al.* in the 1990s [39][40][41]. Imagine two SPDC processes that produce photon pairs in the

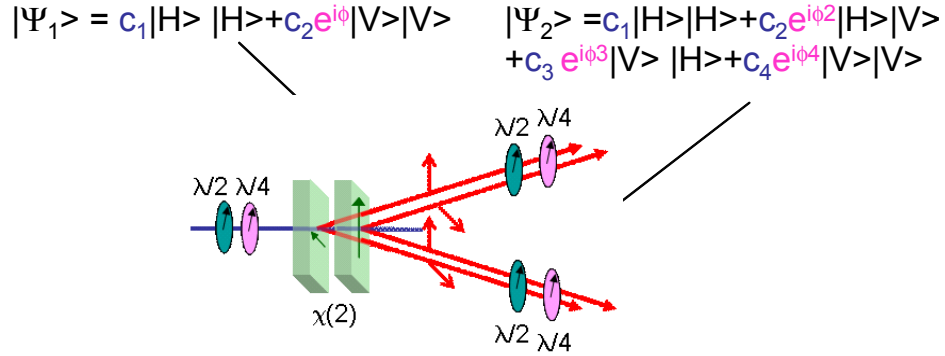


Figure 1.2: In a scheme proposed and demonstrated by Paul Kwiat and coworkers [41], two crystals are oriented to have perpendicular axes. The outgoing idler and signal beams from each are overlapped to produce an entangled two-photon state. The amplitude of the two terms in  $\Psi_1$  are controlled by a quarter and a half waveplate before the crystals. Identical pairs of waveplates in the output beams transform  $\Psi_1$  into an arbitrary output state,  $\Psi_2$ .

same two directional modes but in opposite polarization modes. For instance, one crystal is aligned to produce two horizontally polarized photons in the idler and signal modes, and the other aligned to produce two vertically polarized photons in the same modes (i.e. the same spatial, spectral and temporal modes). If both crystals are pumped coherently (e.g. by a pump beam in a coherent state and separated by less than the coherence length) then, to first order, the following state is produced:

$$|\psi\rangle = |0\rangle - \frac{igt}{\hbar} (|1_H1_H\rangle + |1_V1_V\rangle). \quad (1.10)$$

Each crystal has only a small chance of producing a photon pair but since they are produced coherently the two possible polarizations exist in a superposition.

Compare this to what the resulting state is if each crystal produced exactly one pair (i.e. with 100% certainty) within a given time interval,

$$|\psi\rangle = |1_H1_H1_V1_V\rangle. \quad (1.11)$$

This state contains two pairs of photons and, more importantly, it is a product state, and thus, is unentangled. On the other hand, the second term in Eq. 1.10 is a Bell

state in the photon polarization basis, and is therefore maximally entangled. (This means that a measure of the entanglement of the polarizations of the two photons, such as the *Tangle* [23] is at a maximum. Alternately and equivalently this means the state is pure and tracing over polarization state of one photon leaves the other in a density matrix  $\rho = I$ ). In contrast, a product state, the vacuum term, dominates the total state in Eq. 1.10 in magnitude. For the total state, the entanglement of the two optical modes (as opposed to photon polarizations) is small. Tracing over the degrees of freedom of one optical mode leaves the other in state that is approximately the vacuum state (to within  $\frac{gt}{\hbar}$ ).

Nonetheless, if we restrict ourselves to those instances where a photon pair is subsequently detected, we are, in effect, filtering out the vacuum component. Although this technique has been second nature to experimentalists in the field and might seem quite straightforward, it is an example of the use of *post-selection* to drive transformations that would normally require single-photon-level nonlinear optics. This has been a rapidly developing area of research since the pioneering theoretical work of Knill, Laflamme and Milburn in 2001 [21] and is the subject of the next section.

### 1.3 Post-selection

Post-selection has been commonly used in experimental quantum optics since the first Bell's inequality tests were performed in the 1970s [42], and it has recently taken on a more prominent role in the field as a mechanism for driving nonlinear interactions. In general, the process of post-selection is the selection of some subset of an ensemble of identical systems based on the result of a measurement. In its simplest form, this measurement is a projective measurement that collapses the system into an eigenstate of the observable. The post-selection measurement is often destructive and so this ends the evolution of the quantum system under consideration. But it is possible to do the post-selection by a quantum non-demolition (QND) measurement, which

does not destroy the system. For example, in optical quantum-information a QND measurement determines the polarization of the photon without the photon being absorbed (unlike with the absorption that occurs in polarizer). However, an ideal QND measurement of this type requires strong single-photon nonlinearities and is therefore not usually considered [43] (although, there has been a recent theoretical proposal using only weak Kerr nonlinearities [44]). More generally, the measurement on which the post-selection is based is a POVM (Positive-Operator Valued Measure). POVMs are a generalized description of quantum measurement that encompasses non-projective measurements, and the action on a system of projective measurements in expanded Hilbert spaces (See [24]). In the case of post-selection based on a POVM, the collapse need not be complete and, even in optics, the entire system need not be destroyed. Consequently, as with the QND-based post-selection, the quantum system will still be available for subsequent applications. In the following, I describe four examples of post-selection that demonstrate its prevalence and usefulness in quantum optics and information.

### 1.3.1 Photodetector post-selection

Photodetection-based post-selection is the most common type used in experimental quantum optics, though it is rarely mentioned in the literature and even more rarely considered theoretically. As mentioned above, QND measurements in quantum optics typically require strong single-photon-level nonlinearities that are generally unavailable in the laboratory. Since this includes QND measurements of photon number, we instead rely on photo-detection (typically single-photon detectors) to discriminate between the presence or absence of a photon or photon pairs. This is exactly how entangled states are produced in downconversion. When experimentalists post-select on the part of the total state that contains a photon pair, the subset that was in a

maximally entangled polarization state is selected,

$$|\psi\rangle = |0\rangle - \frac{igt}{\hbar} (|1_H 1_H\rangle + |1_V 1_V\rangle) \rightarrow (|1_H 1_H\rangle + |1_V 1_V\rangle) / \sqrt{2}. \quad (1.12)$$

It is remarkable that although the original state contained very little entanglement, one is now able to violate Bell’s inequality with the part of the state that results in a “coincidence count.” In this sense, photodetector post-selection has the ability to amplify the amount of entanglement in a state. This process is accompanied by loss; only a small fraction of the ensemble results in a successful post-selection and the entire ensemble is destroyed by the post-selection itself. When there is no loss, this type of entanglement amplification can only be accomplished with nonlinear interactions [45].

In a real experiment the initial state might contain information in other photon variables such as frequency, or direction that distinguishes the two photons (or even the two polarizations of a single photon). Filtering to eliminate this spurious information, and the existence of optical imperfections, cause photon loss. Due to this, almost every experiment using downconverted photons relies on post-selecting on coincident detection in both modes to eliminate cases where a photon is lost.

### 1.3.2 Hong-Ou-Mandel plus photodetector post-selection

Two photons have no appreciable interaction, yet they must be in a quantum state with bosonic symmetry. This is enforced by the commutation relation of the associated creation and annihilation operators,

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (1.13)$$

Perhaps the most striking effect of this is the Hong-Ou-Mandel (HOM) effect [46], from which many other more complicated quantum optical processes have been assembled. In HOM interference, two identical photons enter opposite ports (1,2) of a 50-50 beamsplitter. Each photon randomly exits from one of the two output ports

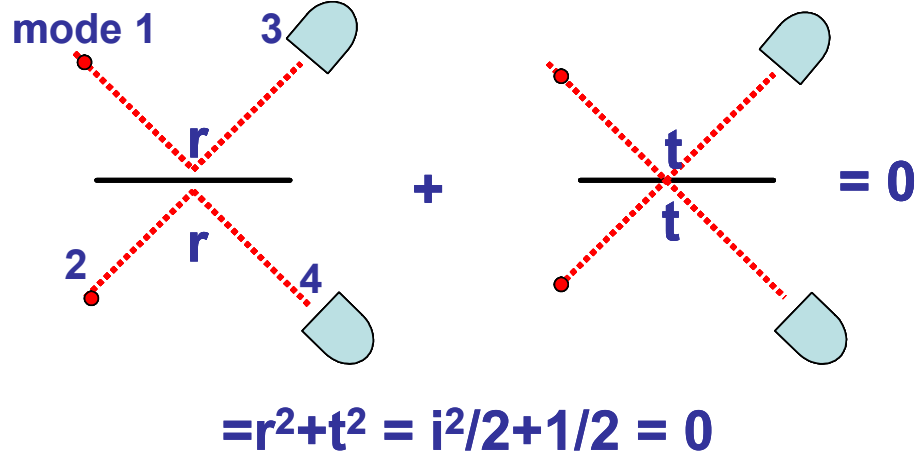


Figure 1.3: In the Hong-Ou-Mandel (HOM) effect, two photons enter opposite ports of a 50/50 beamsplitter (BS). There are two paths that lead to one photon in each BS output port. Both photons can be transmitted or both photons can be reflected. The sum of these two amplitudes vanishes. Consequently, the photon pair always exits together through one or the other output port.

(3,4) of the beamsplitter. However, the exits are correlated in that the two photons always leave from the same port, as if there were an attractive force between them.

$$\hat{a}_1^\dagger \hat{a}_2^\dagger |0\rangle \rightarrow (\hat{a}_3^\dagger + \hat{a}_4^\dagger) (\hat{a}_3^\dagger - \hat{a}_4^\dagger) / 2 |0\rangle \quad (1.14)$$

$$= (\hat{a}_3^{\dagger 2} - \hat{a}_4^{\dagger 2}) |0\rangle / 2 \quad (1.15)$$

$$= (|2_3 0_4\rangle + |0_3 2_4\rangle) / \sqrt{2} \quad (1.16)$$

This is an example of quantum interference; the detection statistics at the two photon counters cannot be simulated with classical electromagnetic theory. Assuming the beamsplitter is non-polarizing, if oppositely polarized photons, say H and V, enter from opposite ports, a similar calculation shows that the following state will be produced,

$$(|1100\rangle + |0011\rangle) / 2 + (|1001\rangle - |0110\rangle) / 2, \quad (1.17)$$

where the modes in the ket are 3V, 3H, 4V, and 4H. In the first term, both photons leave via the same beamsplitter output port, whereas in the second term the photons are split between the output ports. The second term is the maximally entangled singlet-state  $(|VH\rangle - |HV\rangle)/\sqrt{2}$ . Consequently, if we subsequently post-select on there being one photon in each output port we have transformed two previously completely unentangled photons into an entangled state. This was the first method used to produce entangled states from downconverted photon pairs [47]. The combination of the HOM effect and post-selection can create entanglement, a process that ordinarily requires strong nonlinearities. However, since to detect the presence of the photons we need to absorb them, the entangled state is destroyed in the very process that creates it.

### 1.3.3 Hong-Ou-Mandel plus POVM

The above examples show that measurement in the form of post-selection, along with linear-optical elements such as beamsplitters and waveplates, has the ability to enact state transformations that typically require nonlinear optics. These transformations occur probabilistically: In the second example the singlet state is produced fifty percent of the time, whereas in the first, the entangled state is produced only a few percent of the time (for typical values of  $gt$ ). More problematically, the quantum system is destroyed in the process. The first problem is inherent to the process of post-selection. The solution to the second problem is to expand the Hilbert space of our initial system and then to make a projective measurement in only part of the larger space. This type of measurement is described by a POVM [24].

A good example of this type of measurement is the teleportation of the polarization state of a photon. In this now famous device, a Bell-state measurement is performed on a photon pair consisting of one photon of a singlet state pair and an input photon with an arbitrary polarization [9][8]. Upon this measurement, the other photon comprising the singlet state will be projected into the same polarization state as the

input photon up to a rotation that depends on which Bell-state the two photons are found. In this way, the state of the input photon is teleported to the possibly distant remaining photon from the singlet state. Consider what would happen if no photons are in the input mode when the teleportation protocol is run. Now, only one photon will be detected in the Bell-state measurement device. These considerations show that the teleportation protocol not only teleports the state of a particle, but also distinguishes the presence of zero or one photons in the input mode. P. Kok *et al.* recognized this is a QND measurement of the photon's presence (without destroying the polarization) [43]. The QND protocol is as follows: If there are zero photons in the input mode, then the output mode of the teleportation is simply blocked. If there is instead one photon, the teleportation goes ahead as usual and the output photon takes the place of the input photon. Perfect Bell-measurements require, once again, nonlinear interactions between the photons [48], whereas with linear optics only imperfect detection is possible. Specifically, it has been demonstrated that one can detect up to two of the four Bell-states with an optical circuit consisting primarily of polarizing beamsplitters [9].

At this point, it might seem that nothing has been gained from this recognition of the dual purpose of the teleportation protocol; we have used one quantum optical process – teleportation – that requires nonlinear optics to implement another – a QND measurement. The crucial point is that although the teleportation only succeeds probabilistically, we know which cases have succeeded without destroying the output state. In other words, the success is heralded. It follows that probabilistic but heralded QND measurements are also possible with only linear optics. Consider if probabilistic QND measurements are used in the place of direct photo-detection in the first two post-selection examples described above. The resulting probabilistic heralded linear-optical circuit entangles previously unentangled photons without destroying them. Although this circuit can create entanglement, it is not suitable for many nonlinear applications or for quantum information. In the following example, we



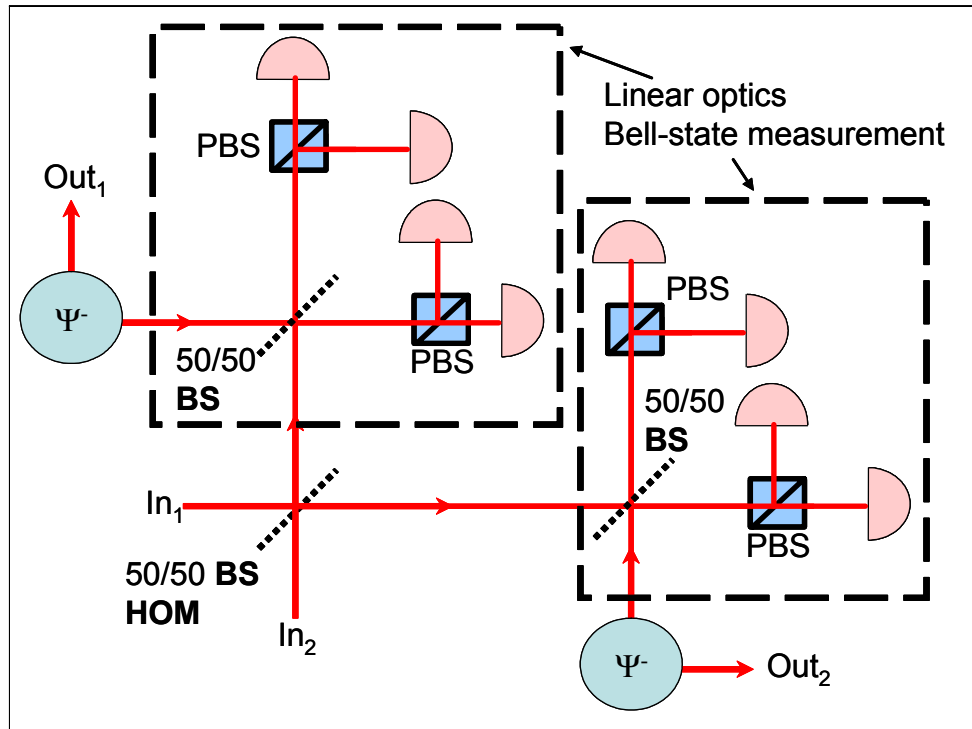


Figure 1.4: Two linear-optics based Bell-state measurements teleport the output modes of a HOM interferometer onto modes  $Out_1$  and  $Out_2$ . Two detectors firing in the Bell-state measurement circuit signals a successful teleportation. Each teleportation is also a QND measurement, projecting onto the part of the input state with one and only one photon. This takes  $|In_1 In_2\rangle = |HV\rangle \rightarrow |Out_1 Out_2\rangle = |\Psi^-\rangle$ , demonstrating that a probabilistic linear optics circuit can entangle two photons without destroying them in the process.

take the polarization basis to be the computational basis. If two horizontal photons enter opposite ports of the HOM beamsplitter they always leave through the same exit port. If we take the polarization basis to be the computational basis, then our circuit takes the  $|HH\rangle = |11\rangle$  state out of the computational space. In this sense, the gate is not unitary, which is a requirement for quantum information and for many nonlinear quantum optical applications.

### 1.3.4 Nonlinear unitary transformations with post-selection

Knill, Laflamme, and Milburn (KLM) introduced the idea of linear optics quantum computation (LOQC) in a seminal paper in 2001 [21]. They proved that one can build the gates for a quantum computer using only linear optics and detection. In particular, they proposed a probabilistic implementation of an elementary two qubit quantum logic gate, the conditional-phase gate (C-PHASE). This gate performs the transformation  $|11\rangle \rightarrow -|11\rangle$  while leaving all other basis states unchanged. Hence, the C-PHASE gate has the potential to entangle previously unentangled qubits. Translated to optical quantum computing (e.g. the dual rail scheme presented in [12]), this gate creates a phase shift only when one photon is present in each of two particular modes. Consequently, it explicitly needs nonlinear optics to function non-probabilistically. KLM's linear-optics scheme for this gate is, as of yet, the most impressive use of post-selection in the field: The gate process is unitary, while the input states are not destroyed during the post-selection. Moreover, the scheme revived the possibility of an optical implementation of a quantum computer with resources that could be available to the field in the near future. The KLM circuit is built from two identical sub-circuits that produce the nonlinear transformation

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle + \gamma |2\rangle \rightarrow \alpha |0\rangle + \beta |1\rangle - \gamma |2\rangle = |\psi'\rangle, \quad (1.18)$$

for arbitrary coefficients  $\alpha$ ,  $\beta$ , and  $\gamma$ . The linear-optical circuit combines one mode containing  $|\psi\rangle$  with two additional modes, containing one ancilla photon. The gate

functions by HOM interference at one or more beamsplitters with carefully chosen reflection and transmission coefficients. (Subsequent to the original paper many variations have been proposed. For a review see [49].) A successful transformation is implemented by a post-selection on zero and one photons in two output modes, which occurs with probability  $1/4$ , leaving the last output mode in state  $|\psi'\rangle$ . Unfortunately, the functioning of the optical circuit has no simple explanation (other than an explicit mathematical derivation). Hopefully, the previous examples of post-selection have suggested that such a device is possible. The linear-optics C-PHASE gate probabilistically entangles photons in a heralded manner like the above example. However, unlike the above example it operates unitarily on every state in the two-photon polarization basis.

### 1.3.5 Conclusion

The experiments in this thesis involve post-selection in a variety of ways. The first experiment uses destructive post-selection to introduce nonlinear absorption for two photons. The second experiment, an implementation of Hardy's Paradox, is based around a single-photon switch developed in our lab that subtly relies on post-selection. The last two experiments are investigations of the functioning of quantum systems and processes that are subsequently post-selected. A typical measurement of an observable will disturb or collapse the quantum system and subsequently change the implications of the post-selection. The next introductory section describes the technique of weak measurement, which circumvents this problem by reducing the strength of the coupling associated with the measurement. We demonstrate the use of this technique to measure the momentum transfer induced by a which-way measurement in Young's double-slit experiment. We also return to Hardy's Paradox and use weak measurement to test the set of contributing logical statements while post-selecting on the paradoxical result.

## 1.4 Weak Measurement

### 1.4.1 Introduction

Weak measurement is a generalization of standard quantum measurement where the precision of the measurement is traded off in exchange for a reduction of the measurement-induced disturbance. This current definition expands on the original proposal by Aharonov, Albert, and Vaidman (AAV) in the context of pre- and post-selected quantum systems, where the limit of an infinitely weak measurement was considered [50][29]. Under the expanded definition, weak measurement is sometimes used to measure a quantum system for the purpose of feedback and control. For example, continuous weak measurement was used to guide an ultracold atomic gas into a chosen quantum state [51]. The expanded definition also covers a type of measurement that is often considered – a measurement of a macroscopic observable composed of the microscopic observables of many identical subsystems, each of which only minimally interacts with the measuring device. The measurement of the magnetization of a large ensemble of spins is a natural example. Another common example is radio-frequency measurements in liquid state Nuclear Magnetic Resonance experiments.

Within its original context of post-selection, weak measurement has had two areas of application. The first is the simplified analysis of pre-existing phenomena or experiments in which it is realized that a weak measurement is already occurring. Examples of this application are the calculation of optical network properties in the presence of polarization-mode dispersion [52][53], modelling of slow- and fast-light effects in birefringent photonic crystals [54], and theoretical predictions [55] for the cavity QED experiment described in [56]. The second area of application is the theoretical investigation of phenomena that are not amenable to characterization by standard measurement. These investigations have had several achievements, which include bringing a new, unifying perspective to the tunneling-time controversy

[57][58][59][60], reconciling the contradictory predictions in Refs. [30][31] on the role of which-path information and the Heisenberg uncertainty principle in the double-slit experiment [32], and offering a consistent resolution for Hardy's Paradox [33]. In Chapters 3 and 4 of this thesis, I describe experimental weak measurements based on the latter two theoretical investigations. Therefore this thesis creates a third area of application for weak measurement: As an *experimental* technique for the characterization of quantum phenomena that would be disrupted by standard measurement.

The process of weak measurement was originally described by AAV using the von Neumann measurement model [61]. This led to criticism that their conclusions were not universal to all types of measurement and, in particular, that their predictions were simply artifacts of the simplistic von Neumann model [62][63][64]. Since those early days, weak measurement has been extended to a wider variety of measurement types [65][66][67], so that there is now convincing, although not conclusive, evidence that weak measurement is indeed universal [68][69]. I will use the von Neumann model for derivations within this thesis since it remains the archetype and the most widely known model of measurement. This model contains two parts, the measured system  $S$  and an ancilla system  $M$  consisting of a meter and a pointer. Through its momentum  $P$ , the pointer is coupled to an observable  $A$  of the measured system for a time  $t$  with strength  $g$ . This coupling shifts the position of the pointer by an amount proportional to the value of  $A$ , which we read off the meter. Reciprocity implies that the measured system will also be changed by an amount proportional to the momentum  $P$  of the pointer. A weak measurement is performed by reducing the pointer momentum spread or the coupling strength so that the associated disturbance becomes asymptotically small. The trade-off is that the induced shift in the pointer relative to its initial position-uncertainty also becomes asymptotically small. At first glance it may appear that this trade-off achieves nothing since each measurement of the pointer shift is now imprecise. However, by repeating this procedure on a large ensemble of identical systems, one can measure the average shift of the pointer

to any desired precision. This average shift is called the weak value. A surprising characteristic of weak values is that they need not lie within the eigenvalue spectrum of the observable and can even be complex. On the other hand, an advantage of weak measurements is that they do not disturb the measured system nor any other simultaneous weak measurements or subsequent strong measurements, even in the case of non-commuting observables. This makes weak measurements ideal for examining the properties and evolution of systems before post-selection and might enable the study of new types of observables.

### 1.4.2 The weak value

I begin by deriving AAV's formula for the weak value of a single particle observable. The von Neumann interaction,

$$\mathcal{H} = g\hat{A}\hat{P}, \quad (1.19)$$

couples an observable  $\hat{A}$  of the measured system to the momentum  $\hat{P}$  of the pointer, where  $g$  is the coupling constant, which is assumed to be real to keep  $\mathcal{H}$  Hermitian. I also assume the interaction is constant over some interaction time  $t$ . Since  $\hat{A}$  and  $\hat{P}$  act in different Hilbert spaces they commute. Following Von Neumann's original measurement model and AAV's derivation of weak values, the measurement pointer is initially in a Gaussian wavefunction centered at zero,

$$\langle x|\phi\rangle = \phi(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{\frac{1}{2}} \exp\left(-\frac{x^2}{4\sigma^2}\right), \quad (1.20)$$

where  $\sigma$  is the rms width of  $|\phi(x)|^2$ . Generalized pointer shapes are considered in Appendix 1.

In most experiments, quantum mechanical systems are initially prepared in a known initial state  $|I\rangle$ . This preparation is called pre-selection since it usually involves measuring an ensemble of systems and selecting the subensemble with the correct

outcome, similar to post-selection. The initial system state can be represented as a superposition of the eigenstates of  $\hat{A}$ ,

$$|I\rangle = \sum_n \langle a_n | I \rangle |a_n\rangle, \quad (1.21)$$

where the  $n$ th eigenstate is  $|a_n\rangle$  with eigenvalue  $a_n$ . The von Neumann interaction in Eq. 1.19 transforms the initially separable state  $|\psi(0)\rangle = |\psi_S\rangle |\psi_M\rangle = |I\rangle |\phi\rangle$  into an entangled state of the pointer and measured system,

$$|\psi(t)\rangle = \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) \sum_n \langle a_n | I \rangle |a_n\rangle |\phi\rangle \quad (1.22)$$

$$= \sum_n \exp\left(\frac{-igt a_n \hat{P}}{\hbar}\right) \langle a_n | I \rangle |a_n\rangle |\phi\rangle \quad (1.23)$$

$$= \sum_n \langle a_n | I \rangle |a_n\rangle |\phi(gta_n)\rangle, \quad (1.24)$$

where  $|\phi(x_0)\rangle$  is the pointer state shifted by  $x_0$ . Thus the shift in the pointer state is correlated with the observable eigenstate  $|a_i\rangle$  that the measured system is in. Tracing over the state of the measured system, we find the average pointer position to be,

$$\langle \psi(t) | \hat{X} | \psi(t) \rangle = \sum_n |\langle a_n | I \rangle|^2 \langle \phi(gta_n) | \hat{X} | \phi(gta_n) \rangle \quad (1.25)$$

$$= \sum_n |\langle a_n | I \rangle|^2 gta_n \quad (1.26)$$

$$= gt \langle I | \hat{A} | I \rangle. \quad (1.27)$$

In a similar way the average momentum is found to be unchanged:  $\langle \hat{P} \rangle = 0$ . In summary, the action of the von Neumann interaction is to shift the pointer by an average amount proportional to the expectation value of  $\hat{A}$ . This is true regardless of  $gt$ , the strength of the interaction.

The advantage of reducing  $gt$  becomes apparent when we consider the state of the measured system after interaction. In standard measurement, also called *strong measurement*,  $gt$  is sufficiently large that the pointer states are shifted by an amount much greater than their width,  $\sigma$ ,  $|\langle \phi(gta_m) | \phi(gta_n) \rangle| \approx \delta_{mn}$ . A subtle idea is that,

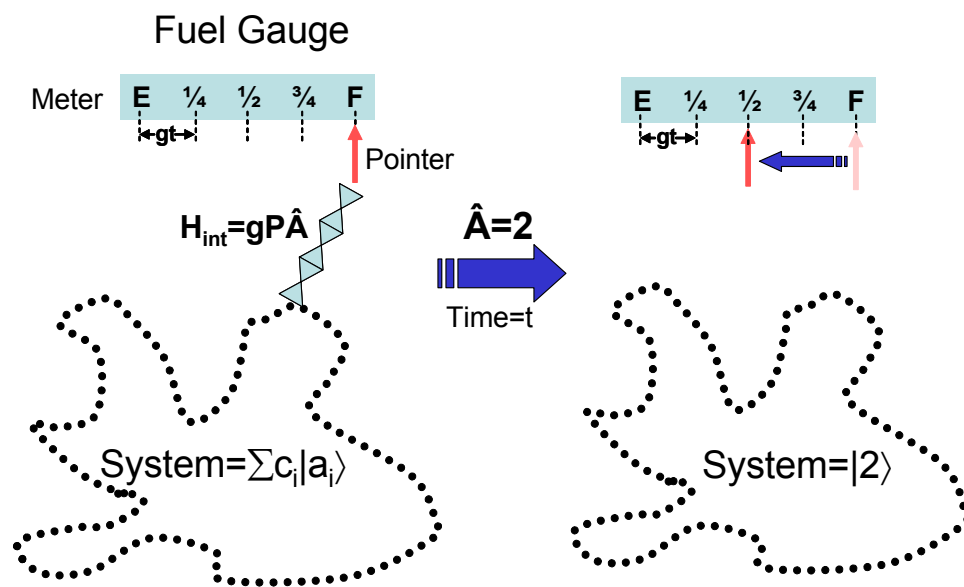


Figure 1.5: In this schematic of a von Neumann measurement of observable  $\hat{A}$ , the meter is a fuel gauge. The system begins in a superposition of eigenstates of  $\hat{A}$ . The system couples with strength  $g$  to the momentum  $P$  of the pointer for time  $t$ . The measurement results in an eigenvalue of 2, collapsing the system to the corresponding eigenstate and moving the pointer by two units.



in this limit, not only is the pointer *shift* correlated with the observable eigenstates, but also the pointer *position*. A measurement of the position of the pointer then collapses the system into one of the observable eigenstates  $|a_i\rangle$ . The measured system is left in a completely mixed state after the interaction,

$$\hat{\rho}_S = \sum_n |\langle a_n | I \rangle|^2 |a_n\rangle \langle a_n|. \quad (1.28)$$

In the limit of  $gt \gg \sigma$ , the coherence in the measured state is destroyed by the large entanglement between the measured system and the pointer.

On the other hand, in the limit of  $gt \rightarrow 0$  the pointer state shifts are small, making  $|\langle \phi(gta_m) | \phi(gta_n) \rangle| \approx 1$ . In this limit, the entanglement goes to zero; tracing over the pointer state leaves the measured system in a pure state,

$$|\psi_S\rangle \approx \sum_n \langle a_n | I \rangle |a_n\rangle = |I\rangle. \quad (1.29)$$

The measured system is undisturbed by the weak measurement and thus remains in the initial state.

AAV considered the case where we further restrict ourselves to the subensemble of system states that are found to be in  $|F\rangle$  after the measurement of  $\hat{A}$ . The previous sections have shown that post-selection is an important and powerful process relevant to driving nonlinear interactions and quantum logic. However, we have seen that if  $\hat{A}$  is measured in the limit of  $gt \gg \sigma$ , the measured system is radically disturbed, changing the action of any subsequent post-selection. For example, if the post-selection drives a probabilistic quantum logic gate (as described in the previous section), then the intermediate measurement of  $\hat{A}$  will generally cause the gate to fail. As we have seen, a weak measurement of  $\hat{A}$  limits this disturbance and avoids this problem. The average pointer shift due to a weak measurement, performed between pre- and post-selection, can be surprisingly different than in Eq. 1.27, as we will see. To consider this scenario, we once again write the state of the combined system after the von Neumann interaction:

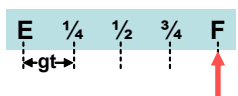
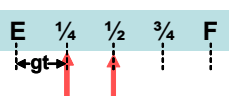
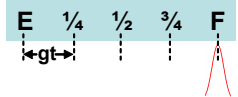
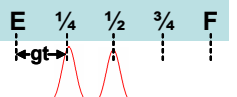
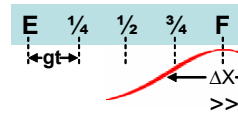
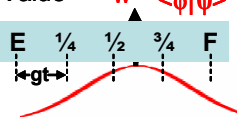
	Initial Pointer Position Uncertainty	Pointer shift due to interaction
<b>Ideal</b>	Dirac Delta 	$H_{\text{int}} = gP\hat{A}$ 
<b>Real</b>	Width $\ll$ Change in Position 	$H_{\text{int}} = gP\hat{A}$ 
<b>Weak</b>	Width $\gg$ Change in Position 	$H_{\text{int}} = gP\hat{A}$ Weak Value = $A_W = \frac{\langle \phi   A   \psi \rangle}{\langle \phi   \psi \rangle}$ 

Figure 1.6: Weak measurement is a generalization of standard measurement in the ideal model, where the pointer is a Dirac delta function, and the real model, where the pointer has a small width. In the leftmost column we show example pointer probability distributions in each case. When there is no post-selection, the average position of the pointer is always equal to the expectation value of the observable. In contrast, with post-selection weak measurement rigidly shifts the pointer by an amount equal to the weak value  $A_W$ .

$$|\psi(t)\rangle = \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |I\rangle |\phi\rangle = \left(1 - \frac{i\mathcal{H}t}{\hbar} - \dots\right) |I\rangle |\phi\rangle \quad (1.30)$$

$$= |I\rangle |\phi\rangle - \frac{igt}{\hbar} \hat{A} |I\rangle \hat{P} |\phi\rangle - \dots \quad (1.31)$$

We project out the part of the  $|\psi(t)\rangle$  that is post-selected in state  $|F\rangle$ , leaving the state of pointer after the interaction and post-selection,

$$\langle F | \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |I\rangle |\phi\rangle = \langle F | I \rangle |\phi\rangle - \frac{igt}{\hbar} \langle F | \hat{A} | I \rangle \hat{P} |\phi\rangle - \dots \quad (1.32)$$

In the absence of any disturbance to the initial state,  $|\langle F | I \rangle|^2 = \text{Prob}_{\text{success}}$  is the probability the post-selection succeeds [59]. If we renormalize the state and then truncate the amplitude of each term to lowest order in  $gt$  we get

$$|\phi_{fi}\rangle = |\phi\rangle - \frac{igt}{\hbar} \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \hat{P} |\phi\rangle - \dots, \quad (1.33)$$

which is just equivalent to dividing by  $\langle F | I \rangle = \sqrt{\text{Prob}_{\text{success}}}$ , as one would expect for an ideal weak measurement ( $gt \rightarrow 0$ ). The subscript *fi*, corresponding to final state  $|F\rangle$  and initial state  $|I\rangle$ , labels the final pointer state, with which we can now calculate the expectation value of  $\hat{X}$  of the pointer. The terms that contain an expectation value of an odd number of operators go to zero, since the pointer is initially an even function about zero. To first order in  $gt$ , the remaining terms give us

$$\langle \hat{X} \rangle_{fi} = \langle \phi_{fi} | \hat{X} | \phi_{fi} \rangle = \frac{-igt}{\hbar} \text{Re} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right) \langle \phi | (\hat{X} \hat{P} - \hat{P} \hat{X}) | \phi \rangle \quad (1.34)$$

$$+ \frac{gt}{\hbar} \text{Im} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right) \langle \phi | (\hat{X} \hat{P} + \hat{P} \hat{X}) | \phi \rangle \quad (1.35)$$

$$= gt \text{Re} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right), \quad (1.36)$$

since, for a Gaussian pointer,

$$\langle \phi | \hat{X} \hat{P} | \phi \rangle = i\hbar/2 \quad (1.37)$$

$$\langle \phi | \hat{P} \hat{X} | \phi \rangle = -i\hbar/2 \quad (1.38)$$

$$\langle \phi | \hat{P}^2 | \phi \rangle = \frac{\hbar^2}{4\sigma^2} \quad (1.39)$$

$$\langle \phi | \hat{X}^2 | \phi \rangle = \sigma^2. \quad (1.40)$$

Here,  $\langle \rangle_{fi}$  is used to signify the expectation value of a pointer observable only in the subensemble of measured systems that start in state  $|I\rangle$  and are later post-selected in state  $|F\rangle$ . Similarly, the momentum expectation value is given by

$$\langle \hat{P} \rangle_{fi} = \langle \phi_{fi} | \hat{P} | \phi_{fi} \rangle = \frac{-igt}{\hbar} \operatorname{Re} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right) \langle \phi | (\hat{P}^2 - \hat{P}^2) | \phi \rangle \quad (1.41)$$

$$+ \frac{gt}{\hbar} \operatorname{Im} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right) \langle \phi | (\hat{P}^2 + \hat{P}^2) | \phi \rangle \quad (1.42)$$

$$= \frac{\hbar gt}{2\sigma^2} \operatorname{Im} \left( \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right). \quad (1.43)$$

The shifts from zero in both the  $\hat{X}$  and  $\hat{P}$  expectation values are proportional to the real and imaginary parts, respectively, of the weak value  $\langle \hat{A} \rangle_w$  which is defined as

$$\langle \hat{A} \rangle_w \equiv \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle}. \quad (1.44)$$

Re-examining Eq. 1.33 we see that we can write the final pointer state as  $|\phi_{fi}\rangle \approx \hat{T}(gt \langle \hat{A} \rangle_w) |\phi\rangle$ , where  $\hat{T}(\epsilon) = \hat{I} + i\epsilon \hat{P}/\hbar$  is the first-order translation operator. Thus, for sufficiently weak coupling the pointer is rigidly shifted by an amount equal to the weak value,  $\langle x | \phi_{fi} \rangle = (\sqrt{2\pi}\sigma)^{-\frac{1}{2}} \exp\left(-\left(x - gt \langle \hat{A} \rangle_w\right)^2 / 4\sigma^2\right)$ .

Note that when applying the formula for the weak value in Eq. 1.44, the initial and final states are assumed to correspond to the quantum system immediately before and after the weak measurement. Any evolution of the system of the system between the actual weak measurement (at time  $t_o$ ) and the pre-selection

(at time  $t_i$ ) or post-selection (at time  $t_f$ ) must be incorporated in the states (e.g.  $|I\rangle = \hat{U}_{t_i \rightarrow t_o} |\text{pre - selected}\rangle$ ).

### 1.4.3 Properties of weak values

Weak values have a number of properties in common with standard expectation values [70]:

1. If there is no post-selection, the weak value is equal to the standard expectation value of the weakly measured observable:

$$\langle \hat{A} \rangle_w = \frac{\langle I | \hat{A} | I \rangle}{\langle I | I \rangle}. \quad (1.45)$$

Since the initial state is undisturbed by the weak measurement and there is no post-selection,  $|F\rangle = |I\rangle$ .

2. If either the pre- or post-selected state is an eigenstate of the weakly measured observable then the weak value is equal to the corresponding eigenvalue:

$$\langle \hat{A} \rangle_w = \frac{\langle a_i | \hat{A} | I \rangle}{\langle a_i | I \rangle} = \frac{\langle a_i | a_i | I \rangle}{\langle a_i | I \rangle} = a_i. \quad (1.46)$$

A strong measurement of  $\hat{A}$  after pre-selection in  $|a_i\rangle$  would return  $a_i$  with certainty, regardless of what post-selection is subsequently performed. Similarly, if the state is post-selected in  $|a_i\rangle$  then a previous strong measurement of  $\hat{A}$  must have returned  $a_i$  and collapsed the state to  $|a_i\rangle$ . Consequently, the weak value is equal to the standard expectation value of  $\hat{A}$  in this situation.

3. Weak values are linearly related in the same manner as the operators describing the observables:

$$\langle \hat{C} \rangle_w = \langle \alpha \hat{A} + \beta \hat{B} \rangle_w = \frac{\langle F | (\alpha \hat{A} + \beta \hat{B}) | I \rangle}{\langle F | I \rangle} \quad (1.47)$$

$$= \alpha \langle \hat{A} \rangle_w + \beta \langle \hat{B} \rangle_w. \quad (1.48)$$

Standard expectation values are related in the same manner.

4. Like standard expectation values, the weak value of the product of two observables is not necessarily equal to the product of the weak values for the two observables:

$$\langle \hat{A}\hat{B} \rangle_w = \frac{\langle F | \hat{A}\hat{B} | I \rangle}{\langle F | I \rangle} \neq \quad (1.49)$$

$$\langle \hat{A} \rangle_w \langle \hat{B} \rangle_w = \frac{\langle F | \hat{A} | I \rangle \langle F | \hat{B} | I \rangle}{\langle F | I \rangle \langle F | I \rangle}. \quad (1.50)$$

Taken individually, each of these four properties is not surprising since they match those of standard expectation values. However, since weak measurements do not disturb the measured system, all of these properties must hold simultaneously (unlike the strong measurements used to measure standard expectation values). For instance, if  $|I\rangle = |b\rangle$  and  $|F\rangle = |a\rangle$ , then  $\langle \hat{A} \rangle_w = a$  and  $\langle \hat{B} \rangle_w = b$  (property 2) and  $\langle \hat{C} \rangle_w = a + b$  where  $\hat{C} = \hat{A} + \hat{B}$  (property 3). This is surprising, since if  $\hat{A}$  and  $\hat{B}$  do not commute,  $a + b$  will in general lie outside the range of the eigenvalues of  $\hat{C}$ , which is impossible for  $\langle \hat{C} \rangle_w$ . Moreover,  $\hat{A}$ ,  $\hat{B}$ , and  $\hat{C}$  can be *weakly* measured simultaneously even though they do not commute. Thus, we have a fifth property that diverges from the properties of standard expectation values.

5. In general, the weak value can be anywhere in the complex plane:

$$\langle \hat{A} \rangle_w \equiv \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle}. \quad (1.51)$$

The numerator gives the transition amplitude for the operator  $\hat{A}$  to take the initial to the final state. The magnitude of such a transition amplitude must be less than the largest eigenvalue of  $\hat{A}$ . The denominator, on the other hand, will cause the weak value to blow up if the overlap of the initial and final states becomes small. Thus, weak values that are far outside the range of  $\hat{A}$  are accompanied by a small probability for successful post-selection,  $\text{Prob}_{\text{success}} = |\langle F | I \rangle|^2$  [71]. Another possibility for the weak value to lie outside the normal range is for it to be complex, since the denominator and the numerator will, in general, be complex. In this way, the weak value can diverge

from the range of standard expectation values even if the post-selection probability is substantial.

#### 1.4.4 Experimental weak measurement

The fifth property above is what first drew attention to weak values and weak measurement. Aharonov, Albert and Vaidman published a Physical Review Letter claiming that the measurement of a spin component of a  $S=1/2$  particle could result in a spin of 100 [29]. In 1991, Richie, Storey, and Hulet performed an optical implementation testing this claim, realizing the first measurement of a weak value [72]. This section will briefly describe this experiment and, based on an analysis by Duck, Stevenson and Sudarshan in 1989 [73], will explicitly show how interference leads to the strange result.

The experiment is performed with light and uses the photon polarization in the place of a spin  $1/2$  system. The photons that come from a laser are prepared in an initial state of  $|I\rangle = |45^\circ\rangle$ , and post-selected in an almost orthogonal state  $|F\rangle = |(-45 + \delta)^\circ\rangle$ , with polarizers. In between these two strong, projective, measurements, a weak measurement of  $\hat{\sigma}_z = (|H\rangle\langle H| - |V\rangle\langle V|)/2$  is performed, where H and V are the horizontal and vertical polarizations. The large weak value for  $\hat{\sigma}_z$  is due to the post-selection success probability  $|\langle F|I\rangle|^2 = \sin^2(\delta^\circ)$ , which is close to zero when  $\delta$  is small. It also follows from the linearity of weak values, property 3; from the pre- and post-selected states we know that the weakly measured photons have a spin projection of 1 along both  $|45^\circ\rangle$  and  $|(-45 + \delta)^\circ\rangle$ , two vectors that point in almost opposite directions in the x-z plane of the Bloch sphere. By the rules of geometry, any projection of the spin along a third direction in the x-z plane will necessarily have a magnitude greater than one. Substituting the initial and final states into Eq. 1.44, the exact weak value for  $\hat{\sigma}_z$  is

$$\langle \hat{\sigma}_z \rangle_W = \frac{\sin(45^\circ + \delta) + \cos(45^\circ + \delta)}{2\sqrt{2}\sin(\delta)}. \quad (1.52)$$

The interaction Hamiltonian (given in Eq. 1.19) in the von Neumann measurement model couples the measured system to an ancilla system. However, the ancilla system need not be an entirely separate entity (i.e. another particle). More frequently, the pointer is an internal parameter of the measured system that is outside the space of  $\hat{A}$  and so can be treated as independent system. However, any measurement of the pointer must typically wait until after the post-selection, as now the measurement usually destroys the measured system,  $|\psi_S\rangle$ . This variation on the von Neumann model is particularly important for experiments with light since photons do not interact strongly with themselves or with other quantum systems. In Hulet *et al.*'s experiment, the transverse position of the photons is used as a pointer. The transverse profile of the laser beam is Gaussian and gives the wavefunction, and thus the position uncertainty of the photons. A measurement of  $\hat{\sigma}_z$  should displace the photons by an amount dependent on their polarization in the  $H - V$  basis. This is accomplished with a birefringent piece of quartz, cut so that one polarization walks off from the other as the light travels through the crystal. The displacement has to be small compared the initial position uncertainty of the photons for the measurement to be considered weak. After the post-selection of  $|(-45 + \delta)^\circ\rangle$  by the second polarizer, the pointer, and the weak value it carries, is finally measured. In this case, the position of the photon is detected with a CCD, an act which necessarily destroys the photon by absorbing it.

In the Hulet experiment, the weak value is proportional to a shift in the pointer's position since Eq. 1.52 is purely real. A different choice of initial or final states or the measured observable could lead to a complex or purely imaginary weak value. For example, if photons were instead prepared in an initial state of  $|I\rangle = (|H\rangle + i|V\rangle)/2$ , right-hand circular, and post-selected  $|F\rangle$  from above, then the weak value  $\langle\hat{\sigma}_z\rangle$  equals  $i/\sqrt{2}$ . The imaginary portion of the weak value would be equal to the average momentum shift of the photons, which could be measured by moving the CCD to the focal plane of a lens placed after the second polarizer. The imaginary portion of the



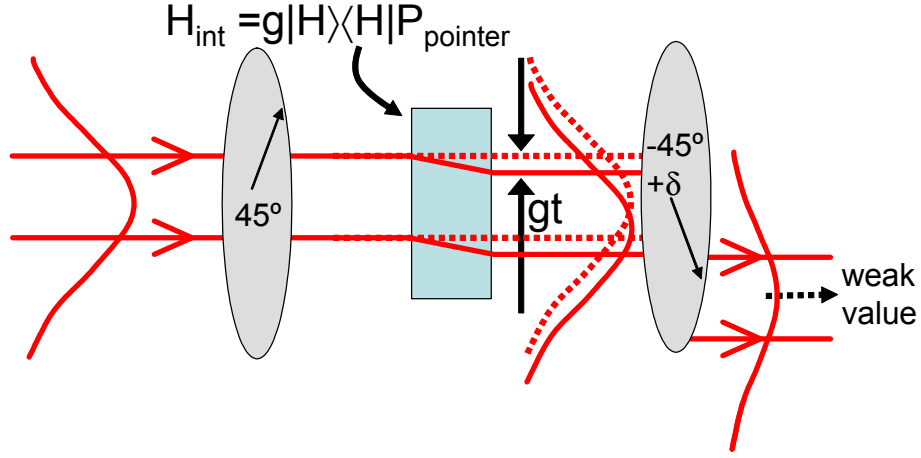


Figure 1.7: After the crystal, there are two orthogonally polarized copies (actually states in a superposition) of the original pointer wavefunction. At the final polarizer, these two copies interfere with a relative phase and weight that depends on the polarizer angle. The result is a pointer that is displaced by the weak value.

weak value has not been measured in any experiment, although the measurement of a complex optical tunneling time by [74] is closely related.

We now outline Duck *et al.*'s explanation for the emergence of the weak value [73]. Before the post-selection the state of the full system is given by

$$|\psi(t)\rangle = \sum_n \langle a_n | I | a_n \rangle |\phi(gta_n)\rangle. \quad (1.53)$$

The post-selection collapses the state of the pointer into the following state,

$$|\phi_{fi}\rangle = \frac{1}{\text{Norm}} \sum_n \langle a_n | I \rangle \langle F | a_n \rangle |\phi(gta_n)\rangle, \quad (1.54)$$

where *Norm* is just a normalization constant, calculated by the usual procedure.

This is a superposition of  $n$  Gaussians, each displaced by an amount proportional to an eigenvalue of  $\hat{A}$ . If the Gaussians have a small overlap (i.e.  $|\langle \phi(gta_m) | \phi(gta_n) \rangle| \approx \delta_{mn}$ ), then the position probability distribution of the pointer is a comb of peaks of different heights, equally spaced. The average position of the pointer is just the

weighted average of the individual displaced pointer positions,

$$\langle \hat{X} \rangle_{fi} = \sum_n \left| \frac{\langle a_n | I \rangle \langle F | a_n \rangle}{Norm} \right|^2 gta_n. \quad (1.55)$$

The weights are just the respective probabilities for the pointer to be shifted by  $gta_n$  (they also give the relative peak heights). Conversely, when the displacements  $gta_i$  are small relative to the width of the Gaussians we can no longer sum their positions probabilistically as in the above or Eq. 1.27. Instead, the Gaussians interfere constructively or destructively according to the amplitudes in Eq. 1.54. Yet, as long as the measurement is weak enough, the resulting probability distribution somehow ends up as a Gaussian shifted by  $\langle \hat{A} \rangle_W$ .

If we substitute the initial and final states of the Hulet experiment into Eq. 1.54 we can explicitly see how this rigid shift happens:

$$\begin{aligned} \langle x | \phi_{fi} \rangle &\approx (\langle H | 45^\circ \rangle \langle -45^\circ + \delta | H \rangle \phi(0) + \langle V | 45^\circ \rangle \langle -45^\circ + \delta | V \rangle \phi(gt)) / Norm \quad (1.56) \\ &= \frac{1}{\sqrt{2} Norm} (\sin(-45^\circ + \delta) \phi(0) - \cos(-45^\circ + \delta) \phi(gt)) \\ &\approx \frac{1}{\sqrt{2} Norm} \left( \sin(-45^\circ + \delta) \phi(0) - \cos(-45^\circ + \delta) \langle x | \hat{T}(gt) | \phi(0) \rangle \right) \\ &= \frac{1}{\sqrt{2} Norm} \left( \sin(-45^\circ + \delta) \phi(0) - \cos(-45^\circ + \delta) \left( \phi(0) - gt \frac{\partial \phi(0)}{\partial x} \right) \right) \\ &= \frac{1}{\sqrt{2} Norm} (\sin(-45^\circ + \delta) - \cos(-45^\circ + \delta)) \phi(0) + \cos(-45^\circ + \delta) gt \frac{\partial \phi(0)}{\partial x} \\ &= \langle x | \hat{T} \left( \frac{\cos(-45^\circ + \delta) gt}{\sin(-45^\circ + \delta) - \cos(-45^\circ + \delta)} \right) | \phi(0) \rangle \\ &\approx \left\langle x \left| \phi \left( \frac{\cos(-45^\circ + \delta) gt}{\sin(-45^\circ + \delta) - \cos(-45^\circ + \delta)} \right) \right. \right\rangle \\ &= \langle x | \phi(\langle \hat{\sigma}_z \rangle_W + 1/2) gt \rangle. \end{aligned}$$

If  $gt$  is small enough, the interference results in a pointer unchanged in shape but shifted by  $\langle \hat{\sigma}_z \rangle_W$ . (In contrast to Eq. 1.19, only the V polarized photon is shifted. Consequently,  $\langle \hat{\sigma}_z \rangle = 0$  is now at  $gt/2$ .) The denominator in the penultimate line is close to zero for  $\delta$  small, which causes the shift, and therefore the measurement result,

to become large. In particular, in the limit  $\delta \rightarrow 0$  the pointer shift equals  $k/\delta$  and hence blows up. However, the pointer will only retain its shape if the shift is much smaller than its width. For larger shifts, the above first-order expansions are invalid. In Appendix A, we consider the exact conditions for weakness.

### 1.4.5 The meaning of the weak value

Over the fifteen years since its introduction, the exact meaning of the weak value is still an open and controversial subject [75][76][77][59][71]. In particular, how should experimentalists interpret the results of weak measurements? And what do weak values tell us about the measured observable  $\hat{A}$  or the initial and final states? Because the weak value can lie outside the range of the eigenvalues of  $\hat{A}$ , interpreting it as equivalent to the expectation value of  $\hat{A}$  leads people to odd and surprising conclusions, such as that an electron has a spin of 100. Perhaps even more troublesome is the imaginary part of the weak value, which is even more distant from standard measurement theory. Some suggest that weak values are not the results of “true” or “real” measurements and consequently tell us little that is meaningful about the measured system [63][62][64]. Nonetheless, both the procedure for measuring the weak value and the formula itself suggest some straightforward and compatible interpretations that we discuss below.

#### Classical Interpretation

The process of measurement arguably does not have a full model contained entirely within quantum mechanics. In particular, the von Neumann measurement model really only models the first stage of measurement, the interaction with the measuring apparatus. This is sufficient to model the measurement-induced decoherence of the measured system, but it is not sufficient to model collapse, which invariably requires the intervention of a classical, deterministic system to measure the position

of the pointer [78]. In this sense, all quantum measurement procedures at some point reduce to classical measurement procedures. Consequently, we first look to classical measurement for the meaning of the weak value by asking how a classical physicist would understand a weak measurement.

Consider a classical physicist who measures parameter  $a$  and then  $b$  for each member in an ensemble of systems. We describe the ensemble by a probability distribution,  $P(a, b)$ , which corresponds to the initial state of the system in weak measurement. The first measurement follows the von Neumann measurement model, which is equally valid in classical physics. (However, in classical physics the ensemble of pointers is described by a phase-space probability distribution rather than by a wavefunction.) If the measurement of  $a$  created an unwanted disturbance (for example one that changes the parameter  $b$ ), then it would be natural for the classical physicist to attempt to reduce that disturbance by similarly reducing the interaction with the pointer, just as in weak measurement. The disturbance from the measurement of  $b$  is of no concern since it is the final measurement, and so it is not a weak measurement. Due to their imprecision, the classical physicist would not be surprised that individual measurements of  $a$  are now impossible. He or she would expect that, given enough measurements, he or she could determine the average value of  $\bar{a}$  from the average pointer shift. If this average was restricted to the subset of measurements in which  $b$  was subsequently found to be equal to  $F$ , then this value, which we call the weak value  $\langle \hat{A} \rangle_W$ , would be interpreted as the average value of  $a$  given  $b = F$ ,  $\langle a \rangle_F$ . If the classical physicist were instead interested in a specific value of  $a$ , say  $a_i$ , then he or she would need to measure the projector,  $\pi_{a_i}$ . In general the average value of a projector (e.g.  $\pi_{a_i}$ ) is the probability for  $a_i$ ,  $P(a_i)$ , given the initial ensemble. Therefore, a classical physicist would interpret the weak value of such a projector as  $P(a_i|F)$ , the conditional probability of  $a_i$  given that  $F$ .

Furthermore, these weakly measured averages and probabilities will be related in exactly the way a classical physicist would expect them to be. An operator can be

expressed as a sum of projectors:

$$\hat{A} = \sum_n a_n \cdot \hat{\pi}_{a_n}. \quad (1.57)$$

And by the linearity of weak values, property 3, it follows that

$$\langle \hat{A} \rangle_W = \sum_n a_n \cdot \langle \hat{\pi}_{a_n} \rangle_W \quad (1.58)$$

$$= \sum_n a_n \cdot P(a_n|F) \quad (1.59)$$

$$= \langle a \rangle_F. \quad (1.60)$$

Just as the classical physicist would expect, the weak values for the projectors sum in the standard way to give the conditional average for  $a$ ,  $\langle a \rangle_F$ . The weak values also sum correctly to give the unconditional probability for  $a_i$ ,

$$P(a_i) = \sum_n P(b_n) \cdot P(a_i|b_n). \quad (1.61)$$

This also follows from property 1: When there is no post-selection, the weak value and the expectation value of  $\hat{\pi}_{a_i}$  are equal.

One formal route to the concept of probability is through Kolmogorov's three theorems. The latter two features of weakly measured probabilities also follow from his third axiom, which requires that probabilities satisfy  $\sigma$ -additivity (which is defined as: The probability of  $A$  and/or  $B$  is equal to  $P(A) + P(B)$  if  $A$  and  $B$  are mutually exclusive.). Another intuitive property of weakly measured conditional probabilities is that they sum to unity. This follows from property 3:

$$\hat{I} = \sum_n \hat{\pi}_{a_n} \quad (1.62)$$

$$\langle \hat{I} \rangle_W = \sum_n \langle \hat{\pi}_{a_n} \rangle_W \quad (1.63)$$

$$1 = \sum_n P(a_n|F). \quad (1.64)$$

This last property is a variation of Kolmogorov's second axiom. In summary, weakly measured conditional probabilities will satisfy many of the properties a classical physicist would expect them to (see also [59][32][71]). One property they will generally not

satisfy, however, is Kolmogorov's first axiom, which requires that all probabilities are greater than zero. In contrast, property 5 implies that the weakly measured probability could be negative or even complex.

This is a strange concept but it is not without precedent in quantum mechanics. Starting with Feynman [79][80], people have considered and developed the possibility of negative and complex probabilities in classical probability theory and in quantum mechanics [81][82][83][84][85][86]. The Wigner phase-space quasi-probability distribution is an attempt to assign a probability to every point in phase-space for a particle [87]. Particles in a classical-like state (e.g. a coherent state or a mixture of different coherent states) will have a Wigner distribution that is positive everywhere, whereas certain quantum states (e.g. a *superposition* of two coherent states) will be negative at certain points in phase-space. In this sense, negative probabilities in the Wigner function are sometimes used to decide whether a state requires a quantum description [88]. As we will see in Chapter 4 of this thesis, negative weakly measured probabilities can be used for a similar purpose. (However, in contrast to the weak value the Wigner distribution is bilinear in the wavefunction. The two are not equivalent or simply related. In addition, as will be shown at the end of this chapter, the weak value gives the phase of the wavefunction directly.) In Bell's theorem, the probabilities predicted by quantum mechanics to violate the inequality also imply that some unmeasured probabilities are negative [80] (this is also true of Hardy's Paradox, as we will see in Chapter 6). An important distinction is that, unlike the latter two examples, weakly measured negative probabilities are not implied or derived but are instead directly read from the measurement apparatus.

### **Bayesian Interpretation**

The standard Copenhagen interpretation of quantum mechanics prohibits us from assigning a value to an observable that has not been measured. This assignment is not a problem exclusive to quantum mechanics. It can also be problematic in classical

physics as well as in classical logic and philosophy. Basing predictions on events in the past is called retrodiction, a useful tool for fields like astrophysics where predictions can only be tested against the past or present. The use of retrodiction for events that actually did not occur or which had a different outcome than reality leads to *counterfactual* statements. Straying from the subject of quantum mechanics for a moment, take the statement “If Archduke Franz Ferdinand had not been assassinated, World War One would not have started.” These two events are sometimes thought to share a cause and effect relationship [89]. However, it is also possible that the assassination was merely a pretext and that the war would have occurred regardless. Since these events occurred in the past it is impossible to test whether the counterfactual statement is true, which is why these statements can be problematic, even though they are commonly made in ordinary conversation.

In quantum mechanics the problem is compounded by the fact that we can only simultaneously measure *commuting* observables (see also [90][91][92]). And once we have made these measurements, we can only make counterfactual statements about the value of other, non-commuting, observables, perhaps guided by previous observations on identical systems. A good example of this is Hardy’s Paradox (which is the subject of Chapter 3) and another is the EPR paradox. This problem makes it difficult to apply the tools of statistical inference to quantum mechanics. Statistical inference concerns the task of inferring properties of an unknown distribution from data generated by that distribution. At its heart is Bayes’ theorem, which relates conditional and unconditional probabilities:

$$P(a|b) = \frac{P(a,b)}{P(b)}. \quad (1.65)$$

Steinberg considered the form of Bayes’ theorem in quantum mechanics [58][59]. As we saw in the last section, in standard (classical or quantum) measurement theory,  $P(b) = \langle \hat{\pi}_b \rangle$ . In classical measurement theory this can be extended to a joint probability,  $P(a,b) = \langle \hat{\pi}_a \hat{\pi}_b \rangle$ . In quantum mechanics, however,  $\hat{\pi}_a \hat{\pi}_b = |a\rangle \langle a| b\rangle \langle b|$  is

generally not Hermitian and so cannot be measured directly with standard measurement. Nonetheless, if we write the classical joint probability formula in terms of the quantum states and observables, we arrive at the following form for Bayes' theorem,

$$P(a|F) = \frac{\langle I|a\rangle \langle a|F\rangle \langle F|I\rangle}{\langle I|F\rangle \langle F|I\rangle} \quad (1.66)$$

$$= \langle \hat{\pi}_a \rangle_W. \quad (1.67)$$

Once again, the weak value appears to have a close relationship to the classical conditional probability of  $a$  given  $b = F$ .

Johansen applied more formal methods of statistical inference in Ref. [93]. He took the initial ensemble corresponding to state  $|I\rangle$  as our prior information about a statistical ensemble. With this information we would like to estimate the parameter  $\hat{A}$ , given that we measured  $\hat{B}$  and found value  $F$ . To do this we estimate  $\hat{A}$  with a function that takes  $F$  as its input,  $\theta(F)$ . The actual form of the function is chosen so that it minimizes a loss (or cost) function,  $L$ . Such a function is known as a Bayes' estimator. With Johansen's choice of a loss function,  $L = \left\langle \left( \theta(F) - \hat{A} \right)^2 \right\rangle$  (apparently this quadratic form is the most common loss function), he found that the form of the Bayes' estimator was the real part of the weak value. In other words, the weak value can be considered to be Bayes' estimator, also known as the most efficient estimate, on a pre- and post-selected ensemble.

### The Imaginary Component of the Weak Value

While the meaning of the real part of the weak value is somewhat uncertain, most discussions, including those described above, agree that it is a measure of  $\hat{A}$  given  $F$ . In contrast, there appear to be as many different interpretations of the imaginary part of the weak value as there are discussions of it [94][59][32]. Quite often the real part is termed the "weak value" while the imaginary part, and the proportional pointer momentum shift, is simply ignored. This is motivated by a classical interpretation of a weak measurement; a classical physicist would not expect a momentum shift and



consequently would not consider it part of the measurement's result. In addition, the magnitude of the momentum shift depends not only on the observable  $\hat{A}$  or the measured system, but also on the shape of the pointer function. For instance, for a Gaussian pointer, as the measurement gets weaker,  $\langle \hat{P} \rangle_{fi}$  decreases with  $gt/\sigma^2$  (the shift depends on the width of the pointer function). Since the width  $\Delta\hat{P}$  decreases as  $1/\sigma$ , this also means it is more difficult to determine the imaginary part of the weak value than the real part.

Still, even if the meaning of  $\text{Im} \langle \hat{A} \rangle_w$  is unknown,  $\langle \hat{P} \rangle_{fi}$  should have a physical significance. After all, the momentum shift arises due to the standard von Neumann measurement interaction Eq. 1.19. Steinberg argued  $\langle \hat{P} \rangle_{fi}$  is due to the “back-action” of the measurement [59]. Backaction is the reaction of the system to the measurement. In the limit of  $gt/\sigma \rightarrow 0$ , a perfect weak measurement, there should be no backaction and therefore a measured system will remain unchanged in state  $|I\rangle$ . However, with finite coupling the probability for the system to subsequently be found in  $|F\rangle$  is changed depending on the momentum of the pointer. Conversely, this weak correlation means that the post-selection of  $|F\rangle$  also increases the weight of certain momentum components in the pointer, which shifts average momentum in one direction or other. With this insight, we would expect that the size and direction of the shift will depend on  $|I\rangle, |F\rangle$ , and the nature of the coupling (i.e.  $\hat{A}$ ), which is consistent with the formula for  $\text{Im} \langle \hat{A} \rangle_w$ .

Johansen also considered the role of the imaginary part of the weak value [93]. He found it to be equal to the minimum of the quadratic loss function discussed in the last section. The estimator that gives this minimum is equal to the real part of the weak value. The minimum loss and hence the imaginary part of the weak value can be considered as a measure of how good or, conversely, how uncertain, the estimate of  $\hat{A}$  is.

Let us consider the example of  $\langle \hat{\sigma}_y \rangle_w = i/\sqrt{2}$ , from the experimental weak measurement section above. Since the weak value is completely imaginary, a classical

physicist would interpret this as a null result. We could also interpret this as an indication that there is relatively large backaction for this measurement. A complementary and perhaps related interpretation is that there is relatively large uncertainty in the measurement result  $\langle \hat{\sigma}_y \rangle_W$ .

Nonetheless, as the next example will show, sometimes these physical interpretations are unnecessary. Consider the following scenario: We begin with an arbitrary wavefunction as our initial state,  $\langle x | I \rangle = \langle x | \psi \rangle = \psi(x)$ . A weak measurement of the position projector  $\hat{\pi}_x = |x\rangle\langle x|$  is performed, followed by post-selection on momentum eigenstate  $|p\rangle$ . According to Eq. 1.44 the weak value will be

$$\langle \hat{\pi}_p \rangle_W = \frac{\langle \psi | x \rangle \langle x | p \rangle}{\langle \psi | p \rangle} \quad (1.68)$$

$$= \psi(x) \cdot C(p) \cdot e^{ipx/\hbar}. \quad (1.69)$$

Surprisingly, a direct measurement results in the complex-valued wavefunction itself, which is usually assumed to be only indirectly observable (e.g. in tomography, or interference experiments). The proportionality constant  $C(p)$  is inconsequential since it is independent of  $x$  (as well, the global phase of the wavefunction is arbitrary) and  $e^{ipx/\hbar}$  is a trivial phase factor that is fixed if  $p = 0$ . Consequently, one could map out the wavefunction for all  $x$  by weakly measuring  $|x\rangle\langle x|$ , varying  $x$  over some range while keeping  $p$  fixed. In this case, the imaginary part of the weak value has a very straightforward meaning as the imaginary part of the wavefunction.

The wavefunction is typically assigned an indirect role in quantum measurements. We require it to make measurement predictions but, as Bell's inequalities show, we should not take it to be a real object (e.g. a field) unless we are willing to allow superluminal cause and effect in our description of the world. Nonetheless, a classical physicist would interpret the result of the measurement in the example above as the conditional probability for  $x$  given  $p$ ,  $P(x|p)$ . He or she would see the wavefunction directly emerging from his or her measurements and thus assign it some sense of reality. Similarly, we can reverse roles of  $x$  and  $p$  to give  $P(p|x) \propto \tilde{\psi}(p)$ . In light

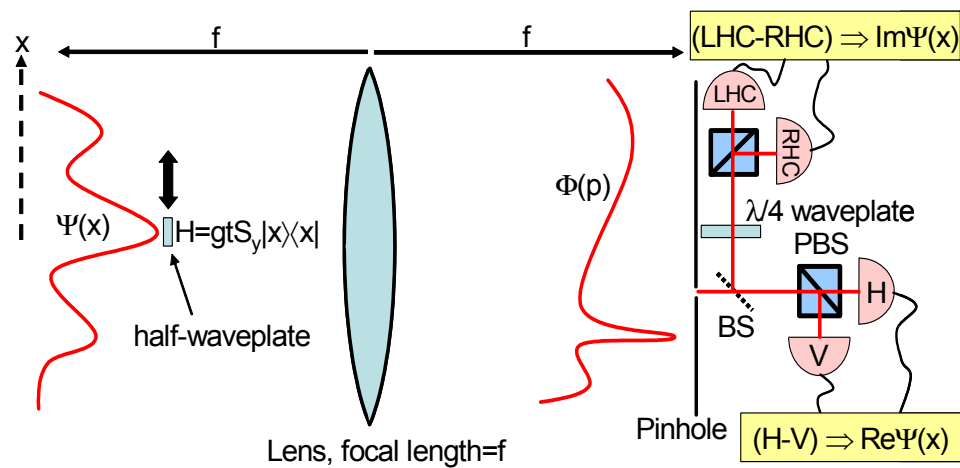


Figure 1.8: A schematic for the weak measurement based characterization of the transverse wavefunction of a photon. The initial state is an arbitrary wavefunction  $\Psi(x)$ . The position projector is measured with a small rotation induced by a translatable waveplate at  $x$ . One focal length after a lens, the wavefunction is now Fourier transformed to  $\Phi(p)$  and a pinhole post-selects on specific and fixed momentum state. Detectors after a polarizing beamsplitter (PBS) allow us to find the result of the position measurement, which is the weak value. The weak value is proportional to  $\Psi(x)$ .

of these generally complex conditional probabilities we once again consider Bayes' theorem,

$$P(p|x) \cdot P(x) = P(x|p) \cdot P(p) \quad (1.70)$$

$$\frac{\langle \psi | p \rangle \langle p | x \rangle}{\langle \psi | x \rangle} \cdot |\langle \psi | x \rangle|^2 = \frac{\langle \psi | x \rangle \langle x | p \rangle}{\langle \psi | p \rangle} \cdot |\langle \psi | p \rangle|^2 \quad (1.71)$$

$$\langle \psi | p \rangle \langle p | x \rangle \langle x | \psi \rangle = \langle \psi | x \rangle \langle x | p \rangle \langle p | \psi \rangle. \quad (1.72)$$

Remarkably, Bayes' theorem is satisfied by these two weakly measured conditional probabilities up to an overall phase, supporting the classical physicist's interpretation of the measurement results as complementary conditional properties.

### 1.4.6 Conclusion

Weak measurement features in three of the next five chapters of this thesis. At the present time, it is still largely overlooked as a tool for experimentalists. Specifically, it has never been used to do anything but produce strange and odd results. Moreover, partly because of this history, it is still very controversial. This thesis takes the first steps towards the development of weak measurement as a viable and advantageous technique for characterizing complicated quantum systems and processes, such as multi-particle quantum states and quantum processes. It will also be demonstrated that the weak values resulting from these characterizations can be analyzed and understood, and relate to each other in much the same way as the results from corresponding classical measurements. Therefore, although the exact meaning of complex weak values is still being debated we can immediately use them to understand the system under study.

# Chapter 2

## The photon-exchange effect

### 2.1 Introduction

In a creative and controversial theoretical work, Jim Franson [95] showed how photon-exchange interactions might give rise to a very strong Kerr-like nonlinearity (see also [96][97][98][99]). Such an effect would be of enormous importance in areas such as optical quantum computation, where the absence of photon-photon interactions has long been seen as a major stumbling block to producing two-qubit quantum logic gates. Nonlinear effects at the two-photon level have been demonstrated in cavity-QED systems [16][17], but such systems are extremely cumbersome, expensive, and not easily scalable. Quantum interference has been shown to enhance nonlinearities by many orders of magnitude (see the discussion of the switch in Chapter 3 [19][100]). However, the dependence on interference limits this type of nonlinearity to work on only a subset of possible input states [101]. Another popular proposal [18] for generating huge nonlinearities involves using electromagnetically-induced transparency and slow light [102][103]. Also, it has been shown that one can mimic the action of a strong nonlinearity using only linear-optical elements and conditional measurements [21][104] – such schemes rely on next-generation technology including single-photon sources “on demand” and near-perfect photon-counting

detectors. While these ideas are very important for quantum information, having “actual” nonlinearities at the quantum level open up an exciting new area of physics – *quantum* nonlinear optics (see Ref. [105] for new work in this direction).

The nonlinear phase-shifts that typically result from  $\chi^{(3)}$  nonlinearities in atomic systems grow linearly with the number of atoms in the region of interest [13], and rarely exceed  $10^{-10}$  for the interaction between a pair of photons. Franson performed a perturbation theory calculation that suggested a different type of mechanism for producing Kerr-like nonlinearities. If pairs of photons 1 and 2 of frequencies  $\omega_1$  and  $\omega_2$  interact with a pair of atoms A and B, then the exchange process whereby atom A nonresonantly “absorbs” the photon 1 and emits photon 2 while atom B absorbs photon 2 and emits photon 1 gives rise to an extra dispersive nonlinearity. However, unlike the usual Kerr effect, which grows linearly with the number of atoms in a medium,  $N$ , this additional nonlinear phase shift was predicted to grow with the number of *pairs* of atoms (of order  $N^2$ ). With a dense enough sample, this phase-shift could be strong enough to act at the quantum level. As yet, however, the effect has not been confirmed.

In a more recent effort to understand photon-exchange effects, John Sipe looked at real transitions and the absorption of photon pairs instead of such virtual transitions and dispersive effects (theory to be published by John Sipe). He predicted a cooperative effect in an atomic gas with a narrow absorption feature in which the two-photon absorption rate is larger than the product of single photon absorption rates. In this sense, this effect mimics standard nonlinear absorption in a nonlinear medium. For researchers who deal with quantum interference with photons, this cooperative effect is easiest to understand as a generalized Hong-Ou-Mandel (HOM) effect. From this perspective there is a close connection to the linear-optics quantum computing (LOQC) concept of Knill, Laflamme and Milburn [21]. As the thesis introduction demonstrates, HOM interference at a beamsplitter, which is a *linear* optical device, can be considered a nonlinear (and thus entangling) interaction between the input

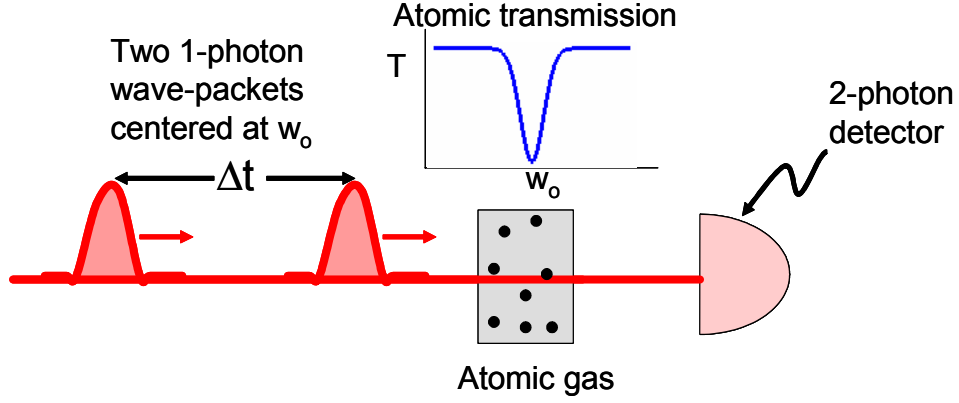


Figure 2.1: The analogous setup to Franson’s proposal with the difference that we now examine real absorption rather than phase shifts in the two photon component of the input. Two wavepackets, containing a single photon each and separated by  $\Delta t$ , pass through an atomic sample with a narrow absorption feature. The cases where both photons are transmitted are registered with a two photon detector.

photons. It follows that the cooperative absorptive effect described in this section that occurs at the atomic cloud (which exhibits only *linear* absorption) can also be understood in the context of LOQC-style nonlinear transformations. For a more general community, the cooperative effect can be understood in a photon wavefunction picture, in which photons act like localized particles. In this picture the effect described in this paper is an exchange effect similar to those observed and predicted in condensed matter systems (e.g. an effective force is often attributed to Pauli exchange in fermions [106]).

## 2.2 Theory

Using a perturbative approach, John Sipe (unpublished) calculated the two-photon absorption probability for two like-polarized photons in the same spatial mode in the state,

$$|A, B\rangle = \mathcal{N}(A, B) \iint d\omega d\omega' f_A(\omega) f_B(\omega') a_A^\dagger(\omega) a_B^\dagger(\omega') |0\rangle, \quad (2.1)$$

where  $\mathcal{N}(A, B)$  is a normalization constant,  $f_{A(B)}(\omega^{(l)})$  is the amplitude for frequency  $\omega^{(l)}$ , and  $a_{A(B)}^\dagger$  is the raising operator for a photon of type  $A(B)$ . Under these conditions, the probability of absorbing both photons,  $P_{AB}$ , can be expressed in terms of the single-photon absorption probabilities  $P_A$  and  $P_B$ ,

$$P_{AB} = P_A P_B \left( \frac{1 + \xi_{AB}}{1 + v_{AB}} \right), \quad (2.2)$$

where  $v_{AB}$  is the square of the overlap integral,  $|\int d\omega f_A^*(\omega) f_B(\omega)|^2 \leq 1$ , and

$$\xi_{AB} = \frac{|\int d\omega g(\omega) f_A^*(\omega) f_B(\omega)|^2}{[\int d\omega g(\omega) |f_A(\omega)|^2] [\int d\omega g(\omega) |f_B(\omega)|^2]}, \quad (2.3)$$

where  $g(\omega)$  is the absorption spectrum of the medium. For independent absorption events we expect  $P_{AB} = P_A P_B$ ; therefore correlated absorption probabilities come from cases where  $v_{AB} \neq \xi_{AB}$ . Such a case can be set up if the two photons are separated in time, but pass through a medium with a narrow absorption feature of width,  $\Delta\omega_a$ . Since the photons do not overlap,  $v_{AB} \approx 0$ . Provided that the absorber has a coherence time,  $1/\Delta\omega_a$ , longer than the delay between the photons  $\tau$ , we may have  $\xi_{AB} \neq 0$ . In other words, if the photons are distinguishable before absorption, but become (mostly) indistinguishable if absorbed, we may have  $P_{AB} > P_A P_B$ , an enhancement of the two-photon absorption. In cases where the photon delays are much longer than the coherence time of the absorber, or when the photons are perfectly overlapped,  $P_{AB} = P_A P_B$ , as expected in the absence of any nonlinear effects. Using a photon wavefunction picture, Geoff Lapaire and John Sipe moved beyond the perturbative approach for calculating the two-photon absorption probability. Precise expressions were derived for  $P_{AB}$  as a function of the delay for general  $f_A(\omega)$  and  $f_B(\omega)$  and also for a specific case. The conditions under which correlated absorption events occur are, nevertheless, the same as found in John Sipe's original perturbation theory.

In the photon wavefunction picture, anomalous absorption effects are due to exchange terms. Enhancement of the two-photon absorption probability occurs when



one has constructive interference between the amplitude for photon A to be absorbed by atom 1 and photon B to be absorbed by atom 2 and the reverse. One can illustrate the photon exchange effect in the photon wavefunction picture, even for two-photon states  $|A, B\rangle$  which are not restricted to one-dimension (as in Eq. 2.1). But in that one-dimensional limit the coincidence detection rate of two ideal detectors at positions  $z_1, z_2$  and at times  $t_1, t_2$  is proportional to,

$$w^{(2)}(z_1, z_2, t_1, t_2) = |\mathcal{N}(A, B)|^2 \left[ \begin{aligned} &|\psi_A(z_1, t_1)|^2 |\psi_B(z_2, t_2)|^2 + |\psi_A(z_2, t_2)|^2 |\psi_B(z_1, t_1)|^2 \\ &+ \psi_A^*(z_1, t_1) \psi_B(z_1, t_1) \psi_B^*(z_2, t_2) \psi_A(z_2, t_2) \\ &+ \psi_A^*(z_2, t_2) \psi_B(z_2, t_2) \psi_B^*(z_1, t_1) \psi_A(z_1, t_1). \end{aligned} \right]. \quad (2.4)$$

The first-order photon wavefunctions,  $\psi_{A(B)}(z, t)$ , satisfy  $E^+(z, t) |A(B)\rangle = \psi_{A(B)}(z, t) |0\rangle$ , where  $E^+(z, t)$  is the positive frequency component of the electric field operator, and  $|A(B)\rangle = \int d\omega f_{A(B)}(\omega) a^\dagger(\omega) |0\rangle$  is a single-photon state. The last two terms on the right-hand side of Eq. 2.4 are the exchange terms. In the absence of these terms, the coincidence detection rate is proportional to the sum of the product of the individual photon detection rates at either detector, which is characteristic of independent detection events. Applying this wavefunction description to a Hong-Ou-Mandel interferometer [46], one can show that the Hong-Ou-Mandel dip is simply a manifestation of photon exchange. Considering now the presence of an absorbing medium, the photon wavefunctions allow one to see the underlying physics in which, in this case, destructive interference occurs between the detection events. The wavefunctions, perhaps initially non-overlapping before passage through the absorbing medium, may no longer be after their passage through it. Thus there can be interference effects in the subsequent detection, and a corresponding reduction in the two-photon transmission probability below the uncorrelated-absorption prediction. This is the experimental signature we seek. Note that this signature is nonclassical (in the sense that it relies on the quantum statistics of the particles involved) since the analogous calculation for fermions (with their anticommutation relation) predicts an increase in

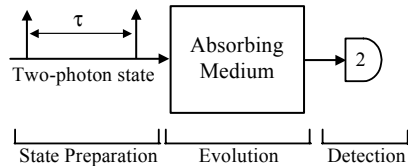


Figure 2.2: The system of interest. A two-photon state with a variable time delay between the like-polarized photons impinges on an absorbing medium. The light that passes through the absorption region is detected by a photon-pair detector. Under the right conditions, the probability of absorbing a photon pair can depend very strongly on the delay,  $\tau$ .

the two-particle transmission probability where there is a decrease for bosons.

To investigate the anomalous two-photon transmission we used the basic system sketched in Fig. 2.2. There are essentially three parts to both the experiment and the theory: state preparation, evolution, and measurement. The quantum state of interest is a normalized two-photon state where both photons have the same polarization and spectrum and the time delay between the photons can be controlled. This state evolves as it passes through an absorptive medium which has some resonant absorption feature in the center of the photon spectrum. Finally, the photon pairs emerging from the medium are counted.

## 2.3 Experiment

The two-photon states were created in the setup shown in Fig. 2.3. Such a setup is essentially a polarization-based Hong-Ou-Mandel (HOM) interferometer [46][107]. We use coincidence detection to post-select the state of interest. Specifically, we used a collinear type-II phase-matched parametric downconversion source (0.1-mm thick BBO) pumped by the second harmonic of a Ti:Sapphire laser. The second harmonic was centred at 405nm (with a bandwidth of 7nm FWHM) and created

photon pairs each with a center wavelength of 810nm. The photon pairs exit from such a source with orthogonal polarizations (one photon has horizontal polarization and the other has vertical). We control the relative time delay between the photons by passing them through a modified Babinet compensator. After the variable delay, a polarizer is placed in the photons' path at  $45^\circ$ . With the polarizer in the system, any photon pairs transmitted through this polarizer are thus polarized at  $45^\circ$ , with a time separation determined by the Babinet. This serves as the source of two-photon states required by the theoretical work. From HOM interference [46], we expect an increase in the number of photon pairs created near zero delay as the photons tend to pair up. We measure this increase when there is no absorber, and divide it out of our signal.

The theory also required an absorber with a narrower absorption feature than the bandwidth of the photon pairs. From our downconversion source, we obtain photon pairs with a FWHM power spectrum of over 100nm. The reflection from a dielectric interference filter (CVI F10-810-4-1.00) was used as our absorption medium. This filter has an 10nm wide absorption feature centred at 810nm. Light transmitted through the filter was discarded. It should be noted that a gaseous atomic sample could be used if one had a downconversion source with a much narrower bandwidth (on the order of the width of an atomic resonance); such narrowband sources of photon pairs have been demonstrated in optical parametric oscillators pumped below threshold [108][109].

The experimental setup is shown in Fig. 2.3. We first measured the number of photon pairs reflected from the broadband dielectric mirror as a function of the delay using a cascaded pair of SPCMs (Perkin Elmer SPCM-AQR-13). This data set shows any changes in the efficiency of two-photon state production which will be divided out. Then the interference filter (with the blacked out back) is placed directly in front of a broadband dielectric mirror. When the filter is in place we measured the number of remaining photon pairs. We monitor the detectors' singles rates and their

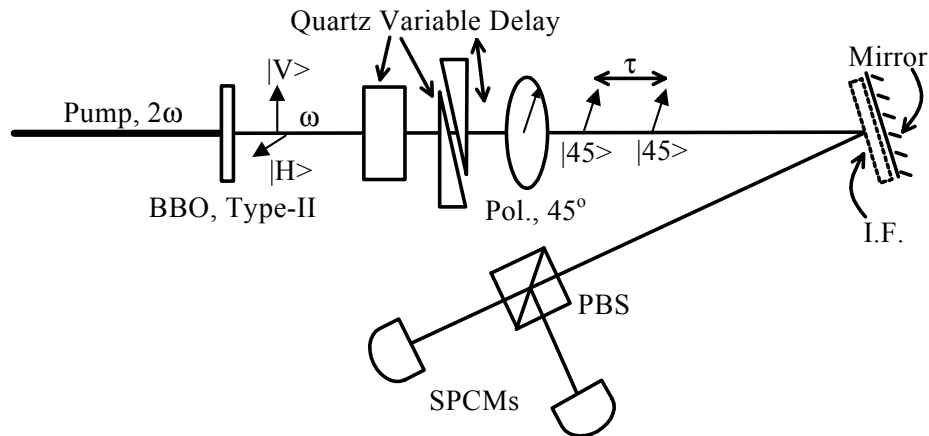


Figure 2.3: The experimental setup. The state preparation is accomplished using the output of a polarization-based HOM interferometer. BBO is a  $\beta$ -barium borate nonlinear crystal phase-matched for type-II down-conversion; PBS is a polarizing beamsplitter; SPCMs are single photon counters; Pol. is a polarizer. The pump laser is separated from the down-conversion beams using a fused silica prism (not shown). The two-photon state is prepared conditioned on successful post-selection of a photon pair after the polarizer. Once prepared in the right quantum state, the light passes reflects off either a removable interference filter (I.F.) (CVI F10-810-1.00-4) (to simulate a medium with a very broad absorption line) or a broadband dielectric mirror. In either case, the  $45^\circ$  polarized light is then split at a polarizing beamsplitter which has a single-photon detector in each output. The PBS in this case has no special function other than as a 50/50 beamsplitter since the photons are both  $45^\circ$  polarized.

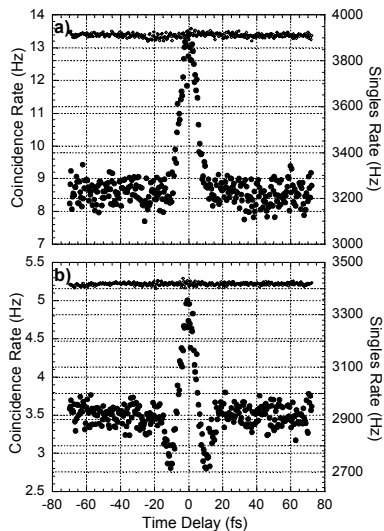


Figure 2.4: Experimental data. The coincidence rate and singles rate as a function of the inter-photon delay are shown in the case where a) the “absorber” (interference filter) is removed and when b) the “absorber” is in place.

coincidence rate in the experiment.

## 2.4 Results

Fig. 2.4 shows the rates of photon-pair detection (solid circles) and singles rate (small open diamonds) at one of the detectors as a function of the time delay between the photons. Fig. 2.4a shows the data taken while the interference filter was removed from the beam path and the light reflected from a broadband mirror; Fig. 2.4b shows data taken when the filter was in place. Fig. 2.4a clearly shows that the two-photon state preparation becomes much more efficient at zero delay. This increased pairing is due to HOM interference [46]. In our experimental results the rate of photon pair production at zero delay is 55% larger than that at large time delays; perfect HOM interference leads to a doubling of the rate. As one would expect from a HOM interferometer with low collection or detection efficiencies, the singles rate is

featureless at the 1% level as a function of the time delay between the photons [110]. Fitting the data in Fig 2.4a. under the assumption of identical Gaussian power spectra for the two photons yielded a FWHM of 129nm. Fig. 2.4b shows the data taken with the interference filter in place. The most striking difference from the previous figure is the drop in the number of photon pairs detected at a delay of approximately  $\pm 10$  fs. There is a second, more subtle, difference in that the number of photon pairs at zero delay is enhanced by only 42% over the rate at large time delays with the filter in place. The singles rate in Fig. 2.4b also shows no dependence on the time delay at the 1% level [111].

The ratio of the data in Fig. 2.4a to Fig. 2.4b is shown in Fig. 2.5 as solid circles. This ratio normalizes the data and shows the photon pair reflection probability. To reduce the noise on the data points a 5-point average was taken. The photon pair detection rate at  $-10$  fs is 15% less than the ratio at large delays and 17% less at  $+10$  fs. There is also a drop in the number of pairs measured at zero delay which is not predicted by the theory. This discrepancy can be explained by the imperfect visibility of our HOM interference, which we attribute to imperfect mode-matching. If the interference were perfect, then the rate at zero delay would always be twice that at long enough time delays. In the case of imperfect interference, one often uses the *transmission* through narrow band interference filters to remove distinguishing spectral information and raise interference visibility. The *reflected* light therefore carries away more distinguishing information. Since we use the reflected light instead of the transmitted, we expect that using the interference filter actually hurts the visibility, as observed. This drop in visibility not only keeps the experimental data from returning to a ratio of 1 at zero delay, but also serves to exaggerate the suppression in the photon pair transmission. This makes obtaining an accurate enhancement factor difficult.

Fig. 2.5 also shows the theoretical predictions for the coincidence rate ratio as a function of the time delay between the photons in two different regimes using

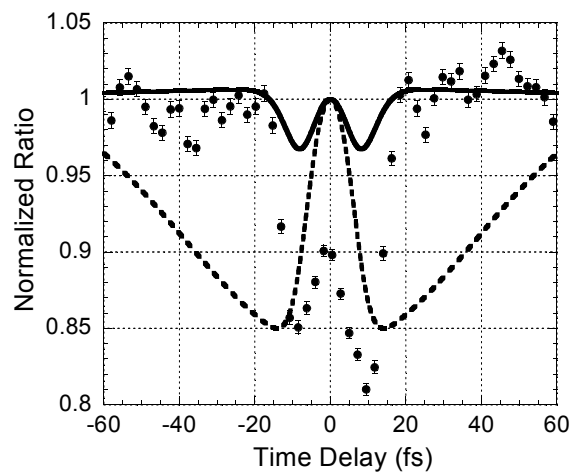


Figure 2.5: Experimental and theoretical normalized ratio of coincidence rates. The data is the ratio of the coincidence rates from Fig. 2.4b to Fig. 2.4a. The maximum drop in the coincidence rate ratio occurs for delays of  $\pm 10$  fs and is approximately 15% less than the rate at long times. The data has been normalized to the average rate at times further than 0 from  $\pm 20$  fs. The theoretical predictions are also shown. The case where there are initially no frequency correlations (NFC) between the two photons is shown as a solid line. The case where the frequencies sum to a well-defined value (PFC) is shown as a dotted line.

Gaussian spectra and a Gaussian absorption feature. The two different curves correspond to photon spectra that are, at least initially, uncorrelated, as in Eq. 1, and where the sum of the photon energies is a constant. The first regime describes the light produced from a downconversion source with a broadband pump laser after sufficiently narrow bandpass filtering, whereas the second describes downconversion created by a CW laser [112][113]. The curves are both scaled to their coincidence rate at very large time delays. In both situations, the enhanced region of photon absorption occurs at approximately  $\pm 10$  fs, in good agreement with the data. The experimentally observed suppression of 15% is in better agreement with the theory for perfect frequency correlations than with that for no frequency correlations. With perfect HOM interference we expect that the suppression could double. The shape of our experimental curve is, however, in better agreement with the theoretical curve with no frequency correlations, where the transmission reaches its long delay value at about  $\pm 20$  fs. The theory for perfect frequency correlations shows suppressed transmission over much longer times. In our experiment we are actually between these two extreme regimes; our femtosecond pump laser has a bandwidth of about 7nm. It is clear from the theory that these correlations can greatly influence the photon pair transmission probability. Such a striking dependence on frequency correlations makes this technique useful for measuring them. This will be the subject of future work.

## 2.5 Is this a quantum effect?

An ongoing question regarding multiple-photon interference experiments is which phenomena are truly quantum and which could be described in a classical model. A number of entangled photon systems were at first thought to have a significant advantage in certain applications when compared to conceivable classical light systems (examples include ghost imaging [114][115][116], quantum polarimetry [117][118] and



coherence tomography[119][120]). In particular, Hong-Ou-Mandel (HOM) interference came into question as a truly quantum effect (see the thesis introduction for a review of the HOM effect). The characteristic experimental signature of the HOM effect was that as the path-length difference  $\tau$  to the beamsplitter (BS) was varied, the coincidence detection rate at the two BS output ports would monotonically drop from a flat background to zero at  $\tau = 0$ . Conversely, it was obvious that two classical light beams would only display first-order interference at the beam-splitter with the associated interference fringes, absent from the HOM effect. However, given that single photons have a completely random phase as does each mode in spontaneous parametric downconversion (which was used as a source of photon pairs for the first demonstrations of the HOM effect), perhaps the phase of the two classical beams should also be randomized. If one were solely measuring the intensity at one output port, as is typical in a first-order interferometer, this randomization would wash out the interference, leaving no visible effect at all. This is not the case when measuring a coincidence rate because in a classical model the coincidence rate is proportional to the *product* of the intensities at the two detectors. The coincidence rate for the randomized classical beams will also drop at  $\tau = 0$ . However, the predicted rate will not drop to zero but rather to half of the background (at large path-length difference). In summary, both the quantum and classical models correctly predict a flat singles rate (intensity at one detector) as a function of  $\tau$ , but the classical coincidence rate dip is limited to 50% visibility, whereas the quantum model correctly predicts 100% visibility.

The origin of the effect in this paper also has conceivable classical explanations. In particular, in a classical model our setup could be considered a first-order interferometer (i.e. replace the polarization HOM with a Mach-Zehnder interferometer) followed by a filter. A first-order interferometer with a finite path-length difference acts as a sinusoidal spectral filter, with a period  $\propto 1/\tau$ . With subsequent filtering by the atomic absorption this possibly could create some effect in the coincidence rate

that varies with  $\tau$ . In fact, our setup is even closer to a standard first-order interferometer than the standard HOM setup discussed above since we only have a detector (albeit a two-photon one) at one of the two output ports of the HOM beamsplitter (which is the polarizer in our case). We test this idea by considering interference between two classical beams with a power spectrum of

$$PS(\sigma, w_o) = \left( \frac{1}{2} \sqrt{2} \sigma \sqrt{\pi} \right)^{-1} e^{\left( -\frac{2(w_o-w)^2}{\sigma^2} \right)}, \quad (2.5)$$

where  $w_o$  is the center frequency and  $\sigma$  is the spectral width. One of these beams has delay of  $\tau$  imposed on it. We calculate the resulting intensity as a function of  $\tau$ , with and without the narrow filter, whose transmission function is given by,

$$filter(w, \delta) = 1 - e^{\left( -\frac{w^2}{\Delta\omega_a^2} \right)}. \quad (2.6)$$

Without the filter, the intensity is,

$$I_{nofilter}(\tau) = \int PS(\sigma, w_o) (1 + \cos((w_o - w)\tau)) dw \quad (2.7)$$

$$= \frac{\left( -\sqrt{\sigma^2 + 2\delta^2} - \cos(w_o\tau) e^{\left( -\frac{1}{8}\sigma^2\tau^2 \right)} \sqrt{\sigma^2 + 2\Delta\omega_a^2} \right.}{\sqrt{\sigma^2 + 2\delta^2}} \left. + \sqrt{2}\delta + \sqrt{2}\cos(w_o\tau)\Delta\omega_a e^{\left( -\frac{1}{4}\frac{\sigma^2\Delta\omega_a^2\tau^2}{\sigma^2+2\Delta\omega_a^2} \right)} \right) \quad (2.8)$$

$$I_{nofilter}(\tau) = \int PS(\sigma, w_o) (1 + \cos((w_o - w)\tau)) dw \quad (2.9)$$

$$= 1 + \cos(w_o\tau) e^{\left( -\frac{1}{8}\sigma^2\tau^2 \right)}. \quad (2.10)$$

With the filter, the intensity is (see Ref. [121] for details),

$$I_{filter}(\tau) = \int PS(\sigma, w_o) \cdot filter(w, \delta) \cdot (1 + \cos((w_o - w)\tau)) dw \quad (2.11)$$

$$= \frac{\left( -\sqrt{\sigma^2 + 2\delta^2} - \cos(w_o\tau) e^{\left( -\frac{1}{8}\sigma^2\tau^2 \right)} \sqrt{\sigma^2 + 2\Delta\omega_a^2} \right.}{\sqrt{\sigma^2 + 2\delta^2}} \left. + \sqrt{2}\delta + \sqrt{2}\cos(w_o\tau)\Delta\omega_a e^{\left( -\frac{1}{4}\frac{\sigma^2\Delta\omega_a^2\tau^2}{\sigma^2+2\Delta\omega_a^2} \right)} \right). \quad (2.12)$$

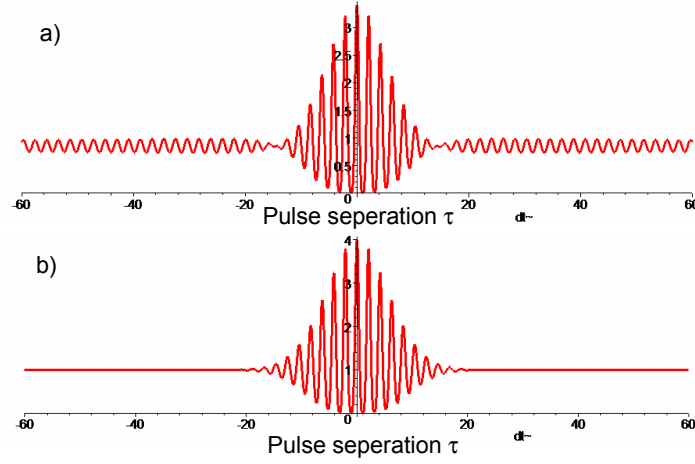


Figure 2.6: Two-photon detection rate before phase-averaging as a function of the path-length difference (or pulse separation)  $\tau$ . Plots were calculated for filter and pulse parameters used in the experiment, a) with filter and b) without filter.

In our setup, we measure the coincidence (two-photon) rate in one output of the HOM. Consequently, we first square the intensity calculated above and then average the phase  $w_o\tau$  to find the coincidence rate  $C$ ,

$$C_{nofilter}(\tau) = 1 + \frac{1}{2}e^{(-\frac{1}{4}\sigma^2\tau^2)}. \quad (2.13)$$

As expected in the classical model, the two-photon coincidence rate only increases by half at  $\tau = 0$ , a 50% visibility. With the filter, the coincidence rate is:

$$C_{filter}(\tau) = \frac{1}{2} \frac{\left( \begin{array}{l} -e^{(-\frac{1}{8}\sigma^2\tau^2)} \sqrt{\sigma^2 + 2\Delta\omega_a^2} \\ + \sqrt{2}\Delta\omega_a e^{(-\frac{1}{4}\frac{\sigma^2\Delta\omega_a^2\tau^2}{\sigma^2+2\Delta\omega_a^2})} \end{array} \right)^2}{\sigma^2 + 2\Delta\omega_a^2} \quad (2.14)$$

$$+ \frac{\left( -\sqrt{\sigma^2 + 2\Delta\omega_a^2} + \sqrt{2}\Delta\omega_a \right)^2}{\sigma^2 + 2\Delta\omega_a^2}. \quad (2.15)$$

The ratio of these two coincidence rates gives the two-photon transmission rate (the

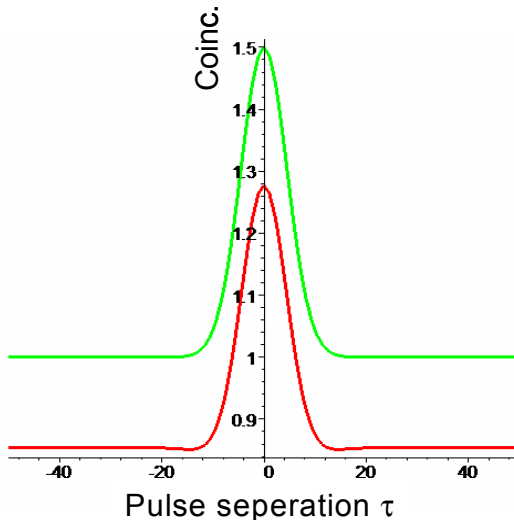


Figure 2.7: Phase averaged coincidence rate for classical interference without the filter (upper curve) and with the filter (lower curve) as a function of the path-length difference  $\tau$ .

quantity we plot in Fig. 2.5),

$$T_{2\text{-photon}}(\tau) = \frac{C_{\text{filter}}(\tau)}{C_{\text{nofilter}}(\tau)}. \quad (2.16)$$

In summary, in the classical model the anomalous transmission effect has much the same shape as in the full quantum model, but has half the size.

We can understand the existence of this classical analog to the photon-exchange effect in another way. It is well known that photon-bunching, which at first look

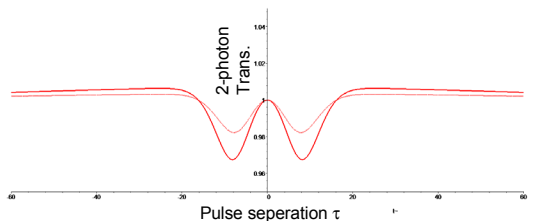


Figure 2.8: Classical (dotted curve) and quantum (solid curve) for the anomalous two-photon transmission rate through the filter. The dip in transmission in the classical case is roughly half the size of the dip in the quantum case.

appears to be a quantum effect, can be described in terms of constructive interference of classical waves [122]. Add to this the fact that the increased two-photon rate in one output port of a HOM interferometer at  $\tau = 0$  can be partly described by photon bunching (see thesis introduction). Since the photon-exchange effect is a generalized HOM effect, we should expect an analogous phenomena in classical interference.

## 2.6 Conclusion

There are inter-photon time delays for which the photon pair tends to be “transmitted” (reflected from the interference filter) with lower probability than at very long time delays where each photon is “transmitted” independently. This suppression is due to photon-exchange terms in the photon wavefunction theory. These exchange terms become important when the photons are initially distinguishable but are, at least partially, indistinguishable to a medium. In our case, this happens for intermediate delays when the photons are separated by more than their coherence length but by less than the coherence length of the absorber. The anomalous absorption occurs regardless of the method by which the two-photon state is created; in that sense it is a property of the state and not of the method of state preparation. Viewed another way, the experiment can be understood in terms of the generalization of Hong-Ou-Mandel interference to nondegenerate photons (see Fig. 2.9).

We have observed that photon-exchange effects can give rise to nonlinear effects on the pair-transmission probability, in analogy with Franson’s proposal [95][96][97][98][99]. Like-polarized photon pairs with a variable interparticle delay have been shown to exhibit suppressed two-photon transmission through an absorbing medium – an experimental signature of enhancement of photon pair absorption. This suppression occurs for delays that are longer than the photons’ coherence times but shorter than the coherence time of the absorber. The nonlinearity created between photons in a HOM interferometer (see thesis introduction) only occurs in the post-selected cases

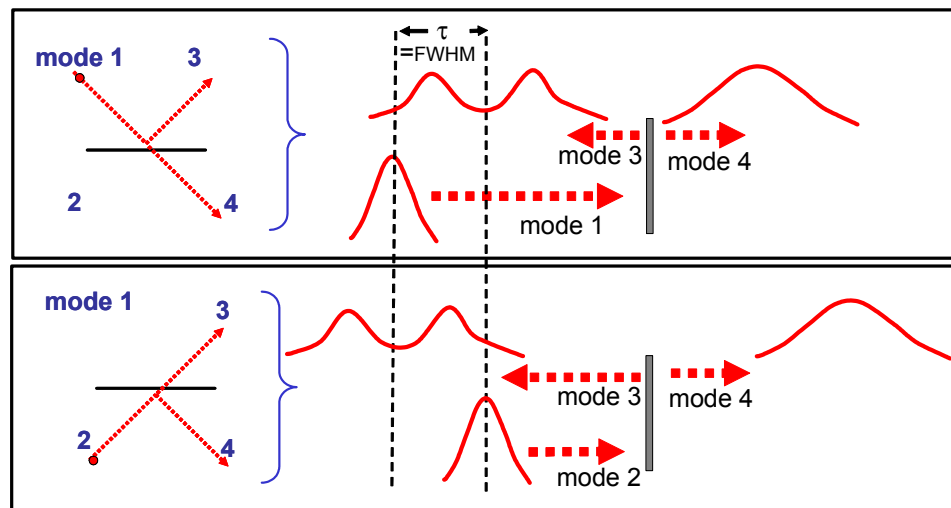


Figure 2.9: The photon-exchange effect explained in terms of Hong-Ou-Mandel (HOM) interference. Reflection from the interference filter in the experiment takes the place of the 50/50 beamsplitter in HOM interference. If the delay  $\tau$  is approximately equal to the photon wavepacket width  $\Delta t$ , there are four temporal modes corresponding to the four modes in the HOM interference. For larger delays, modes 3 and 4 do not overlap temporally. For shorter delays, modes 1 and 2 are not orthogonal.

in which one photon is detected at each output port of the beamsplitter. Much like this, the anomalous absorption only occurs in the post-selected events in which two photons are detected. In particular, the singles rates remain unchanged for all  $\tau$ . In a recent short note [123], Franson admitted that there was error in his original calculations. In particular, this implied that his Kerr-like effect was impossible without some loss. Specifically, like in the photon-exchange effect we describe, Franson's nonlinear phase-shift only occurs in the photon pairs that are detected, while there are necessarily other photons that are scattered and absorbed by the atomic gas. Franson concludes that this necessary absorption makes this type of scheme unsuitable for quantum computation. However, this might not be the case; A LOQC C-NOT gate has been observed under the same conditions (i.e. the nonlinear transformation only occurs in post-selected photon pairs [124]). Moreover, the photon-exchange effect can be described as generalized HOM effect in which there are many input and output modes (e.g. frequency components). Consequently, this type of post-selected nonlinearity might be suitable for processing systems with higher-dimensional entanglement [125][126].

# Chapter 3

## Hardy's Paradox

### 3.1 Introduction

Since its conception, quantum mechanics has confounded scientists who seek to add a clear and consistent interpretation to its entirely successful mathematical framework. Among others, Einstein and Bohr famously argued at length, volleying gedanken experiments back and forth. Some of the more paradoxical of these gedanken experiments later transformed into the pedagogical examples now used to teach quantum mechanics, such as Young's double-slit experiment. One particular gedanken experiment, the EPR paradox, was used by John Bell in the 1960s to prove that quantum mechanics cannot be interpreted in the same intuitive manner that classical mechanics is [27]. More specifically, he showed that if quantum mechanics is correct, it is not a local realistic theory [127]. It took until the 1970s for this controversial result to be adapted for experimentation [128], resulting in a statistical inequality, which was tested with entangled pairs of photons in the singlet state [42]. In the 1980s, entangled three-particle states called GHZ (Greenberger, Horne, Zeilinger) states were used to show a direct, as opposed to statistical, contradiction between local realism and quantum mechanics [26][14]. At roughly the same time, consideration of the unique properties of quantum mechanics, such as entangle-



ment, led to the development of quantum computation and cryptography and other information-related schemes [5][4]. Finally, in 1992, Lucien Hardy created a scheme, now known as Hardy’s Paradox, in which just two particles (and only two observers) can be used to prove Bell’s theorem without using inequalities [129][28]. Later, by changing from particles in different paths to polarized particles, Hardy generalized his scheme and maximized the number of events in which the contradiction occurs [130]. Since Hardy’s original two papers there have been a few implementations of this generalized scheme [131][132][133]. However, the paradoxical nature of Hardy’s original scheme remains the simpler concept to understand. It clearly shows that our classical reasoning is wrong in many ways. To this point, David Mermin poetically stated that Hardy’s Paradox “stands in its pristine simplicity as one of the strangest and most beautiful gems yet to be found in the extraordinary soil of quantum mechanics” [134][135]. This section describes an experiment that implements Hardy’s Paradox in its original simple form. Moreover, this implementation is the first application of a single-photon-level absorptive switch that was developed in the Steinberg lab prior to this thesis. While performing this experiment another group posted similar results on the xxx.lanl.org archive, which have now been published in Ref. [136]. In the place of our switch, they follow an early proposal by Hardy [137] and use the nonlinearity introduced by the Hong-Ou-Mandel effect [46], as described in the thesis introduction.

## **3.2 Interaction-free measurements.**

The setup for Hardy’s Paradox consists of two interaction-free measurements (IFM). These are a type of measurement proposed by Elitzur and Vaidman in 1991 [138] in which the presence of an object can be discerned without it being disturbed. The simplest example of this type of measurement is a Mach-Zehnder interferometer (shown in Fig. 3.1) that is aligned so that all the photons entering the interferometer

leave through the “bright” output port C and none through D, the “dark” output port. Both the first and last beamsplitters (BS1 and BS2) reflect and transmit with 50% probability. Consider what would happen if an object, say a highly light-sensitive bomb, were positioned in the right-hand path (I). This bomb will absorb any photon that hits it and subsequently explode. If we let only one photon enter the interferometer, it has a 50% chance of taking the left path (O), thus avoiding triggering the bomb. It then has a 50% chance of exiting through port C, which conveys no information about the presence of the bomb, and a 50% chance of exiting through the previously dark port, D. The latter then indicates the presence of the bomb without the photon triggering the bomb. It is in this sense that the measurement is free of interaction (a closer examination of what is meant by “interaction-free” is given in [139]). In total, one has a 25% chance of success (detecting the bomb without detonating it), a 25% chance of receiving no knowledge, and a 50% chance of failure (an explosion).

The above analysis naturally leads one to the question of whether the photon was in the bomb arm in trials which were successful. If not, how could the bomb’s presence influence which output port the photon took? If so, why did the bomb not explode, and how did the photon get to the detector? As the next step towards Hardy’s Paradox, consider what would happen if the bomb were in an equal superposition of two position states, one in arm I and one outside of the interferometer? In this scenario, if a photon is detected in the dark port D, the state of the bomb collapses to a state in the arm I. On the other hand, if the photon exits from the bright port C, no collapse occurs. Evidently, IFMs also function properly when the object is inherently quantum mechanical.

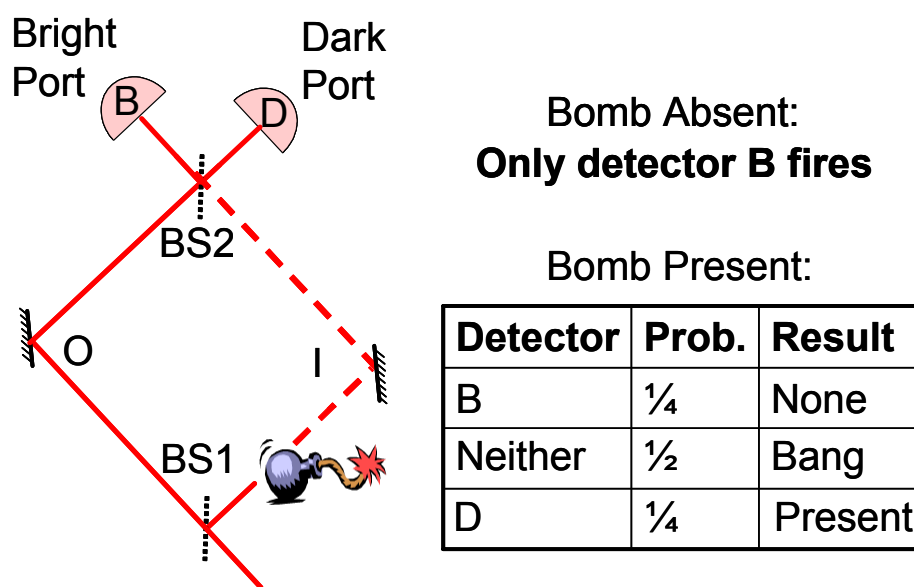


Figure 3.1: An interaction-free measurement (IFM). A single particle enters an empty interferometer aligned so all the particles leave through the bright port (B). A click at the dark port detector (D) signals the presence of an object in one of the arms.

### 3.3 Intuitive version of Hardy's Paradox

Hardy's Paradox takes this scenario one step further. Hardy envisioned using an electron as the test particle in the IFM and a positron as the bomb (see Figure 2). The positron is placed into a superposition by a 50-50 beamsplitter. The extra step is that after the electron and positron would have annihilated each other in region  $W$ , the positron superposition is recombined at another 50-50 beamsplitter. We now have a symmetric experiment where both the electron and positron enter Mach-Zehnder interferometers (distinguished by the subscripts  $-$  and  $+$  respectively). Each interferometer performs an IFM on the particle in the other interferometer. If one detects the electron at  $D_-$  this implies that the positron was in  $I_+$  and conversely, a positron at  $D_+$  implies the electron was in  $I_-$ . Surprisingly, it is possible for both the electron and positron to exit their respective  $D$  port in the same trial, implying they were each in their respective  $I$  arm. Consequently, this appears to answer the question in the last section of which arm the electron was in before a successful measurement: The arm with the object in it. But then how did the electron and positron avoid annihilation and arrive at our detectors? These two contradictory conclusions arise from the false classical idea that a particle has determinate properties, such as position, before observation. This is the essence of Hardy's Paradox.

### 3.4 Quantum mechanical model

A quantum mechanical model of the Hardy's Paradox is quite simple:

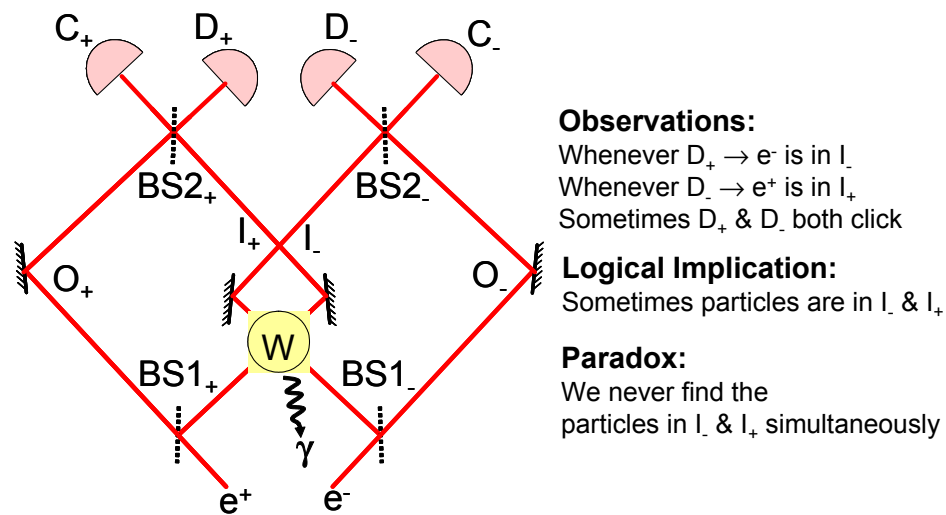


Figure 3.2: The scheme for Hardy's Paradox. There are two IFMs. One uses an electron and the other a positron. Each IFM is positioned to detect the presence of the particle in the inner (I) arm of the other IFM. Simultaneous dark port (D) detections at both IFMs indicate that the electron and positron were at the overlap region  $W$  simultaneously. Here they should have annihilated, but clearly they did not since they arrived at the D detectors.

$$|In_-\rangle |In_+\rangle \quad (3.1)$$

$$\xrightarrow{\text{BS1}} \frac{1}{2} (|O_-\rangle + i |I_-\rangle) (|O_+\rangle + i |I_+\rangle) \quad (3.2)$$

$$= \frac{1}{2} (|O_-\rangle |O_+\rangle + i |O_-\rangle |I_+\rangle + i |I_-\rangle |O_+\rangle - |I_-\rangle |I_+\rangle)$$

$$\xrightarrow{\text{Annihil.}} \frac{1}{2} (|O_-\rangle |O_+\rangle + i |O_-\rangle |I_+\rangle + i |I_-\rangle |O_+\rangle) - |\gamma\rangle \quad (3.3)$$

$$\xrightarrow{\text{BS2}} -\frac{3}{4} |B_-\rangle |B_+\rangle + \frac{i}{4} |D_-\rangle |B_+\rangle + \frac{i}{4} |B_-\rangle |D_+\rangle - \frac{1}{4} |D_-\rangle |D_+\rangle - \frac{1}{2} |\gamma\rangle \quad (3.4)$$

The particles annihilate each other 1/4 of the time (produces light  $\gamma$ ), and either or both bright port detectors click 11/16 of the time, leaving the last 1/16 for the paradoxical coincident dark port detector clicks. This last probability can also be derived by considering the two IFMs: The electron is in the inner arm 1/2 the time, and when it is in the inner arm, the positron IFM has a 1/4 chance of detecting it (meaning the positron exits through its dark port). After this detection, the electron's superposition has collapsed and the interference is destroyed, so it leaves through its dark port 1/2 of the time. Thus, the probability for both particles to leave their dark ports is  $1/2 \times 1/4 \times 1/2 = 1/16$ . The goal of Hardy's generalization was to maximize this probability by altering the state after the annihilation [130]. This required switching the calculation to the polarization basis (as opposed to interferometer arms), which allows for more general transformations than with 50-50 beamsplitters. However, in this generalization the simple intuitive interpretation in terms of IFMs was lost. Moreover, the experiments that implemented this generalized version ([131][132][133]) created Hardy's generalized post-annihilation state directly, eliminating half of the original experiment.

It is important to note that the post-annihilation state is non-maximally entangled, in contrast with the Bell-states, for which the correlations between the particles are perfect. This demonstrated that, surprisingly, maximal entanglement is not necessary to demonstrate Bell's theorem.

### 3.5 Rigorous version of Hardy's Paradox

The paradox can be analyzed more rigorously in terms of logical statements that summarize the essential elements from the discussion above. All symbols in the following refer to the Hardy's Paradox figure. Each of the logical statements can be verified by making the appropriate measurements on the two interferometers. In particular, the final beamsplitters BS2+ and BS2- can be removed so that the two detectors now directly measure which path the particle is in. With one beamsplitter in and the other removed we can verify that both IFMs appear to work correctly. This gives two logical statements: A click at D+ indicates e- is in I- with certainty ( $D+ \rightarrow I-$ ), and by symmetry,  $D- \rightarrow I+$ . Assume that the measurements are independent (this could also be directly instituted by making the measurements in non-overlapping space-time cones). With both beamsplitters in place, simultaneous clicks at D+ and D- logically implies I- and I+, the electrons were in the inner arms simultaneously and, as pointed out above, they should have annihilated. More rigorously, in those instances where we measure the paradoxical D+D- clicks, we could just as well have measured with the beamsplitters removed. In this case, we would expect to find particles in the I+ and I- arms simultaneously. The paradox is that, if we do measure with both beamsplitters removed then we never do find them in the I+ and I- arms simultaneously. This is a direct logical contradiction, in contrast with the statistical contradiction in the more familiar Bell's inequality.

Of course, in a real experiment there are imperfections in the apparatus, which necessitates moving to statistical inequality. To start with, we expect to find, on average, as many pairs of particles in the I arms as we register coincident D detector clicks or,

$$R(D_+, D_-) \leq R(I_+, I_-). \quad (3.5)$$

If there is imperfect interference in the IFMs then clicks at D will sometimes indicate the corresponding target particle is in arm O, as opposed to I. Imperfect annihilation

of the particles introduces a background of pairs of particles in the inner arms but this is omitted since it makes the right-hand side smaller and thus easier to violate. Consequently, we should add two additional terms to the right-hand side of the inequality,  $R(D_+, O_-)$  and  $R(O_+, D_-)$ . For a real apparatus we now have the statistical inequality,

$$R(D_+, D_-) \leq R(I_+, I_-) + R(D_+, O_-) + R(O_+, D_-). \quad (3.6)$$

This inequality, which arises naturally from considerations of the apparatus, is the famous Clauser-Horne inequality [140]. It differs in appearance from the standard form only because it pertains to a different set of measurements (typically, the measurements are polarization projections at a set of angles – see Ref. [141] for example). Just as with the Clauser-Horne inequality, a violation of Eq. 3.6 signifies that no local hidden variable theory can properly predict the outcome of this set of measurements.

This points to the crux of the paradox: We attempt to assign a location, which is an example of a simple hidden variable, to each of the particles. We do this despite the fact that it is in principle impossible to directly measure the location in those cases where we register the paradoxical result,  $D_+ D_-$ . They are complementary measurements. If we tried to directly measure the location, we would collapse the particle to one of the arms and this would destroy the IFM interference and, in turn, negate the logical implications of the paradox. Instead, we attempt to make counterfactual statements about the location of the particles (in  $I_+$  and  $I_-$ ) based on the logical implications of IFM results. The IFMs appear to be perfectly good measurements of the presence (or location) of either classical or quantum objects, yet they are indirect measurements. This paradox is a striking example of the perils of indirect measurement in quantum mechanics.



### 3.6 The switch

Lucien Hardy’s original scheme involved electrons and positrons, which are prohibitively difficult particles to make interferometers with. Photons would be the ideal particles to use but they do not annihilate each other. In fact they interact very weakly, even in nonlinear crystals. However, in work done before this thesis, Kevin Resch, myself and Aephraim Steinberg developed an absorptive switch for single-photon-level fields that makes this experiment possible [19]. The switch ideally allows a photon pair passing through it to up-convert or “annihilate” with a probability of close to 100%. It consists of a nonlinear  $\chi^2$  crystal with a strong pump at 405 nm passing through it. As we saw in the thesis introduction, through SPDC this creates photon pairs in two output modes. The two input modes to the switch, 1 and 2, pass through the crystal and overlap with the output modes, 3 and 4. If these input modes occasionally contain photons then there are two Feynman paths that can lead to a pair of photons, one in each output mode. Each of the input modes can contribute a photon to make up the pair, or the pair can come from downconversion. If the two Feynman paths share a phase relationship and are indistinguishable they will exhibit interference [142]. In particular, if they are out of phase they will destructively interfere, resulting in an absence of photon pairs in the output modes. In this situation, any paired photons in the input modes are selectively upconverted in the pump beam while the unpaired photons pass through unimpeded. In this way, one photon has the ability to control the transmission of the other, as in a switch. From another viewpoint, should two photons arrive at the crystal at the same time then they “annihilate,” just as we desire for Hardy’s Paradox.

Experiments have shown that the phase information of the pump is not lost during downconversion [143][144], and that, even though the phase of one of the downconverted beams is random, the phase of both together must sum to the phase of the pump. We use downconversion as the source of photon pairs for the two IFMs

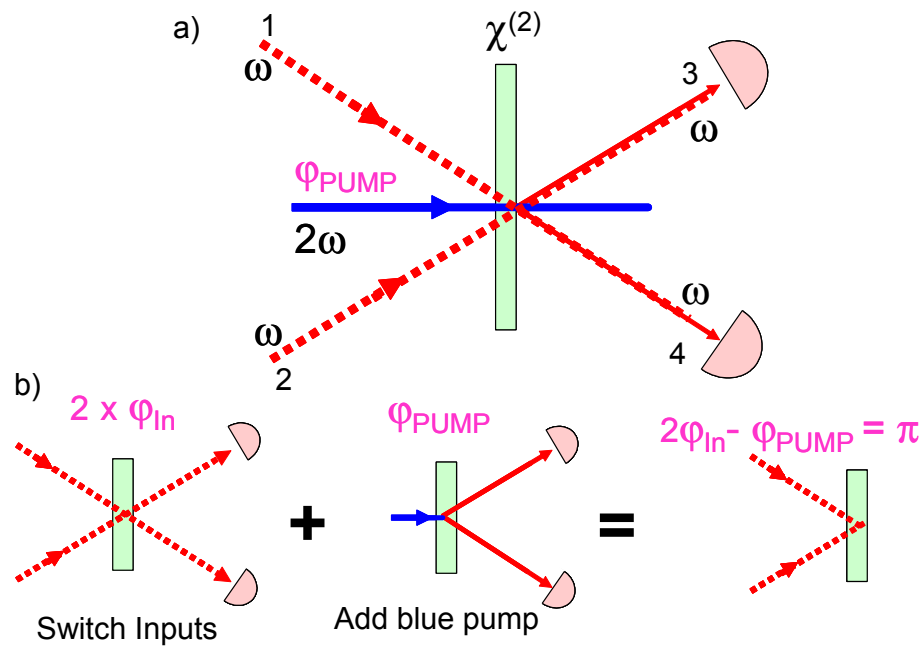


Figure 3.3: a) The setup for the switch effect. Two input beams enter the back of  $\chi^{(2)}$  nonlinear crystal in modes 1 and 2. A pump beam passes through the crystal so that downconversion pairs of photons are produced in the output modes 3 and 4. b) Two Feynman paths lead to a photon pair in the output modes. When these are adjusted for destructive interference then no photon pairs emerge in the output modes. However, unpaired photons pass through the crystal unimpeded.

in Hardy's Paradox. Consequently, if we use the same pump for the production of these pairs as we do for the pumping the crystal in the switch, then the two Feynman paths will be related by the phase of the pump beam. Moreover, if instead of using a different crystal for the switch and photon pair source we retro-reflect the pump through the same crystal, then the two paths will be indistinguishable; the photons will have identical spectral, temporal, and spatial modes.

In fact, since the switch relies on phase, one of its limitations is that single-photon states cannot be switched. Instead, to be rigorous, it is the amplitude to have a single photon that is switched. Furthermore, the actual size and phase of the amplitude of the photon pair term must be known in advance to balance the two Feynman paths. While these features might limit the usefulness of this effect for a controlled-not gate for a quantum computer, conditional-phase operations have also been demonstrated [20]. Fortunately, the switch is well suited to Hardy's Paradox as we know the photon pairs encounter each other at constant location (the crossing of the I arms), and at a constant rate, since the reflectivities of the first beamsplitters are also constant. This means that the photon pair amplitude and phase should ideally remain constant throughout the experiment, as is necessary for the switch.

### 3.7 Experimental implementation

The experimental setup is shown in Fig. 3.4. A diode laser produces a 30 mW 405 nm beam, which is filtered by blue glass and sent through a dichroic mirror. This beam produces 810 nm photon pairs in a 4 mm long BBO (Beta Barium Borate) crystal through the process of Type II spontaneous parametric downconversion. These pairs, consisting of a collinear horizontal (E) and vertical (P) photons, take the place of the electron and positron from the introduction. The pump passes through a dichroic mirror at 45 degrees, to later be retro-reflected back along the same path it travelled before. The photon pair reflects off the same dichroic and then passes through another

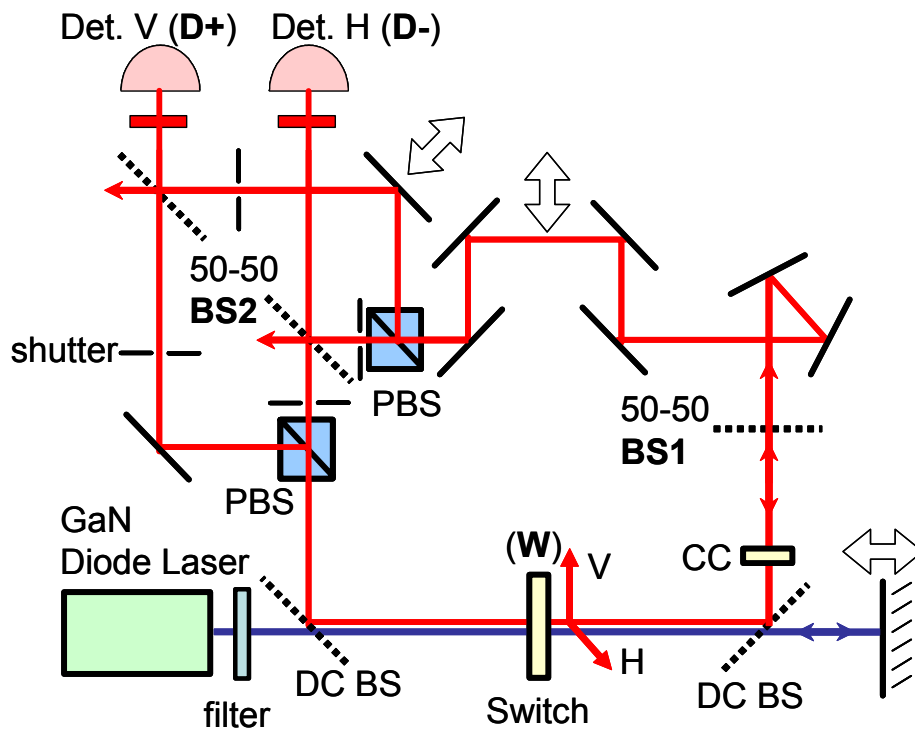


Figure 3.4: The experimental setup for Hardy's Paradox. See text for details. Beam-splitter (BS), dichroic beamsplitter (DC BS), polarizing beamsplitter (PBS), compensating crystal (CC).

dichroic, to remove any residual pump light, followed by a 2mm-thick BBO crystal, to compensate for the birefringent temporal delay in the first crystal. The photon pair then meets a 50-50 beamsplitter, which acts as the first beamsplitter in both the electron and positron interferometers. At this point, each photon can either be retroreflected and enter the Inner (I) arm or be transmitted and enter the Outer (O) arm.

Reflected photons pass back through the downconversion crystal along with the retroreflected 405 mW pump. The amplitude for the retro-reflected pump to create a pair of photons in the crystal is set to interfere destructively with amplitude for a photon pair in the reflected 810 nm beams. Thus, if both the E and P photons enter

their Inner arms they are removed from the beams, whereas if only a single photon enters, it passes through the crystal unimpeded. This is the absorptive switching that takes the place of electron-positron annihilation.

Transmitted photons enter the Outer arms. They have their spatial orientation flipped by two mirrors and are then reflected from a translatable mirror arrangement that can vary the path length. Both the Inner and Outer paths encounter polarizing beamsplitters (PBS) so that the E and P photons are split into their own spatially separate interferometers. The P interferometer contains an additional translation stage so that both interferometers can be adjusted to have the same path-length difference. The Inner and Outer paths of the two interferometers are recombined at two non-polarizing beamsplitters (NPBS), taking the place of the final Mach-Zehnder beamsplitters for the electron and positron. Behind each of the NPBSs is a single-photon detector (a module from Perkin-Elmer that detects using a Silicon Avalanche Photodiode operated above the breakdown voltage).

A temperature-stabilized 780 nm laser diode is used as a phase reference for the adjustment and stabilization of both Mach-Zehnders and the absorptive switch (which is based on an interference effect and hence needs to be stabilized). The stabilization apparatus is shown in Fig. 3.5. A computer adjusts the phase by tilting 0.1mm thick glass slides in all three interferometers, each approximately at Brewster's angle. It first searches for the point of maximum destructive interference and then uses the 780 nm laser to stabilize the interferometers there. At this phase, the two single-photon detectors are at the dark ports of the two Mach-Zehnders and the switch ideally removes any photon pairs passing through it.

To test Hardy's Paradox we need to make two complementary measurements for each interferometer; the IFM, with the beamsplitter in place, and the photon location, with the beamsplitter removed. Removing the beamsplitter while still maintaining interferometric alignment is problematic. Instead, we shutter the opposite arm from the one we are interested in. For example, if we want to measure the photon rate in

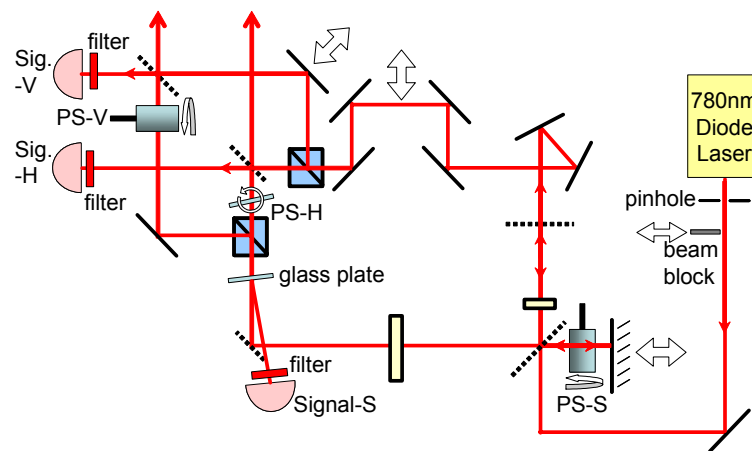


Figure 3.5: All three interferometers in the experiment are stabilized using one 780 nm laser as a reference. The laser first enters through the back of the dichroic beamsplitter in the switch interferometer. Following the path of the downconverted photons, the reference laser then enters the Mach-Zehnder interferometers. Glass slides 0.1 mm thick labelled PS-S, PS-V, and PS-H are placed in the switch interferometer, and the vertical, and horizontal Mach Zehnder interferometers, respectively. The reference signals are measured by detectors Sig.-S, Sig.-V, and Sig.-H, respectively. The signals are processed in computer, which then rotates the glass slides in response to stabilize the phase. A stability of  $\lambda/20$  over one hour was achieved.

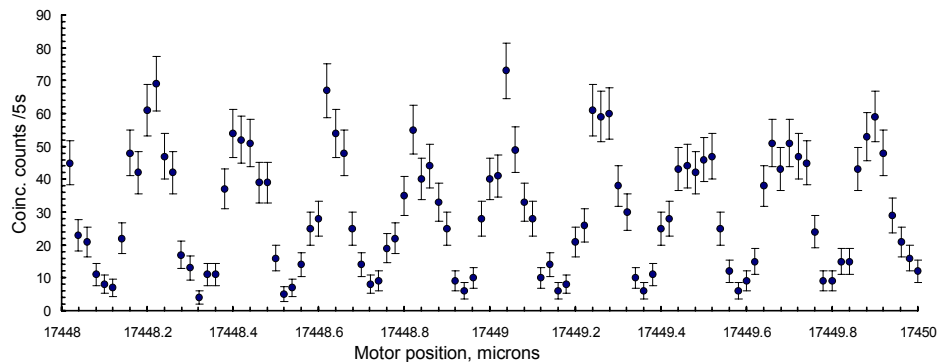


Figure 3.6: The switch. As we change the pump phase (through the retro-reflecting mirror position) the number of coincidences oscillates with a visibility of 82%. These rates are measured with the outer arms blocked and are thus a measure of the number of photon pairs in the inner arms. The switching effect occurs at the minimum of the curve.

arm I then we block arm O. All the photons at the detector will be from arm I and rate will be known up to the transmission of the beamsplitter. To test the functioning of the switch, we measure the rate of remaining photon pairs in the inner arms (by shuttering the outer arms).

### 3.8 Results

There are a number of technical factors that prevent the experiment from performing ideally. Using the same crystal as a source and a switch for the photon pairs eliminates most distinguishing information between the two Feynman paths that destructively interfere to give the switch effect. However, the crystal axis (at some angle to the pump beam) breaks the reflection symmetry (which we rely on since the modes are retro-reflected back to crystal). This causes the photons to have some spectral shift across their transverse momentum profile, and upon reflection this shift changes in sign. Consequently, only those transverse momentum components near zero will interfere effectively. We limit ourselves to these components with 1mm irises in front

of the detectors. Nonetheless, as Fig. 3.6 shows, some distinguishability remains, causing the switch to remove 82%, rather than all, of the photon pairs on which it acts. Stabilizing all three interferometers, we measure a rate of  $24.9 \pm 1.12/5s$  coincidence counts arising from the inner arms (when adjusted for the beamsplitter transmissions). We can also test the efficiency of the two IFMs. For the E Mach-Zehnder, a click at the dark port detector indicates the P photon is in the Inner path  $95 \pm 3\%$  of the time, whereas for the P Mach-Zehnder a similar measure gives  $94 \pm 4\%$ , which is roughly consistent with the visibilities of the two interferometers shown in Fig 3.7. With all three interferometers aligned we observe  $19.1 \pm 0.5$  coincidence counts every 5 s between the E and P dark ports. This is the paradoxical result from Hardy's Paradox in which both IFMs simultaneously succeed. However, although all three interferometers have fairly high visibilities, they are *not* large enough to violate the inequality in Eq. 3.6. To violate this inequality the switch visibility ( $V_S$ ) has to be sufficiently large to satisfy the following relation:

$$V_S \geq \frac{11}{8} - \frac{V_{MZ}^2}{2}, \quad (3.7)$$

which assumes the two Mach-Zehnder visibilities ( $V_{MZ}$ ) are equal. Even if the Mach-Zehnders were perfect ( $V_{MZ} = 1$ ), the switch would still need to remove greater than 87.5% of the photon pairs to be able to violate Eq. 3.6 and thus prove the impossibility of a local realistic description of the world. Nonetheless, the essential elements of the paradox are present. Later in the thesis we directly test the locations of the photons using weak measurement.



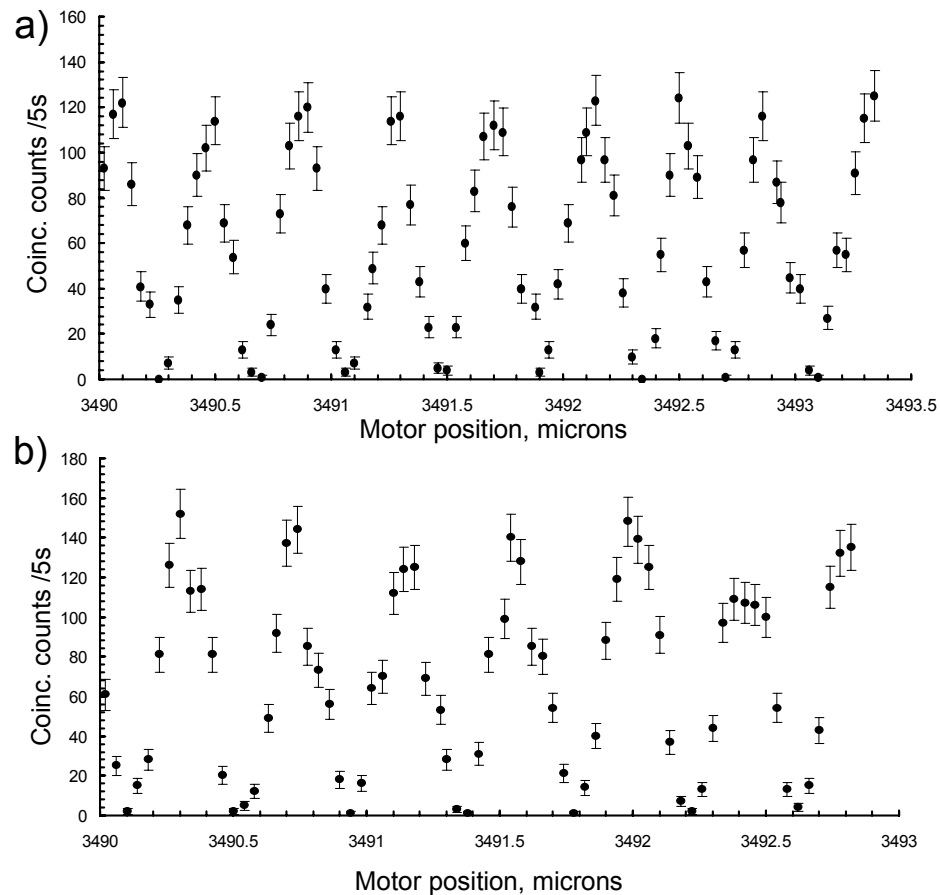


Figure 3.7: The interference fringes for the a) E (vertical) and b) P (horizontal) Mach-Zehnder interferometers. These plots show the coincidence rate at the two detectors in the experimental setup figure, measured while the pump retroreflection is blocked. The motor in the outer arms (while they overlap) is translated. The E coincidence curve is measured while the outer arm of the P interferometer is blocked and vice versa. To align the IFMs, each interferometer is stabilized at the point of maximum destructive interference.

## Chapter 4

# Momentum transfer and Which-Way Measurement

### 4.1 Introduction

An interference pattern forms when it is impossible to tell through which of two slits a particle travelled to a screen. Performing a which-way measurement (WWM) to determine which path the particle took destroys this pattern. Thus, there is a trade-off between wave- and particle-like behaviour, which Bohr named complementarity. However, the exact mechanism that enforces this trade-off has been controversial. According to Bohr [145], this destruction can be ascribed to an “uncontrollable change in the [transverse] momentum,” enforced by the Heisenberg (position-momentum) Uncertainty Principle,  $\Delta p \Delta x \geq \hbar/2$ . A good example of this is Feynman’s Microscope [146]. In this gedanken experiment, there is an electron double-slit interferometer. A WWM is performed by illuminating the slits with light and distinguishing which slit the electron travelled through by collecting the scattered photons with a microscope. During the scattering, the photons impart a momentum to the electrons of roughly  $\hbar/\lambda$ . The resolution of the microscope is limited to the wavelength of the light  $\lambda$ , which then must be smaller than the slit separation  $s$ . Consequently, the

scattering process transfers a random momentum  $q \sim \hbar/s$  to the particle, which is enough to shift the pattern by one fringe spacing, thus washing out the interference pattern. This was accepted for 60 years until, in 1991, Scully, Englert and Walther invented a WWM scheme that appeared to have no physical mechanism for transferring momentum [30]. They claimed, “it is simply the information... and not the uncontrollable alterations of the spatial wave function” that destroy the interference. They concluded, “the principle of complementarity is much deeper than the uncertainty relation.”

Scully *et al.* considered a double-slit interferometer that worked with atoms rather than the photons used in Young’s original apparatus [147]. Their WWM scheme consisted of an optical cavity placed behind each of the slits, in which the excited atoms would deposit a photon as they passed through. If a photon is found in one of the two cavities, it is known which slit the corresponding atom travelled through. Since the cavities were much wider than the widths of the slits, the fields inside them varied slowly over each of the slits, suggesting that there is no mechanism for them to transfer significant momentum. Scully *et al.*’s more detailed calculations of the interaction of the atom with the cavity seemed to confirm this.

Although the above conclusions fit well with the emerging field of quantum information, not everyone was convinced. After all, the WWM still differentiates between two positions a distance  $s$  apart, so why does Heisenberg uncertainty not apply in this case? Furthermore, in the far-field the distribution of particles on a screen is proportional to the probability distribution of their momenta. Thus, we can identify the interference pattern observed without a WWM as the initial momentum distribution. Likewise, with the WWM one observes the final momentum probability distribution,  $P(p_f)$ . The actual shape of the initial and final distributions are different – one has fringes and the other does not, implying that there must have been momentum transfer [148] to change the initial distribution to the final distribution. However, if one attempted to quantify this difference by doing a detailed calculation of the variance of

the two distributions for the Scully *et al.* WWM, one would be surprised to find they were equal. In fact, all the moments of the initial and final momentum distributions were equal, which was not the case for the Feynman Microscope. Clearly, the WWM of Scully *et al.* is different from the WWMs that have been considered in the past. Nonetheless, in response to Scully *et al.*'s claims, Storey *et al.* [31] proved a general theorem showing, they claimed, that there is always momentum transfer of order  $\hbar/s$ . They also disagreed with Scully *et al.*'s detailed interaction calculations.

That both sides in the debate ([30][31][149][148]) had valid claims was first pointed out in Ref. [150]. The disagreement came from the fact that the two groups were using different concepts of momentum transfer. These concepts agreed only for *classical* transfers (i.e. random kicks) such as those induced by WWMs of the type discussed by Bohr [145][151] and also Feynman [146]. The WWM of Scully *et al.* is not of this type. In particular, Scully *et al.* showed that a *single-slit wavefunction* would be unchanged by their WWM. This is what makes their WWM unique and what defines Scully *et al.*'s concept of momentum transfer (or lack there of it). The theorem of Storey *et al.*, on the other hand, meant that if the initial state were a *momentum eigenstate* then after any WWM the final momentum distribution would have a width of at least  $\hbar/s$  [152]. (We define the “distribution width of at least  $\chi$ ” to mean that it is nonzero somewhere outside the interval  $[-\chi, \chi]$ .) It is unsatisfying that these physical predictions require experiments (with a single-slit wavefunction and momentum eigenstate, respectively) that are incompatible with each other and with the *double-slit* experiment that they are supposed to illuminate. In contrast, the weak measurement technique that we will now outline allows us to directly measure the momentum transfer while carrying out the original double-slit experiment.

## 4.2 The Idea

Before we describe that technique, we begin with a straightforward procedure to measure the momentum transfer. We could tag each momentum component in the initial double-slit wavefunction and then, after the WWM, observe at which final momenta they are detected. The difference between the two is the momentum transfer,  $q = p_f - p_i$ . As a tag, we would want to use a parameter of the particle uninvolved in the interference effect. In this experiment, we tag particles with an initial momentum  $p_i$  by inducing a vertical displacement  $D$  in them. After the WWM, we would select the subset of particles with final momentum  $p_f$  and count the number of these that are tagged. In this subset,  $(\#tagged)/(total\#) = P(p_i|p_f)$ , the probability that a particle began with  $p_i$  given that we subsequently found it had momentum  $p_f$ . If we repeat this for every combination of  $p_i$  and  $p_f$ , then we can find the unconditional probability distribution  $P(p_i \& p_f) = P(p_i|p_f) \cdot P(p_f)$ . And from this, one can find the probability for a momentum transfer  $q$  regardless of the initial or final momentum of the particle,

$$P(q) \equiv \sum_{p_i} P(p_i \& p_i + q). \quad (4.1)$$

Alternately, we could repeat this procedure with the substitution that we measure the *average displacement* of the whole subset,  $d = [(\#tagged) \cdot D + (\#untagged) \cdot 0]/(total\#)$ . It follows that the conditional probability  $P(p_i|p_f) = d/D$ . In this experiment, we use this alternate procedure because it does not require us to know whether or not a particular detected particle was tagged and began with  $p_i$ . This turns out to be important. If we can determine a particular particle's initial momentum with certainty, then it must have begun in the corresponding eigenstate, a plane-wave wavefunction, as opposed to beginning in the double-slit wavefunction we wanted. In effect, our tagging procedure is a “strong” measurement of initial momentum, and the ensuing collapse into a plane-wave disturbs the very process we wish to investigate, the WWM. Naturally, one would desire a way of arbitrarily reducing this disturbance.

The solution is to similarly reduce our ability to discriminate whether or not a particular particle had momentum  $p_i$ . We do this by making the induced displacement  $D$  small compared to the vertical width of the particle's wavefunction. The trade-off in reducing the disturbance is that we are now unable to count the number of tagged particles, even in principle. Fortunately, we can still determine the average displacement of the subset to arbitrary accuracy simply by running more trials. In short, this solution amounts to reducing the relative strength of the interaction (and hence, disturbance) that generates the measurement. Consequently, a classical physicist would still interpret the result  $d/D$  as giving  $P(p_i|p_f)$ .

This is a weak measurement (see the thesis introduction) and  $d/D$  is the average result, the weak value (in the limit of  $D \rightarrow 0$ ). A slightly more general formula for the weak value than the one given in the introduction to the thesis is:

$${}_{\phi}\langle X_w \rangle_{\psi} = \text{Re} \frac{\langle \phi | \hat{U} \hat{X} | \psi \rangle}{\langle \phi | \hat{U} | \psi \rangle}. \quad (4.2)$$

The parameters  $|\psi\rangle$ ,  $|\phi\rangle$ , and  $\hat{X}$  are respectively the initial state of the system, the post-selected state, and the weakly measured observable. In the above procedure, these are the double-slit wavefunction, the final momentum state  $|p_f\rangle$ , and the initial momentum  $|p_i\rangle$   $\langle p_i|$ . Evolution after the weak measurement is given by  $\hat{U}$ , typically unitary, but in our case, an operation describing the measurement of the particle by the WWM device [32][153].

Remember that the results of weak measurements can have strange values that would be impossible for standard measurements. To interpret these strange results in the context of the measurement procedure described above let us first recount what the weak value is equal to when there is no post-selection. In this case, we set  $|\phi\rangle$  equal to  $|\psi\rangle$ , and Eq. 4.2 just equals the expectation value of  $\hat{X}$ . If  $\hat{X}$  is a projector  $\hat{\pi}$ , this expectation value, and hence the weak value, is generally interpreted as a probability,  $P(\pi)$ . However, in the case of post-selection on state  $|\phi\rangle$ , the weak value may lie outside the eigenvalues of  $\hat{X}$  [29]. In particular, if we weakly measure a projector,

the weak value can lie outside the range  $[0,1]$ , impossible for a standard probability (see the thesis introduction for a longer discussion of non-standard probabilities). To make this distinction, the weak value of a projector is called a *weak-valued* probability (WVP). The fact that the WVP can be negative enables it to describe states and processes which *require* a quantum description, similar to other quasi-probabilities such as the Wigner function. Reference [32] applied WVPs to momentum transfer in WWMs. The basic idea is to find the conditional WVP  $P_{\text{wv}}(p_i|p_f)$ , or the WVP for initial momentum to be  $p_i$  given that the final momentum is found to be  $p_f$ . Our weak value,  $d/D$ , is exactly  $P_{\text{wv}}(p_i|p_f)$ , although a classical physicist would label it as  $P(p_i|p_f)$ . Therefore we manipulate this result according to Eq. 4.1 just as a classical physicist would, but with one difference: We refer to the resulting quantity as the weak-valued momentum transfer distribution,  $P_{\text{wv}}(q)$ . It is in this manner that we directly observe a momentum transfer distribution: It is derived via a simple prescription, with no reference to quantum physics, from measurements a classical physicist would understand.

A minor complication is that the experimental resolution of the initial momentum measurement is limited to  $\Delta$ . In the place of  $|p_i\rangle\langle p_i|$  we use the projector,

$$\hat{\pi}(p_i) = \int_{p_i-\Delta/2}^{p_i+\Delta/2} dp|p\rangle\langle p|, \quad (4.3)$$

weakly measuring it for  $p_i = n\Delta$ , where  $n$  is an integer. The resolution  $\Delta$  should be  $\ll h/s$ , the interference pattern fringe spacing. The weak value  $d/D = \int_{p_f} \langle \pi_{\text{w}}(p_i) \rangle_{\psi}$  is now the WVP for the initial momentum to lie within  $\Delta/2$  of  $p_i$ , given that the final momentum is  $p_f$ , which we still write as  $P_{\text{wv}}(p_i|p_f)$  for simplicity. Our prescription for calculating  $P_{\text{wv}}(q)$  from  $P_{\text{wv}}(p_i|p_f)$  remains unchanged. If there is no WWM, then  $P_{\text{wv}}(p_i|p_f) = 1$  if  $|p_i - p_f| < \Delta/2$ , and 0 otherwise. Any deviation from this represents a momentum disturbance.

Much like a standard probability distribution,  $P_{\text{wv}}(q)$  integrates to unity. Moreover, its mean and variance exactly reflect the change in the mean and variance of the

momentum distribution that occurs as a result of the WWM [153]. For WWMs that produce random momentum kicks [150],  $P_{\text{wv}}(q)$  is guaranteed positive. However, in some cases quantum processes *must* be described as a distribution of amplitudes, not probabilities [150]. If  $P_{\text{wv}}(q)$  goes negative, it indicates one of these cases; a WWM that induces a “nonclassical momentum transfer”. Reflecting the theorem of Storey *et al.*,  $P_{\text{wv}}(q)$  must have a width of at least  $\hbar/s$ . On the other hand, in the Scully *et al.* scheme where the WWM has no effect on a single-slit wavefunction, the mean and variance of  $P_{\text{wv}}(q)$  are exactly zero (in the limit  $\Delta \rightarrow 0$ ), reflecting the no-disturbance calculation of Storey *et al.*

There have been a couple of elegant experiments that purport to address this issue. However, the question of momentum transfer by WWMs in a double-slit apparatus, as treated by Bohr, Feynman, Scully *et al.*, and Storey *et al.*, has yet to be investigated. The key element of the original Scully *et al.* scheme was the preservation of the single-slit wavefunction. There are many ways to accomplish this more simply than with atomic cavities. In particular, an experiment by Rempe and co-workers [154] creates a WWM that rotates the atomic spin. However, they say, “In our experiment, no double-slit is used and no position measurement is performed, so that the results of [Reference 9] do not apply.” Specifically, the two paths for the atoms are separated by much less than the width of the wavefunction of the atoms. Consequently, a double-slit could not be inserted into their experiment at any point, even in principle. Moreover, although they argue convincingly that there exists no theoretical physical mechanism for momentum transfer in their setup, they make no direct measurement of the transfer itself or lack thereof.

### 4.3 Experiment

The experimental apparatus is shown in Fig. 4.1. Since photons are non-interacting particles, it unnecessary to send only one through the apparatus at a time. Instead,



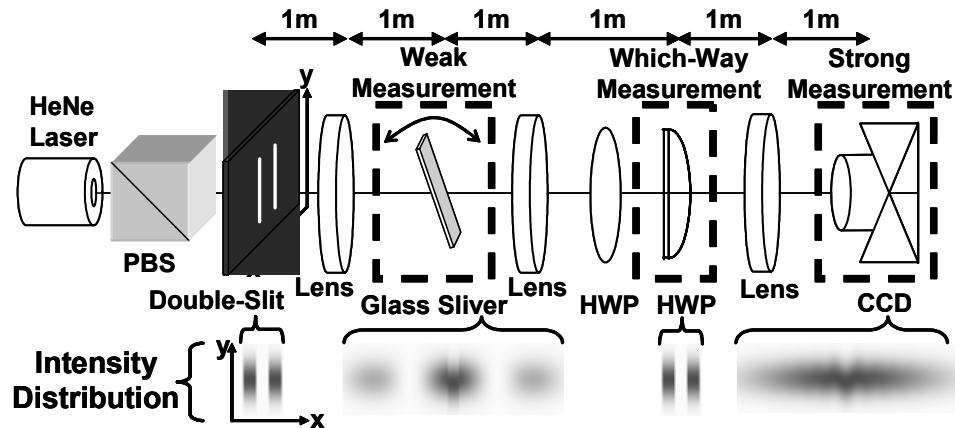


Figure 4.1: Diagram of apparatus. After the photons are prepared in the initial state by a polarizing beamsplitter (PBS) and a double-slit aperture, the experiment can be divided into three stages: The weak measurement of the initial momentum  $p_i$ , the which-way measurement consisting of a half-wave plate (HWP) that rotates the polarization of one slit by  $90^\circ$ , and the post-selection consisting of a strong measurement of  $p_f$  by a CCD camera. Below are shown representative intensity distributions at the double-slit, after the glass sliver, upon reimaging the slits, and at the CCD.

we use a large ensemble simultaneously prepared with the same wavefunction. It follows that the intensity distribution of the light is proportional to the probability distribution for each photon.

The photon ensemble is produced by a 2 mW  $\lambda = 633\text{nm}$  HeNe laser that first passes through a polarizing beamsplitter and then illuminates a double-slit aperture with a slit width of  $w = 40\mu\text{m}$  and a center-to-center separation of  $s = 80\mu\text{m}$ . The slits were laser machined from a thin aluminum sheet and the edges of the slit had a surface smoothness limited to  $1\mu\text{m}$ . The slit width was chosen such that the entire interference pattern fit within the two inch optical aperture of our system. In addition, the slit separation was chosen so that only three fringes fit within the envelope of the pattern, thus maximizing  $h/s$  (so  $\Delta \ll h/s$ ). We call the long (vertical) axis of the slits  $y$  and the axis joining their centers  $x$ . We use  $f = 1\text{m}$  focal-length lenses to switch back and forth between transverse position and momentum space

for the photons. For example, the first lens is placed 1m after the slits, resulting in the Fourier transform of the double-slit wavefunction occurring 1m after the lens. Consequently, at this point the intensity distribution in the  $x$ -direction is that of the expected double-slit interference pattern with a fringe spacing of  $8.2 \pm 0.1\text{mm}$ . In the  $y$ -direction, the intensity distribution is Gaussian with a  $1/e^2$  half-width  $\sigma = 1.01 \pm 0.01\text{mm}$ , which is the weak measurement pointer uncertainty.

We tag the photons with a  $y$ -displacement  $D = 0.14 \pm 0.01\text{mm}$  ( $\ll \sigma$  ensures weakness) in a range of momenta  $\Delta$  centered on  $p_i$ . This weakly measures  $\hat{\pi}(p_i)$  by coupling it to the  $y$  degree of freedom of the photon, which we will later measure directly. This displacement is induced by tilting an optically flat glass sliver placed at  $x_i$  with a width of  $\delta = 1.77 \pm 0.02\text{mm}$  in the  $x$ -direction and a thickness of  $1.00 \pm 0.25\text{mm}$ . An optically flat glass sliver with a width of  $\delta = 1.77 \pm 0.02\text{mm}$  in the  $x$ -direction and a thickness of  $1.00 \pm 0.25\text{mm}$  is placed at  $x_i$  and tilted so that it creates a  $y$ -displacement of  $0.14 \pm 0.01\text{mm}$  ( $\ll \sigma$  ensures weakness) in a range of momenta  $\Delta$  centered around  $p_i$ . Fine adjustment of the tilt ensured that upon passing through the glass the photons accumulated no additional phase (modulo  $2\pi$ ). The glass sliver was cut with a diamond saw so that its edges would be relatively flat ( $\approx 10\mu\text{m}$ ). Nonetheless, black lines along the edges of the image of the sliver (which occurs further along in the optical setup) indicated that some light was scattered.

To implement the WWM we switch back to position space with a second  $f = 1\text{m}$  lens placed 1m from the interference pattern, in essence imaging the slits with a magnification factor of unity. Just before the image plane, which is 1m from the lens, the photons pass through a half-waveplate for fine alignment of their polarization. A second half-waveplate is positioned in front of the image of just one of the slits. This second waveplate was a one-inch uncoated round plate that we cut in half with a diamond saw so that its edge would lie entirely between the two slit edges (separated by  $40\mu\text{m}$ ). The slant of the cut relative to the plate's surface combined with the surface flatness was measured with a microscope to be  $\pm 10\mu\text{m}$ . It was positioned

in between the slits with a micrometer and its edge angle was well-aligned with the long slit axis using a goniometer. This waveplate performs the WWM by flipping the polarization of all the photons in one of the imaged slits. Unlike Scully's original WWM scheme where an ancilla system (the cavity) carries the WWM result, here it is an internal parameter of the photon – polarization. The waveplate is uniform over the slit and thus leaves the spatial wavefunction unaltered while still destroying the interference pattern, making this exactly the type of WWM Scully *et al.* considered.

A third  $f = 1\text{m}$  lens is placed 1m from the double-slit image to transform the wavefunction back into momentum space, so that finally 1m from this lens the transverse position is proportional to the final momentum,  $x_f = f \cdot (p_f \lambda / h)$ . Here we record the intensity distribution with a CCD camera ( $640 \times 480$  pixel,  $3.60 \times 2.70\text{mm}$  sensor). A sequence of images are taken as the camera is shifted horizontally on a translation stage to record the entire intensity distribution, an  $x$ - $y$  region of size  $27.5\text{mm} \times 2.70\text{mm}$ . This was repeated for the glass sliver at a series of positions,  $x_n = n\delta$  for  $n$  running from  $-7$  to  $7$ . Reza Mir was responsible for taking the data and processing the images, including performing the correction for nonlinearities and spatial inhomogeneity in the CCD sensitivity.

## 4.4 Theoretical modelling

The results of the experiment can be modelled in two ways. First, they can be derived with the formula for the weak value in Eq. 4.2. This approach was taken by Josh Garretson, who expanded on the simple theory in [32] by including finite slit widths and a finite weak measurement resolution  $\Delta$  [153]. He calculated both  $P_{\text{wv}}(p_i|p_f)$  and  $P_{\text{wv}}(q)$  for the experimental parameters given above, which resulted in experimental predictions for the case of an infinitely weak measurement and ideal experimental conditions. Secondly, as with all weak measurements the average result can also be predicted by modelling the entire experiment including the von Neumann

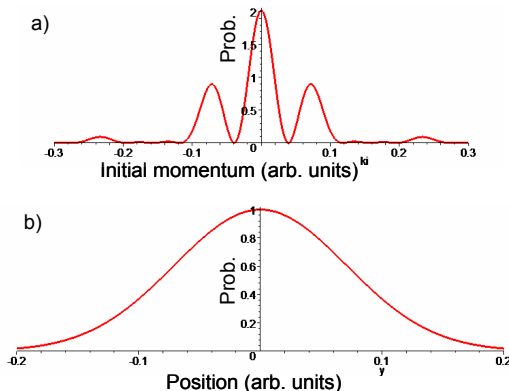


Figure 4.2: a) The initial momentum distribution  $P(k_i)$ . b) The pointer (the vertical  $y$ -direction of the photon) probability distribution.

interaction, which in our case creates a vertical displacement. I took this approach, which amounts to modelling the system optically. This can be done analytically since we are simply applying and inverting Fourier transforms of the optical field. Unlike the weak value calculation, this second modelling method can be used to estimate the effect of finite measurement coupling and other experimental realities. For instance, from this modelling we learned that our system aperture should be enlarged to two inches (we subsequently changed the setup to incorporate this adjustment). One challenge in the experiment was to precisely angle the glass-slide so that the transmitted light received no phase-shift. This model enabled us to determine a physical signature for a suitable angle. The general steps in the calculation correspond to sequential locations in the experiment and are outlined below:

1. Double-slit: We start with Heaviside functions summed to give a flat-topped two-slit field distribution in the  $x$ -direction multiplied by a Gaussian profile in the  $y$ -direction:

$$F_1(x, y) = \text{twoslit}(x, w, s) \cdot e^{-y^2/\sigma^2}. \quad (4.4)$$

2. At the glass-silver: The field distribution here is divided into two terms. The first corresponds to the double-slit interference ( $\text{DBSLIT}(k, w, s) = \mathcal{F}(\text{twoslit}(x, w, s))$ ),

the Fourier transform of the distribution in the first step. From this we subtract a vertical slice of width  $w$  at  $x_i$ . We add another vertical slice ( $\text{slit}(\Delta/\hbar, k_n, k)$ ) displaced by  $D$  in the  $y$ -direction. Dividing the field expression into two terms in this way allows us to perform separate (as opposed to nested) Fourier transforms of the  $x$  and  $y$  components of the field through the subsequent steps. Of course, the Fourier transform of a Gaussian is simple (we neglect Fourier transforms in the  $y$ -direction in the simplified analysis presented here), leaving only the  $x$ -dependence to deal with. Furthermore, the first term is just the undisturbed field distribution and from here on it will evolve as in the regular double-slit experiment (with WWM and without weak measurement):

$$F_2(k, y) = \text{DBSLIT}(k, w, s) \cdot e^{-y^2/\sigma^2} \quad (4.5)$$

$$+ \text{DBSLIT}(k_n, w, s) \cdot \text{slit}(\Delta/\hbar, k_n, k) \cdot (e^{-(y-D)^2/\sigma^2} - e^{-y^2/\sigma^2}). \quad (4.6)$$

3. At the WWM: To derive the field before the WWM we calculate the inverse Fourier transform of the field from each of the terms in the last step. The first term just gives the original two-slit field distribution. The second term gives a single-slit diffraction pattern ( $\text{SSLIT}(x, \Delta/\hbar) = \mathcal{F}^{-1}(\text{slit}(\Delta/\hbar, k_n, 0))$ ) multiplied by Gaussians centered on  $y = 0$  and  $y = D$ , respectively. The overall phase of these latter two terms (which have a relative phase of  $\pi$ ) depends on the position  $x_n = n\delta = \hbar k_n$  of the sliver, since a position-shift changes to a phase-shift under a Fourier transform:

$$F_3(x, y) = \text{twoslit}(x, w, s) \cdot e^{-y^2/\sigma^2} \quad (4.7)$$

$$+ e^{ixk_n} \cdot \text{DBSLIT}(k_n, w, s) \quad (4.8)$$

$$\cdot \sqrt{\Delta/\hbar} \text{SSLIT}(x, \Delta/\hbar) \cdot (e^{-(y-D)^2/\sigma^2} - e^{-y^2/\sigma^2}). \quad (4.9)$$

We incorporate the action of the WWM in the next step.

4. At the CCD: We take the Fourier transform of the field  $x > 0$  from step 3, calculate the absolute value squared, and add this to the same calculation for  $x < 0$  to give the intensity distribution at the CCD,  $I_4(k_f, y)$ . Consequently, we add the

probabilities, as opposed to the amplitudes, for a photon to pass through each of the two slits.

$$I_4(k, y) = |\mathcal{F}(F_3(x, y)\theta(-x))|^2 + |\mathcal{F}(F_3(x, y)\theta(x))|^2 \quad (4.10)$$

The first term in  $F_3(x, y)$  (i.e. for  $x > 0$ ) is just a displaced single-slit and so under Fourier transform it will become a single-slit diffraction pattern. The Fourier transforms for the second term can be found with convolution theorem. They are equal to the convolution of the field at step 2 with the Fourier transform of a Heaviside function ( $\theta(x)$ ). The center of mass of the intensity distribution in the  $y$ -direction at each  $x_f$  ( $\propto k_f$ ) is the prediction for the average result of our measurement, which we take as  $P_{\text{wv}}(p_i|p_f)$ . Here is the  $x > 0$  direction, for example:

$$\mathcal{F}(F_3(x, y)\theta(x)) \quad (4.11)$$

$$= \mathcal{F}(F_3(x, y)) * \mathcal{F}(\theta(x)) \quad (4.12)$$

$$= F_2(k, y) * \mathcal{F}(\theta(x)) \quad (4.13)$$

$$= \text{SSLIT}(k_f, w) \cdot e^{iks/2} \cdot e^{-y^2/\sigma^2} \quad (4.14)$$

$$+ \text{DBSLIT}(k_n, w, s) \cdot \sqrt{2\pi}(e^{-(y-D)^2/\sigma^2} - e^{-y^2/\sigma^2}) \quad (4.15)$$

$$\cdot \left( \begin{array}{l} \sqrt{\frac{\pi}{2}} \text{slit}(\Delta/\hbar, k_n, k_f) \\ + \frac{i}{\sqrt{2\pi}} \ln \left( \frac{2k_f - 2k_n - \Delta/\hbar}{2k_f - 2k_n + \Delta/\hbar} \right) \end{array} \right) \quad (4.16)$$

$$= (a + ib)f_1(y) + (c + id)f_2(y), \quad (4.17)$$

where  $f_1(y) = e^{-y^2/\sigma^2}$  and  $f_2(y) = e^{-(y-D)^2/\sigma^2} - e^{-y^2/\sigma^2}$ , and  $a \rightarrow d$  are functions of the final transverse momentum  $k_f$  (defined by the equation above). At every point  $x_f$  in the intensity distribution we have an expression that only contains terms that are products of identical Gaussians, some shifted by  $D$ ,

$$I_4(k_f, y) = |\mathcal{F}(F_3(x, y)\theta(-x))|^2 + |\mathcal{F}(F_3(x, y)\theta(x))|^2 \quad (4.18)$$

$$= 2(a^2 + b^2)f_1^2 + 2(ac - 2bd)f_1f_2 + 2d^2f_2^2 + c^2f_2^2. \quad (4.19)$$

The  $y$ -expectation value of each term is weighted by the corresponding function (i.e.

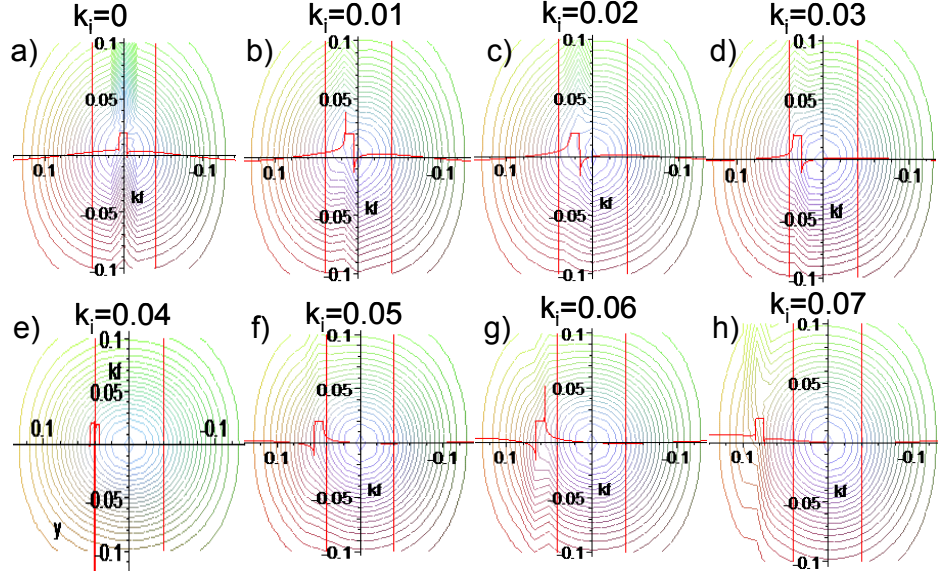


Figure 4.3: Example photon probability distributions (contour plots) at the CCD for the glass-slide at an initial momentum of  $k_i$ . The superposed curve on each distribution is the predicted displacement in the  $y$ -direction at each  $k_f$ . The induced displacement  $D = 0.04$  is relatively large (compared to the vertical width  $\sigma = 0.1$ ) so that the effect of the weak measurement on the distribution is visible. The vertical lines correspond to the first zeroes in the initial momentum distribution given in Fig. 4.2.

$a \rightarrow d$  evaluated at  $x_f$ ), to find the  $y$ -expectation value of the whole distribution,

$$\text{Center}(I_4(k_f, y)) \quad (4.20)$$

$$= \int y \cdot I_4(k_f, y) dy \quad (4.21)$$

$$= \frac{\left( 2(a^2+b^2)\text{Center}(f_1^2) + 2(ac-2bd)\text{Center}(f_1, f_2) \right) + (2d^2+c^2)\text{Center}(f_2^2)}{2(a^2+b^2) + 2(ac-2bd) + 2d^2 + c^2}. \quad (4.22)$$

In the limit of  $D \rightarrow 0$ , this is equal to  $P_{\text{wv}}(p_i|p_f)$ .

The following Figures show contour plots of the  $x - y$  intensity distribution at the CCD for a few values of  $x_i$ , along with the prediction for  $\text{Center}(I_4(k_f, y))$ .

## 4.5 Results

The inset of Fig.4.4 shows the momentum distribution of the photons at the CCD, with and without the WWM, giving  $P(p_f)$  and  $P(p_i)$ , respectively. When there is no WWM, the intensity distribution at the CCD is an image of the double-slit interference pattern (modified by the glass sliver), which appears 1m after the first lens. Consequently, if the glass-sliver displaces light at  $x_i$  an exact image of this displaced light appears at  $x_f = x_i$ , signifying no momentum disturbance. We use this fact to accurately measure  $D$ , the induced displacement. To find  $P_{\text{wv}}(p_i|p_f)$ , at  $x_f$  we measure the average displacement  $d$  in the  $y$ -direction of the intensity distribution while the glass sliver is at  $x_i$  and the WWM is in action. As an example, in Fig. 4.4 we show  $P_{\text{wv}}(p_i|p_f)$  for  $p_i = -1.8\text{mm} \cdot h/(\lambda f)$ , along with a theoretical curve derived from the model described above. Notice that  $P_{\text{wv}}(p_i|p_f)$  becomes negative for certain values of  $p_f$ . We sum the conditional probabilities for all fifteen  $p_i$  according to Eq. (4.1) to obtain the unconditional WVP of a momentum transfer  $P_{\text{wv}}(q)$ , plotted in Fig. 4.5 along with a theoretical curve. This shows that with the Scully *et al.*-type WWM,  $P_{\text{wv}}(q)$  is nonzero outside the range  $[-\hbar/s, \hbar/s]$ , substantiating Storey *et al.*'s theorem.

Nonetheless, we also expect  $P_{\text{wv}}(q)$  to have zero variance, consistent with Scully *et al.*'s conclusions. Unfortunately, the discontinuity of the double-slit wavefunction  $\psi(x)$  results in a theoretical  $P_{\text{wv}}(q)$  that falls off slowly enough to require apodization to evaluate the variance [153]. Since experimental data sets are finite we instead calculate the variance over the range  $[-q_{\text{max}}, q_{\text{max}}]$  without apodization. The theory predicts a variance that diverges as a function of  $q_{\text{max}}$ , oscillating between positive and negative values as it must for the apodized variance to evaluate to zero. This is plotted in the inset of Fig. 4.5, along with the experimental variance, which agrees well. This is the experimental signature of the absence of momentum-disturbance (in the sense of Scully *et al.*).



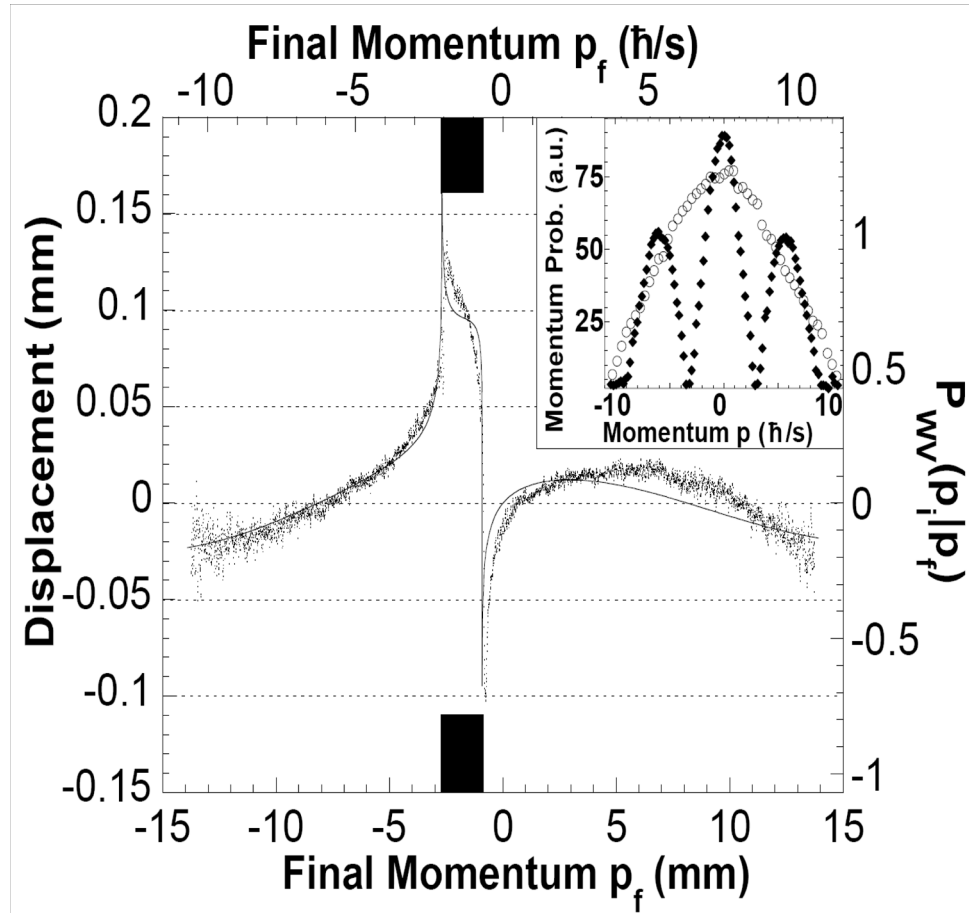


Figure 4.4: Weak measurement of  $p_i = -1.8\text{mm} \cdot h/(\lambda f)$ . The dots indicate the  $y$ -displacement  $d$  of the intensity distribution at each  $x$ -position, and hence  $p_f$ , on the CCD. Dividing by  $D$  gives the weak-valued probability  $P_{wv}(p_i|p_f)$ . The thin solid line is a theoretical curve calculated from Eq. (4.1) similar to Ref. [153] using only the independently measured parameters,  $w$ ,  $s$ , and  $\Delta$ . The solid black rectangles indicate the region of weak measurement (bounded by  $p_i \pm \Delta/2$ ). The inset shows the measured intensity at each  $x$ -position integrated over the  $y$ -dimension of CCD with (solid diamonds) and without (empty circles) the WWM. The light levels outside the  $x$ -range were below the sensitivity of our detector.

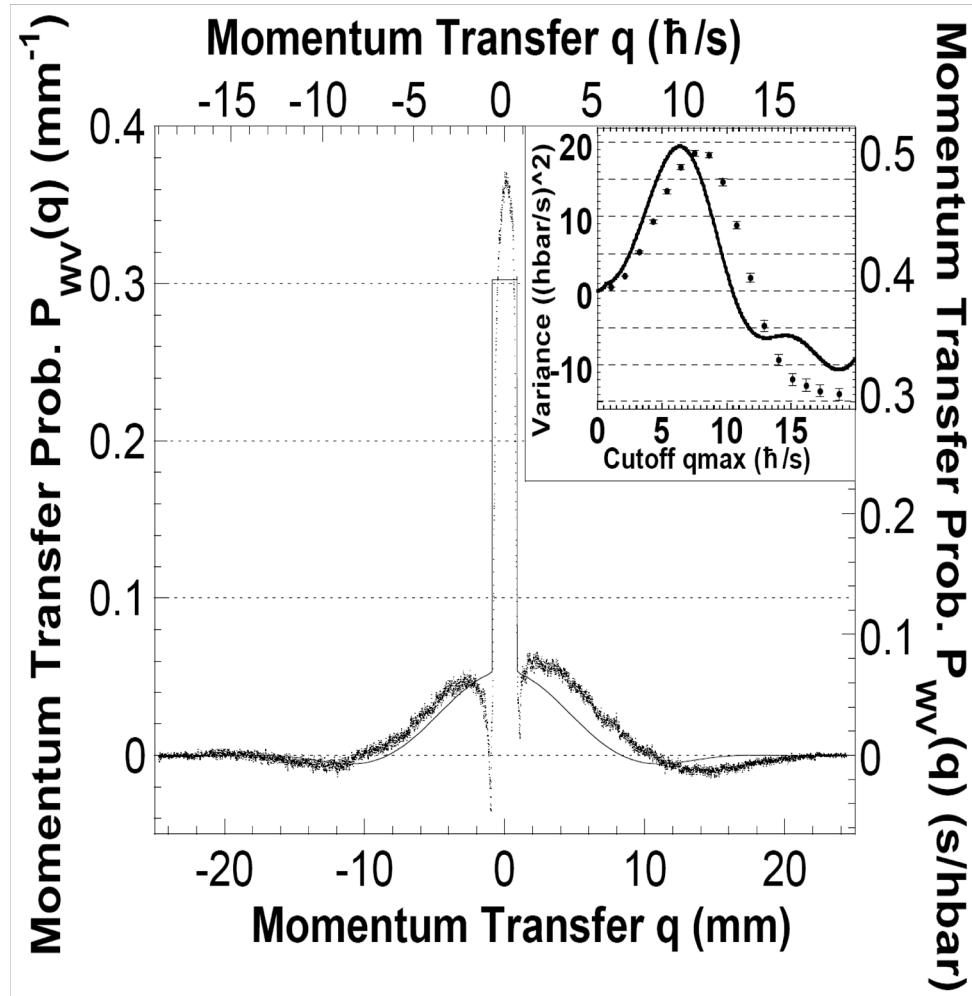


Figure 4.5: The weak-valued probability distribution for the momentum transfer  $P_{wv}(q)$ . The dots are experimental points and the thin solid line is a theoretical curve calculated from Eq. 4.2 and Eq. (4.1) similar to Ref. [153] using only the independently measured parameters,  $w$ ,  $s$  and  $\Delta$ . The inset is the variance of experimental data (solid circles) in the range  $[-q_{\max}, q_{\max}]$  as a function of  $q_{\max}$  along with the theoretical prediction (solid line).

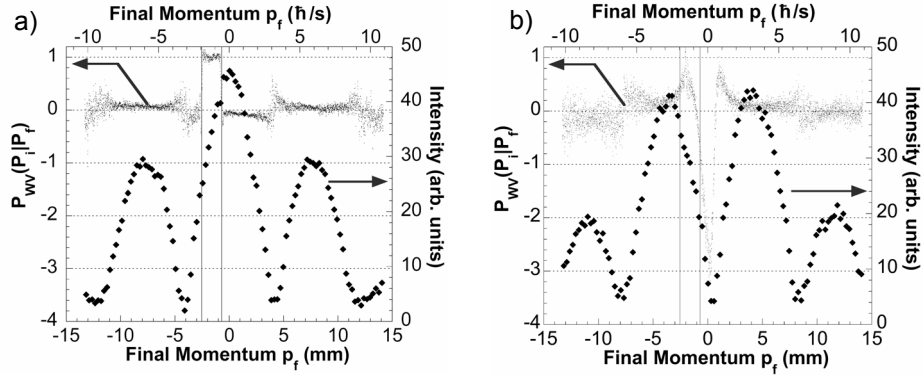


Figure 4.6: The weak-valued probability  $P_{wv}(p_i|p_f)$  for  $p_i = -1.8\text{mm} \cdot h/(\lambda f)$  with a quantum eraser consisting of a) a  $45^\circ$  polarizer and b) a  $-45^\circ$  polarizer, both placed after the which-way measurement. The dots indicate the  $y$ -displacement of the intensity distribution at each  $x$ -position, and hence  $p_f$ , on the CCD. The diamonds indicate the intensity at each  $p_f$ . The vertical lines indicate  $p_i \pm \Delta/2$ , the region of the weak measurement.

## 4.6 Quantum eraser

Scully *et al.* also considered the retrieval of interference through the use of a quantum eraser, considering this to be additional proof that a random momentum transfer does not occur. Implicit in this is the idea that a random process leads to an irretrievable loss of coherence. This is incorrect. Even in the case of the Feynman microscope, an appropriate measurement on the scattered photons (that illuminate the slits and perform the WWM) will perform an erasure and allow one to observe interference [155][156]. However, this leads to another question: In the case of quantum erasure, what becomes of the randomly kicked particles? In this section, we investigate the change a quantum eraser makes to the nonclassical momentum transfer we presented above.

A quantum eraser [157] is a measurement on the WWM apparatus that retrieves the double-slit interference pattern destroyed by the WWM. That is, one sorts the particles into bins according to the results of a measurement performed in a conjugate

basis to the one that carries the WWM result. Even though the entire set of particles will still not form an interference pattern, the subset of particles in each bin will. For a WWM with classical momentum transfer, the different bins contain identically shaped interference patterns, but are shifted in the  $x$ -direction by varying amounts [155][156]. By contrast, for a WWM such as that of Scully *et al.* [30], the interference patterns in the different bins all have the same envelope; only the phases of the patterns differ [158][159].

Since our WWM result is carried in the Horizontal/Vertical basis of the photon polarization, we implement a quantum eraser with a polarizer in the  $45^\circ / -45^\circ$  basis. This sorts the photons into two bins:  $45^\circ$ , which form the original interference pattern (fringes); and  $-45^\circ$ , which form a pattern with the opposite phase (antifringes). In Fig. 4.6, we plot  $P_{\text{wv}}(p_i|p_f)$  with  $p_i = -1.8\text{mm} \cdot h/(\lambda f)$  for both polarizer settings, along with the measured interference patterns. For the photons transmitted at  $45^\circ$ , the data shows that, to a good approximation,  $P_{\text{wv}}(p_i|p_f) = 1$  if  $|p_i - p_f| < \Delta/2$  and 0 otherwise, indicating that there is no momentum transfer. On the other hand, for the  $-45^\circ$  photons,  $P_{\text{wv}}(p_i|p_f)$  is substantially different from zero even for  $p_f$  outside the range  $p_i \pm \Delta/2$ , reflecting a momentum disturbance. This result is found for all values of  $p_i$ , showing clearly that the momentum transfer only appears in the  $-45^\circ$  photons. Since these are the anti-fringe photons, this indicates an intimate connection between the nonclassical momentum transfer and the phase between the slits induced by the quantum eraser.

## 4.7 Conclusion and discussion

With the technique of weak measurement we directly observe momentum transfers greater than  $\hbar/s$  accompanying the type of WWM that Scully *et al.* considered. This supports Storey *et al.*'s claim that complementarity is a consequence of Heisenberg's position-momentum Uncertainty Principle. However, our observations also support

Scully *et al.*'s claim of no momentum transfer since the variance of the observed momentum transfer is consistent with zero. These seemingly contradictory observations are compatible only because the weak-valued probability distribution we measure takes negative values, indicating that a nonclassical momentum transfer takes place.

More precisely, the random momentum transfer induced by our WWM is a quantum process and therefore must be represented in terms of amplitudes versus probabilities. This leads to an additional consideration, or caveat, regarding the above conclusions: Although the weak-valued probability distribution for the momentum transfer has a width of  $\hbar/s$ , meaning support outside  $[-\hbar/s, \hbar/s]$ , the Heisenberg uncertainty relation is given in terms of the variances of real, as opposed to quasi, probability distributions and so is not directly relevant to our results. There are, however, proposals for generalized uncertainty relations that may be applicable [160].

Furthermore, given that we have directly measured that some momentum transfer does occur (with the caveat that it is a quantum process), one is drawn to look for the physical mechanism that induces the transfer. A simple answer is that the waveplate changes the relative phase of the wavefunction at two positions (the slits) from being equal to non-existent. Since the particle momentum is proportional to the phase gradient of the wavefunction, one should expect that the momentum is changed when the relative phase is changed or destroyed. We can consider a similar situation in optics: If a waveplate flips the polarization of half of a laser beam spot, we would expect roughly twice the diffraction as in the original beam. The extra momentum components come from the electromagnetic interaction with dipoles in the edge of the waveplate. However, for an ideal WWM of Scully's type, there is no light (or wavefunction) in between the slits or at the edge of the waveplate. So where does the physical interaction occur? At this point in time, this is an unanswered question.

Another interesting question is, how is the total system momentum distributed so that it is conserved in the WWM process? In Feynman's Microscope, the total momentum of the scattered photon and electron is conserved. In our WWM, we

should similarly conclude that the waveplate's momentum is changed by the WWM process in such a way that when added to the photon momentum, the total is conserved. Remember that the waveplate must have an initial momentum width of  $\hbar/s$  to be localized to within the slit separation, so this momentum transfer might result in a relatively small change to the wavefunction.

# Chapter 5

## Measurement of Joint Weak Values

### 5.1 Introduction

In modern quantum mechanics, we are increasingly interested in a different class of observables than were discussed in the thesis introduction to weak measurement, in which only a single particle is involved [94]. Often, one would like to measure correlations between joint observables. These are products of observables of distinct particles, like  $\hat{S}_1^z \hat{S}_2^z$ , the z-spin component of particle one times particle two. Any experiment that utilizes or directly measures properties of entanglement is based on such observables and therefore quantum information and quantum optics frequently deal with these composite or joint observables. The exciting results and complex, rich range of features discovered by studies of entanglement suggest that weak measurement of joint observables should also produce valuable and interesting results. In fact, a few theoretical ideas for weak measurements that center around joint observables have already been published, including nonlocality of a single particle [57], extensions of the Quantum Box Problem [161][162], and Hardy's Paradox [33]. (In particular, we later describe an implementation of the proposal in the last paper, applying the techniques developed in this chapter.) We call the weak value of a joint observable the "joint weak value." If the composite observable is a product of  $N$

single particle observables, the weak value is called the “ $N$ th-order joint weak value”.

Remember from the thesis introduction that in simple weak measurements the measured system itself is typically used as part of the measurement device. When measuring  $\hat{A}$  of a particle, an independent degree of freedom of the particle can be used as the pointer. For example, a birefringent crystal can be oriented so that it will displace the position of a photon by an amount that depends on the photon’s polarization [72]. Here,  $\hat{A}$  is the polarization observable and the pointer is the position of the photon. (Another example is the Stern-Gerlach apparatus, where  $\hat{A}$  is the spin of the particle and the pointer is the momentum of the particle.) If such a measurement strategy were not available, one would require a strong controllable (von Neumann) interaction between the quantum system and a separate pointer system. This is typically far too technically difficult to implement.

Joint observables are extremely difficult to measure directly with either strong or weak types of measurement. The difficulty lies in the fact that the necessary von Neumann interaction (in Eq. 1.19) couples two separate observables, and hence particles, to a single pointer. One can therefore no longer use transformations of the state of a single particle to implement von Neumann interaction, and so one requires multiparticle interactions. An approach using multiparticle interactions was outlined in a proposal for a weak measurement experiment with ions [163], but so far there have been no experimental weak measurements of joint observables.

On the other hand, indirect strong measurements of joint observables are feasible and even commonplace. This is made possible by employing a different measurement strategy. Instead of measuring the joint observable directly, each single particle observable is measured simultaneously but separately. For example, instead of measuring  $\hat{S}_1\hat{S}_2$  directly we can measure  $\hat{S}_1$  and  $\hat{S}_2$  separately and then multiply the results trial by trial. If one wants to strongly measure the joint observable  $\hat{A}_1\hat{A}_2\dots\hat{A}_N = \hat{M}$ , instead of using the multiparticle von Neumann Hamiltonian  $\mathcal{H} = g\hat{M}\hat{P}$ , the general strategy is to simultaneously apply  $N$  standard single-particle von Neumann interac-



tion Hamiltonians:

$$\mathcal{H} = g_1 \hat{A}_1 \hat{P}_1 + g_2 \hat{A}_2 \hat{P}_2 + \dots \quad (5.1)$$

$$= \sum_{j=1}^N g_j \hat{A}_j \hat{P}_j. \quad (5.2)$$

Given that we can already perform each of the single-particle Hamiltonians, it is straightforward to implement the total Hamiltonian in Eq. 5.2. In summary, this strategy allows one to make projective measurements of  $\hat{M}$  – all that is required to measure the expectation value of  $\hat{M}$ .

Alternately, one could equivalently calculate  $\langle \hat{M} \rangle$  directly from the position distributions of the pointers after the measurement, since these carry the results. Specifically,

$$\langle \hat{M} \rangle = \langle \hat{A}_1 \hat{A}_2 \dots \hat{A}_N \rangle \propto \langle \hat{X}_1 \hat{X}_2 \dots \hat{X}_N \rangle, \quad (5.3)$$

where  $\hat{X}_i$  is the position operator of the pointer, and provided all  $\hat{A}_i$  commute. In other words, the expectation value of  $\hat{M}$  is related to the correlation between the positions (specifically, the average of the product of the positions) of all  $N$  pointers. This variation has the advantage that the measurement coupling could be weak, (i.e. it shifts a pointer by an amount less than its width). Although projective measurements of  $\hat{M}$  would now be impossible, the average product of the positions can still be determined given enough data. In analogy with single particle (or observable) weak measurements, one might expect that in the weak coupling limit of this measurement  $\langle \hat{X}_1 \hat{X}_2 \dots \hat{X}_N \rangle$  will be proportional to the weak value. However, in the case of subsequent post-selection this is not the case. To measure the joint weak value of  $\hat{M}$ , one requires a version of Eq. 5.3 that is generalized to weak measurement. This chapter derives this generalization and motivates it as a simple extension of single particle (or observable) weak measurement.

This chapter builds on two papers in which an analogous strategy was applied to weak measurements [164][165]. The Hamiltonian in Eq. (5.2) is utilized in the weak regime to create correlations in the deflections of the  $N$  pointers proportional to the weak value. Specifically, the  $N$ th-order joint weak value was related to two correlations between all  $N$  pointer deflections and a complicated combination of lower-order joint weak values. In this chapter, we show that the  $N$ th-order joint weak value takes on an elegant and simple form, closely related to the strong measurement formula in Eq. (5.3), when expressed entirely in terms of  $N$ -pointer correlations. This new and simplified form lends itself to a new way of thinking about single and joint weak measurements in terms of expectation values of products of annihilation operators. As mentioned in the thesis introduction, many people who have considered weak measurement regard the imaginary part as extraneous to the measurement result. In contrast, the new form for the joint weak value gives greater importance to the momentum shift of the pointer and, along with it, the imaginary part of the weak value.

## 5.2 Theory

### 5.2.1 Background

We begin with the standard weak measurement theory presented in the thesis introduction. For convenience, we repeat a few equations from that section, including the von Neumann interaction Hamiltonian,

$$\mathcal{H} = g\hat{A}\hat{P}, \quad (5.4)$$

where  $\hat{A}$  is the observable of the measured system,  $\hat{P}$  is the momentum of the pointer, and  $g$  is the coupling constant. The measurement pointer is initially in a Gaussian

wavefunction centered at zero:

$$\langle x|\phi\rangle = \phi(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{\frac{1}{2}} \exp\left(-\frac{x^2}{4\sigma^2}\right). \quad (5.5)$$

After the weak measurement (and post-selection) the position and momentum shifts of the pointer are found to be,

$$\langle \hat{X} \rangle_{fi} = gt \operatorname{Re} \left( \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right), \quad (5.6)$$

$$\langle \hat{P} \rangle_{fi} = \frac{\hbar gt}{2\sigma^2} \operatorname{Im} \left( \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right), \quad (5.7)$$

where the weak value is

$$\langle \hat{A} \rangle_w \equiv \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle}. \quad (5.8)$$

## 5.2.2 The annihilation operator and weak measurement

One can express the full weak value in terms of the two expectation values of the pointer:

$$\langle \hat{A} \rangle_w = \operatorname{Re} \langle \hat{A} \rangle_w + i \operatorname{Im} \langle \hat{A} \rangle_w \quad (5.9)$$

$$= \frac{2\sigma}{gt} \left\langle \frac{1}{2\sigma} \hat{X} + i \frac{\sigma}{\hbar} \hat{P} \right\rangle_{fi}. \quad (5.10)$$

In their derivation of weak values, Aharonov, Albert and Vaidman (AAV) made the natural choice of a Gaussian for the initial pointer state, as do we. This state also happens to be the ground state  $|0\rangle$  of a harmonic oscillator with mass  $m$  and frequency  $\omega$ . For illustration, if one reparameterizes the width of the Gaussian in terms of  $m\omega$ , such that  $\sigma = \sqrt{\hbar/2m\omega}$ , it becomes apparent that the operator in the expectation value in Eq. (5.10) is just the familiar lowering operator,

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \hat{X} + i \sqrt{\frac{1}{2m\omega\hbar}} \hat{P}. \quad (5.11)$$

The operator in Eq. (5.10) will transform the pointer just as the lowering operator does, even though the pointer is not actually in a harmonic potential. This fact will

simplify some of the following calculations. Furthermore, now the weak value can be re-expressed as:

$$\langle \hat{A} \rangle_w = \frac{2\sigma}{gt} \langle \hat{a} \rangle_{fi}. \quad (5.12)$$

This is the first time in the literature that this simple but important relationship between the annihilation operator and weak measurement has been described. The reason the annihilation operator is related to the weak value is not obvious. At the present time, we are limited to an explanation based on the mathematical action of the annihilation operator. When the coupling is sufficiently weak, the expansion in Eq. (1.30) shows that the largest pointer amplitude is left unchanged in the ground state. The interaction Hamiltonian shifts some of the pointer state into the first excited state by creating a small amplitude, proportional to  $gt\hat{A}$ , for the  $|1\rangle$  state. If we restrict ourselves to the post-selected subensemble, as in Eq. (1.33), this small amplitude changes to be proportional to  $gt \langle \hat{A} \rangle_w$ . The annihilation operator removes the part of the state that is left unchanged by the coupling, leaving only the shifted component. In other words, the annihilation operator isolates only that part of the pointer state that is changed by the interaction.

### 5.2.3 Joint weak values

We now move on to a derivation of  $N$ th-order joint weak values. In this section, we combine the strategy outlined in the chapter introduction for measuring joint observables with the use of the annihilation operator to extract the weak value. As in previous works, to measure the operator  $\hat{M} = \prod_{j=1}^N \hat{A}_j$  we apply  $N$  separate von Neumann interactions, coupling each  $\hat{A}_j$  to its own pointer, as in Eq. (5.2) [164][165]. To simplify the expressions to come, we set all  $g_j$  to be equal and rewrite the momentum operators  $\hat{P}_j$  in terms of the respective raising and lowering operators,

$\hat{a}_j^\dagger$  and  $\hat{a}_j$ , for each of the pointers,

$$\mathcal{H} = i \frac{\hbar g}{2\sigma} \sum_{j=1}^N \hat{A}_j (\hat{a}_j^\dagger - \hat{a}_j). \quad (5.13)$$

Now we require  $N$  different pointers, all beginning in an initial state defined by Eq. (1.20). The total initial pointer state can be described by the ground state of  $N$  harmonic oscillators:

$$|\Phi\rangle = \prod_{j=1}^N |\phi_j\rangle = |0\rangle^{\otimes N}. \quad (5.14)$$

Continuing, using the number-state notation to describe the pointer, we calculate the state of the combined system after the interaction Hamiltonian is applied,

$$|\Phi\rangle |I\rangle \rightarrow \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |0\rangle^{\otimes N} |I\rangle = \left(1 - \frac{i\mathcal{H}t}{\hbar} + \dots\right) |0\rangle^{\otimes N} |I\rangle \quad (5.15)$$

$$= \left(1 + \frac{gt}{2\sigma} \sum_{j=1}^N \hat{A}_j (\hat{a}_j^\dagger - \hat{a}_j) + \dots\right) |0\rangle^{\otimes N} |I\rangle \quad (5.16)$$

$$= |0\rangle^{\otimes N} |I\rangle + \frac{gt}{2\sigma} \sum_{j=1}^N \hat{A}_j |1_j\rangle |I\rangle + \dots, \quad (5.17)$$

where  $|1_j\rangle$  is the state in which the  $j$ th pointer is in the first-excited state and all the other pointers are in the ground state (e.g.  $|0_1 1_2 0_3 \dots 0_N\rangle$ ). Here, we have expanded the state in powers of  $gt$ . Eq. (5.15) shows that to first-order, the interaction Hamiltonian coupling the measuring device to the system can displace only one of the  $N$  pointers at a time. Simultaneous shifts of multiple pointers come from higher-order terms in the propagator. We are particularly interested in the  $N$ th term in the expansion,

$$\frac{1}{N!} \left(\frac{-i\mathcal{H}t}{\hbar}\right)^N = \frac{1}{N!} \left(\frac{gt}{2\sigma} \sum_{j=1}^N \hat{A}_j (\hat{a}_k^\dagger - \hat{a}_k)\right)^N. \quad (5.18)$$

This term is the lowest-order one in the expansion which can simultaneously transfer all  $N$  pointers into the first excited state (e.g.  $|1_1 1_2 1_3 \dots 1_N\rangle$ ). This state, which we label as  $|1\rangle^{\otimes N}$ , is created when each term in the above sum supplies one raising

operator. The terms in the sum can contribute the  $N$  distinct raising operators in any order, and so the portion of Eq. (5.18) that creates the  $|1\rangle^{\otimes N}$  state is equal to,

$$\frac{1}{N!} \frac{gt}{2\sigma} \wp \left\{ \hat{A}_k \hat{a}_k^\dagger \right\}_N, \quad (5.19)$$

where  $\wp \left\{ \hat{L}_k \right\}_N$  denotes the sum of all  $N!$  orderings of the set of  $N$  operators  $\left\{ \hat{L}_k \right\}$ . Note that these different orderings are only distinct when the operators do not commute. The remaining portions of Eq. (5.18) create states where at least one pointer is left in the initial state (e.g.  $|2_1 0_2 1_3 \dots 1_N\rangle$ ). Projecting onto  $\langle F|$  completes the post-selection and leaves us with,

$$\langle F| \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |0\rangle^{\otimes N} = |0\rangle^{\otimes N} \langle F|I\rangle + \frac{gt}{2\sigma} \sum_{j=1}^N \langle F| \hat{A}_j |I\rangle |1_j\rangle + \dots \quad (5.20)$$

$$+ \left(\frac{gt}{2\sigma}\right)^N \frac{1}{N!} \langle F| \wp \left\{ \hat{A}_k \right\}_N |I\rangle |1\rangle^{\otimes N} + \dots \quad (5.21)$$

We renormalize the resulting  $N$ -pointer state  $|\Phi_{\text{fi}}\rangle$  and then truncate the amplitude of each term at the lowest nonzero order in  $gt$ ,

$$|\Phi_{fi}\rangle = |0\rangle^{\otimes N} + \frac{gt}{2\sigma} \sum_{j=1}^N \frac{\langle F| \hat{A}_j |I\rangle}{\langle F|I\rangle} |1_j\rangle + \dots + \left(\frac{gt}{2\sigma}\right)^N \frac{1}{N!} \frac{\langle F| \wp \left\{ \hat{A}_k \right\}_N |I\rangle^{\otimes N}}{\langle F|I\rangle} + \dots \quad (5.22)$$

This is equivalent to dividing by  $\langle F|I\rangle$ , the renormalization constant in the limit of no coupling. In analogy with Eq. (5.12), we now wish to take the expectation value of the product of the annihilation operators for all  $N$  pointers,

$$\hat{O} \equiv \prod_{j=1}^N \hat{a}_j. \quad (5.23)$$

In Eq. (5.22), the  $|1\rangle^{\otimes N}$  state is the lowest-order term that does not go to zero when acted on by  $\hat{O}$ ; this term becomes,

$$\hat{O} |\Phi_{fi}\rangle = \left(\frac{gt}{2\sigma}\right)^N \frac{1}{N!} \frac{\langle F| \wp \left\{ \hat{A}_j \right\}_N |I\rangle}{\langle F|I\rangle} |0\rangle^{\otimes N} + O\left((gt)^{N+1}\right). \quad (5.24)$$

Clearly, to lowest nonzero-order the expectation value then becomes,

$$\langle \hat{O} \rangle_{fi} = \langle \Phi_{fi} | \hat{O} | \Phi_{fi} \rangle \quad (5.25)$$

$$= \langle 0 | \left( \frac{gt}{2\sigma} \right)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_j \right\}_N | I \rangle}{\langle F | I \rangle} | 0 \rangle \quad (5.26)$$

$$= \left( \frac{gt}{2\sigma} \right)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_j \right\}_N | I \rangle}{\langle F | I \rangle}. \quad (5.27)$$

The next lowest-order term in the expectation value corresponds to any of the  $N$  pointers undergoing an extra pair of transitions (i.e., a pointer is raised to  $|2\rangle$  and subsequently lowered back to  $|1\rangle$ ). Consequently it will be reduced in size by a factor of  $2 \left( \frac{gt}{2\sigma} \right)^2$  compared to the lowest-order term. Using Eq. (5.27) the  $N$ th-order joint weak value can now be expressed in the simple formula,

$$\frac{1}{N!} \langle \wp \left\{ \hat{A}_j \right\}_N \rangle_W = \left\langle \prod_{j=1}^N \hat{a}_j \right\rangle_{fi} \left( \frac{2\sigma}{gt} \right)^N. \quad (5.28)$$

#### 5.2.4 The $N$ th-order joint weak value

It is often the case that each operator  $\hat{A}_j$  acts on a different particle, ensuring that all  $\hat{A}_j$  commute. This allows the further simplification of the  $N$ th-order joint weak value to,

$$\left\langle \prod_{j=1}^N \hat{A}_j \right\rangle_W = \left\langle \prod_{j=1}^N \hat{a}_j \right\rangle_{fi} \left( \frac{2\sigma}{gt} \right)^N. \quad (5.29)$$

For commuting observables, the magnitude of the simultaneous shift in the  $N$  pointers, which results from concurrent kicks from all  $N$  terms in the Hamiltonian in Eq. (5.2), is proportional to the shift in one pointer created by a single von Neumann Hamiltonian for measuring operator  $\hat{M}$ . The role of the annihilation operators are to isolate this simultaneous pointer shift from the total uncorrelated shifts of the  $N$  pointers, and thus duplicate the action of  $\mathcal{H} = g\hat{M}\hat{P}$ , without the need for multiparticle interactions.

Since Eq. (5.29) requires the measurement of the annihilation operator, which is not Hermitian, one might think the expression is, in principle, unmeasurable. In fact, if one expands the annihilation operator in terms of  $\hat{X}$  and  $\hat{P}$  for each pointer, one is simply left with expectation values of products of  $\hat{X}$  or  $\hat{P}$  for each pointer. It is then possible to measure  $\hat{X}$  in one ensemble of pointers and  $\hat{P}$  in an identically-prepared ensemble.

The expression in Eq. (5.29) is the central result of this chapter. As in previous papers [164][165], this result shows how one can practically measure a joint weak value even without the multiparticle interactions the AAV method requires. However, this expression is much more elegant and makes it clear that the annihilation operator plays a key role in joint weak measurements. Specifically, with the use of the annihilation operator the similarity to the strong measurement expectation value in Eq. (5.3) is apparent. For strong measurement, the equivalent expectation value to the  $N$ th-order joint weak value is:

$$\left\langle \prod_{j=1}^N \hat{A}_j \right\rangle = \left\langle \prod_{j=1}^N \hat{X}_j \right\rangle \left( \frac{1}{gt} \right)^N. \quad (5.30)$$

The similarity is striking and makes a good case for the use of the annihilation operator in the understanding of weak values.

Let us compare Eq. (5.29) to the previous results for the  $N$ th-order joint weak value in Ref. [164]. In this reference, it was expressed recursively in terms of two  $N$ th-order correlations between the pointers, and to  $N$  different joint weak values of order  $N - 1$ . Utilizing this recursive formula, the  $N$ th-order joint weak value can be expressed purely in terms of the expectation value of position and momentum correlations. This expression includes  $2^{N+1} - 2$  distinct correlations of various orders, although most will be close to the  $N/2$  order as the number of distinct expectation values at each order follows the binomial distribution. In comparison, Eq. (5.29) relates the  $N$ th-order joint weak value to  $2^N$  correlations in the positions and momenta of all  $N$  pointers, and so requires roughly half the number of expectation values as



the final result from Ref. [164] (but those expectation values are of higher order).

### 5.2.5 An example

As a specific example of the use of Eq. (5.29), the weak value of the product of two spin components,  $S_{1x}S_{2y}$ , would be,

$$\langle S_{1x}S_{2y} \rangle_W = \left( \frac{2\sigma}{gt} \right)^2 \langle \hat{a}_1 \hat{a}_2 \rangle_{fi} \quad (5.31)$$

$$= \left( \frac{2\sigma}{gt} \right)^2 \left\langle \left( \frac{1}{2\sigma} \hat{X}_1 + i \frac{\sigma}{\hbar} \hat{P}_1 \right) \left( \frac{1}{2\sigma} \hat{X}_2 + i \frac{\sigma}{\hbar} \hat{P}_2 \right) \right\rangle_{fi}. \quad (5.32)$$

The real and imaginary parts of the weak value are then,

$$\text{Re} \langle S_{1x}S_{2y} \rangle_W = \left( \frac{1}{gt} \right)^2 \left( \langle \hat{X}_1 \hat{X}_2 \rangle_{fi} - \frac{4\sigma^4}{\hbar^2} \langle \hat{P}_1 \hat{P}_2 \rangle_{fi} \right) \quad (5.33)$$

$$\text{Im} \langle S_{1x}S_{2y} \rangle_W = \frac{2\sigma^2}{\hbar} \left( \frac{1}{gt} \right)^2 \left( \langle \hat{X}_1 \hat{P}_2 \rangle_{fi} + \langle \hat{P}_1 \hat{X}_2 \rangle_{fi} \right). \quad (5.34)$$

The importance and relevance of the pointer momentum shift is demonstrated in the above example. With our measurement technique even the real part of the weak value is related to the pointers' momenta,  $\hat{P}_1$  and  $\hat{P}_2$ . In general, the momentum and position observables for each of the  $N$  pointers will appear in the expression for the real part of the  $N$ th-order joint weak value.

## 5.3 Spin pointers

Note that, like single weak measurements, this method for measuring the  $N$ th-order joint weak value is not limited to the particular interaction or pointer used in our measurement model [66]. For example, one can perform a derivation very similar to the one presented here in which a spin pointer, as opposed to position pointer, is used. For a spin pointer, the Hamiltonian would be  $\mathcal{H} = -g\hat{A}\hat{S}_y = ig\hat{A}(\hat{S}_z^+ - \hat{S}_z^-)/2$ , where  $\hat{S}_i^+$  and  $\hat{S}_i^-$  are the raising and lowering operators for the  $\hat{S}_i$  basis. The initial pointer state would be the lowest eigenstate of  $\hat{S}_z$ , with eigenvalue  $-\hbar s$ . In this

case, the expression for the  $N$ th-order joint weak value in terms of  $N$  spin pointers is,

$$\left\langle \prod_{j=1}^N \hat{A}_j \right\rangle_W = \left\langle \prod_{j=1}^N \hat{S}_{jz}^- \right\rangle_{fi} \left( \frac{1}{g\hbar s} \right)^N, \quad (5.35)$$

where  $\hat{S}_{jz}^-$  is the  $z$ -basis lowering operator for the  $j$ th pointer and all  $\hat{A}_j$  are assumed to commute. An explicit derivation of this formula is given in Appendix 2. An important advantage of using spin is the absence of unequal coefficients in the expression for the lowering operator. This puts the shifts in the pointer observable and its conjugate on equal footing; using such a pointer means that the physical shift in the conjugate observable does not become smaller as the measurement becomes weaker. Expectation values are also particularly easy to measure for spins (and polarizations), especially spin 1/2 systems since there are only two basis states which need to be projected onto. For instance, the  $N$ th-order joint weak value requires  $2^{2N}$  measurements in total if  $N$  spin 1/2 pointers are used. The advantage of spin pointers is particularly striking when compared to a continuous pointer (e.g. position) for which the  $N$ -pointer distribution must be mapped out in a multidimensional space defined by the  $N$   $x$  and  $p$  pointer coordinates.

Since this is the type of pointer that is used in the next chapter for experimental weak measurements in Hardy's Paradox, we will spend some more time considering it. In the next chapter, we use the polarization of each of two photons as two pointers. As we have just seen, a polarization ( $s=1/2$ ) pointer is easily treated by roughly the same theoretical derivation that we used for a position pointer. However, there is an even closer mathematical analogy that can be made. The position pointer is defined by  $\langle x|\phi\rangle$  to be a delocalized distribution with a width  $\sigma$ . The polarization pointer can be similarly defined by  $\langle \theta|\phi\rangle$ , where  $|\theta\rangle = \cos(\theta)|H\rangle + \sin(\theta)|V\rangle$  is a linear polarization state at angle  $\theta$ , except now the width is fixed to  $\sigma = \Delta\theta = 1$ . The shape of the pointer is not Gaussian but rather a  $\cos^2(\theta)$  function, which, if we limit ourselves to the range  $[-\pi/2, \pi/2]$ , can still be considered a function peaked

at a certain angle. Initially it points to zero degrees and after the weak coupling it points to the measurement result. More formally, the equivalent of  $\langle \phi | \hat{X} | \phi \rangle$  is,

$$\langle \phi | \hat{\theta} | \phi \rangle \equiv \frac{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \theta \cdot |\langle \theta | \phi \rangle|^2 d\theta}{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\langle \theta | \phi \rangle|^2 d\theta}, \quad (5.36)$$

and the equivalent of  $\langle \phi | \hat{P} | \phi \rangle$  is the same equation with the substitution  $\theta \rightarrow \bar{\theta}$ , defined by  $|\bar{\theta}\rangle = \cos(\bar{\theta}) |H\rangle + i \sin(\bar{\theta}) |V\rangle$ . With these definitions, the weak value is given by,

$$\langle \hat{A} \rangle_w = \frac{\langle \theta \rangle_{fi} + i \langle \bar{\theta} \rangle_{fi}}{\theta_D} \quad (5.37)$$

$$= \frac{\langle \hat{\theta} + i\bar{\theta} \rangle_{fi}}{\theta_D}, \quad (5.38)$$

where  $\theta_D$  is the induced spin rotation. The generalization of this result to the Nth-joint weak value is,

$$\left\langle \prod_{j=1}^N \hat{A}_j \right\rangle_w = \left\langle \prod_{j=1}^N (\hat{\theta}_j + i\bar{\theta}_j) \right\rangle_{fi} \left( \frac{1}{\theta_D} \right)^N. \quad (5.39)$$

(Note that the approach we have used is reminiscent of attempts in the past to create a phase operator in quantum optics.) This approach allows us to visualize the pointer correlations that develop as a result of nonzero joint weak values. To this end, we present some contour plots of the equivalent of the position distribution of two polarization pointers (i.e. as a function of  $\theta_1$  and  $\theta_2$ ) in Fig. 5.1. In particular, Fig. 5.1b) and c) correspond to the weak values for the occupation of the two inner arms, and the two outer arms, respectively, in Hardy's Paradox, as discussed in the next chapter. In addition, this Figure demonstrates an intuitive way of understanding our joint weak measurement procedure. Consider a system where the weak values of some arbitrary pair of observables satisfy  $\langle \hat{A}_1 \rangle_w = 1$ ,  $\langle \hat{A}_2 \rangle_w = 1$ ,  $\langle \hat{A}_1 \hat{A}_2 \rangle_w = 1$ . In this case, pointer 1 and pointer 2 always move together and the uncertainty in their difference  $\Delta(\theta_1 - \theta_2)$  does not change, which is what we see in Fig. 5.1c). Conversely,

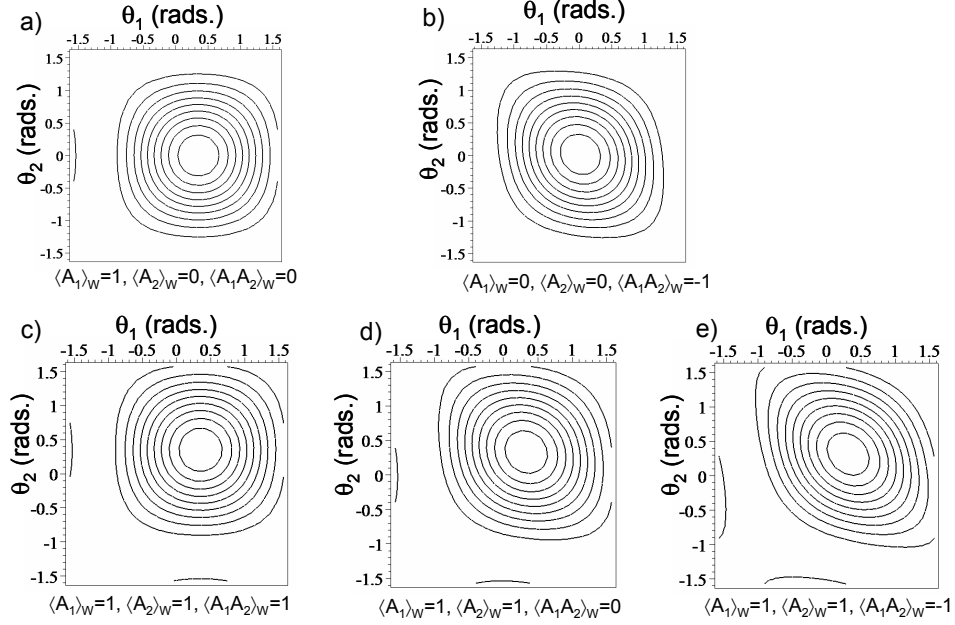


Figure 5.1: Pointer distributions after a weak measurement. These contour plots show the detection probability for the two polarization pointers to be transmitted through polarizers at angles  $\theta_1$  and  $\theta_2$  respectively. The coupling parameter of the weak measurement  $gt = \theta_D = 20^\circ = 0.35 \text{ rad}$  to match the parameters used in the next chapter. In plot a) all the weak values are zero except that of particle one. The resulting pointer distribution is rigidly shifted by 0.35 rad along the  $\theta_1$  axes, indicating a weak value of  $\langle \hat{A}_1 \rangle_W = 1$ . b) When  $\langle \hat{A}_1 \rangle_W \langle \hat{A}_2 \rangle_W > \langle \hat{A}_1 \hat{A}_2 \rangle_W$ , correlations of the form  $\theta_1 = -\theta_2$  are created between the pointer polarizations. Plots c) through d) show that as the joint weak value is decreased while the single particle weak values are kept constant, the correlations become stronger.

if  $\langle \hat{A}_1 \rangle_w = 1$ ,  $\langle \hat{A}_2 \rangle_w = 1$ ,  $\langle \hat{A}_1 \hat{A}_2 \rangle_w = 0$ , then pointer 1 and pointer 2 both move, but never together. This implies that  $\Delta(\theta_1 - \theta_2)$  must increase, which is what we see in Fig. 5.1d). Thus, correlations between the pointers carry the joint weak value. In summary, although the joint weak value formula in Eq. 5.29 might seem strange, it simply quantifies the correlations implied by  $\langle \hat{A}_1 \rangle_w \langle \hat{A}_2 \rangle_w \neq \langle \hat{A}_1 \hat{A}_2 \rangle_w$ .

## 5.4 Conclusion

This new formula for the joint weak value greatly simplifies a recent extension of weak measurement, which makes the experimental investigation of composite or joint observables possible [165][164]. We have shown that when single and joint weak values are expressed as expectation values of annihilation operators, they take on a surprisingly elegant form very similar to that seen in standard strong measurement. This form is easily generalized to any measurement device in which the initial pointer state is the eigenstate of an appropriate lowering operator. With the extension, the weak measurement of joint observables only requires the same apparatus needed to weakly measure each of the component observables separately. Joint observables are central to the detection and utilization of entanglement in multiparticle systems. The weak measurement of these observables should be particularly useful for investigating post-selected systems such as those that have been used to produce novel multiparticle entangled states or those that implement quantum logic gates [21][166][124][167][168][169][170].

## Chapter 6

# Weak measurement and Hardy's Paradox

### 6.1 Introduction

In this chapter, we return to Hardy's Paradox [28] for a closer experimental investigation of the logical statements that lead to it. The essential elements of Hardy's Paradox are three logical statements based on distinct combinations of measurements: When the electron Interaction Free Measurement (IFM) fires (a click at the dark port), the positron was in its inner arm; when the positron IFM fires, the electron was in its inner arm; and the electron and positron never pass through the inner arms at the same time since they annihilate each other. The paradoxical measurement result is that sometimes both IFMs fire, leading us to conclude that both the electron and the positron were in the inner arms, thus contradicting the third logical statement. However, an astute observer (perhaps a philosopher) would view this conclusion and the resulting paradox with suspicion because it arises from counterfactual statements.

A counterfactual statement is a statement or inference based on events that did not actually occur. In this case, when both IFMs are aligned to look for the paradoxical result, we do not simultaneously perform the three other measurements. More-

over, the three logical statements are based on both measurements performed in the past and on other electron positron pairs besides the one detected in the paradoxical result. Including the paradoxical result, there are a total of four measurements, or experiments, one has to perform to demonstrate Hardy's Paradox. These correspond to removing or leaving the final beamsplitter in the electron interferometer, the positron interferometer, or both, or neither. Complementarity ensures that these experiments are mutually exclusive since the IFM mechanism functions on wave-like behavior, whereas particle-like behaviour is inherent in measuring the particle locations. Specifically, when one attempts make one type of measurement, the induced quantum disturbance precludes the measurement of the other. In the classical world counterfactual statements are problematic, but one could conclude that they are even less reliable when pertaining to this quantum system since it is not only impossible to establish them simultaneously in practice, but also in principle.

Nevertheless, the goal of the experiment described in this chapter is to measure which arms the electrons and positrons travelled through in the very cases where they are simultaneously detected at the interferometer dark ports – the paradoxical result. This might seem to contradict the arguments of the last paragraph. The distinction is that, while we measure the particle locations in every trial that leads to a paradoxical result, we only extract the *average* location of the particles. If we perform a large number of trials, we can determine the average location with little error. We obtain minimal information per trial but also cause minimal disturbance. This technique is weak measurement. The advantage of weak measurements is that they do not disturb the measured system nor any other simultaneous weak measurements or subsequent strong measurements, even in the case of non-commuting observables or post-selection. It was on this basis that Aharonov *et al.* used the theory of weak measurement to find the locations of the particles in Hardy's Paradox [33], with subsequent post-selection of only those trials resulting in the paradoxical detections. Using the theoretical foundation and techniques from the last chapter, we perform

the corresponding experiment. It is significant that this experiment follows from what might be the first reaction of a lay-person (or astute observer) when told of Hardy's Paradox and the related problem in the second paragraph; a demand that we reduce the disturbance induced by our measurements so that we *can* perform them simultaneously.

## 6.2 Theory

Aharonov *et al.* calculated the weak value (the weak measurement result) of the occupation of each of the interferometer arms [33]. Specifically, they found the weak value of the following observables:  $\hat{N}(O_+) = |O_+\rangle\langle O_+|$ ,  $\hat{N}(I_+) = |I_+\rangle\langle I_+|$ ,  $\hat{N}(O_-) = |O_-\rangle\langle O_-|$ , and  $\hat{N}(I_-) = |I_-\rangle\langle I_-|$ , where O and I indicate the outer and inner arms, respectively, of the electron (-) and positron (+) Mach-Zehnder interferometers. Aharonov *et al.* first use the standard formula in Eq. 1.44 [94] to calculate the corresponding weak values, in which they set the initial state to be the quantum state after the particle annihilation,

$$|\text{Initial}\rangle = \frac{1}{\sqrt{3}} (|O_-\rangle|O_+\rangle + i|O_-\rangle|I_+\rangle + i|I_-\rangle|O_+\rangle), \quad (6.1)$$

and the final state consists of each pair of interferometer arms in the superposition that leads to a click at the dark port detector,

$$|\text{Final}\rangle = \frac{1}{2} (|O_-\rangle - i|I_-\rangle)(|O_+\rangle - i|I_+\rangle) \quad (6.2)$$

$$\xrightarrow{\text{BS}_2} |D_-\rangle|D_+\rangle. \quad (6.3)$$

By choosing these initial and final states, Aharonov *et al.* account for the evolution of the quantum system before and after the weak measurement.

However, this is not the best way to understand the weak values of the arm occupations. Aharonov *et al.* point out that it is more informative to derive them using the intuitive properties of weak values that we presented in the thesis introduction.



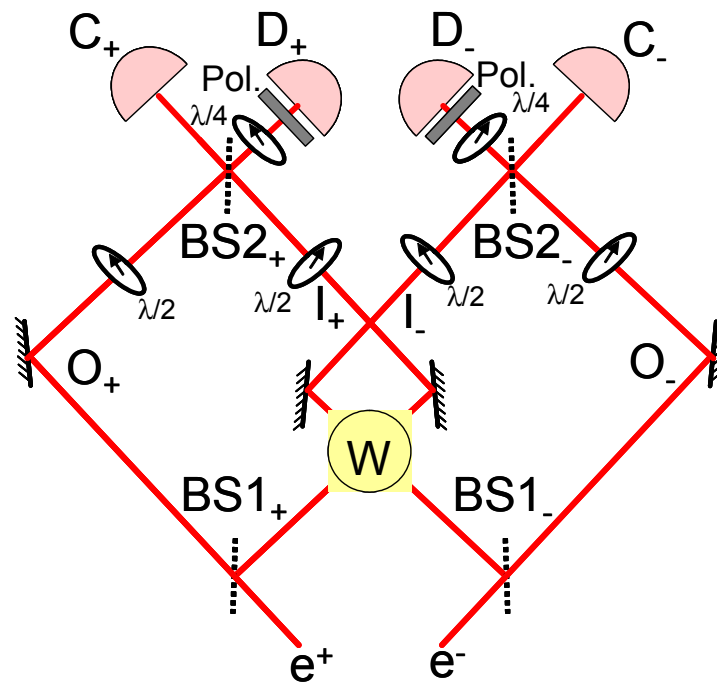


Figure 6.1: Schematic of weak measurement of the arm occupations in Hardy's Paradox using half-waveplates to induce small polarization rotations. Polarization analyzers in front of the dark port detectors, consisting of a quarter-waveplate and a polarizer, are used to find the resulting weak values.

We will follow this route, and include a review of these properties. To begin with, consider a single strong measurement performed in between the preparation of the initial state and the post-selection of the final state. If this measurement were 100% certain to return a particular value, the weak value of the same observable will possess the same value. This property (#2 from the thesis introduction) connects our counterfactual statements with corresponding weak values [171]. For example, if the electron IFM fires then a strong measurement of  $\hat{N}(I_+)$  will result in the eigenvalue of 1 every time, indicating the positron is in the inner arm. This strong measurement collapses the superposition in the positron interferometer, after which there is a 50% chance of a click at the positron dark port, giving a finite probability for post-selection on a detection at both dark ports. It follows that the corresponding weak value,  $\langle \hat{N}(I_+) \rangle_w$ , also equals 1, but in contrast to the strong measurement case, the positron dark port detection is still meaningful and indicates a positive IFM result (and therefore the paradoxical result). Similarly,

$$\langle \hat{N}(O_+) \rangle_w = 0, \langle \hat{N}(I_+) \rangle_w = 1 \quad (6.4)$$

$$\langle \hat{N}(O_-) \rangle_w = 0, \langle \hat{N}(I_-) \rangle_w = 1, \quad (6.5)$$

confirming the IFM results. At this point it might seem that we are done, in that we know the locations of each particle in the case of the paradoxical result. However, like normal expectation values, weak values do not obey the product rule (property #4):  $\langle \hat{N}(I_+) \hat{N}(O_-) \rangle_w \neq \langle \hat{N}(I_+) \rangle_w \langle \hat{N}(O_-) \rangle_w$ . In other words, although we might weakly measure that the positron is in the inner arm, this gives us absolutely no information about the joint location of the particles. Consequently, we must also weakly measure the joint observables  $\hat{N}(OO)$ ,  $\hat{N}(IO)$ ,  $\hat{N}(OI)$ , and  $\hat{N}(II)$ , where  $\hat{N}(JK) \equiv \hat{N}(J_+) \hat{N}(K_-)$ . To find these we once again use property #2: Since a strong measurement will never find the two particles in the inner arms simultaneously,

$$\langle \hat{N}(II) \rangle_w = 0. \quad (6.6)$$

In summary, all three counterfactual statements contributing to Hardy's Paradox are substantiated by corresponding weak measurements. Not only that, but they can be substantiated at the same time and while post-selecting on the paradoxical result.

A third property of weak values, linearity (property #3), allows us to find the remaining joint observables. For example,

$$\hat{N}(IO) + \hat{N}(II) = \hat{N}(I_+) \quad (6.7)$$

$$\langle \hat{N}(IO) \rangle_w + \langle \hat{N}(II) \rangle_w = \langle \hat{N}(I_+) \rangle_w \quad (6.8)$$

$$\langle \hat{N}(IO) \rangle_w + 0 = 1. \quad (6.9)$$

Therefore,

$$\langle \hat{N}(IO) \rangle_w = 1. \quad (6.10)$$

This weak value corroborates the electron IFM results: The positron is in the inner arm and, because it did not annihilate, the other particle must be in the outer arm. But by symmetry, we can exchange the roles of the two particles to find,

$$\langle \hat{N}(OI) \rangle_w = 1, \quad (6.11)$$

seemingly indicating that the particle pair is in two places at the same time. To find the remaining weak value we use linearity once again,

$$\hat{N}(II) + \hat{N}(OI) + \hat{N}(IO) + \hat{N}(OO) = \hat{1} \quad (6.12)$$

$$0 + 1 + 1 + \langle \hat{N}(OO) \rangle_w = 1. \quad (6.13)$$

It follows that the joint weak value for the particle pair to be in the outer arms is,

$$\langle \hat{N}(OO) \rangle_w = -1. \quad (6.14)$$

Given that individually all of the other weak values merely agree with the corresponding counterfactual statements that lead to Hardy's Paradox, it is this strange negative result that resolves the paradox. This point is reinforced by the fact that we can now fill in a classical logic truth table regarding the particle positions:

Weak Value	$N(I_+)$	$N(O_+)$	
$N(I_-)$	0	1	1
$N(O_-)$	1	-1	0
	1	0	

Table 6.1: Truth table for weak values in Hardy's Paradox

The third column and row give the weak values for the individual particle locations and the four middle entries give the weak values for their joint location.

Earlier in this thesis, we showed that weak values of projectors can be interpreted as conditional probabilities. For example,  $\langle \hat{N}(I_-) \rangle_w = P(I_-|D_-D_+)$ ; the conditional probability that the electron was in the inner arm given that both detectors at the dark ports clicked. Furthermore, if a probability is equal to one then we can conclude that the corresponding statement is certain, or always true (e.g.  $P(I_-|D_-D_+) = 1$ , so we say the electron *was* in the inner arm). But how should we interpret a measured probability of -1? The magnitude suggests that this should also correspond to a true statement about the location of the photon pair, but what about the phase? In the thesis introduction, we considered the properties of weakly measured probabilities and showed that they share a number of crucial properties with standard probabilities. Others have more thoroughly considered exotic probabilities that can be negative or complex [81][82][83][84][85][86]. In particular, it has been shown that a violation of Bell's inequality logically implies that certain unmeasured probabilities are negative [80]. The same is true for Hardy's Paradox; the four contributing logical statements, combined with the fact that probabilities for a set of mutually-exclusive events must sum to unity, actually imply that  $P(I_-I_+|D_-D_+) = -1$ . This is the essence of the truth table above. The difference here is that not only is the negative probability implied, but it is also predicted to be the result of a direct measurement.

Another interpretation of the negative weak value follows from an example in the thesis introduction, in which there is a clear meaning for the magnitude and phase of the weak value. A brief review of this example: When post-selecting on a par-

ticular momentum eigenstate  $|p\rangle$ , the weak value of the position projector  $|x\rangle\langle x|$  is proportional to the initial complex-valued spatial wavefunction  $\psi(x)$ . By measuring  $|x\rangle\langle x|$  for all  $x$  we completely determine the wavefunction  $\psi(x)$ , including its phase. An analogous situation occurs in this experiment, although for a two-particle wavefunction and for a discrete two-dimensional Hilbert position space (one dimension per particle). For each particle, we weakly measure an arm occupation such as  $|I_+\rangle\langle I_+|$  (which is like a position projector) and subsequently post-select on a particular state in the complementary basis (on  $|O_-\rangle - i|I_-\rangle$ , the arm superposition that leads to a dark port detection). It follows that our joint weak values for all the two-particle occupations  $\hat{N}(II)$ ,  $\hat{N}(OI)$ ,  $\hat{N}(IO)$ , and  $\hat{N}(OO)$ , completely determine the initial state,  $|\text{initial}\rangle$ . Consequently, the negative sign is directly proportional to the phase in the initial two-particle wavefunction, and in this way it is a characteristic of the initial entanglement in the system. This type of in situ characterization is a possible application for weak measurement in multiparticle entangled states and processes, such as those that occur in quantum information.

In summary, one can interpret the negative weak value in Hardy's Paradox as a quasi-probability or as a direct measurement of the initial wavefunction. But putting interpretations aside, the negative weak value has a very clear *operational* meaning: The theory of weak measurement shows that it comes directly from the response of the measuring apparatus. Specifically, the measurement pointer will be shifted in the opposite direction than in the case of a positive result.

## 6.3 Experiment

There have been two experimental proposals for testing Aharonov *et al.*'s weak value predictions in [33]. Neither has actually been implemented. The first was based on ions, which enabled the measurement of two-particle observables through the ions' mutual repulsion [163]. The second proposal was, like our experiment, based solely on

linear optics [172]. Unfortunately, it was erroneous in a few ways and consequently, offered no method for measuring the joint weak values that Aharonov *et al.* predicted [173][174]. Our experiment builds on the implementation of Hardy's Paradox described in Chapter 3. We make a small switch in notation here to avoid confusion later. We use a horizontal (E) photon and a vertical (P) photon in the place of the electron and positron from the introduction ( $+ \rightarrow P$ ,  $- \rightarrow E$ ). The primary addition to the previous experimental setup is the weak measurement mechanism. The horizontal and vertical interferometers overlap until the arms encounter polarizing beamsplitters (PBS). After this point each of the two photons is in its own spatially separated interferometer, freeing up the polarization to be used as our pointer for the measurement of the occupation of a single arm. For example, consider the placement of a half-waveplate in one of the outer arms, aligned so as to flip the polarization of a photon passing through it. The polarization of the photon arriving at the dark port detector then precisely indicates whether it was in the outer path or not. At the same time, the polarization rotation completely destroys the interference in the interferometer. To avoid this radical disturbance, rather than flipping the polarization, we rotate it by only  $20^\circ$ . The cost is that we can no longer tell which arm a particular photon detected at the dark port went through with certainty. Instead, we measure the average polarization rotation at the detector over many trials (i.e. for many detected photons) to find what fraction of photons passed through that particular arm. If no rotation is observed then we conclude that there was zero probability of the photon having been in the arm with the waveplate. Conversely, if we measure an average rotation of  $20^\circ$ , we conclude that every photon passed through the waveplate. This procedure is a single-particle weak measurement of the occupation of a particular interferometer arm.

We are also interested in the simultaneous location of the photons. To measure this we need to perform a two-particle measurement, which, in the von Neumann measurement model, requires a nonlinear interaction at the single-particle level (see

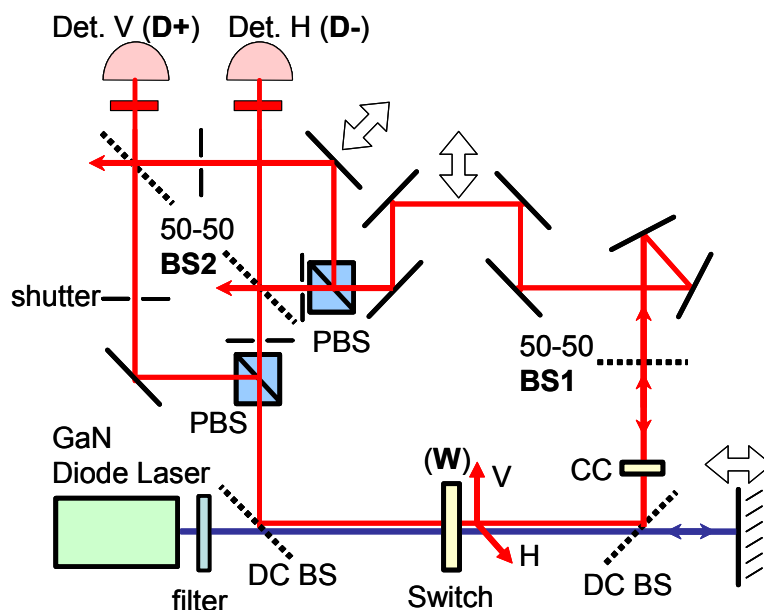


Figure 6.2: Experimental setup for weak measurements in Hardy’s Paradox. The basic setup is the same as in Chapter 3. Modifications are shown in Fig. 6.3.

Chapter 5). This type of interaction is feasible (but still difficult) for particles such as ions, which have a natural Coulomb interaction. This is the basis of proposal by Klaus Mølmer in Ref. [163], that as of yet has not been implemented. On the other hand, in our optical setup a Kerr-type nonlinearity could be used to rotate a “read-out” polarization by  $20^\circ$  if, and only if, the E and P photon simultaneously travelled through a particular pair of arms. Unfortunately, even in nonlinear crystals single photons have insignificant nonlinear interactions and so a strategy without a strong nonlinearity is required. In the last chapter, we showed that if one performed the simple single-particle weak measurement procedure described above for both the E and P photons simultaneously, specific correlations in the two polarization rotations are equal, in general, to the results of the corresponding two-particle weak measurement. With this measurement strategy, we only require linear optics, namely waveplates,

and the ability to measure the following polarization correlation:

$$\text{Re} \langle \hat{\sigma}_{Ez}^- \hat{\sigma}_{Pz}^- \rangle = \langle \hat{\sigma}_{Ex} \hat{\sigma}_{Px} \rangle - \langle \hat{\sigma}_{Ey} \hat{\sigma}_{Py} \rangle, \quad (6.15)$$

where  $\hat{\sigma}_{jz}^- = (\hat{\sigma}_{jx} - i\hat{\sigma}_{jz})$  is the  $z$ -basis lowering operator for the  $j = E$  or  $P$  photons. All four arms contain half-waveplates, two of which we rotate by  $10^\circ$  (ideally inducing a  $20^\circ$  rotation) in order to weakly measure the joint occupation of those two arms. At both dark port detectors we have polarization analyzers, consisting of a quarter-waveplate and a polarizer. These allow us to measure  $\langle \hat{\sigma}_{Ex} \rangle$  and  $\langle \hat{\sigma}_{Px} \rangle$ , which we use to quantify the average polarization rotations, as well as the correlations specified in Eq. 6.15. Due to polarization-dependent losses and residual polarization phase-shifts, the actual rotation each waveplate induced varied slightly from  $20^\circ$ . We calibrated the waveplates in situ by independently measuring the rotation,  $\langle \hat{\sigma}_{Ex} \rangle$ , induced by each (e.g.  $\langle \hat{\sigma}_{Ex} \rangle_{IE}$  indicates the calibration for the waveplate in the inner arm of the E photon interferometer). Another technical complication is that the final non-polarizing beamsplitters (NPBS) in the two interferometers create extraneous polarization phase-shifts. To compensate for this we also added carefully angled quartz pieces before and after the NPBSs.

As in all weak measurements, the signal to noise ratio is purposefully small and requires a large ensemble of photon pairs. In our joint weak measurement procedure, this is even more of a problem since the signal is now proportional to the coupling squared. The main purpose of the interferometer stabilization system, described in the Chapter 3, is to allow for a long period of data accumulation in order to counter this problem. When measuring the occupation of a particular pair of arms, we first minimize each of the interferometers, including the switch, and lock them at that point. This sets up the system so that dark ports will now be looking at the paradoxical result. Then with the E and P polarization analyzers in one of two positions,  $+45^\circ$  or  $-45^\circ$ , we measure the four coincidence rates,  $R_{P+E+}$ ,  $R_{P-E+}$ ,  $R_{P+E-}$ , and  $R_{P-E-}$ ,



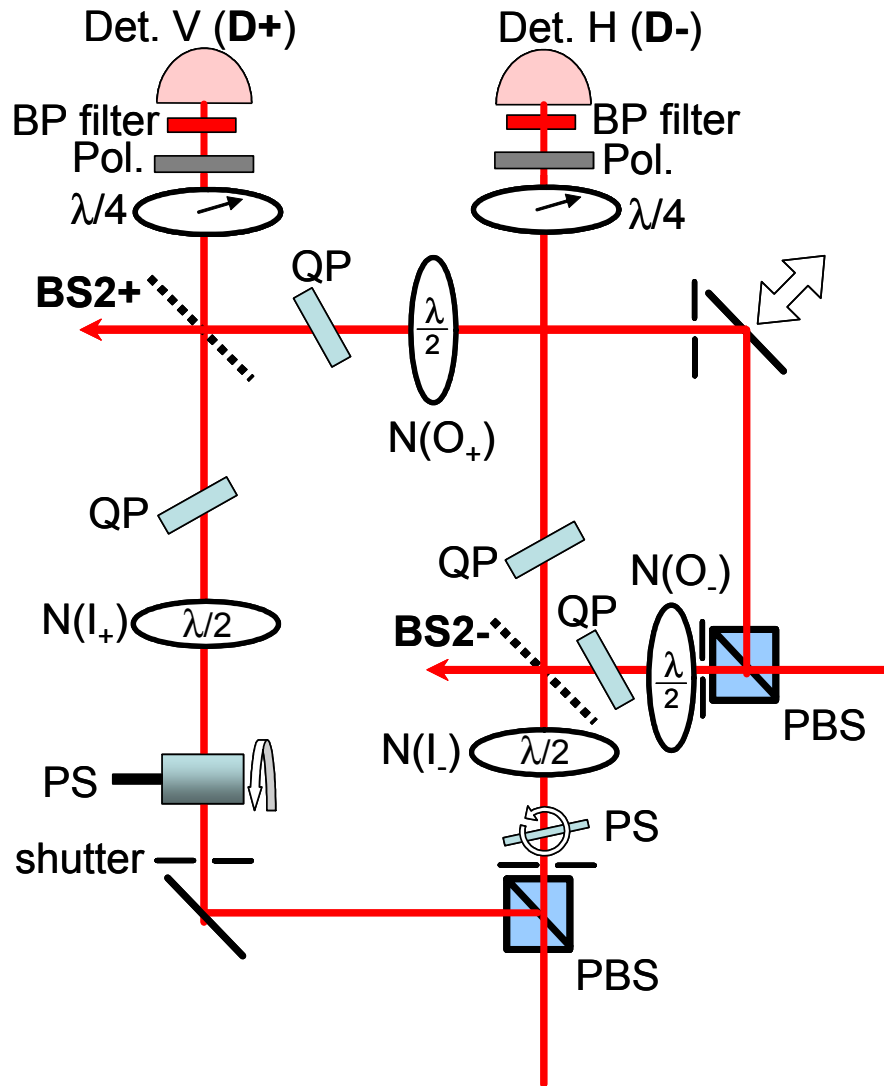


Figure 6.3: A enlarged version of part of the experimental setup in Fig. 6.2. Weak measurements of the arm occupations  $\hat{N}(I(O)_{+(-)})$  using half-waveplates. Quartz pieces (QP) are used to correct for phase-shifts at the the final beamsplitters (BS2). The interferometers are stabilized by rotating thin glass pieces (PS), described later.

Coincidence Counts	$P+$	$P-$	Sum
$E+$	2261	772	3033
$E-$	746	115	861
Sum	3007	887	

Table 6.2: Raw data for the weak measurement of the occupation of the inner arms

that are necessary to calculate the various expectation values. For example,

$$\langle \hat{\sigma}_{Ex} \rangle = \frac{R_{P+E+} + R_{P-E+} - R_{P+E-} - R_{P-E-}}{R_{P+E+} + R_{P-E+} + R_{P+E-} + R_{P-E-}}, \quad (6.16)$$

and,

$$\langle \hat{\sigma}_{Ex} \hat{\sigma}_{Px} \rangle = \frac{R_{P+E+} + R_{P-E-} - R_{P-E+} - R_{P+E-}}{R_{P+E+} + R_{P-E+} + R_{P+E-} + R_{P-E-}}. \quad (6.17)$$

For every pair of arms, we measure the set of rates four times, to reduce any possible bias from either any drift of the interferometer locks or from pump laser power. To find  $\langle \hat{\sigma}_{Ey} \hat{\sigma}_{Py} \rangle$  we use the quarter-waveplate in the analyzer to measure the analogous rates for the circular polarization basis. We repeat procedure for all four pairs of arms. The weak values are calculated from the measured polarization expectation values by normalizing by the induced polarization rotation. For example, for the E photon inner arm,

$$\langle \hat{N}(I_E) \rangle_W = \frac{\langle \hat{\sigma}_{Ex} \rangle}{\langle \hat{\sigma}_x \rangle_{IE}}, \quad (6.18)$$

and the joint weak value to find both particles in the inner arms is,

$$\langle \hat{N}(II) \rangle_W = \frac{\langle \hat{\sigma}_{Ex} \hat{\sigma}_{Px} \rangle - \langle \hat{\sigma}_{Ey} \hat{\sigma}_{Py} \rangle}{\langle \hat{\sigma}_x \rangle_{IE} \langle \hat{\sigma}_x \rangle_{IP}}. \quad (6.19)$$

## 6.4 Results

We first examine the raw data and discuss what it suggests about the veracity of the logical statements in Hardy's Paradox. In Table 6.2, we present the four coincidence counts recorded over 420s each in the weak measurement of the occupation of both inner arms.

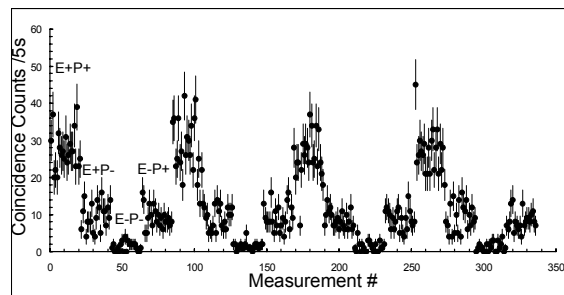


Figure 6.4: Raw data for the weak measurement of the occupation of the inner arms. The four polarizer settings are cycled through repeatedly to minimize the effect of drift on the results. Data is taken in 5s intervals, in between which, the interferometer is stabilized.

Coincidence Counts	$P+$	$P-$	Sum
$E+$	556	834	1390
$E-$	730	583	1313
Sum	1286	1417	

Table 6.3: Raw data for the weak measurement of the occupation of the outer arms

The sum of the counts in each row or column is given in the entries at the right side or bottom of the table, respectively. These sums are equal to what would be detected at one dark port if an analyzer were not present at the other dark port detector. Consider these sums for the P photon interferometer. Since there are roughly three times more counts at P+ than at P-, we can say the P photons experienced a large average rotation. Similar analysis reveals the photons at E also experienced a large rotation. So far, this is consistent with the IFM results; the photons emerged rotated so they must have been in the inner arms.

In Table 6.3, we present the analogous coincidence rates for the weak measurement of the occupation of both outer arms.

If we once again look at the summed coincidences for P, disregarding the position of E's analyzer, we see the rates are roughly equal, indicating no polarization rotation. The same is true for E. Thus, both measurements indicate that each photon was not

Weak Value	$N(I_P)$	$N(O_P)$	
$N(I_E)$	$0.243 \pm 0.068$	$0.663 \pm 0.083$	$0.882 \pm 0.015$
$N(O_E)$	$0.721 \pm 0.074$	$-0.758 \pm 0.083$	$0.087 \pm 0.021$
	$0.925 \pm 0.024$	$-0.039 \pm 0.023$	

Table 6.4: A truth table containing the joint and individual experimental weak values in Hardy’s Paradox

in its respective outer arm, in agreement with the two IFM results. Since neither photon was rotated, one naturally expects  $\langle \hat{\sigma}_{Ex} \hat{\sigma}_{Ex} \rangle$  should be zero. Surprisingly, it is instead negative, as if both photons *were* rotated. Moreover, one of them had to have been rotated in the negative direction, which is at odds with the fact that both waveplates are aligned to only induce a rotation in the positive direction.

In Table 6.4, we present the product of the full analysis, the weak measurement results for the various arm occupations. Note that the final weak values are calculated for each of the four times the set of rates is measured and then these four results are averaged. The error is a quadratic sum of the standard deviation and error to due to counting statistics. The weakly measured occupation of each of the inner arms is close to one, and close to zero for each of the outer arms, verifying the IFM results presented in the Hardy’s Paradox chapter. Unlike in the ideal setup considered by Aharonov, roughly 15% of the photon pairs are not annihilated in our switch effect. Interpreting it as a probability, the joint weak value for the inner arms reflects this by indicating the photon pair was present  $24\% \pm 7\%$  of the time. Furthermore, the joint weak value for the outer arms is negative, as is necessary to resolve the paradox. Unfortunately, the operational meaning of this negative value is less clear with our joint weak measurement strategy than with a direct two-particle von Neumann measurement. For example, for a joint weak value of negative one, the Kerr nonlinearity mentioned in the experimental section of this chapter would result in a “readout” polarization rotated in the negative direction by twenty degrees, whereas for our strategy only certain correlations are negative. Nonetheless, the negative weak

value we measure ensures that, within the measurement error, the weakly measured joint occupations sum to give the single arm occupations in the outer cells. In this sense, the weak measurement results appear to not only verify Hardy's Paradox, but also to provide it with a consistent resolution.

## 6.5 Conclusion

The heart of Hardy's Paradox is the conflicting classical logic statements about the location of the particle in each of two Mach-Zehnder interferometers. It is impossible to simultaneously verify these statements with typical measurements, since testing one statement disturbs the system and consequently negates the others. We attempt to minimize this disturbance by minimizing the interaction used to perform the measurement. The results of these weak measurements indicate that each of the logical statements are correct. The results also provide a self-consistent, if unusual, resolution to the paradox. Since they do not disturb subsequent post-selection of the systems under study, weak measurements are ideal for the interrogation and characterization of post-selected multiparticle states, such as GHZ or Cluster states, and processes such as Linear-Optics Quantum Computation. This experiment demonstrates a new technique that, for the first time, allows for the weak measurement of general multiparticle observables in these systems.

# Chapter 7

## Conclusion

The field of quantum optics is advancing rapidly, partly due to the drive to create systems and processes suitable for quantum information processing, and partly due to new insights from the theoretical study of quantum information. In the past few years, entangled systems of six, ten, and even one hundred photons have been reported in the literature [175][176][177], and LOQC-style transformations are now a common way to create novel entangled states [168][169][166]. Most experiments in photonic quantum information still use downconversion as a source of photons. But a number of alternatives are becoming viable, such as correlated Stokes and anti-Stokes radiation from atomic clouds [178][179], single atoms in cavities [180][181], and quantum dots [182]. Many significant goals are within reach in the immediate future. The complete KLM proposal has not yet been implemented, and a demonstration of a basic quantum computing protocol with uses beyond the realm of physics would be an important milestone.

Through a series of experiments, this thesis outlines several contributions to the field of quantum optics and its applications. In Chapter 2, in an attempt to investigate a proposal for strong optical nonlinearities in an atomic medium, we demonstrated effective nonlinear absorption between two photons. This photon-exchange effect, an example of a LOQC-style transformation, could be used in high-dimensional quan-

tum information systems, such as a qudit based on frequency or time bins of single photons. In Chapter 3 we described an implementation of Hardy's Paradox, a logical contradiction that arises from relying on the results of indirect measurements, namely interaction-free measurements (IFMs). In this implementation, we demonstrated a novel version of an IFM in which we detected the presence of a photon (only a macroscopic object had been detected in earlier attempts). The experiment was the first application of our innovative single-photon-level absorptive switch, which functions via quantum interference and post-selection.

In Chapter 4, we investigated the momentum transfer induced by a which-way measurement in Young's double-slit experiment. We used weak measurement to record the initial momentum while post-selecting on a final momentum. This experiment demonstrates how weak measurement can be used to characterize quantum processes in situ, without disruption. In Chapter 5, we theoretically extended the weak measurement technique to multiparticle systems, which will allow further research into the large entangled systems now demonstrated in quantum optics. A first application of this theory is given in Chapter 6, where we used the multiparticle weak measurement technique to test the contradictory statements in Hardy's Paradox. This is the first example of weak measurement of an entangled system, and the results provide a consistent resolution to the logical contradiction while characterizing the entangled state created in the paradox.

Quantum information is moving toward more and more complicated quantum systems. In optics, it appears that post-selection-driven quantum information processing will be key to advances in the production of these systems. The research presented in this thesis has demonstrated post-selection-based devices that could play a role in future schemes, and constitutes the first steps toward the theoretical and experimental development of weak measurement as a tool for investigating quantum systems.

# Bibliography

- [1] J. F. Clauser and A. Shimony, “Bell’s theorem: experimental tests and implications,” *Rep. Prog. Phys.* **41**, 1881 (1978).
- [2] A. Aspect, P. Grangier, and G. Roger, “Experimental tests of realistic local theories via Bell’s theorem,” *Physical Review Letters* **47**, 460–463 (1981).
- [3] R. P. Feynman, “Simulating physics with computers,” *Int. J. Theor. Phys.* **21**, 467–488 (1982).
- [4] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* pp. 175–179 (1984).
- [5] D. Deutsch, “Quantum Theory, the Church-Turing principle and the universal quantum computer,” *Proceedings of The Royal Society of London, Series A: Mathematical and Physical Sciences* **400**, 97–117 (1985).
- [6] D. Deutsch and R. Jozsa, “Rapid solution of problems by quantum computation,” *Proc. R. Soc. London, Ser. A* **439**, 553–558 (1992).
- [7] P. W. Shor, “Algorithms for quantum computation: Discrete logarithms and factoring,” *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science* pp. 124–134 (1994).



- [8] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels,” *Physical Review Letters* **70**, 1895–1899 (1993).
- [9] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, “Experimental quantum teleportation,” *Nature* **390**, 575–579 (1997).
- [10] K. Mattle, H. Weinfurter, P. Kwiat, and A. Zeilinger, “Dense coding in experimental quantum communication,” *Physical Review Letters* **76**, 4656–4659 (1996).
- [11] D. P. DiVincenzo, “The physical implementation of quantum computation,” *Fortschritte der Physik* **48**, 771–783 (2000).
- [12] G. J. Milburn, “Quantum optical Fredkin gate,” *Physical Review Letters* **62**, 2124–2127 (1989).
- [13] A. L. Gaeta and R. W. Boyd, in *Atomic, Molecular and Optical Physics Handbook* (1996), p. 809.
- [14] J.-W. Pan, M. Daniell, A. Zellinger, D. Bouwmeester, and H. Weinfurter, “Experimental test of quantum nonlocality in three-photon Greenberger-Horne-Zeilinger entanglement,” *Nature* **403**, 515–519 (2000).
- [15] J.-W. Pan, M. Daniell, S. Gasparoni, G. Weihs, and A. Zeilinger, “Experimental Demonstration of Four-Photon Entanglement and High-Fidelity Teleportation,” *Physical Review Letters* **86**, 4435–4438. (2001).
- [16] Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble, “Measurement of conditional phase shifts for quantum logic,” *Physical Review Letters* **75**, 4710–4713 (1995).

- [17] A. Rauschenbeutel, G. Nogues, S. Osnaghi, P. Bertet, M. Brune, J. M. Raimond, and S. Haroche, “Coherent Operation of a Tunable Quantum Phase Gate in Cavity QED,” *Physical Review Letters* **83**, 5166–5169 (1999).
- [18] S. E. Harris and L. V. Hau, “Nonlinear Optics at Low Light Levels,” *Physical Review Letters* **82**, 4611–4614 (1999).
- [19] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, “Nonlinear Optics with Less Than One Photon,” *Physical Review Letters* **87**, 123603 (2001).
- [20] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, “Conditional-Phase Switch at the Single-Photon Level,” *Physical Review Letters* **89**, 037904 (2002).
- [21] E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics,” *Nature* **409**, 46–52 (2001).
- [22] R. Raussendorf and H. J. Briegel, “The one-way quantum computer,” *Physical Review Letters* **85**, 910–913 (2001).
- [23] M. A. Nielson, “Cluster-state quantum computation,” [quant-ph/0504097](https://arxiv.org/abs/quant-ph/0504097) (2005).
- [24] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (2000).
- [25] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?,” *Physical Review* **47**, 777 (1935).
- [26] D. M. Greenberger, M. Horne, and A. Zeilinger, in *Bell’s theorem, quantum theory, and conceptions of the universe*, M. Kafatos, ed., (Kluwer, Dordrecht, 1989), pp. 69–72.
- [27] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics* **1**, 195 (1964).

- [28] L. Hardy, “Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories,” *Physical Review Letters* **68**, 2981–2984 (1992).
- [29] Y. Aharonov, D. Z. Albert, and L. 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,” *Physical Review Letters* **60**, 1351–1354 (1988).
- [30] M. O. Scully, B.-G. Englert, and H. Walther, “Quantum optical tests of complementarity,” *Nature* **351**, 111–116 (1991).
- [31] E. P. Storey, S. M. Tan, M. J. Collett, and D. F. Walls, “Path detection and the uncertainty principle,” *Nature* **367**, 626–628 (1994).
- [32] H. M. Wiseman, “Directly observing momentum transfer in twin-slit "which-way" experiments,” *Physics Letters A* **311**, 285–291 (2003).
- [33] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, “Revisiting Hardy’s paradox: Counterfactual statements, real measurements, entanglement and weak values,” *Physics Letters A* **301**, 130–138 (2002).
- [34] B. Y. Zel’dovich and D. N. Klyshko, “Field statistics in parametric luminescence,” *Soviet Physics: Letters to the Editor* **9**, 40 (1969).
- [35] D. Burnham and D. Weinberg, “Observation of Simultaneity in Parametric Production of Optical Photon Pairs,” *Physical Review Letters* **25**, 84 (1970).
- [36] R. Boyd, *Nonlinear Optics* (Academic Press, New York, 2003).
- [37] M. H. Rubin, D. N. Klyshko, Y. H. Shih, and A. V. Sergienko, “Theory of two-photon entanglement in type-II optical parametric down-conversion,” *Physical Review A* **50**, 5122 (1994).
- [38] C. Hong and L. Mandel, “Experimental realization of a localized one-photon state,” *Physical Review Letters* **56**, 58 (1986).

- [39] P. G. Kwiat, P. H. Eberhard, A. M. Steinberg, and R. Y. Chiao, “Proposal for a loophole-free Bell inequality experiment,” *Physical Review A* **49**, 3209–3220 (1994).
- [40] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Z. A. V. Sergienko, and Y. Shih, “New high-intensity source of polarization-entangled photon pairs,” *Physical Review Letters* **75**, 4337 (1995).
- [41] P. G. Kwiat, E. Waks, A. G. White, and P. H. Eberhard, “Ultrabright source of polarization-entangled photons,” *Physical Review A* **60**, R773 (1999).
- [42] S. J. Freedman and J. F. Clauser, “Experimental test of local hidden-variable theories,” *Physical Review Letters* **28**, 938 (1972).
- [43] P. Kok, H. Lee, and J. P. Dowling, “Single-photon quantum-nondemolition detectors constructed with linear optics and projective measurements,” *Physical Review A* **66**, 063814 (2002).
- [44] W. J. Munro, K. Nemoto, R. G. Beausoleil, and T. P. Spiller, “A high-efficiency quantum non-demolition single photon number resolving detector,” *Physical Review A* **71**, 033819 (2005).
- [45] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin, “Experimental entanglement distillation and ‘hidden’ non-locality,” *Nature* **409**, 1014–1017 (2001).
- [46] C. K. Hong, Z. Y. Ou, and L. Mandel, “Measurement of subpicosecond time intervals between two photons by interference,” *Physical Review Letters* **59**, 2044–2046 (1987).
- [47] Y. H. Shih and C. O. Alley, “A new type of Einstein-Podolsky-Rosen-Bohm experiment using pairs of light quanta produced by optical parametric down-conversion,” *Physical Review Letters* **61**, 2921 (1988).

- [48] N. Lutkenhaus, J. Calsamiglia, and K. A. Suominen, “Bell measurements for teleportation,” *Physical Review A* **59**, 3295 (1999).
- [49] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, “Linear optical quantum computing,” [quant-ph/0512071](https://arxiv.org/abs/quant-ph/0512071) (2005).
- [50] Y. Aharonov, D. Z. Albert, A. Casher, and L. Vaidman, “Surprising quantum effects,” *Physics Letters A* **124**, 199–203 (1987).
- [51] G. A. Smith, S. Chaudhury, A. Silberfarb, I. H. Deutsch, and P. S. Jessen, “Continuous Weak Measurement and Nonlinear Dynamics in a Cold Spin Ensemble,” *Physical Review Letters* **93**, 163602 (2004).
- [52] N. Brunner, A. Acin, D. Collins, N. Gisin, and V. Scarani, “Optical Telecom Networks as Weak Quantum Measurements with Postselection,” *Physical Review Letters* **91**, 180402 (2003).
- [53] N. Brunner, V. Scarani, M. Wegmuller, M. Legre, and N. Gisin, “Direct measurement of superluminal group velocity and of signal velocity in an optical fiber,” *Physical Review Letters* **93**, 203902 (2004).
- [54] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, “Fast Light, Slow Light, and Phase Singularities: A Connection to Generalized Weak Values,” *Physical Review Letters* **92**, 043601 (2004).
- [55] H. M. Wiseman, “Weak values, quantum trajectories, and the cavity-QED experiment on wave-particle correlation,” *Physical Review A* **65**, 032111 (2002).
- [56] G. T. Foster, L. A. Orozco, H. M. Castro-Beltran, and H. J. Carmichael, “Quantum state reduction and conditional time evolution of wave-particle correlations in cavity QED,” *Physical Review Letters* **85**, 3149–3152 (2000).
- [57] A. M. Steinberg, “Can a falling tree make a noise in two forests at the same time?,” *Causality and Locality in Modern Physics* pp. 431– (1997).

- [58] A. M. Steinberg, “How much time does a tunneling particle spend in the barrier region?,” *Physical Review Letters* **74**, 2405–2409 (1995).
- [59] A. M. Steinberg, “Conditional probabilities in quantum theory and the tunneling-time controversy,” *Physical Review A* **52**, 32–42 (1995).
- [60] A. M. Steinberg, S. Myrskog, H. S. Moon, H. A. Kim, J. B. Kim, and J. Fox, “An atom optics experiment to investigate faster-than-light tunneling,” *Annalen der Physik (Leipzig)* **7**, 593–601 (1998).
- [61] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, 1955).
- [62] A. Peres, “Quantum measurements with postselection,” *Physical Review Letters* **62**, 2326–2326 (1989).
- [63] A. J. Leggett, “Comment on “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100”,” *Physical Review Letters* **62**, 2325–2325 (1989).
- [64] Y. Aharonov and L. Vaidman, “Aharonov and Vaidman reply,” *Physical Review Letters* **62**, 2327–2327 (1989).
- [65] L. Vaidman, in *Quantum Interferometry*, F. de Martini, G. Denardo, and Y. H. Shih, eds., (VCH Publishers, 1996), quant-ph/9607023.
- [66] L. M. Johansen, “Weak measurements with arbitrary probe states,” *Physical Review Letters* **93**, 120402 (2004).
- [67] G. J. Pryde, J. L. O’Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, “Measurement of Quantum Weak Values of Photon Polarization,” *Physical Review Letters* **94**, 220405 (2005).

- [68] Y. Aharonov, S. Massar, S. Popescu, J. Tollaksen, and L. V. L. Vaidman, “Adiabatic measurements on metastable systems,” *Physical Review Letters* **77**, 983–987 (1996).
- [69] O. Oreshkov and T. A. Brun, “Weak measurements are universal,” *Physical Review Letters* **95**, 110409 (2005).
- [70] Y. Aharonov and L. Vaidman, “Complete description of a quantum system at a given time,” *Journal of Physics A* **24**, 2315–2328 (1991).
- [71] Y. Aharonov and A. Botero, “Quantum averages of weak values,” *Physical Review A* **72**, 0521111 (2005).
- [72] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, “Realization of a measurement of a “weak value”,” *Physical Review Letters* **66**, 1107–1110 (1991).
- [73] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, “The sense in which a “weak measurement” of a spin-1/2 particle’s spin component yields a value 100,” *Physical Review D* **40**, 2112–2117 (1989).
- [74] P. Balcou and L. Dutriaux, “Dual optical tunneling times in Frustrated total internal reflection,” *Physical Review Letters* **78**, 000851 (1997).
- [75] Y. Aharonov and E. Y. Gruss, “Two-time interpretation of quantum mechanics,” [quant-ph/0507269](https://arxiv.org/abs/quant-ph/0507269) (2005).
- [76] L. Vaidman, “Weak-Measurement Elements of Reality,” *Foundations of Physics* **26**, 895–906 (1996).
- [77] R. E. Kastner, “Weak values and consistent histories in quantum theory,” *Studies in History and Philosophy of Modern Physics* **35B**, 57–71 (2004).
- [78] J. A. Wheeler and W. H. Zurek, *Quantum theory and measurement* (Princeton Univ. Press, 1983).

- [79] R. Feynman, in *Quantum Implications*, B. J. Hiley and F. D. Peat, eds., (Routledge and Kegan Paul, 1987).
- [80] M. O. Scully, H. Walther, and W. Schleich, “Feynman’s approach to negative probability in quantum mechanics,” *Physical Review A* **49**, 1562–1566 (1994).
- [81] W. Muckenheim, G. Ludwig, C. Dewdney, P. R. Holland, A. Kyprianidis, J. P. Vigiier, N. C. Petroni, M. S. Bartlett, and E. T. Jaynes, “A review of extend probabilities,” *Physics Reports* **133**, 337–401 (1986).
- [82] Y. D. Han, W. Y. Hwang, and I. G. Koh, “Explicit solutions for negative-probability measures for all entangled states,” *Physics Letters A* **221**, 283–286 (1996).
- [83] S. Youssef, “Physics with exotic probability theory,” hep-th/0110253 (2001).
- [84] S. Youssef, “Is complex probability theory consistent with bell’s theorem?,” *Physics Letters A* **204**, 181–187 (1995).
- [85] S. Youssef, “Quantum mechanics as complex probality theory,” *Modern Physics Letters* **A9**, 2571–2586 (1994).
- [86] A. Caticha, “Consistency, amplitudes, and probabilities in quantum theory,” *Physical Review A* **57**, 1572–1582 (1998).
- [87] E. P. Wigner, “On the quantum correction for thermodynamic equilibrium,” *Physical Review* **40**, 749–759 (1932).
- [88] M. Hillery, R. F. O’Connell, M. O. Scully, and E. P. Wigner, “Distribution functions in physics: fundamentals,” *Physics Reports* **106**, 121–167 (1984).
- [89] D. Fromkin, *Europe’s last summer: Who started the Great War in 1914?* (Knopf, New York, 2004).



- [90] H. P. Stapp, “Nonlocal character of quantum theory,” *American Journal of Physics* **65**, 300–304 (1997).
- [91] H. P. Stapp, “Meaning of counterfactual statements in quantum physics,” *American Journal of Physics* **66**, 924–926 (1998).
- [92] W. Unruh, “Nonlocality, counterfactuals, and quantum mechanics,” *Physical Review A* **59**, 126–130 (1999).
- [93] L. M. Johansen, “What is the value of an observable between pre- and postselection?,” *Physics Letters A* **322**, 298–300 (2004).
- [94] Y. Aharonov and L. Vaidman, “Properties of a quantum system during the time interval between two measurements,” *Physical Review A* **41**, 11–20 (1990).
- [95] J. D. Franson, “Cooperative enhancement of optical quantum gates,” *Physical Review Letters* **78**, 3852–3855 (1997).
- [96] J. D. Franson and T. B. Pittman, “Nonlocality in quantum computing,” *Fortschritte der Physik* **46**, 697–705 (1998).
- [97] J. D. Franson and T. B. Pittman, “Quantum logic operations based on photon-exchange interactions,” *Physical Review A* **60**, 917–936 (1999).
- [98] T. Opatrny and G. Kurizki, “On the possibility of quantum computation based on photon exchange interactions,” *Fortschritte der Physik* **48**, 1125–1131 (2000).
- [99] J. D. Franson, “Reply to a review of photon-exchange interactions by Opatrny and Kurizki,” *Fortschritte der Physik* **48**, 1133–1138 (2000).
- [100] K. Resch, J. Lundeen, and A. Steinberg, “Electromagnetically induced opacity for photon pairs,” *Journal of Modern Optics* **49**, 487 (2002).

- [101] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, in *The Physics of Communication*, Antoniou, Sadovnichy, and Walther, eds., (World Scientific, 2003), pp. 437–451, quant-ph/0302003.
- [102] L. V. Hau, Z. Dutton, C. H. Behroozi, and S. E. Harris, “Light speed reduction to 17 metres per second in an ultracold atomic gas,” *Nature* **397**, 594–598 (1999).
- [103] M. M. Kash, V. A. Sautenkov, A. S. Zibrov, G. R. Welch, Y. Rostovtsev, E. S. Fry, M. O. Scully, L. Hollberg, and M. D. Lukin, “Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas,” *Physical Review Letters* **82**, 5229–5232 (1999).
- [104] T. C. Ralph, A. G. White, W. J. Munro, and G. J. Milburn, “Simple scheme for efficient linear optics quantum gates,” *Physical Review A* **65**, 012314 (2002).
- [105] G. G. Lapaire, P. Kok, J. P. Dowling, and J. E. Sipe, “Conditional linear-optical measurement schemes generate effective photon nonlinearities,” *Physical Review A* **68**, 042314 (2003).
- [106] W. J. Mullin and G. Blaylock, “Quantum statistics: Is there an effective fermion repulsion or boson attraction?,” *American Journal of Physics* **71**, 1223 (2003).
- [107] G. D. Giuseppe, L. Haiberger, F. D. Martini, and A. V. Sergienko, “Quantum interference and indistinguishability with femtosecond pulses,” *Physical Review A* **56**, R21–R24 (1997).
- [108] J. H. Shapiro and N. C. Wong, “An ultrabright narrowband source of polarization-entangled photon pairs,” *Journal of Optics B* **2**, L1–L4 (2000).
- [109] Y. J. Lu and Z. Y. Ou, “Optical parametric oscillator far below threshold: experiment versus theory,” *Physical Review A* **62**, 033804–033801 (2000).

- [110] R. J. Glauber, "Optical coherence and photon statistics," *Quantum Optics and Electronics* pp. 63–185 (1965).
- [111] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, "Experimental observation of nonclassical effects on single-photon detection rates," *Physical Review A* **63**, 020102 (2001).
- [112] W. P. Grice, R. Erdmann, I. A. Walmsley, and D. Branning, "Spectral distinguishability in ultrafast parametric down-conversion," *Physical Review A* **57**, 2289 (1998).
- [113] M. Atature, A. V. Sergienko, B. M. Jost, B. E. A. Saleh, and M. C. Teich, "Partial Distinguishability in Femtosecond Optical Spontaneous Parametric Down-Conversion," *Physical Review Letters* **83**, 1323–1326 (1999).
- [114] D. Strekalov, A. Sergienko, D. Klyshko, and Y. Shih, "Observation of two-photon "ghost" interference and diffraction," *Physical Review Letters* **74**, 3600–3603 (1995).
- [115] M. D'Angelo, M. Chekhova, and Y. Shih, "Two-photon diffraction and quantum lithography," *Physical Review Letters* **87**, 013602 (2001).
- [116] R. Bennink, S. Bentley, and R. Boyd, "'Two-photon" coincidence imaging with a classical source," *Physical Review Letters* **89**, 113601 (2002).
- [117] A. F. Abouraddy, K. C. Toussaint, A. V. Sergienko, B. E. A. Saleh, and M. C. Teich, "Ellipsometric measurements by use of photon pairs generated by spontaneous parametric downconversion," *Optics Letters* **26**, 1717–1719 (2001).
- [118] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, "Ancilla-assisted quantum process tomography," *Physical Review Letters* **90**, 193601 (2003).

- [119] M. B. Nasr, B. E. A. Saleh, A. V. Sergienko, and M. C. Teich, “Demonstration of dispersion-canceled quantum-optical coherence tomography,” *Physical Review Letters* **91**, 083601 (2003).
- [120] K. Banaszek, A. S. Radunsky, and I. A. Walmsley, “Blind dispersion compensation for optical coherence tomography,” [quant-ph/0404054](#) (2004).
- [121] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley, New York, 1991).
- [122] R. Hanbury-Brown and R. Twiss, “Correlation between photons in two coherent beams of light,” *Nature* **177**, 27–29 (1956).
- [123] J. D. Franson, “Photon exchange interactions and quantum information processing,” *Physical Review A* **70**, 054301 (2004).
- [124] J. L. O’Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, “Demonstration of an all-optical quantum controlled-NOT gate,” *Nature* **426**, 264–267 (2003).
- [125] M. Atature, G. D. Giuseppe, M. D. Shaw, A. V. Sergienko, B. E. A. Saleh, and M. C. Teich, “Multiparameter entanglement in quantum interferometry,” *Physical Review A* **66**, 023822 (2002).
- [126] J. Howell, R. Bennink, S. Bentley, and R. Boyd, “Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum and Position-Entangled Photons from Spontaneous Parametric Down Conversion,” *Physical Review Letters* **92**, 210403 (2004).
- [127] J. S. Bell, “Introduction to the hidden variable question,” In *Proceedings of the International School of Physics ‘Enrico Fermi’, Course II, Foundations of Quantum Mechanics*, pp. 171–181 (1971).

- [128] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” *Physical Review Letters* **23**, 880–884 (1969).
- [129] L. Hardy, “A new way to obtain Bell inequalities,” *Physics Letters A* **161**, 21–25 (1991).
- [130] L. Hardy, “Nonlocality for two particles without inequalities for almost all entangled states,” *Physical Review Letters* **71**, 1665–1668 (1993).
- [131] J. R. Torgerson, D. Branning, C. H. Monken, and L. Mandel, “Experimental demonstration of the violation of local realism without Bell inequalities,” *Physics Letters A* **204**, 323–328 (1995).
- [132] D. Boschi, F. D. Martini, and G. D. Giuseppe, “Test of the violation of local realism in quantum mechanics without Bell inequalities,” *Physics Letters A* **228**, 208–214 (1997).
- [133] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat, “Nonmaximally Entangled States: Production, Characterization, and Utilization,” *Physical Review Letters* **83**, 3103–3107 (1999).
- [134] N. D. Mermin, “Hidden variables and the two theorems of John Bell,” *Reviews of Modern Physics* **65**, 803 (1993).
- [135] N. D. Mermin, “Quantum mysteries refined,” *American Journal of Physics* **62**, 880 (1994).
- [136] W. T. M. Irvine, J. F. Hodelin, C. Simon, and D. Bouwmeester, “Realization of Hardy’s Thought Experiment with Photons,” *Physical Review Letters* **95**, 030401 (2005).
- [137] L. Hardy, “A quantum optical experiment to test local realism,” *Physics Letters A* **167**, 17–23 (1992).

- [138] A. C. Elitzur and L. Vaidman, “Quantum mechanical interaction-free measurements,” *Found. Phys.* **23**, 987–997 (1993).
- [139] L. Vaidman, “The meaning of the interaction-free measurements,” [quant-ph/0103081](https://arxiv.org/abs/quant-ph/0103081) (20001).
- [140] J. F. Clauser and M. A. Horne, “Experimental consequences of objective local theories,” *Physical Review D* **10**, 526–535 (1974).
- [141] A. Garuccio, “Hardy’s approach, Eberhard’s inequality, and supplementary assumptions,” *Physical Review A* **52**, 2535 (1995).
- [142] T. Herzog, J. Rarity, H. Weinfurter, and A. Zeilinger, “Frustrated two-photon creation via interference,” *Physical Review Letters* **72**, 629–632 (1994).
- [143] X. Zou, L. Wang, and L. Mandel, “Induced coherence and indistinguishability in optical interference,” *Physical Review Letters* **67**, 318–321 (1991).
- [144] Z. Ou, L. Wang, and L. Mandel, “Vacuum effects on interference in two-photon down conversion,” *Physical Review A* **40**, 1428–1435 (1989).
- [145] N. Bohr, in *Albert Einstein: Philosopher scientist*, P. A. Schlipp, ed., (Library of Living Philosophers, Evaston, 1949), pp. 200–241.
- [146] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics Vol. III* (Addison Wesley, Reading MA, 1965).
- [147] T. Young, *Course of Lectures on Natural Philosophy and the Mechanical Arts* (Royal Institution, 1807).
- [148] E. P. Storey, S. M. Tan, M. J. Collett, and D. F. Walls, “Reply: Complementarity and Uncertainty,” *Nature* **375**, 368–368 (1995).
- [149] B. G. Englert, M. O. Scully, and H. Walther, “Complementarity and uncertainty,” *Nature* **375**, 367–368 (1995).

- [150] H. Wiseman and F. Harrison, “Uncertainty over complementarity?,” *Nature* **377**, 584 (1995).
- [151] N. Bohr, in *Quantum theory and measurement*, J. A. Wheeler and W. H. Zurek, eds., (Princeton Univ. Press, 1983), pp. 9–49.
- [152] H. M. Wiseman, F. E. Harrison, M. J. Collett, S. M. Tan, D. F. Walls, and R. B. Killip, “Nonlocal momentum transfer in welcher Weg measurements,” *Physical Review A* **56**, 55–75 (1997).
- [153] J. L. Garretson, H. M. Wiseman, D. T. Pope, and D. T. Pegg, “The uncertainty relation in ‘which-way’ experiments: how to observe directly the momentum transfer using weak values,” *Journal of Optics B* **6**, S506–S517 (2004).
- [154] S. Durr, T. Nonn, and G. Rempe, “Origin of quantum-mechanical complementarity probed by a ‘which-way’ experiment in an atom interferometer,” *Nature* **395**, 33–37 (1998).
- [155] W. Wootters, W. K. & Zurek, “Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr’s principle,” *Physical Review D* **19**, 473–484 (1979).
- [156] M. S. Chapman, T. D. Hammond, A. Lenef, J. Schmiedmayer, R. A. Rubenstein, E. Smith, and D. E. Pritchard, “Photon scattering from atoms in an atom interferometer - Coherence lost and regained,” *Physical Review Letters* **75**, 3783–3787 (1995).
- [157] M. O. Scully and K. Drohl, “Quantum eraser: A proposed photon correlation experiment concerning observation and “delayed choice” in quantum mechanics,” *Physical Review A* **25**, 2208–2213 (1982).
- [158] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, “Observation of a quantum

- eraser and revival of coherence in a 2-photon interference experiment,” *Physical Review A* **45**, 7729–7739 (1992).
- [159] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, “3 proposed quantum erasers,” *Physical Review A* **49**, 61–68 (1994).
- [160] T. Curtright and C. Zachos, “Negative probability and uncertainty relations,” *Mod. Phys. Lett* **A16**, 2381–2385 (2001).
- [161] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, “Experimental realization of the quantum box problem,” *Physics Letters A* **324**, 125–131 (2004).
- [162] A. M. Steinberg, in *Science and ultimate reality: Quantum theory, cosmology and complexity*, J. D. Barrow, P. C. W. Davies, and C. L. H. Jr., eds., (Cambridge Univ. Press, Cambridge, 2004), quant-ph/0302003.
- [163] K. Molmer, “Counterfactual statements and weak measurements: an experimental proposal,” *Physics Letters A* **292**, 151–155 (2001).
- [164] K. J. Resch, “Practical weak measurement of multiparticle observables,” *Journal of Optics B* **6**, 482–487 (2004).
- [165] K. J. Resch and A. M. Steinberg, “Extracting Joint Weak Values with Local, Single-Particle Measurements,” *Physical Review Letters* **92**, 130402 (2004).
- [166] S. Gasparoni, J.-W. Pan, P. Walther, T. Rudolph, and A. Zeilinger, “Realization of a photonic controlled-NOT gate sufficient for quantum computation,” *Physical Review Letters* **93**, 020504–1 (2004).
- [167] K. Sanaka, T. Jennewein, J.-W. Pan, K. Resch, and A. Zeilinger, “Experimental Nonlinear Sign Shift for Linear Optics Quantum Computation,” *Physical Review Letters* **92**, 017902 (2004).



- [168] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, "Super-resolving phase measurements with a multiphoton entangled state," *Nature* **429**, 161–164 (2004).
- [169] Z. Zhao, A.-N. Zhang, Y.-A. Chen, H. Zhang, J.-F. Du, T. Yang, and J.-W. Pan, "Experimental Demonstration of a Nondestructive Controlled-NOT Quantum Gate for Two Independent Photon Qubits," *Physical Review Letters* **94**, 030501 (2005).
- [170] Z. Zhao, Y.-A. Chen, A.-N. Zhang, T. Yang, H. J. Briegel, and J.-W. Pan, "Experimental demonstration of five-photon entanglement and open-destination teleportation," *Nature* **430**, 54–58 (2004).
- [171] L. Vaidman, "Time-symmetrized counterfactuals in quantum theory," *Foundations of Physics* **29**, 755–765 (1999).
- [172] S. E. Ahnert and M. C. Payne, "Linear optics implementation of weak values in Hardy's paradox," *Physical Review A* **70**, 042102 (2004).
- [173] J. S. Lundeen, K. J. Resch, and A. M. Steinberg, "Comment on "Linear optics implementation of weak values in Hardy's paradox",," *Physical Review A* **72**, 016101 (2005).
- [174] S. E. Ahnert and M. C. Payne, "Reply to "Comment on 'Linear optics implementation of weak values in Hardy's paradox' ",," *Physical Review A* **72**, 016102 (2005).
- [175] K. Resch, K. Pregnell, R. Prevedel, A. Gilchrist, G. Pryde, J. O'Brien, and A. White, "Time-reversal and super-resolving phase measurements," [quant-ph/0511214](https://arxiv.org/abs/quant-ph/0511214) (2005).
- [176] D. Achilles, C. Silberhorn, C. Sliwa, K. Banaszek, I. A. Walmsley, M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, "Photon-number-resolving

- detection using time-multiplexing,” *Journal of Modern Optics* **51**, 1499–1515 (2004).
- [177] H. S. Eisenberg, G. Khoury, G. Durking, C. Simon, and D. Bouwmeester, “Quantum entanglement of a large number of photons,” *Physical Review Letters* **93**, 193901 (2004).
- [178] V. Balic, D. A. Braje, P. Kolchin, G. Y. Yin, and S. E. Harris, “Generation of paired photons with controllable waveforms,” *Physical Review Letters* **94**, 1836011 (2005).
- [179] C. H. van der Wal, M. D. Eisaman, A. Andr<sup>?</sup>, R. L. Walsworth, D. F. Phillips, A. S. Zibrov, and M. D. Lukin, “Atomic memory for correlated photon states,” *Science* **301**, 196 (2003).
- [180] A. Kuhn, M. Hennrich, and G. Rempe, “Deterministic Single-Photon Source for Distributed Quantum Networking,” *Physical Review Letters* **89**, 067901 (2002).
- [181] J. McKeever, A. Boca, A. D. Boozer, R. Miller, J. R. Buck, A. Kuzmich, and H. J. Kimble, “Deterministic Generation of Single Photons from One Atom Trapped in a Cavity,” *Science* **303**, 1992–1994 (2004).
- [182] D. Fattal, K. Inoue, J. Vuckovic, C. Santori, G. S. Solomon, and Y. Yamamoto, “Entanglement Formation and Violation of Bell’s Inequality with a Semiconductor Single Photon Source,” *Physical Review Letters* **92**, 037903 (2004).

# Appendix A

## Conditions necessary for weak measurement

A complete set of conditions that need to be satisfied for a measurement to be considered weak are still unknown. However, a few simple cases, such as the basic von Neumann model are well understood and are reviewed below.

### A.1 Pointer width

In the derivation of the weak value, we truncated the post-selected pointer state  $|\phi_{fi}\rangle$  in Eq. 1.33 to first-order in  $gt$ . This is only valid if all the higher order terms in  $|\phi_{fi}\rangle$  are small compared to the first-order term. As well, the first-order term must be small compared to the zero-order term. It follows that the measurement is weak if

$$(gt)^n \left| \frac{\langle F | \hat{A}^n | I \rangle}{\langle F | I \rangle} \frac{\partial^n \phi(x)}{\partial x^n} \frac{1}{\phi(x)} \right| \ll gt \left| \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \frac{\partial \phi(x)}{\partial x} \frac{1}{\phi(x)} \right| \ll 1. \quad (\text{A.1})$$

A similar condition was derived in [73], as a correction to the condition given in an early paper by Aharonov *et al.* [94]. For a Gaussian pointer, the condition on the right becomes

$$\frac{gt}{\sigma} \left| \langle \hat{A} \rangle_w \right| \ll 1. \quad (\text{A.2})$$

Whether a measurement is weak or not depends not only on the uncertainty of the measurement but also on the average result of the measurement,  $\langle \hat{A} \rangle_w$ . A surprising corollary of this is that the coupling  $g$  can be large as long as  $\langle \hat{A} \rangle_w$  is small. The condition on the left is more complicated to understand but we can consider two special cases:

1. If any of the derivatives  $\partial^i \phi(x) / \partial x^i$  of the pointer wavefunction are discontinuous then the measurement cannot be weak no matter how small the coupling is.

2. If  $\hat{A}$  is a projector then  $\hat{A}^n = \hat{A}$ , and the left condition is independent of the weak value. Now, if we assume the pointer is in a minimum uncertainty state then we can approximate  $(\partial^n \phi(x) / \partial x^n) / \phi(x)$  by  $\sigma^{-n}$ , and we see that in this case the left condition is redundant.

## A.2 Pointer shape

Although in the calculation of  $\langle \hat{X} \rangle_{fi}$  and  $\langle \hat{P} \rangle_{fi}$  we assumed a Gaussian pointer shape, we also showed  $|\phi_{fi}\rangle \approx \hat{T}(gt \langle \hat{A} \rangle_w) |\phi\rangle$ , which is valid for any pointer shape as long as the corresponding wavefunction and its derivatives are continuous. The pointer interference observed in the Hulet experiment is independent of the actual pointer shape if these conditions are met: As long as the measurement is weak enough, the pointer will be rigidly shifted by the weak value.

## A.3 Pointer coherence

We now consider the coherence of the pointer. Strong measurement in the von Neumann model does not require that the pointer is in a pure state. For example,

consider a pointer that is in a mixture  $\hat{\rho}_A$  of pure states. The result of the von Neumann interaction is that each pure state will be shifted according to Eq. 1.24. If the pointer position probability distribution,  $P(x) = \langle x | \hat{\rho}_A | x \rangle$ , is smaller than the shift induced by the measurement ( $\sqrt{\text{Tr}(\hat{X}^2 \hat{\rho}_A)} \ll g t a_i$ ) one can deduce with certainty the result of the measurement for each trial, which is the necessary requirement for a strong measurement. In contrast, in weak measurement we have seen that the weak value arises from interference between different shifts of the pointer, and consequently depends on the coherence of the pointer. Furthermore, a mixed state whose position probability distribution has a width  $< \sigma$  will necessarily contain a larger momentum spread than a pure state with the same width. Since the interaction Hamiltonian Eq. 1.19 couples the system to the pointer's momentum, the measurement will inevitably lead to a larger disturbance to the measured system than would occur from a standard weak measurement. It follows that for mixed pointer states a sufficient condition for weak measurement is that

$$\sqrt{\text{Tr}(\hat{X}^2 \hat{\rho}_A)} \ll g t \langle \hat{A} \rangle_w \ll \hbar / \sqrt{\text{Tr}(\hat{P}^2 \hat{\rho}_A)}. \quad (\text{A.3})$$

Of course, if each of the pure states in the mixture satisfy Eq. A.1 then this condition will also be satisfied.

Surprisingly, although this condition is sufficient it may not be necessary. In Ref. [66], Lars Johansen considered mixed-state pointers in more detail and found that the right side of Eq. A.3 was unnecessary if the current density of the pointer vanishes:

$$\langle x | \hat{P} \hat{\rho}_s | x \rangle + \langle x | \hat{\rho} \hat{P} | x \rangle = 0, \quad (\text{A.4})$$

or equivalently,

$$\langle x | \hat{P} \hat{\rho}_A | x \rangle = \frac{\hbar}{2i} \frac{\partial}{\partial x} \langle x | \hat{\rho}_A | x \rangle. \quad (\text{A.5})$$

He found that if this condition is met, then the probability to find the measured system in  $|F\rangle$  is unchanged by the weak measurement. To understand what Johansen's condition implies we consider two simple examples of pointers. For a first example

we take a pure state pointer,  $\hat{\rho}_A = |\psi_A\rangle\langle\psi_A|$ . In this case, the condition requires that  $|\psi_A\rangle$  has a constant phase as a function of  $x$ ,  $\theta(x)$  (e.g.  $\psi_A(x)$  is positive for all  $x$ ). Since momentum is proportional to the derivative of phase, this requirement roughly reduces to the requirement that the pointer is in a minimum uncertainty state. Now we consider the next simplest mixed pointer state,  $\hat{\rho}_A = (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)/2$ . With this state, Johansen's condition becomes

$$\begin{aligned} \langle x | \hat{P} (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|) | x \rangle &= \frac{\hbar}{2i} \frac{\partial}{\partial x} \langle x | (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|) | x \rangle \quad (\text{A.6}) \\ \langle x | \hat{P} |\psi_1\rangle\langle\psi_1| | x \rangle + \langle x | \hat{P} |\psi_2\rangle\langle\psi_2| | x \rangle &= \frac{\hbar}{2i} \frac{\partial}{\partial x} (\langle x | |\psi_1\rangle\langle\psi_1| | x \rangle + \langle x | |\psi_2\rangle\langle\psi_2| | x \rangle) \\ \psi_1^*(x) \frac{\partial}{\partial x} \psi_1(x) - \frac{1}{2} \frac{\partial}{\partial x} |\psi_1(x)|^2 &= \psi_2^*(x) \frac{\partial}{\partial x} \psi_2(x) + \frac{1}{2} \frac{\partial}{\partial x} |\psi_2(x)|^2 \\ |\psi_1(x)|^2 \frac{\partial}{\partial x} \theta_1(x) &= - |\psi_2(x)|^2 \frac{\partial}{\partial x} \theta_2(x), \end{aligned}$$

where  $\psi_i(x) = |\psi_i(x)| \exp(i\theta_i(x))$ . Surprisingly, each pure state composing the mixed state need not be real (nor in a minimum uncertainty state!). The current density from one state can cancel that from the other pure states in the mixture. For example, the pointers could be identical in shape but moving with equal and opposite velocities of arbitrary magnitude. Indeed, for a strong measurement this pointer would on average point to  $gt \langle \hat{A} \rangle$  and its shape would be unchanged. However, since the pointer has arbitrarily large momentum variance, it clearly need not satisfy the right side of Eq. A.3.

## Appendix B

# The $N$ th-order joint weak value with $N$ spin pointers

The following derivation shows how  $N$  spin pointers can be used to measure the  $N$ th joint-weak value of a system. The formalism is essentially the same as the annihilation-operator based derivation given in Chapter 5 for  $N$  position pointers. Any variables that we do not define here are defined in Chapter 5. We begin with a von Neumann-style interaction chosen to induce a rotation in a spin pointer polarized along  $z$ -direction,

$$\mathcal{H} = -g\hat{A}\hat{S}_y \tag{B.1}$$

$$= i\frac{g}{2}\hat{A}(\hat{S}_z^+ - \hat{S}_z^-), \tag{B.2}$$

where  $\hat{S}_z^+ = \hat{S}_x + i\hat{S}_y$ , and  $\hat{S}_z^- = \hat{S}_x - i\hat{S}_y$  are the spin raising and lowering operators.

We begin with a single pointer that is in the lowest  $s_z$  spin state,

$$|\phi\rangle = |s_z = -s\rangle \equiv |0\rangle$$

$$|s_z = -s + 1\rangle \equiv |1\rangle.$$

These latter two states are the only states we need to consider in the following calculations. We call them our basis states. Under the action the von Neumann interaction the system evolves to,

$$|\psi\rangle \rightarrow \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |I\rangle |\phi\rangle = \left(1 - \frac{i\mathcal{H}t}{\hbar} - \dots\right) |I\rangle |\phi\rangle \quad (\text{B.3})$$

$$\approx |I\rangle |0\rangle + \frac{gt}{2\hbar} \hat{A} |I\rangle (\hat{S}_z^- - \hat{S}_z^+) |0\rangle + \dots \quad (\text{B.4})$$

The action of the spin raising and lowering operators on the basis states is,

$$\begin{aligned} \hat{S}_z^+ |0\rangle &= \hbar\sqrt{2s} |1\rangle \\ \hat{S}_z^- |1\rangle &= \hbar\sqrt{2s} |0\rangle. \end{aligned}$$

We project out the part of the system that is post-selected in state  $|F\rangle$ ,

$$\langle F | \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |I\rangle |\phi\rangle = \langle F | I \rangle |0\rangle + gt\sqrt{\frac{s}{2}} \langle F | \hat{A} | I \rangle |1\rangle + \dots \quad (\text{B.5})$$

This leaves the state of pointer after the interaction and post-selection which we renormalize by dividing by  $\langle F | I \rangle$ ,

$$|\phi_{fi}\rangle = |0\rangle + gt\sqrt{\frac{s}{2}} \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} |1\rangle + \dots \quad (\text{B.6})$$

With this we calculate the expectation value of the spin lowering operator to find the weak value of  $\hat{A}$ ,

$$\begin{aligned} \langle \hat{S}_z^- \rangle_{fi} &\simeq gt\hbar s \frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \\ &\simeq gt\hbar s \langle \hat{A} \rangle_w. \end{aligned}$$

Extending this analysis to a joint-weak measurement, the von Neumann interaction becomes,



$$\mathcal{H} = -g_1 \hat{A}_1 \hat{S}_{1y} - g_2 \hat{A}_2 \hat{S}_{2y} - \dots \quad (\text{B.7})$$

$$= -\sum_{j=1}^N g_j \hat{A}_j \hat{S}_{jy} \quad (\text{B.8})$$

$$= \frac{ig}{2} \sum_{j=1}^N \hat{A}_j (\hat{S}_{jz}^+ - \hat{S}_{jz}^-). \quad (\text{B.9})$$

The initial state of the  $N$  spin pointers is,

$$|\Phi\rangle = \prod_{j=1}^N |\phi_j\rangle = |0\rangle^{\otimes N}. \quad (\text{B.10})$$

We calculate the state of the combined system after the interaction Hamiltonian is applied,

$$|\Phi\rangle |I\rangle \rightarrow \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |0\rangle |I\rangle = \left(1 - \frac{i\mathcal{H}t}{\hbar} + \dots\right) |0\rangle |I\rangle \quad (\text{B.11})$$

$$= \left(1 + \frac{gt}{2\hbar} \sum_{j=1}^N \hat{A}_j (\hat{S}_{jz}^+ - \hat{S}_{jz}^-) + \dots\right) |0\rangle |I\rangle. \quad (\text{B.12})$$

Projecting out the part of the measured system that we post-select, we get:

$$\langle F | \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) |0\rangle = |0\rangle \langle F | I \rangle + gt \sqrt{\frac{s}{2}} \sum_{j=1}^N \langle F | \hat{A}_j | I \rangle |1_j\rangle + \dots \quad (\text{B.13})$$

$$+ \left(gt \sqrt{\frac{s}{2}}\right)^N \frac{1}{N!} \langle F | \wp \left\{ \hat{A}_k \right\}_N | I \rangle |1\rangle + \dots, \quad (\text{B.14})$$

where  $|1_j\rangle$  is the state where the  $j$ th pointer is in the  $|s_z = -s + 1\rangle$  and all other pointers are in the bottom state  $|s_z = -s\rangle$ ; and  $|1\rangle$  is the state with all  $N$  pointers in the  $|s_z = -s + 1\rangle$  state (eg.  $|1_1 1_2 1_3 \dots 1_N\rangle$ ). We renormalize the resulting pointers state  $|\Phi_{fi}\rangle$  by dividing by  $\langle F | I \rangle$ ,

$$|\Phi_{fi}\rangle = |0\rangle + gt \sqrt{\frac{s}{2}} \sum_{j=1}^N \frac{\langle F | \hat{A}_j | I \rangle}{\langle F | I \rangle} |1_j\rangle + \dots + \left(gt \sqrt{\frac{s}{2}}\right)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_k \right\}_N | I \rangle |1\rangle}{\langle F | I \rangle} + \dots \quad (\text{B.15})$$

In addition we define the product of all  $N$  spin lowering operators to be,

$$\hat{O} \equiv \prod_{j=1}^N \hat{S}_{jz}^- \quad (\text{B.16})$$

which we call the  $N$ -lowering operator. The rightmost term in Eq B.15 is the lowest order term that does not go to zero when acted on by the  $N$ -lowering operator; this term becomes,

$$\hat{O} |\Phi_{fi}\rangle = (gt\hbar s)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_j \right\}_N | I \rangle}{\langle F | I \rangle} | 0 \rangle. \quad (\text{B.17})$$

Clearly, to lowest nonzero order the expectation value then becomes,

$$\left\langle \hat{O} \right\rangle_{fi} = \langle \Phi_{fi} | \hat{O} | \Phi_{fi} \rangle \quad (\text{B.18})$$

$$= \langle 0 | (gt\hbar s)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_j \right\}_N | I \rangle}{\langle F | I \rangle} | 0 \rangle \quad (\text{B.19})$$

$$= (gt\hbar s)^N \frac{1}{N!} \frac{\langle F | \wp \left\{ \hat{A}_j \right\}_N | I \rangle}{\langle F | I \rangle}. \quad (\text{B.20})$$

In the usual case, each operator  $\hat{A}_j$  acts on the a different particle, allowing all the  $\hat{A}_j$  to commute. We can thus write the Nth joint-weak value as,

$$\left\langle \prod_{j=1}^N \hat{A}_j \right\rangle_w = \left\langle \prod_{j=1}^N \hat{S}_{jz}^- \right\rangle_{fi} (gt\hbar s)^{-N}. \quad (\text{B.21})$$