On Weak Measurements & Foundational Experiments in Quantum Mechanics



Surya Narayan Sahoo

Supervisor : Prof. Urbasi Sinha

Light & Matter Physics
Raman Research Institute, Bengaluru

A thesis submitted for the degree of
Doctor of Philosophy
to Jawaharlal Nehru University



Declaration

The scientific work reported in this thesis is based on the experimental research carried out at Quantum Information and Computing lab, Light and Matter Physics Group, Raman Research Institute,Bangalore, India with various collaborations. No part of this thesis has been submitted elsewhere for any other degree or qualification. This thesis has been run through the plagiarism checker software Turnitin.

Supervisor: Prof. Urbasi Sinha
Light & Matter Physics
Raman Research Institute
Bengaluru

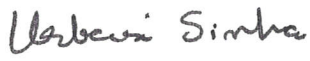
Student: Surya Narayan Sahoo
Light & Matter Physics
Raman Research Institute
Bengaluru

Certificate

This is to certify that the thesis titled "On Weak Measurement & Foundational Experiments in Quantum Mechanics" submitted by Surya Narayan Sahoo for the award of the degree of Doctor of Philosophy of Jawaharlal Nehru University is his original work. This work has not been submitted to any other University for the award of any other degree or diploma.



Director
Raman Research Institute
Bengaluru



Supervisor: Prof. Urbasi Sinha
Light & Matter Physics
Raman Research Institute
Bengaluru

Synopsis

On weak measurement & foundational experiments in quantum mechanics

Surya Narayan Sahoo

Quantum Information and Computating Lab

Light and Matter Physics

Raman Research Institute, Bengaluru

under the guidance of Prof. Urbasi Sinha

30 September, 2019

Quantum Foundations Despite the remarkable success of quantum theory in explaining black-body and atomic spectra as well as in paving the foundation for future technologies like Quantum Computing and Quantum Cryptography, interpretation of many of its concepts has led to a lot of debates. At the heart, the problem always reduces to asking when does a quantum object behave as a wave and when does it behave as a particle. This, in essence, is the wave-particle duality which gets manifested in two-path interferometers like the double-slit or the Mach Zehnder interferometer. If we assume that the propagation is of wave nature then we have to explain how the wave reduces or collapses to exhibit particle nature in individual events of detection, which is popularly known as the Measurement Problem. To

avoid explaining the collapse, one can assume the particle nature to begin with, so that no new physics is necessary to explain the detection mechanism. However, one has to then explain how the ensemble of particles collectively exhibits wave nature i.e. gives us interference pattern in a double-slit experiment. In its essence, what we are asking is the answer to the question 'which slit the particle went through' provided it formed a part of the ensemble that gives us the interference pattern (and are not blocked). Generalizing this, we ask what is the property of a particle (e.g. which slit it went through) between two events (e.g. emission and detection).

Weak Value & Weak Measurement The property of a particle given its initial and final states, in general, gets affected if we perform any intermediate strong measurement. The effect can be quantified from the change in the pointer state to which the particles get coupled during the interaction in a von-Neumann measurement scheme. In the limit where the pointer state before the interaction is indistinguishable from pointer state after the interaction, we can claim that property between pre- and post-selection is minimally disturbed. This can be achieved when the interaction strength is very weak. But, from a single measurement, we would not get any information about the system. However, if we take an ensemble of identical particles subject to the same pre-selection and post-selection, the average of the pointer shift gives some information about the property of particle to which the pointer was weakly coupled to. This pointer shift corresponds to (the real part of) what is known as the weak value of the observable A between pre-selection and post-selection.

$$\langle A \rangle^{(w)} = \frac{\langle \phi | A | \psi \rangle}{\langle \phi | \psi \rangle} \tag{1}$$

Here, $|\psi\rangle$ is the pre-selected state (the state in which the particle is prepared) and $|\phi\rangle$ is the post-selected state (the state that guarantees a successful detection). The weak value is, in general, a complex number and its magnitude can go beyond the largest eigenvalue. What exactly the weak value physically corresponds to has been a matter of much debate. However, when the weak value equals one of the extreme eigenvalues, we can claim that all the particles in the ensemble must have the same value of the observable between pre and post-selection.

The first part of this thesis will be regarding the study of certain seemingly paradoxical conclusions that can be drawn consistently from weak values in a pre and post selected system.

Quantum Cheshire Cat One such apparent pre- and post-selection paradox is the Quantum Cheshire Cat effect where the weak value suggests that there exists σ_x component of particle's spin in the arm of a two-path interferometer where the particle does not exist. The particle is pre-selected with the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle |A\rangle + |\downarrow\rangle |B\rangle)$ where A and B refer to the two arms of a Mach Zehnder interferometer. Here, $|\uparrow\rangle$ and $|\downarrow\rangle$ refers to spin-up and spin-down, respectively, for the spin-1/2 particle. If we post-select the particle in the state $|\phi\rangle = \frac{1}{\sqrt{2}} |\uparrow\rangle (|A\rangle + |B\rangle)$, it seems that the only consistent description of 'which path the particle took' that satisfies both pre-selection and post-selection is that it went through path A. But what can we say about its spin? Clearly, the spin of the photon in arm A is up $|\uparrow\rangle$. But, can we ask the question, what is the spin of the particle in the arm B? Conventionally, this question may seem ill-posed because we don't expect the particle in arm B. But if we compute the weak value of the σ_x spin component we get the value 1. This being an eigenvalue indicates that the σ_x spin component exists in the arm B where the photon does not exist subject to both pre- and post-selection.

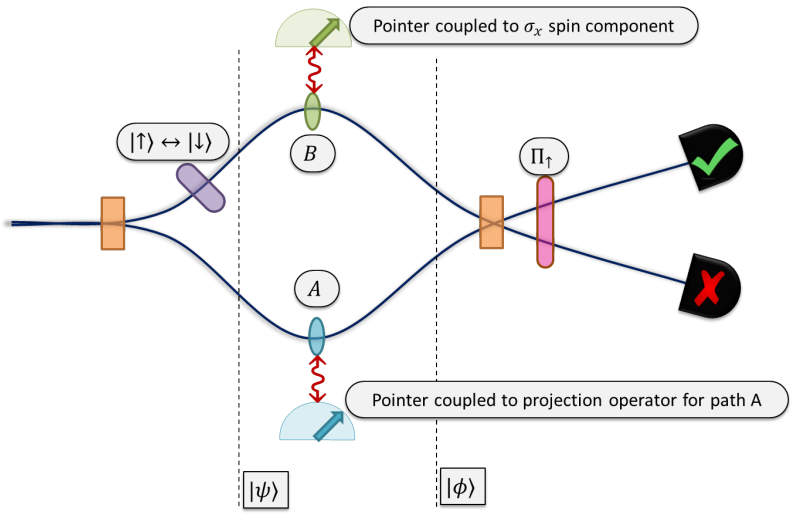


Figure 1: Schematic of the Quantum Cheshire Cat effect in a two-path interferometer for a spin-1/2 particle

Few experiments claiming to obtain signatures of Quantum Cheshire Cat have already been performed but all of them have one common limitation, besides few others depending on the particular experiment. They use absorbers to comment on which path the particle took. If the particle is absorbed within the interferometer then it does not belong to the ensemble that satisfies pre- and post-selection simultaneously.

In this work, the aim was to study whether or not we can experimentally obtain signatures of Quantum Cheshire Cat in a Mach Zehnder interferometer with a light source from pointer shifts. We implement weak measurements for the spin and 'which-path' of the particle and refrain from using absorbers to obtain loophole-free signatures of the Quantum Cheshire Cat effect. To claim that the signatures of spin and 'which-path' corresponds to the same ensemble, we also perform simultaneous weak measurement of both the observables.

Experimental implementation with laser source & single-photon source The weak value obtained from an ensemble of single photons agrees with the weak value obtained from a coherent source. Therefore, we can perform the experiment with a coherent laser source and still can infer the effect for a photon. In this work, we not only ask which path the photon took and what is the polarization of the photon in individual experiments but also simultaneously, by coupling each operator to a different pointer and measuring the joint weak value. We obtain the pointer shifts which forms the operational signature of the Cheshire Cat. We repeat the experiment with single photons created using SPDC to put any doubts to rest that the inferences obtained from weak values apply to the ensemble of pre- and post-selected single photons as well.

In pursuit of understanding the experimental signatures of pre- and post-selection paradoxes better, it is important to understand when any operation qualifies as an observable in quantum mechanics and how its interaction with the pointer gives us the signatures. Whether the weak value can be obtained only in pre- and post-selected systems or it can be consistently obtained without performing weak measurements is explored further.

Measuring expectation value of a non-Hermitian operator : *Weak value without weak measurements* Hermitian operators have been used as observables in most of Quantum Physics because they have real eigenvalues and they provide their eigen-spectrum as a complete orthogonal basis. One cannot infer the expectation value of non-Hermitian operators from statistics of outcomes due to the same reason. However, previous work by our group has shown that if we measure the weak value of a Hermitian operator to which the non-Hermitian operator is polar decomposed into, we can infer the expectation value of the non-Hermitian operator. We experimentally obtain the weak value of the projector operator and infer the expectation value of the spin ladder operators and compare with theory. We also show that if the optical components corresponding to the polar decomposed operators are available, we can infer the expectation value of the non-Hermitian operator from visibility and phase shift in a Mach Zehnder interferometer where these components are placed in either arm. From the same experiment, we also demonstrate that the weak value of operators can be inferred from experiments that do not involve weak measurements or post-selection. We finally compare and contrast the experimental advantage of obtaining the weak value from the interferometric technique with the conventional weak measurement technique.

Interferometry has been one of the ubiqutous tools used in most experiments studying Quantum Foundations. In this thesis, we have also used Mach Zehnder interferometer along with weak measurements in experiments studying the fundamental implications of the weak value. Even the pointer shifts obtained after weak interaction and post-selection can be viewed as the consequence of interference. The theoretical prediction of intensity pattern in any interference-based experiment involving quantum objects involves the application of superposition principle along with Born rule. Any deviation from these two fundamental principles/rules may require us to investigate theories beyond quantum mechanics. However, the superposition principle is often incorrectly applied in slit-based experiments where the boundary conditions do change. The next chapter of this thesis is devoted towards experimentally measuring the deviation of the interference pattern in a triple slot

experiment in microwave domain from what one would predict by applying the superposition principle incorrectly. .

Superposition principle and higher-order Feynman paths In a double-slit experiment, let the solution to the wave equation be denoted as ψ_A when one slit is open and ψ_B when the other slit is open. It is often assumed that the wavefunction ψ_{AB} when both the slits are open is $\psi_A + \psi_B$. However, the superposition principle strictly applies when the boundary conditions remain the same. When any of the slit is opened or closed, the boundary condition do change. Therefore, the wavefunction ψ_{AB} could deviate from what the naive application of superposition principle gives us. In the scientific literature, we have many slit-based experiments, but hardly any significant deviation has been reported. This indicates that the intensity when two slits are open I_{AB} is well approximated by the naive application of superposition principle i.e. $I_{AB} \approx |\psi_A + \psi_B|^2$. One may attempt to find the difference between experimentally obtained I_{AB} and theoretically computed $|\psi_A + \psi_B|^2$. However, this difference is subject to precise modelling of slits and other experimental parameters. Fortunately, in previous works, our group came up with a theory independent method of measuring the deviation from the naive application of superposition principle by using three slits instead of two. The deviation is captured by the Sorkin parameter κ which the normalized sum and difference of intensities in various combination of slit opening combinations for a triple slit experiment.

$$\kappa = \frac{I_{ABC} - (I_{AB} + I_{BC} + I_{CA}) + (I_A + I_B + I_C)}{\max(I_{ABC})} \tag{2}$$

If we apply the superposition principle naively, we expect κ to be zero. However, if we compute κ from numerical solutions of Maxwell Equations for a triple slit experiment, we find that κ is non-zero. The deviation can be elegantly computed by adding the contribution of higher-order paths to the Feynman path integrals. Using path integral formalism the maximum deviation can be predicted from the slit width w, inter-slit distance d and the wavelength λ.

$$|\kappa_{max}| \approx 0.03 \frac{\lambda^{3/2}}{w\sqrt{d}} \tag{3}$$

In this thesis, the aim is to obtain the non-zero κ experimentally. We design a triple slot experiment in the microwave domain keeping the expression 3 in consideration. We obtain non-zero κ and show that κ as a function of detector position agrees well with the numerical solutions of Maxwell equations and hence is not consistent with traditional incorrect application of superposition principle. This is the first experimental demonstration where deviation from naive superposition principle is measured in the classical domain (with microwave source). Additionally, it is shown that κ decreases if we introduce baffles between the slots which indicate that the higher-order paths in Feynman path integral have major contributions towards a non-zero κ.

Summary In this thesis, experimental signatures of Quantum Cheshire Cat effect is obtained from shifts of two orthogonal pointers, simultaneously coupled to the 'which-path' and polarization degrees of freedom, due to weak interaction followed by post-selection. The insights and experimental methods used in this thesis can be used to address other pre and post-selection paradoxes. The expectation value of spin-ladder operators is inferred in an interferometric setup, along with obtaining weak value without performing weak measurements, which may have future metrological applications. Precision interferometry has been used as a tool in experiments aimed towards investigating the fundamental principles of nature. Our work on measuring the deviation from the naive application of the superposition principle involving a high-precision triple-slot experiment can potentially become a key consideration towards design of such experiments in future.

List of Publications

Journal articles

Published

1. Gaurav Nirala, Surya Narayan Sahoo, Arun K. Pati, and Urbasi Sinha. Measuring average of non-hermitian operator with weak value in a mach-zehnder interferometer. *Phys. Rev. A*, 99:022111, Feb 2019 doi:10.1103/PhysRevA.99.022111
(See Ch. 5)

2. G Rengaraj, U Prathwiraj, Surya Narayan Sahoo, R Somashekhar, and Urbasi Sinha. Measuring the deviation from the superposition principle in interference experiments. *New Journal of Physics*, 20(6):063049, jun 2018 doi:10.1088/1367-2630/aac92c
(See Ch. 6)

Submitted

1. Surya Narayan Sahoo, Sanchari Chakraborti, Som Kanjilal, Dipankar Home, Alex Matzkin, and Urbasi Sinha. Unambiguous joint detection of spatially separated properties of a single photon in the two arms of an interferometer. *arXiv:quant-ph*, 2022 [1] arxiv.org/abs/2201.11425
(See Ch. 4)

Published but not a part of this thesis

1. Surya Narayan Sahoo, Sanchari Chakraborti, Arun K. Pati, and Urbasi Sinha. Quantum state interferography. *Phys. Rev. Lett.*, 125:123601, Sep 2020 doi: 10.1103/PhysRevLett.125.123601
(To be a part of the thesis of Sanchari Chakraborti)

Magazine articles

1. Urbasi Sinha, Surya Narayan Sahoo, Ashutosh Singh, Kaushik Joarder, Rishab Chatterjee, and Sanchari Chakraborti. Single-photon sources. *Opt. Photon. News*, 30(9):32–39, Sep 2019 doi: 10.1364/OPN.30.9.000032

Urbasi Sinha

Prof. Urbasi Sinha
Supervisor
Light & Matter Physics
Raman Research Institute,
Bengaluru

Surya Narayan Sahoo

Surya Narayan Sahoo
Student
Light & Matter Physics
Raman Research Institute,
Bengaluru

[1] Currently, the manuscript is submitted to a peer-reviewed journal. The title of the accepted manuscript subject to change upon revision.

Contents

Introduction 1

On the foundations of Quantum Mechanics 1

Quantum Cheshire Cat Effect 9

Expectation value of Non-Hermitian Operators 11

The Born rule and the superposition principle 12

1 Weak Measurements & Weak Value 23

1.1 Time-reversal symmetry of dynamical laws 26

1.2 Time symmetric approach towards quantum mechanics 31

1.3 Properties of pre- and post-selected ensembles - weak value 39

1.4 Weak Measurements 48

1.5 Weak value 61

2 Optical Implementation of Weak Measurement in Polarization degree of freedom 79

2.1 Measurement in Polarization degree of freedom 80

2.2 Transverse spatial pointer shift in weak measurement of polarization components 82

2.3 Experimental Setup for weak measurement 102

2.4 Weak Measurement of the σ_x observable 118

2.5 Modifications to σ_x setup for single photon experiments 133

3 Quantum Cheshire Cat 143

3.1 The partial quantum chehsire cat effect 144

3.2 Optical setup to realize partial cheshire cat 149

3.3 Weak measurements for the Cheshire cat effect . . . 163
3.4 Has quantum Cheshire cat already been experimentally demonstrated ? . . . 176
3.5 Experimental Implementation with proper weak measurements . . . 185
3.6 Experimental Results & Analysis . . . 191

4 Experiment with Single Photons 203
4.1 Single Photon source . . . 207
4.2 Experimental Setup . . . 215
4.3 Data Analysis . . . 222

5 Inferring Expectation value of non-Hermitian operator 230
5.1 Polar decomposition of Operators . . . 234
5.2 Inferring Expectation values using Mach Zehnder Interferometers . . . 236
5.3 The apparent 'non-unitary' evolution . . . 242
5.4 Experimental Implementation . . . 245
5.5 Results . . . 250
5.6 Outlook . . . 255

6 Measuring deviation from naive application of superposition principle 258
6.1 Superposition principle & Born rule . . . 265
6.2 Multi-order interference . . . 267
6.3 Superposition principle & boundary conditions . . . 270
6.4 κ for the Triple Slot Case . . . 272
6.5 Experiment . . . 274
6.6 Results . . . 278
6.7 Error Types and their effect in the experiment . . . 281
6.8 Sources of error . . . 287
6.9 Comparison of Path integral result with Finite Difference Time Domain simulations . . . 298
6.10 Mutual Inductance in Electromagnetism & κ . . . 303

Summary & Outlook 311

Quantum Cheshire Cats . . . 311

Expectation Value of Non Hermitian Operators . . . 311

Application of Superposition principle . . . 311

Appendix 315

A Determining weak value using two-photon interference . . . 315

B Probability from ABL rule and weak values for observables in cheshire cat effect . . . 332

C Comment of Cheshire Cat using two photon interference . . . 338

D Stabilization of the Mach Zehnder Interferometer . . . 341

E Error Budget . . . 343

F Alignment with Pulsed Laser . . . 353

G Uncertainty and Error Propagation in Pointer Shift . . . 361

H Mechanical Stability of the Mach Zehnder Interferometer . . . 367

I Effect of incoherence on weak values and cheshire cat effect . . . 372

J Stability of photon source and Statistics of centre over Datasets . . . 376

K Comment on other experiments on superposition principle . . . 380

Introduction

On the foundations of Quantum Mechanics

Most introductions to quantum mechanics [I.1, I.2, I.3] begin with the limitations of the then-existing classical theories of mechanics, electrodynamics, thermodynamics, statistical mechanics and optics when applied to phenomena such as the black-body radiation that requires the consistent intersection of various branches of physics. Then follows the discussion on how the introduction of the 'quantum' of light - the particle attribution to what we have so far learned to treat as a wave helps calculate the black-body spectrum, explain the photoelectric effect and the Compton effect to name a few. Next, the Schrödinger equation appears as gospel and with it spells a whole new conceptual world. It is no longer governed by laws of nature that would aid computation of how the 'beables' [I.4] - the properties of the object would evolve over time, but how the probability of the observables would.

The fact that usually very little attempt is made to justify how we arrive at the equation is a testimony that with the ever dominating emphasis on the 'shut up and calculate' approach [I.5, I.6, I.7], what we aim to preserve is the knowledge obtained rather than giving due importance to the processes that leads us there. One may argue that it is more important to know how to use the wheel rather than dig deep into the past to learn how the wheel was invented. That is certainly true as far as technology is concerned, which in turn drives science forward. Galileo seeing the moons of Jupiter required the technology of how to use lenses to make telescopes [I.8]. Science drives technology forward, which allows science to explore territories that otherwise would seem distant. Scientific theories triumph when they are tested in the wildest of conditions and still work. Quantum theory has been at the core of particle physics, a significant part of modern chemistry, modern electronics

etc. to name a few [I.9]. However, much of the issues with the foundation of quantum mechanics are not in its range of applicability. Of course, we have not tested it [I.10, I.11] near a black hole and hence have not been able to rationalize with the foundational theories of spacetime. But irrespective of whether and when we get there, the issues with quantum mechanics are intrinsic in its description [I.12, I.13, I.14].

We use the Schrödinger equation [I.15] and the measurement postulate [I.16] as the wheel that keeps the field driving. Justifying how we arrived at the conclusion that the Schrödinger equation is the dynamical law of nature is almost similar to asking how we came to conclude that the Newton's second law is the correct dynamical law [I.17]. The latter is hard because along with it one needs the concept of inertia[I.18, I.19], the concept of force [I.20] and the concept of mass [I.21]. These are useful concepts but hard to introduce ab-initio, and probably therefore it is best fed to all so early that familiarity makes questioning it tease absurdity. Similarly, the Schrödinger equation requires the notion of the wavefunction - the meaning of which is still debated [I.22, I.23, I.24]. Whether it has ontological existence [I.25] or we need to limit its use for epistemic purpose [I.26, I.27] is still not clear [I.28]. This makes motivating the introduction of Schrödinger's equation extremely difficult. Further, the very ideas of de Broglie (pilot-wave theory) and Schrödinger (interpretation of physical density of charge) with which the quantum wave equations were formulated [I.29] are no longer the part of the standard interpretation of quantum mechanics [I.30]. Today, we have interpretations ranging from de Broglie - Bohm's pilot-wave theory to the many-worlds interpretation [I.31]. Many of the interpretations are aimed to make partial sense of how the wavefunction behaves, but often at the cost of introducing or emphasizing another unsaid assumption that not everyone would agree to be sensible to begin with. Sadly, science seemingly has reached towards its own antithesis – the postulates of quantum mechanics are treated like biblical verse and we, like priests, are searching for its interpretation.

But why question the essence of the theory that works? Probably, it's all fine and consistent, but we are not designed to make sense of it. Why not leave it at that and forgo all interpretations? We must be mostly right, since the theory has worked wonders. Well, not understanding what is life does not make us stop living. And similarly, being able to convert physics described by quantum theory into technology does not mean that we

have understood physics at its roots. The creationism theory also 'worked' during its prevalent times without much inconsistency as far as people can see. Even suggestions of the earth being flat would have been a good scientific hypothesis a millennium ago, which would have later been invalidated by observations as it was indeed the case. Thus, it is not unworthwhile to inspect what works with scrutiny. Further, we must differentiate between what features of a theory works and what are the unnecessary baggage. Surely, the computation of probabilities from knowing the evolution of the wavefunction using Schrödinger's equation is useful. But the concepts like that of ontological state collapsing instantly upon measurement could have easily been forgotten as the theme of a fairy tale had it not been the decades of efforts to understand quantum mechanics that forces us to seriously consider the possibility. It is therefore important to identify whether concepts like these are an integral part of the theory that works or are the unnecessary baggage that we are carrying. One way of making progress in science is by the construction of new unified theories that are applicable in unexplored circumstances. But before they need to be experimentally tested, it is often practical and useful to predict whether the new theory agrees with the earlier one which has been known to be successful in this domain of applicability. Now, if such a post-quantum theory [I.32] is to be developed, should it agree with the probabilities or does it need to have a description of the non-unitary state reduction as well? It is very easy to wish a theory to be minimalist, but in practice, it is often difficult to separate the interpretation. Not only the interpretation provides partial intuition which even if often wrong give a sense of direction to any scientific investigation.

This thesis is aimed at experimental investigation of foundational aspects of quantum mechanics, with careful attention to identifying the baggage that we carry while describing an experiment through the lens of quantum mechanics. There is no attempt made here to propose deeper interpretations. The focus is on having a back and forth conversation from the perspective of theory and experiment so that the conclusions made from the experiment spell out the necessary assumptions in the theory as well. Although not a giant leap, hopefully, this could be one of the many baby steps being taken to better address the foundational aspects of quantum mechanics.

In this introduction chapter, a few foundational concepts of quantum mechanics are revisited with the aim to set the perspectives and premises within which the chapters to

follow hopefully would attain structural coherence and avoid an existential crisis.

The wave-particle duality

At the core of quantum mechanics lies the fact that in any interferometer we can detect what appears to be particles arriving one by one at the detector and collectively the statistics of the ensemble of such identically prepared particles exhibit wave-like interference [I.33]. If we assume the objects to be particles, we need to describe how the ensemble of particles propagates as a wave would. If we assume that what exists is the wave, we have to explain how we detect them one by one in time and space. So what are the objects in a quantum interference experiment – waves or particles?

One may argue, and rightfully so – What's in a name? Both waves and particles are idealizations in physics. We have treated earth to be a point particle revolving around the Sun, when we clearly know it is a giant sphere. We treat ripples on the water surface as waves, ignoring the interactions of many particles at the molecular level. It is not entirely unexpected that at some scale these idealizations would break down. The objects may neither be waves nor particles but say blobs. We may name them "wavicles" or "partwaves" to answer the question of what they are, but how does this help?

But it is not a simple question of classifying these, what we would henceforth be calling quantum particles, as waves or particles. It is about how the properties of quantum particles evolve! How does the quantum particle in a double-slit interferometer get influenced by the two slits at once, which are arguably extended objects, and yet arrive at the detector with spatio-temporal localization? When and how does the transformation of this spatial property occur?

Some of the above questions can be avoided altogether by considering the appropriate interpretation of quantum mechanics. For example, in de Broglie - Bohm's pilot-wave theory [I.34, I.35, I.36], there exists a pilot wave that guides the particle. The waves and particles are separate entities which interact. Because what we can detect is a particle, it is always localized. There is no question of how the spatial property of the particle transformed from being extended to suddenly be localized, because the particle had no extended spatial property to begin with. It is the wave that is extended, goes through the two slits, and guides the particles. The ensemble of such particles statistically follows the

structure of the wave giving rise to the interference pattern. This, so far, is very compelling unless we investigate further to realize that this turns out to be highly non-local[I.37, I.38]. To give up locality would be a huge price to pay in the quest for the foundational consistency of quantum mechanics. Non-locality [I.39, I.40, I.41] should not be considered as the problem that is extrinsic to the quantum mechanics itself, as it is usually discussed in the context of Einsteinian relativity [I.42]. Special Theory of relativity may impose a speed limit but to claim some physical properties to have a non-infinitesimal change instantaneously from another object light-years apart would also make Newton [I.43, I.44] wary who would rather accept something to be very fast rather than instantaneous as demanded by the theory. Therefore, to avoid complications of addressing issues specific to the interpretations, here they shall be posed only considering the postulates in the standard quantum mechanics [I.45, I.3]. The nuances of the Copenhagen interpretation would sometimes spill over to the context under consideration, and hence here we would try to be specific at times at the cost of repeating.

The Measurement problem

If we presume that the quantum particle evolves as a wave, the central mystery boils down to how upon measurement the wave reduces to a localized particle. This may at first appear to be a reiteration of the wave-particle duality, but in essence, applies to anything in general where we have a superposition of probability amplitudes. Waves do superpose, but the concept of superposition is not limited to waves - it applies to vectors, phasors, spinors as well [I.46]. Let us consider a quantum spin-1/2 particle as an example, which is to be made incident on a Stern-Gerlach apparatus [I.47, I.33] that generates the inhomogeneous magnetic field along a given direction (say $\hat{z}$). Let the particle be prepared in a superposition state of the eigenstates corresponding to the Stern-Gerlach apparatus, i.e., say the state is $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$. Upon detection on a screen kept at a distance from the Stern Gerlach apparatus, the particle either appears on the upper side of the screen or on the lower side. Based on which side of the screen the particle appears, we infer that upon measurement the state reduced to either $|\uparrow\rangle$ or $|\downarrow\rangle$. But how did the initial state which was in superposition got reduced to one of the eigenstates? As long as the theory prescribes probabilistic outcomes, we may never be able to answer this question. The very essence

of asking for a mechanism to describe the state reduction would also have to make the theory deterministic. We may very well describe the process statistically by introducing decoherence [I.48, I.49] to evolve a coherent density matrix into an incoherent mixture in a given basis. But that does not answer the evolution of a single particle. We may include the measuring apparatus in our quantum treatment and evolve the joint system-apparatus system, hoping to go forward [I.50]. But as the Schrödinger's cat [I.51] teaches us, this only leads to the superposition of joint system. To explain the state reduction of the system, one has to introduce the state reduction of the apparatus. That is just passing the burden of the question from the system to the apparatus and provides no insight for a possible resolution.

Even if we learn to live with the probabilistic outcome, there are still a few questions which remain unresolved. When does the state reduction exactly occur [I.52]? Does it occur instantaneously or takes finite time [I.53]? In the example above of the spin-1/2 half particle, does the state reduction occur right after the Stern Gerlach magnet or at the screen[I.54] ? One is tempted to treat the Stern Gerlach apparatus as a unitary evolution operator and hence keep the description of the superposition state of the system and apparatus till the particle hits the screen. This may seem consistent till we associate the physical property of the particles with its probability distribution. Just before the screen, if we claim the particle to be in a superposition of two spatial locations, do we also extend the concept to its physical properties like mass or charge. If suddenly the state reduces from a superposition of spatial location to one definite location, does the charge and mass distribution also change immediately? If it were remotely a description of any physical process, should not there be an associated gravitational/electromagnetic current when the mass or charge distribution suddenly collapse? Is there any way to experimentally detect the current associated with the collapse of a single particle?

This begs the fundamental question, do the properties of the particle have the same distribution as the probability of its spatial location ? Silly as it may sound, but as a corollary we may ask, can the properties of the particle be manifested at a separate location where the probability of finding the particle goes to zero? Spoiler Alert! This seemingly silly question shall be pursued for a major part of this thesis, a short introduction to which is provided in a later section in this chapter.

On the completeness of quantum mechanics and hidden variables

Most of the complications while dealing with the measurement problem arises when we try to give an ontological meaning to the wave-function and treat the collapse as a physical process. It may be that the notion of probabilistic description occurs only because the quantum theory is incomplete [I.55]. It may be that, at its core, nature is deterministic [I.56], just that we have not been able to figure out the complete description of nature. We end up with a probabilistic description may be because of the impracticality of accounting for all the details [I.57]. May be, it is just like a coin toss, the outcome of which we are happy to treat in a probabilistic manner, but at the core we understand that if we know the mass distribution of the coin and the exact force applied when we toss the coin, we should be able to predict the outcome of a coin toss. Just that we have not accounted these details, as a result of which we currently have a theory that is indeterministic [I.58]. It is certainly possible that at the fundamental level, the interactions involve non-linearity [I.59, I.60, I.61] and the dynamics becomes chaotic [I.62, I.63] and thus all we can precisely address is the property of the distribution such as the average (expectation value) and the standard deviation (uncertainty) without precisely predicting each outcome. It is like our theories are good enough to predict how much the sea level would rise over the next century, but unable to predict at which minute of the clock it would rain.

Rather than searching for a deeper fundamental theory from scratch, a simpler approach is to hope that we are in an Aristotelian age [I.64] of quantum mechanics where we have not yet figured out the complete description of state. In the physics of Aristotle [I.65], there were two laws of the motion - one for the celestial bodies, which move in perfect orbits without the need of any force. The other one for terrestrial objects for which force was needed to cause unnatural motion. Natural motion, like vertical fall, did not require any force. Similarly, in quantum mechanics we have two dynamical laws – the deterministic unitary evolution of the state under the Schrödinger equation and the probabilistic non-unitary state reduction. For Aristotle, the description of a state of motion involved only position. It took a millennium to realize that the complete description of a state of motion must also include the velocity or momentum along with the position. With this realization, we had the Newton's laws of motion, which unified the laws of terrestrials with that of

the heavenly. In the present standard description of quantum mechanics, we don't know whether and at what scale the Heisenberg's cut [I.66, I.67] exists as we struggle to have a consistent description of quantum to classical transition [I.68, I.69]. The classical and the quantum world can be unified if we make the quantum theory deterministic. This would lead to quantum particles have definite property for all the time, leading to trajectories for observables. For this, the description of the quantum state may be completed by introduction of hidden variables that supplement the wavefunction ψ. For example, in Bohmian mechanics [I.70], the position of the particle is introduced as the hidden variable that completes the description of the quantum state. The coupled dynamical equations for evolution of the wavefunction and the hidden variable makes the quantum state evolve deterministically, thus providing a definite value to the observables over time.

It may be tempting to consider that the experimental violation [I.71, I.72, I.73] of Bell's inequality rule out all the possible hidden variable theories. However, along with local realism another assumption that goes into deriving the Bell expression is the absence of super-determinism [I.74, I.75]. If everything in the universe is deterministic, just like the classical Newtonian world, there would not exist freedom of choice and hence Bell violation would not rule out super-deterministic quantum theory.

However, from the perspective of experiments, if we do not have easy access to measure the hidden variables or modify them during state preparation, the theory would be less useful. But do we have any alternative? What would we have done in classical mechanics if all we could measure is the initial position, and it was impractical to determine the initial momentum? How would we have known the trajectory of the observables? Well, we have already done this using the Lagrangian formulation. Instead of considering the complete initial state consisting of position and momentum, we consider the incomplete state consisting of position only but at two points in time - basically converting the initial value problem to a boundary value problem in time. Similarly, we can use the knowledge of incomplete quantum state at two points in time to describe the property of the quantum particle at all times in between. These properties can be interpreted as weak values, and the measurement schemes to determine them are described in the next two chapters.

The Quantum Cheshire cat

Consider a double-slit interferometer on which quantum particles are made incident one at a time. They are also detected one at a time and at specific locations of the detector. But as a function of detector position, the distribution of particles emerges as an interference pattern. This compels us to consider that every individual particle during its propagation is in superposition state of having gone through both the slits. Do the properties of the particle also are in a superposition state of being located at both the slit along with the particle? But how do we test it? We can place a particle detector in one of the slit and the property detector in the other slit. But both the detectors may never click together, indicating that the property of the particle always resides where the particle is. But it may so happen that, because we placed the detectors, the particle never entered the superposition state. We would never know unless we evolve the particles through to the detector without demolition and then obtain signatures of interference. What we can also do is to place markers along the slits so that the particles are tagged [I.76] of which slit they went through. Say, we start with spin -1/2 particles pointing up $|\uparrow\rangle$. Then along one of the slits (slit B) we place a spin flipper so that the particles going through that slit have spin down $|\downarrow\rangle$. We know that each particle went through a superposition state of going through both the slits because if we project each particle to left $|\leftarrow\rangle$ or right $|\rightarrow\rangle$ spin, we obtain the interference pattern. Thus, the act of tagging the particles as such has not prevented the particles from being in a superposition state of two spatial locations. Now if we consider only the spin up $|\uparrow\rangle$ particles at the time of detection by placing a spin filter before the detector which is placed at the location where constructive interference was expected, we would infer that they must have come from the slit A. Any particle coming from slit B would have spin down $|\downarrow\rangle$ and hence would be blocked by the spin filter. Thus, one would also guess that all the properties of the particle must also have come through slit A. But can some property of the detected particles be in slit B?

If this were the case, we would have properties of the particle separated from the physical location. Just like the Cheshire Cat [I.77] in “Alice in Wonderland" whose grin can exist disembodied from the body of the cat itself, the properties of such a quantum state can be present in a location where the particle itself is not present. Such states can be

referred to as quantum Cheshire cat states.

But how do we obtain signatures of spin in slit B? Well, one may try to rotate the spin of the particles going through slit B. We would observe a change in intensity in the detected particles. Thus, apparatus aimed towards changing the spin property along slit B is apparently affecting the number of detected particles. Surely, this is the case, but now we cannot claim that all the particles that are detected have come only through slit A because the spin of particles coming through slit B is no longer orthogonal to them. Rotating the spin along slit B changes the number of detected particles, but since the detected particles itself can come through slit B, we cannot conclude anything regarding whether the properties of the particle are always associated with the physical location of the particle.

What we need is a measurement scheme which does not disturb the system [I.78], or at least, that does not change the ensemble of particles which are detected. The simplest way is to choose an operation that does not change the spin of the particles. We can place a σ_z Stern-Gerlach magnet after slit A that shifts the particles with spin $|\uparrow\rangle$ upwards. We would notice that the ensemble of detected particles has moved up in the screen. Thus, the σ_z component of the spin of the particles still resides with the particles coming from slit A. But if the upward shift is so much that with a single run we can identify that the particle has spin $|\uparrow\rangle$, then by rotating the spin filter we would not observe any interference as well. This is because now the particles coming from the different slits do not have any spatial overlap anymore. Thus, we cannot presume the quantum particles are in a superposition state of being through the two-slits either.

Thus, we need to devise a measurement scheme in which we can maintain the superposition state and still extract information about the spin of the particle. Once done, we need to measure all the Pauli spin components of the particles coming from slit B. If we find that shifting the particles B with any spin component creates a change in the spatial distribution of detected particles, we would obtain the signature of the quantum Cheshire cat.

But to be able to infer the spin of the particle in a given slit without destroying the coherence, we need to explore the measurement scheme known as *weak measurement* in Ch 1 and its optical implementation in polarization degree of freedom in Ch 2. Then we shall return to the quantum Cheshire cat effect in Ch 3 and Ch 4 to study it in detail

and understand the experimental implementations of it. We shall argue that, despite the existence of many experiments, the Cheshire cat effect has not yet been demonstrated in a loophole-free manner. The chapters would be regarding the experimental implementation of the Cheshire cat effect in semi-classical and quantum optics domain.

Expectation value of Non-Hermitian Operators

In quantum mechanics, it is postulated that the only result of a precise measurement of a dynamical variable is one of the eigenvalues of the linear operator associated with the observable [I.3]. Often it is argued that the eigenvalues of the operator representing the dynamical variable must be real and hence they are required to be Hermitian. However, the use of Hermitian operator also guarantees that, for a non-degenerate spectrum, the eigenvectors are orthogonal and forms a complete set [I.45]. Since a non-Hermitian operator does not assure us of a complete set of eigenbasis, we do not use them as observables. Further, from measurement we expect to gain information from the outcome. Thus, all possible outcomes must be distinguishable from each other. This can be regarded as the manifestation of the orthogonality of the set of eigenstates of the operator representing the observable.

In general, we can measure complex quantities in experiment by measuring the real and imaginary part separately. But in quantum mechanics, those measurements must be compatible. Thus, even if a non-Hermitian operator is decomposed into real and imaginary part, one would need to ensure that the two commute, which often is not the case [I.79]. However, let's consider an example in which they do. For simplicity, we would represent the operators for a qubit in the basis $\{|0\rangle, |1\rangle\}$. The projectors $\Pi_0 = |0\rangle\langle 0|$ and $\Pi_1 = |1\rangle\langle 1|$ trivially commute by the virtue of being orthogonal. Now, let's consider the operator $A = \Pi_0 + i\Pi_1$. Is this a valid observable? It's eigen vectors are orthogonal $\{|0\rangle, |1\rangle\}$ and forms a complete set. But, one of its eigenvalue from the set $\{1, i\}$ is not real. Nevertheless, given that $[\Pi_0, \Pi_1] = 0$, we may in principle measure them without disturbing the system. Thus, would a simultaneous measurement of Π_0 and Π_1 be considered as a measurement of the non-Hermitian operator A ? Let's investigate the same in terms of what information do we get in the process of measurement. Given as state $|\psi\rangle$, the

probability that after the process of measurement, the state reduces to $|0\rangle$ (or/and $|1\rangle$) is the same regardless of whether we describe the operator associated with the measurement to be the Hermitian unitary $\sigma_z = \Pi_0 - \Pi_1$ or the non-Hermitian A. This is because all that the coefficient i in front of Π_1 in the expression for A does is add a relative phase to the $|1\rangle$ component with respect to $|0\rangle$. This information about the relative phase gets lost when we perform projective measurements [I.80, I.81]. Thus, even though we can realize the operator A in an experiment[1], we would not get any different information had we performed measurements associated with σ_z observable. Although the experimental result may not distinguish, we prefer σ_z instead of A, because the description in terms of Hermitian operator, makes the evolution unitary as well [I.84]. Consequently, we cannot measure the expectation value of a non-Hermitian operator from the statistics of outcomes like we prescribe for Hermitian observables.

The average of specific non-Hermitian operators are useful in determination of unknown quantum state [I.85]. Since the average of the non-hermitian operator is in general complex, the initial approach would be to relate to the measurable quantities in the experiment which are complex. The *weak values* which are the result of *weak measurements*, as we shall dwell further in Ch 1, are in general complex. The weak value of an observable R, obtained as the Hermitian part during the polar decomposition of the non-Hermitian operator A, turns out to be the expectation value of the non-Hermitian operator A , as will be shown in Ch 5. However, there exists an interferometric method using which we can obtain the expectation value of A. The concerned experimental methods and results are discussed in Ch 5.

The Born rule & the superposition principle

Most of the phenomenon in quantum mechanics which do not have classical analogue emerges from the quantum state of a particle being in superposition of at least two states realizable in classical physics. The superposition principle is a consequence of the dynamical equation in theory being linear in terms of the quantum state [I.86]. But the quantum

[1]The specific operator A corresponds to the Jones matrix of a quarter-wave plate(QWP). Nevertheless, QWP manifests as an evolution operator. But, we can have a birefringent crystal[I.82, I.83] in which the ordinary and extraordinary ray have spatial separation as well as phase shifts that are analogous to the QWP

state as such is not directly observed in the experiment and needs to be inferred from various measurements. The probability of measurement outcomes is given by the Born rule [I.87]. The interference pattern in a double slit experiment, as per many textbooks, is obtained after the application of the Born rule to the superposition of solutions where individual slits are open [I.33]. However, the two cases of the individual slits being opened correspond to two different boundary conditions for the differential equation. Thus, the superposition principle, strictly speaking, does not hold.

But, we have used the superposition principle and Born rule to compute the interference pattern for almost a century. Thus, the deviation to the intensity in the interference pattern resulting from this naive application of the superposition principle, when the boundary conditions differ, must be negligible. But establishing whether this deviation exists is difficult because one would have to compare the experimental data with predictions from theory using accurate modeling of the setup. One may find a difference between experimental data and theoretical prediction because of uncertainties and inaccuracies in parameters used in the experiment. However, if one uses a triple slit setup, we can express the deviation, which we call κ, from what is known as the third-order sum rule, using the sum and difference of intensities in various combinations of slits being open or closed. If the superposition principle and Born rule were applicable to all the combinations of slits being open or closed, we would expect this deviation to be zero, which is independent of experimental parameters. If the superposition principle were applicable, measuring a deviation in the triple-slit setup would have raised concerns over the Born rule.

Nevertheless, from the theory of differential equations, we know the superposition principle is strictly not applicable. However, since it has been used for such a long time, the deviation would be small. Thus, it is important to figure out parameters on which this deviation depends on, which were earlier reported to be mainly slit parameters and the wavelength [I.88]. By performing an experiment with suitable choice of slit width, inter-slit distance and wavelength, it is possible to measure the deviation to the naive application of the superposition principle, by measuring κ to be non-zero over and above the experimental uncertainties. Such an experiment was performed in microwave domain and the deviation from naive application of superposition principle due to geometrical effects was reported. This is discussed in Ch.6.

References

[I.1] Robert M Eisberg and Robert Resnick. *Quantum physics of atoms, molecules, solids, nuclei, and particles*. en. 2nd ed. Nashville, TN: John Wiley & Sons, 1985.

[I.2] Nouredine Zettili. *Quantum Mechanics*. en. 2nd ed. Hoboken, NJ: Wiley-Blackwell, 2009.

[I.3] B. H. Bransden and C. J. Joachain. *Quantum mechanics*. Pearson, 2017.

[I.4] J. S. Bell and Alain Aspect. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. 2nd ed. Cambridge University Press, 2004. DOI: 10.1017/CBO9780511815676.

[I.5] Anders Johansson et al. ""Shut up and calculate": the available discursive positions in quantum physics courses". In: *Cultural Studies of Science Education* 13.1 (2016), pp. 205–226. DOI: 10.1007/s11422-016-9742-8.

[I.6] N. David Mermin. "Could Feynman Have Said This?" In: *Physics Today* 57.5 (2004), pp. 10–11. DOI: 10.1063/1.1768652.

[I.7] Shabnam Siddiqui and Chandralekha Singh. "How diverse are physics instructors' attitudes and approaches to teaching undergraduate level quantum mechanics?" In: *European Journal of Physics* 38.3 (2017), p. 035703. DOI: 10.1088/1361-6404/aa6131.

[I.8] Stillman Drake. *Galileo at work*. en. Mineola, NY: Dover Publications, 1995.

[I.9] Wilton L Virgo. *Quantum mechanics in everyday life*. Rhodium, 2014.

[I.10] Miguel Navascués and Harald Wunderlich. "A glance beyond the quantum model". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2115 (2010), pp. 881–890. DOI: 10.1098/rspa.2009.0453.

[I.11] Abhay Ashtekar. *Conceptual problems of quantum gravity : based on the proceedings of the 1988 Osgood Hill Conference, North Andover, Massachusetts, 15-19 May 1988*. Boston: Birkhäuser, 1991. ISBN: 9783764334437.

[I.12] G. A. D. Briggs, J. N. Butterfield, and A. Zeilinger. “The Oxford Questions on the foundations of quantum physics”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469.2157 (2013), p. 20130299. DOI: 10.1098/rspa.2013.0299.

[I.13] “SOME CONCEPTUAL ISSUES IN QUANTUM THEORY”. In: *Lectures on Quantum Theory*, pp. 129–160. DOI: 10.1142/9781860944093_0008.

[I.14] Guido Bacciagaluppi and Antony Valentini. *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*. Cambridge University Press, 2009. DOI: 10.1017/CBO9781139194983.

[I.15] E. Schrödinger. “An Undulatory Theory of the Mechanics of Atoms and Molecules”. In: *Phys. Rev.* 28 (6 1926), pp. 1049–1070. DOI: 10.1103/PhysRev.28.1049.

[I.16] Lluis Masanes, Thomas D. Galley, and Markus P. Muller. “The measurement postulates of quantum mechanics are operationally redundant”. In: *Nature Communications* 10.1 (2019). DOI: 10.1038/s41467-019-09348-x.

[I.17] Julian B Barbour. *The Discovery of Dynamics*. en. New York, NY: Oxford University Press, 2001.

[I.18] Julian B Barbour and Herbert Pfister, eds. *Mach’s principle*. en. 1995th ed. Einstein Studies. Secaucus, NJ: Birkhauser Boston, 1995.

[I.19] D. W. Sciama. “On the Origin of Inertia”. In: *Monthly Notices of the Royal Astronomical Society* 113.1 (1953), pp. 34–42. ISSN: 0035-8711. DOI: 10.1093/mnras/113.1.34.

[I.20] Max Jammer. *Concepts of Force*. en. Dover Books on Physics. Mineola, NY: Dover Publications, 2003.

[I.21] Max Jammer. *Concepts of mass in classical and modern physics*. Mineola, NY: Dover Publications, 2003.

[I.22] Maximilian Schlosshauer, Johannes Kofler, and Anton Zeilinger. “A snapshot of foundational attitudes toward quantum mechanics”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern*

Physics 44.3 (2013), pp. 222–230. ISSN: 1355-2198. DOI: https://doi.org/10.1016/j.shpsb.2013.04.004.

[I.23] Y. Aharonov, J. Anandan, and L. Vaidman. "Meaning of the wave function". In: *Phys. Rev. A* 47 (6 1993), pp. 4616–4626. DOI: 10.1103/PhysRevA.47.4616.

[I.24] M. Ringbauer et al. "Measurements on the reality of the wavefunction". In: *Nature Physics* 11.3 (2015), pp. 249–254. DOI: 10.1038/nphys3233.

[I.25] Matthew Saul Leifer. "Is the Quantum State Real? An Extended Review of ψ-ontology Theorems". In: *Quanta* 3.1 (2014), p. 67. DOI: 10.12743/quanta.v3i1.22.

[I.26] Arkady Plotnitsky and Andrei Khrennikov. "Reality Without Realism: On the Ontological and Epistemological Architecture of Quantum Mechanics". In: *Foundations of Physics* 45.10 (2015), pp. 1269–1300. DOI: 10.1007/s10701-015-9942-1.

[I.27] Florian Boge. *Quantum Mechanics Between Ontology and Epistemology*. Cham: Springer, 2018. ISBN: 978-3319957647.

[I.28] Wayne Myrvold. "Philosophical Issues in Quantum Theory". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2018. Metaphysics Research Lab, Stanford University, 2018.

[I.29] D. J. Bohm and B. J. Hiley. "The de Broglie pilot wave theory and the further development of new insights arising out of it". In: *Foundations of Physics* 12.10 (1982), pp. 1001–1016. DOI: 10.1007/BF01889273.

[I.30] Jan Faye. "Copenhagen Interpretation of Quantum Mechanics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2019. Metaphysics Research Lab, Stanford University, 2019.

[I.31] Graham P Collins. "The many interpretations of quantum mechanics". In: *Scientific American* 297.5 (2007), p. 19.

[I.32] Jack Sarfatti. “Progress in Post-Quantum Physics and Unified Field Theory”. In: *Gravitation and Cosmology: From the Hubble Radius to the Planck Scale*. Ed. by Richard L. Amoroso et al. Dordrecht: Springer Netherlands, 2002, pp. 419–430. ISBN: 978-0-306-48052-2.

[I.33] Richard Phillips Feynman, Robert Benjamin Leighton, and Matthew Sands. *The Feynman lectures on physics Vol III; New millennium ed.* Originally published 1963-1965. New York, NY: Basic Books, 2010.

[I.34] David Bohm. “A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I”. In: *Phys. Rev.* 85 (2 1952), pp. 166–179. DOI: 10.1103/PhysRev.85.166.

[I.35] David Bohm. “A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II”. In: *Phys. Rev.* 85 (2 1952), pp. 180–193. DOI: 10.1103/PhysRev.85.180.

[I.36] Peter R. Holland. *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics*. Cambridge University Press, 1993. DOI: 10.1017/CBO9780511622687.

[I.37] Detlef Dürr et al. “Can Bohmian mechanics be made relativistic?” In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2162 (2014), p. 20130699. DOI: 10.1098/rspa.2013.0699.

[I.38] Dylan H. Mahler et al. “Experimental nonlocal and surreal Bohmian trajectories”. In: *Science Advances* 2.2 (2016), e1501466. DOI: 10.1126/sciadv.1501466.

[I.39] A. Einstein, B. Podolsky, and N. Rosen. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: *Phys. Rev.* 47 (10 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[I.40] J. S. Bell. “On the Einstein Podolsky Rosen paradox”. In: *Physics Physique Fizika* 1 (3 1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[I.41] Eric G. Cavalcanti and Howard M. Wiseman. “Bell Nonlocality, Signal Locality and Unpredictability (or What Bohr Could Have Told Einstein at Solvay Had

He Known About Bell Experiments)". In: *Foundations of Physics* 42.10 (2012), pp. 1329–1338. DOI: 10.1007/s10701-012-9669-1.

[I.42] G. C. Ghirardi, A. Rimini, and T. Weber. "A general argument against superluminal transmission through the quantum mechanical measurement process". In: *Lettere al Nuovo Cimento* 27.10 (1980), pp. 293–298. DOI: 10.1007/bf02817189.

[I.43] Isaac Newton. *The principia : mathematical principles of natural philosophy*. Place of publication not identified: Snowball Pub, 2010. ISBN: 978-1607962403.

[I.44] Douglas W. MacDougal. *Newton's Gravity*. Springer New York, 2012. DOI: 10.1007/978-1-4614-5444-1.

[I.45] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. 3rd ed. Cambridge University Press, 2018. DOI: 10.1017/9781316995433.

[I.46] S. Massar and E. S. Polzik. "Generating a Superposition of Spin States in an Atomic Ensemble". In: *Phys. Rev. Lett.* 91 (6 2003), p. 060401. DOI: 10.1103/PhysRevLett.91.060401.

[I.47] J. F. Van Huele and Jared R. Stenson. "Stern-Gerlach experiments : past , present , and future". In: 2004.

[I.48] Wojciech Hubert Zurek. "Decoherence, einselection, and the quantum origins of the classical". In: *Rev. Mod. Phys.* 75 (3 2003), pp. 715–775. DOI: 10.1103/RevModPhys.75.715.

[I.49] Maximilian Schlosshauer. "Decoherence, the measurement problem, and interpretations of quantum mechanics". In: *Rev. Mod. Phys.* 76 (4 2005), pp. 1267–1305. DOI: 10.1103/RevModPhys.76.1267.

[I.50] John Neumann. *Mathematical foundations of quantum mechanics*. Princeton N.J: Princeton University Press, 1955. ISBN: 978-0691028934.

[I.51] John D. Trimmer. "The Present Situation in Quantum Mechanics: A Translation of Schrödinger's "Cat Paradox" Paper". In: *Proceedings of the American Philosophical Society* 124.5 (1980), pp. 323–338. ISSN: 0003049X.

[I.52] P.J. Bussey. “When does the wavefunction collapse?” In: *Physics Letters A* 106.9 (1984), pp. 407–409. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(84)90979-4.

[I.53] A Yu Ignatiev. “How fast is the wave function collapse?” In: *Journal of Physics: Conference Series* 410 (2013), p. 012153. DOI: 10.1088/1742-6596/410/1/012153.

[I.54] Marlan O. Scully, R. Shea, and J.D. McCullen. “State reduction in quantum mechanics: A calculational example”. In: *Physics Reports* 43.13 (1978), pp. 485–498. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(78)90210-7.

[I.55] A. Einstein, B. Podolsky, and N. Rosen. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: *Phys. Rev.* 47 (10 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[I.56] Carl Hoefer. “Causal Determinism”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2016. Metaphysics Research Lab, Stanford University, 2016.

[I.57] Marco Genovese. “Research on hidden variable theories: A review of recent progresses”. In: *Physics Reports* 413.6 (2005), pp. 319–396. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2005.03.003.

[I.58] Brigitte Falkenburg and Friedel Weinert. “Indeterminism and Determinism in Quantum Mechanics”. In: *Compendium of Quantum Physics*. Ed. by Daniel Greenberger, Klaus Hentschel, and Friedel Weinert. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 307–311. ISBN: 978-3-540-70626-7. DOI: 10.1007/978-3-540-70626-7_96.

[I.59] Alexey V. Melkikh. “Nonlinearity of Quantum Mechanics and Solution of the Problem of Wave Function Collapse”. In: *Communications in Theoretical Physics* 64.1 (2015), pp. 47–53. DOI: 10.1088/0253-6102/64/1/47.

[I.60] Gadi Fibich. *The nonlinear Schrödinger equation : singular solutions and optical collapse*. Cham: Springer, 2015. ISBN: 978-3-319-12747-7.

[I.61] Dieter Schuch. *Quantum Theory from a Nonlinear Perspective*. Vol. 191. Springer, 2018. DOI: 10.1007/978-3-319-65594-9.

[I.62] Steven Strogatz. *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering*. Boulder, CO: Westview Press, a member of the Perseus Books Group, 2015. ISBN: 978-0813349107.

[I.63] M. C. Gutzwiller. *Chaos in classical and quantum mechanics*. New York: Springer-Verlag, 1990. ISBN: 978-1-4612-0983-6.

[I.64] Helen Lang. *The order of nature in Aristotle's physics : place and the elements*. Cambridge New York: Cambridge Unviersity Press, 1998. ISBN: 9780521624534.

[I.65] Aristotle. *Aristotle's physics*. Lincoln: University of Nebraska Press, 1961. ISBN: 978-0803250932.

[I.66] Wojciech H. Zurek. "Decoherence and the Transition from Quantum to Classical". In: *Physics Today* 44.10 (1991), pp. 36–44. DOI: 10.1063/1.881293.

[I.67] George F.R. Ellis. "On the limits of quantum theory: Contextuality and the quantum–classical cut". In: *Annals of Physics* 327.7 (2012). July 2012 Special Issue, pp. 1890–1932. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2012.05.002.

[I.68] Maximilian Schlosshauer and Kristian Camilleri. "The quantum-to-classical transition: Bohr's doctrine of classical concepts, emergent classicality, and decoherence". In: *arXiv: Quantum Physics* (2008).

[I.69] Maximilian Schlosshauer. *Decoherence and the quantum-to-classical transition*. Berlin London: Springer, 2007. ISBN: 978-3540357735.

[I.70] Sheldon Goldstein, Roderich Tumulka, and Nino Zanghi. *Bohmian Trajectories as the Foundation of Quantum Mechanics*. 2009.

[I.71] Marissa Giustina et al. "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons". In: *Phys. Rev. Lett.* 115 (25 2015), p. 250401. DOI: 10.1103/PhysRevLett.115.250401.

[I.72] B. Hensen et al. “Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres”. In: *Nature* 526.7575 (2015), pp. 682–686. DOI: 10.1038/nature15759.

[I.73] Alain Aspect. “Experimental Tests of Bell’s Inequalities in Atomic Physics”. In: *Atomic Physics 8*. Ed. by Ingvar Lindgren, Arne Rosén, and Sune Svanberg. Boston, MA: Springer US, 1983, pp. 103–128. ISBN: 978-1-4684-4550-3.

[I.74] Sabine Hossenfelder. “Superdeterminism: A Guide for the Perplexed.” In: *arXiv: Quantum Physics* (2020).

[I.75] Jan-Åke Larsson. “Loopholes in Bell inequality tests of local realism”. In: *Journal of Physics A: Mathematical and Theoretical* 47.42 (2014), p. 424003. DOI: 10.1088/1751-8113/47/42/424003.

[I.76] Yoon-Ho Kim et al. “Delayed “Choice” Quantum Eraser”. In: *Phys. Rev. Lett.* 84 (1 2000), pp. 1–5. DOI: 10.1103/PhysRevLett.84.1.

[I.77] Lewis Carroll and Hugh Haughton. *Alice in Wonderland ; and, Through the looking glass and what Alice found there*. Penguin Classics, 2013.

[I.78] C. S. Unnikrishnan. “Quantum non-demolition measurements: Concepts, theory and practice”. In: *arXiv e-prints*, arXiv:1811.09613 (2018), arXiv:1811.09613.

[I.79] P. A. M. Dirac. *The principles of quantum mechanics*. Oxford England: Clarendon Press, 1981. ISBN: 978-0198520115.

[I.80] Michael Nielsen. *Quantum computation and quantum information*. Cambridge New York: Cambridge University Press, 2010. ISBN: 978-1107002173.

[I.81] John Wheeler. *Quantum theory and measurement*. Princeton, New Jersey: Princeton University Press, 1983. ISBN: 978-0691613161.

[I.82] Max Born. *Principles of optics : electromagnetic theory of propagation, interference, and diffraction of light*. Cambridge: Cambridge University Press, 2019. ISBN: 978-1-108-47743-7.

[I.83] Robert Boyd. *Nonlinear Optics*. London: Academic Press, 2020. ISBN: 978-0128110027.

[I.84] J. von Neumann, R.T. Beyer, and N.A. Wheeler. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 2018. ISBN: 9781400889921.

[I.85] Eliot Bolduc, Genevieve Gariepy, and Jonathan Leach. "Direct measurement of large-scale quantum states via expectation values of non-Hermitian matrices". In: *Nature Communications* 7.1 (2016). DOI: 10.1038/ncomms10439.

[I.86] Peter Olver. *Applied linear algebra*. Cham, Switzerland: Springer, 2018. ISBN: 9783319910413.

[I.87] Daniel Greenberger. *Compendium of quantum physics : concepts, experiments, history and philosophy*. Dordrecht New York: Springer, 2009. ISBN: 978-3-540-70622-9.

[I.88] Aninda Sinha, Aravind H. Vijay, and Urbasi Sinha. "On the superposition principle in interference experiments". In: *Scientific Reports* 5.1 (2015). DOI: 10.1038/srep10304.

Chapter 1

Weak Measurements & Weak Value

Contents

1.1 Time-reversal symmetry of dynamical laws 26

1.2 Time symmetric approach towards quantum mechanics 31

1.3 Properties of pre- and post-selected ensembles - weak value 39

1.3.1 Simultaneous Measurements 40

1.3.2 Imposing Time Symmetry on the value of Observables 42

1.3.3 Discontinuity in potential breaking Time Reversal Symmetry : Classical analogue 43

1.4 Weak Measurements 48

1.4.1 von Neumann Measurement Interaction 50

1.4.2 Weak Interaction 56

1.4.3 Pointer shift after Post-selection 58

1.5 Weak value 61

1.5.1 ABL rule 62

1.5.2 Two-state vector formalism 68

1.5.3 Complex weak values and Pointer Shift - Josza theorem 70

The generalized statistical interpretation of quantum mechanics does not provide us the tools to predict the definite state that would be the outcome of a given measurement process [1.1]. May be nature is fundamentally non-deterministic in the sense that the infinitesimal time evolution is not a one-one map, i.e., the same initial state under the same

conditions may evolve to multiple final states. Or may be our description of the quantum state is not complete [1.2]. Note that, there is a subtle difference in asking whether a theory provides a mechanism to calculate the exact value of an observable versus asking whether it provides a unique quantum state as a result of measurement. Asking whether there exists a definite value of the observable at all times is burdened with a few other presumptions that we make about the system. Firstly, if we insist that a quantum particle has a definite position before the act of measurement, we are presuming that our notion of point particle do not break down in the quantum world. It may be that the quantum object is an extended wave-packet. Therefore, asking what is the exact position of the particle may need additional qualifiers about the structure of the object under consideration. For example, when we say the distance between two stars in a galaxy is certain light years, we implicitly understand that the notion of this distance is valid as long as the precision with which we are inspecting allows us to approximate the stars as point objects. It is often meaningless to attribute the distance between two intergalactic objects with precision down to few millimeters. And even if we do, we respect the fact that the star has a structure and our attribution of the distance measure is only valid as an average. Therefore, extending this limitation to the quantum world is neither surprising nor fundamentally a concern.

What is disturbing is the inability to predict the outcome even if we are not hunting for the precision and at the same time claiming that no further information would be of much use. If we place a quantum particle in a box and then insert a partition in the box. We divide the box into two and take them far apart. In quantum mechanics, we claim that the particle is in a superposition of the state of being in the two boxes. Hence, there is no definite answer to the question of which box the particle is actually in. We can only have the answer once a measurement is made or the box is open. Note that the question is not about the precision of the measurement. We are not asking where exactly the particle is in the box. We are merely asking in which box it is and the boxes are allowed to be as large as possible. Quantum mechanics does not give a definite prediction. But, practically, we would almost always not be able to predict the answer even if it were a classical ball instead of a quantum particle. We would need additional information like how exactly the ball was dropped in the box and how the partitions were inserted. This information can be used to simulate the process using the laws of classical mechanics and then predict the

answer. However, the standard Quantum mechanics would insist that no such additional information would be of any help and the description of the quantum state is complete as there is nothing else to add to it [1.3].

A complete theory must be able to predict, in principle, with the certainty of the outcomes that are guaranteed to be obtained in an experiment [1.2]. This implies that a complete theory must account for enough information at present to compute the uniquely determined future. But is the time evolution in a complete theory needs to be described by a one-one map between the present and the future? Or it is allowed that the possibility that different states at present can evolve into the same state in a future time and hence subsequent evolution would be identical. In classical mechanics, all the dynamical laws happen to create a one-one map between the present and the future. Thus, not only the future can be predicted with the complete information at present, but also the past can be retrodicted with the same information and knowledge of all the forces that affects the dynamics.

In quantum mechanics, all the evolution unless a measurement is made is described by a unitary map. This one-one map maintains the orthogonality of the states. Thus, if we know a quantum state at present undergoing unitary evolution, we can not only compute the future state but also retrodict what state it was in the past. However, quantum mechanics describes measurement differently [1.4]. Often it is treated as a black box that acts as that reduces the state to one of the possible outcomes with a defined probability. This means two orthogonal states are allowed to be reduced to the same state after the process of measurement. For example, a horizontally polarized photon can end up being in a Diagonal polarization state after passing through a Polarizer oriented accordingly. But so can the Vertically polarized photon. This evolution is no longer a one-one map and hence cannot be inverted. Knowing the result of measurement do not give us any clue about what the state was before the measurement. We cannot even assign probability weights to what the state of the photon could have been prior to measurement. This introduces time asymmetry at a fundamental level in quantum mechanics [1.5].

But what is a measurement, and at what time the measurement is said to be complete? Often it is argued that once the photon is detected in the sensor, we can claim that the measurement is completed. Measurement involves the interaction of the system with an

apparatus, and in the quantum description they become entangled [1.6]. The measurement is complete when the entanglement breaks and the system is reduced to a definite state. But where does this process occur physically? Do the Silicon atoms in a sensor carry out this procedure? If so, what makes them undergo non-unitary evolution when placed inside a sensor and undergo unitary evolution when used as a quantum dot? Could it be possible that although the systems are prepared identically, the measurement process was not identical for all the runs of the experiment, therefore leading to different results. In general, the initial state of any detector is described by a state which is classically admissible (e.g., no click) so that upon detection the state changes to another classical state accessible to us (e.g., click) so that the measurement can be recorded. The randomness in preparing the initial state in the detector may explain the difference in outcomes of otherwise identically prepared systems. However, on a statistical level, an ensemble of pure states upon measurement would reduce to a mixed state and the process can be described as Decoherence [1.7]. But at a fundamental level of individual particle, the notion of measurement begs further investigation.

In this chapter [1], we shall be particularly interested in exploring how the notion of weak measuremen can be developed that respects the time symmetry. Weak measurements in the time-symmetric formulation of quantum mechanics [1.11] promise to predict with certainty the 'weak value' of an observable when the system is subjected to boundary conditions in time. Therefore, it provides inquisition towards having a complete description of quantum mechanics with 'weak values' as the element of physical reality [1.12].

1.1 Time-reversal symmetry of dynamical laws

The dynamical law of quantum state reduction upon measurement not only introduces non-determinism but also makes the theory time-reversal asymmetric [1.5]. Given the state of the system at time $t = t_0$ and all the dynamical causal agents at all times, laws of dynamical motion in classical mechanics, such as Newton's 2nd Law of motion, help us deterministically predict the state of the system at any time $t > t_0$ as it evolves into the future. In classical mechanics, there do not seem to be a fundamental direction to time as

[1] The reader may prefer these review articles [1.8, 1.9, 1.10] of the concept of weak measurements and weak values

such [1.13]. The Newton's 2nd law of motion, for example, is identical under the time reversal transformation [1.14] of $t \to -t$. The time reversal transformation in Hamiltonian mechanics involve transformation of other dynamical observables as well i.e., we need to transform $p \to -p$ along with $t \to -t$. Nevertheless, upon the transformation, the laws of motion remains the same . As a result, we can use the same laws to retrodict what the state of the system was in the past at the time $t < t_0$. Both prediction and retrodiction using classical laws of motion gives us uniquely determined states and hence the value of all observables at all times. Nevertheless, there are laws in the classical world that are time asymmetric. The second law of thermodynamics seems to be giving a direction to time [1.15]. Entropy seems to only increase forward in time. However, this law is the statistical law of an emergent phenomenon where an ensemble of large number of particles are involved. Spontaneous ordering is less probable in the sense that the number of possibilities in the configuration space which describes order is very less compared to the number of disordered configurations [1.13]. In principle, if at anytime all the momentum of particles are reversed, one can have the initial configuration. If the second law of thermodynamics gets violated in one specific experiment, the fundamentals of classical mechanics do not need to be changed. In fact, for a small amount of time, there is a reasonable probability of spontaneous decrease of entropy for small enough ensembles [1.16]. Thus, all the seemingly non-reversible processes like burning of paper, mixing of solutions etc. can be described as emergent statistical descriptions of large ensembles where fundamental interactions still involves dynamical laws that respect time reversal symmetry.

One of the dynamical laws of the quantum world – the Schrödinger's equation $H|\psi\rangle = i\partial_t|\psi\rangle$ respects time reversal symmetry [1.17, 1.18] with the time-reversal operator transforming the imaginary number i to $-i$ along with $t \to -t$. Given a state $|\psi\rangle_0$ at time $t = t_0$, and the Hamiltonian H for all values of time t, we can deterministically evolve the quantum state forward to predict the quantum state $|\psi\rangle_f(t > t_0)$ for all future times $t > t_0$. Because the equation is invariant under time-reversal transformation, we can also evolve the quantum state into the past to calculate the state $|\psi\rangle_b(t < t_f)$. Thus, the Schrodinger's equation do not mandate a direction of flow of time, just like the fundamental laws of motion in classical mechanics. In classical mechanics, if we take snapshots of

a ball freely falling under gravity, we can reverse the arrangement of the snapshots and still it would appear as a valid physical phenomenon. It would be impossible to recognize the order of time from the snapshots, i.e, whether the ball is actually rising because somebody threw it upwards or it's falling because someone dropped it. Similarly, from knowing the states evolved using the Schrödinger's equation, it is not possible to recognize whether the time evolution is made in forward or in the backward direction. This may seem straight forward when the state is bound, as is the case for a particle in a box or the harmonic oscillator potential. However, if we are considering free particles and start with a state which is localized in position i.e., say $\psi(t = t_0) = \delta(x - x_0)$, one would expect the uncertainty in position to only increase with time. If we evolve the state backwards in time, the uncertainty would again increase as time is propagated backwards. So it seems to be possible to distinguish the snapshots of quantum states when arranged forward in time vs when they are arranged backwards in time with the clue that uncertainty in position or the spread of the wavefunction can only decrease if the snapshots are arranged in revere time order. However, the free particle problem in quantum mechanics do not have eigensolutions that can be normalized for the whole space. Rather, we first obtain a solution with box normalization and then consider the limiting case the size of the box is infinity. Therefore, when we see a sequence of snapshots of wavefunction in which the spread of the wavefunction is decreasing, it may the case that they are arranged forward in time just that the state evolved infinitely long enough to be reflected from the walls of the infinite box. The limit of the time evolved and the size of the box can be consistently taken. Although, unlike the dynamical laws of classical mechanics the Schrödinger's equation do not deterministically evolve the values of all the observables, but it evolves the quantum state with a unitary operator and hence, as per the Born's rule, the probability of measuring a particular value of the observable deterministically both forward and backward in time.

What makes quantum mechanics time asymmetric is the postulate of non-unitary state reduction baked into the Copenhagen interpretation . If a measurement is made (presuming it is instantaneous) at time $t = t_f$ on a state prepared as $|\psi_i\rangle$ at $t = t_i < t_f$, we can compute the probability of obtaining one of the eigenstates of the operator A that describes the observable for which the measurement is made. But once the measurement is made and the state reduces to $\left|\psi_f\right\rangle = |\psi_a\rangle$ one of the eigenstates of the observable A, can we compute

the retrodict the probability of state before measurement being $|\psi_i\rangle$? The usual description of quantum mechanics is predictive up to the time of measurement i.e., we can compute the probability of obtaining a state $|\psi_f\rangle$ as a result of measurement. But it is not retrodictive in the sense that we do not compute the probability of the state $|\psi_i\rangle$ - the initial state that was before measurement. The measurement seems to have erased all the history of the state evolution.

But is the non-determistic non-unitary state reduction a fundamental law of nature giving a sense of direction of time to the nature? Or it is an emergent law of describing probabilities of ensembles of particles just like the second law of thermodynamics. Or we are just posing the problem in a time-asymmetric way while the fundamental laws still respect time-reversal symmetry. The asymmetry arises from the fact that while we have declared the operator with which the measurement is made at the time t_f, we are not commenting on anything regarding the process of state preparation. How is quantum state preparation fundamentally different from measurement? From an experimental perspective, state preparation often involves the same apparatus as measurement of a quantum state. For example, if we want to prepare the spin of a spin-1/2 particle along a direction $+\hat{z}$, what do we do ? We take the Stern-Gerlach apparatus and rotate it such that the two possible outcomes of the apparatus as $|\uparrow\rangle$ and $|\downarrow\rangle$ - the spin up and spin down along the direction $\hat{z}$. We then ignore the particles in the ensemble that got shifted downwards that we identify as $|\downarrow\rangle$ and only select the ensemble of particles that got shifted upwards. This filtered ensemble is used for the concerned experiment, for which we claim the initial state to be prepared in $|\uparrow\rangle$. We would indeed use the same Stern-Gerlach apparatus to measure the state as well. [2] The only difference would perhaps be in approach. While preparing the state, we are not interested in the probability distribution of outcomes but concentrate on the ensemble of particles that got filtered. While making measurement of spin along $\hat{z}$ direction, we are interested in the probability distribution of outcomes and do not ignore any set of resultant eigenstates. In other words, the prediction is not blind i.e., $\sum_f P_{i\to f} = 1$. Since we know what the initial state is, it would definitely go to some final state and hence the sum of probability of such transition over all final states must be 1. On the other hand,

[2]The state preparation need not always be via a projective measurement. For example, atoms prepared in ground state if left as such for a long time. But this state preparation per atom is probabilistic. Thus, such methods would not guarantee a particular prepared state at a given time.

the retrodiction is blind $\sum_i P_{i\leftarrow f} \neq 1$, i.e., there is no guarantee that all the initial states would definitely transition to the desired final state.

Thus, we can pose the questions in quantum mechanics in a time-symmetric manner by either

- considering the observable with which the state preparation is made just like while computing the probabilities of outcomes from a given state we are provided with the information of what the observable corresponding to the measurement is, or by,
- considering only a part of the ensemble that resulted in a particular eigenstate after the measurement. We can ignore all other eigenstates just like we did during state-preparation.

As an analogy, consider the classical system of ball placed on an inverted bowl. Initially the ball is on a unstable equilibrium and can roll down along any direction. Our aim is to reconstruct the trajectory of the ball by considering only the information available to us at the time of state preparation(placing the ball on the bowl) and measurement(which side the ball rolled down to the ground). Even if we try our best to prepare the system identically by placing the ball exactly at the center of the bowl, still it would fall on different sides in different runs of the experiment. It appears that the trajectory of the balls are not well determined by the state preparation. However, if we select the subset of the balls that ended up at a particular side of the bowl, we can reconstruct the trajectory of the ball. We do not have to resort to a description of physics in which asking the question regarding the position of the ball at an intermediate time between state preparation and measurement is banned. In this case, both the knowledge of the final outcome and the state preparation can help us determine the unique trajectory which is applicable to the subset of balls that rolled down in the particular path. Similarly, the hope is that if we select a subset of particles that satisfy a particular selection at the stage of state preparation as well as another particular selection at the stage of state measurement, we may be able to assign values to the observable at the intermediate times in between [1.19, 1.20, 1.21]. We may be able to reconstruct trajectories of some observable for those subsets of the ensemble [1.22].

1.2 Time symmetric approach towards quantum mechanics

Let's consider a d-dimensional quantum system which is prepared at time t_i by selecting one (m-th) of the eigenstate $|\psi_i^m\rangle$ from the result of the use of measurement apparatus described by the observable A_i. Now, if we measure the system at time t_f with the apparatus corresponding to the observable described by the operator A_f, what is the probability of obtaining the eigenstate $\left|\psi_f^n\right\rangle$? The question is addressed in the standard formulation of quantum mechanics. We allow the forward time evolution of the state at time t_i till the time t_f using the Schrodinger equation. The evolution operator that makes the transformation is given by the following unitary operator

$$U(t_i \to t_f) = \exp\left(-i\int_{t_i}^{t_f} H dt\right) = \exp\big(-iH(t_f - t_i)\big) \tag{1.1}$$

The Hamiltonian H is often time independent. Thus, the initial state $|\psi_i^m\rangle$ at time t_i evolves to the state given by $U(t_i \to t_f)|\psi_i^m\rangle$ at time t_f. Then the probability that the result of measurement carried out at time t_f is the state is $\left|\psi_f^n\right\rangle$ is given by the Born rule.

$$P_{m_i \to n_f} = \left|\left\langle \psi_f^n \middle| U(t_i \to t_f)\, \psi_i^m \right\rangle\right|^2 \tag{1.2}$$

The standard formulation of quantum mechanics allows us prediction of probabilities of obtaining a final state given an initial state.

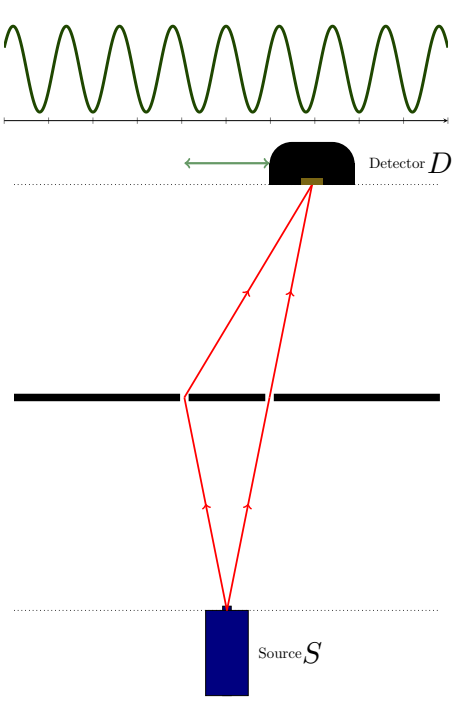
Now, let's consider the reverse problem in which the result of measurement at time t_f is known, but which of the eigenstates of the operator A_i was selected at time t_i is unknown. Can we use quantum mechanics to retrodict the probability of having selected the state $|\psi_i^m\rangle$ at time t_i given the state at time t_f is $\left|\psi_f^n\right\rangle$? We can evolve the final state backwards in time using the evolution operator $U(t_i \leftarrow t_f)$ defined below to obtain what would have been the state at time t_i if $\left|\psi_f^n\right\rangle$ were the only possibility at t_f.

$$U(t_i \leftarrow t_f) = \exp\left(i\int_{t_f}^{t_i} H dt\right) = \exp\big(iH(t_i - t_f)\big) = U(t_i \to t_f) \tag{1.3}$$

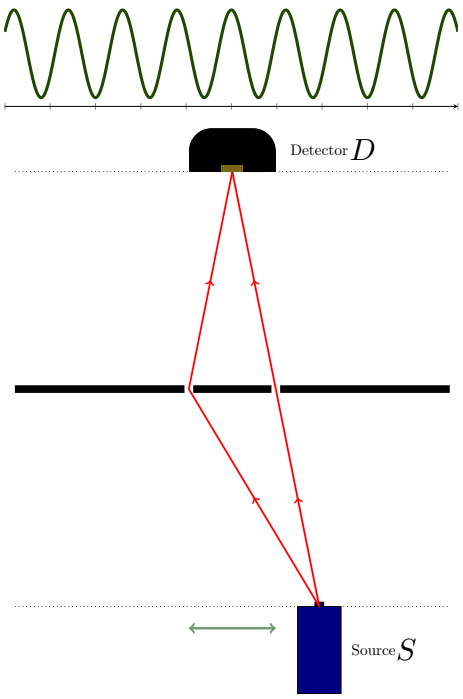
We can then compute the probability of having obtained the state $|\psi_i^m\rangle$ as follows:

$$P_{m_i \leftarrow n_f} = \left| \left\langle \psi_i^m \middle| U(t_i \leftarrow t_f) \psi_f^n \right\rangle \right|^2 \tag{1.4}$$

Note that the forward and backward time evolution operators are the same if the Hamiltonian itself do not have intrinsic time asymmetry or if we consider the usual case where the Hamiltonian is fundamentally time independent. Therefore, we would have the probability of prediction of final state $\left|\psi_f^n\right\rangle$ given an initial state $|\psi_i^m\rangle$ to be same as the probability of retrodiction of the initial state $|\psi_i^m\rangle$ given the final state $\left|\psi_f^n\right\rangle$. That is $P_{m_i \to n_f} = P_{m_i \leftarrow n_f}$. The above could be established only because we treated the initial and final time symmetrically. Just like in predicting probabilities, we know the operator with which the measurement is to be made, we provided the operator with which the state was prepared. We also evolve the final state backward in time disregarding other possible outcomes just like we propagate the initial state forward in time disregarding other possible selections during state-preparation.



(a) In a double-slit interferometer, the Intensity at the detector as function of its transverse position, when the source is fixed.



(b) The detector is fixed. However, as we move the Source along the transverse direction, the intensity at the detector changes.

Figure 1.1: We obtain the same interference pattern, i.e., the intensity as a function of transverse distance of source is the same as that as a function of the detector.

As an example to appreciate that nature indeed is time-symmetric, let us consider the double slit experiment with quantum particles as described in Fig. 1.1a. For simplicity, let's assume that the point source, the point detector lie on the $x-z$ plane. Let's also assume that the two slits are centered at $(x = s_d, y = 0, z = 0)$ and $(x = -s_d, y = 0, z = 0)$. Thus, the set up has a symmetry along the $\hat{y}$ direction which we can ignore. The particles emerge from the source S located at $(x = 0, z = -L)$. We usually ask the question regarding what is the probability of finding the particle at the detector D located at $z = L$ as we move it along x. We know from both theory and experiment that we obtain an interference pattern at the detector as a function of x. Now, if we fix the detector at $x = 0$ but do not provide the information regarding where the source is located, can we ask what is the probability that the particle emerged from the source located at $(z = -L)$ as a function of its transverse position x ? Note that in the above formulation, we have already provided the basis in which the particle was detected and the basis in which the particle emerged. Thus, the formulation is time-symmetric and as described in Fig. 1.1b we would expect the same interference pattern in the probability of the source located at x given the detection is made at $x = 0$. This can easily be verified by fixing the detector and moving the source along x. The intensity detected as the function of source's position is a same interference pattern as what we obtained when the source was fixed and the detector was moving.

Knowing the probabilities to be time symmetric does not however answer how the states actually evolve. More precisely, what is the state of the system between time t_i and t_f ? Does the state $|\psi_i^m\rangle$ unitary evolves till time t_f where the measurement with the observable A_f led to the state reduction as per the standard Copenhagen description? Or just for the sake of speculation, the state right after the interaction with the observable A_i changed to $\left|\psi_f^n\right\rangle$. This speculation seems to be against the postulate that the result of a measurement is the eigenstate of the observable. Since at time t_i, the interaction was with the observable A_i, the state must be one of its eigenstates and by virtue of our selection in the process of state preparation the state should be $|\psi_i^m\rangle$. But may be the process of state preparation and measurement are not instantaneous and hence there is a possibility that the state $|\psi_i^m\rangle$ evolves into $\left|\psi_f^n\right\rangle$. How do we rule this out ? Well, let's measure the state right after t_i, i.e., at time $\lim_{t_h \to 0} t_i + t_h$. We need to take the limiting case because the Hamiltonian may evolve the state. To avoid complications the problem would be

simpler if the eigenstates of A_i would also be eigenstates of the Hamiltonian so that the time evolution is at most by a global phase. For this, we can demand the condition that $[A_i, H] = 0$ for all time $t : t_i < t < t_f$. For a time symmetric approach, let us also assume $\left[A_f, H\right] = 0$ as well [3]. As a practical example we can again consider the spin degree of freedom of an ensemble of identical free particles in absence of electromagnetic fields and other spins. The spin prepared at time t_i should not change to any other state in absence of any other interaction during the time between t_i and t_f. This allows us to carry another measurement with the observable A_i at any time $t_1 : t_i < t_1 < t_f$. The operator A_i has d possible eigenstates and let us block all other eigenstates except $|\psi_i^m\rangle$. If the state after t_i evolves to something else other than $|\psi_i^m\rangle$, then because of the measurement at time t_1 it has non-zero probability of reducing to other states which are blocked. And since we are selectively allowing only $|\psi_i^m\rangle$ through, when it is finally measured with the observable A_f, the number of particles that ended up in the state $\left|\psi_f^n\right\rangle$ would on an average decrease compared to the case when the intermediate measurement and blocking of other paths were not made. But in the experiment, we always find that the intensity remain the same. A simple example in optics is that if we prepare a polarization state along the horizontal $|H\rangle$ and measure the power after a polarizer oriented along the diagonal, then the intensity of light after the polarizer does not change (ignoring loss due to absorption in material) if we insert a polarizer oriented horizontally in between the prepared state and the diagonal polarizer. Therefore, it seems imperative to conclude that the state at time t_1 is indeed $|H\rangle$ or in general the prepared state $|\psi_i^m\rangle$ conditional to the fact that at time t_i and t_f the states are known to be in $|\psi_i^m\rangle$ and $\left|\psi_f^n\right\rangle$ respectively. However, what happens if we keep the intermediate polarizer along the diagonal ? We again observe that the detected intensity do not change. Thus, shall we also conclude that since whenever at time t_1 we make a measurement with the same operator A_f as the final measurement the detected intensity do not change, the state of particles at time $t_f : t_i < t < t_f$ is also $\left|\psi_f^n\right\rangle$? Of course, (it seems) no. How can the state of the particle at time t_1 be described by both the states $|\psi_i^m\rangle$ and $\left|\psi_f^n\right\rangle$! There must be a way to resolve it. May be by performing the intermediate measurement by an observable which is neither A_i or A_f. By detecting the probability

[3] Note that, the commutator of hamiltonian with A_i and A_f being zero simultaneously do not imply that A_i and A_f are restricted in anyway. The Hamiltonian shares a set of eigenstates with the operator A_i and A_f but not necessarily the same eigenstates with both the operators.

distribution of outcomes of that operator, say A_b satisfying $[A_b, H] = 0$, we should be able to deduce what the state is.

Let us first compute the probability of any particular eigenstate $\left|\psi_b^k\right\rangle$ at time t_b given that the state at time t_i and t_f are known to be $|\psi_i^m\rangle$ and at $\left|\psi_f^n\right\rangle$ respectively. We can compute the probability of obtaining $\left|\psi_b^k\right\rangle$ from the knowledge of state at t_i as $P_{m_i \to k_b} = \left|\left\langle \psi_b^k \middle| \psi_i^m \right\rangle\right|^2$. Now, we can approach backwards in time as well and compute the probability of obtaining the same from the knowledge of the state at time t_f as $P_{k_b \leftarrow n_f} = \left|\left\langle \psi_f^n \middle| \psi_b^k \right\rangle\right|^2$. However, the two computed probabilities are in general not necessarily equal i.e. $P_{k_b \leftarrow n_f} \neq P_{m_i \to k_b}$. Then which one is the correct answer ? It is tempting to bet on $P_{m_i \to k_b}$, the probability obtained by forward time evolution of the quantum state. Of course, it tallies with the usual standard description of quantum mechanics and hence must be correct.But then what fundamentally makes $P_{k_b \leftarrow n_f}$ incorrect ? Is it that there is no merit in retrodicting probabilities and it accidentally works in very few specific cases and hence time symmetry is not a generic feature respected by nature ?

The probability $P_{m_i \to k_b}$ describes the overlap of the evolution of the initial state $|\psi_i^m\rangle$ with one of the possible result of intermediate measurement $\left|\psi_b^k\right\rangle$ at time t_b without any constraint on what is the state at time t_f. Similarly, the probability $P_{k_b \to n_f}$ describes the overlap of the backward evolution of the final state $\left|\psi_f^n\right\rangle$ with the possible result of intermediate measurement $\left|\psi_b^k\right\rangle$ at time t_b without any consideration of what the state at time t_i is. Both the probabilities are applicable to two different ensembles of particles where the mentioned constraints at time t_i and t_f are not necessarily satisfied simultaneously. Let us illustrate this with an example.

Consider an ensemble of $E_{i,f}$ of non-interacting spin-1/2 particles whose states at time t_i and t_f are known as $|\psi_i\rangle = |\uparrow\rangle$ and $\left|\psi_f\right\rangle = |\leftarrow\rangle$. Let's consider the measurement of the spin along diagonal direction at any in between time t_b where $t_i < t_b < t_f$. The probability of obtaining the state $|\nearrow\rangle$ at time t_b computed using the forward time evolution is $|\langle\nearrow|\uparrow\rangle|^2 = \cos^2(\pi/8) \approx 85\%$. Thus, about 85% of particles that started with $|\uparrow\rangle$ have intermediate state $|\nearrow\rangle$ and the rest 15% have the state $|\swarrow\rangle$. But do all the particles from the ensemble $E_{i,f}$ now end up being in the state $|\leftarrow\rangle$ at time t_f ? From the set of particles that have the intermediate state $|\nearrow\rangle$, about $|\langle\leftarrow|\nearrow\rangle|^2 = \cos^2(5\pi/8) \approx 15\%$ end up in the state $|\leftarrow\rangle$. Similarly, from the set of particles that have the intermediate state $|\swarrow\rangle$, about

$|\langle\leftarrow|\swarrow\rangle|^2 = \cos^2(\pi/8) \approx 85\%$. Therefore, given that we have performed an intermediate measurement along the diagonal spin, the number of particles from the ensemble $E_{i,f}$ that now satisfy the conditions at t_i and t_f is

$$
\begin{aligned}
&P\left(\left|\psi_f\right\rangle = |\leftarrow\rangle \,\middle|\, |\psi_i\rangle = |\uparrow\rangle, |\psi_b\rangle = |\nearrow\rangle \, or \, |\swarrow\rangle\right) \\
&= P\left(\left|\psi_f\right\rangle = |\leftarrow\rangle \,\middle|\, |\psi_i\rangle = |\uparrow\rangle, |\psi_b\rangle = |\nearrow\rangle\right) \qquad (1.5) \\
&\quad + P\left(\left|\psi_f\right\rangle = |\leftarrow\rangle \,\middle|\, |\psi_i\rangle = |\uparrow\rangle, |\psi_b\rangle = |\swarrow\rangle\right) \\
&= \cos^2(\pi/8)\cos^2(5\pi/8) + \cos^2(-5\pi/8)\cos^2(-\pi/8) \approx 0.25 \qquad (1.6)
\end{aligned}
$$

Without the intermediate measurement, about $|\langle\leftarrow|\uparrow\rangle|^2 = \cos^2(\pi/4) \approx 50\%$ of prepared particles would end up in the desired final state. Hence, the ensemble $E_{i,f}$ is significantly modified when we perform an intermediate measurement. This can be better illustrated with examples where the initial and final states are either same or orthogonal. When the initial and final state are both same say $|\uparrow\rangle$, and the intermediate measurement is performed with σ_x, only 50% of particles would satisfy being $|\uparrow\rangle$ at time t_f compared to 100% without the intermediate measurement. If initial and final state under consideration are orthogonal say $|\uparrow\rangle$ and $|\downarrow\rangle$, no particle would satisfy the condition. However, with intermediate σ_x measurement about 50% of the particles would end up in $|\downarrow\rangle$. In general, any measurement at intermediate time $t_i < t_b < t_f$ creates another ensemble $E_{i,f|b}$ in which the state of particles at t_i and t_f are $|\psi_i\rangle$ and $\left|\psi_f\right\rangle$ respectively even when a measurement is made at time t_b. The ensemble $E_{i,f|b}$ is in general different and is not a subset of the ensemble $E_{i,f}$. Hence, we should compute the probability of getting the intermediate state $\left|\psi_b^k\right\rangle$ for the ensemble $E_{i,f|b}$ and not $E_{i,f}$.

Using forward time evolution, we can calculate the probability of the particle having initial state $|\psi_i^m\rangle$, intermediate state $\left|\psi_b^k\right\rangle$ and final state $\left|\psi_f^n\right\rangle$ as follows:

$$P_{\left(m_i \to k_b \to n_f\right)} = \frac{P_{(m_i \to k_b)} P_{\left(k_b \to n_f\right)}}{\sum_j \left(P_{(m_i \to j_b)} P_{\left(j_b \to n_f\right)} \right)} \tag{1.7}$$

$$= \frac{\left|\left\langle \psi_b^k \middle| \psi_i^m \right\rangle\right|^2 \left|\left\langle \psi_f^n \middle| \psi_b^k \right\rangle\right|^2}{\sum_j \left(\left|\left\langle \psi_b^j \middle| \psi_i^m \right\rangle\right|^2 \left|\left\langle \psi_f^n \middle| \psi_b^j \right\rangle\right|^2 \right)} \tag{1.8}$$

Here, the $P_{(m_i \to k_b)}$ term in the numerator describes the probability of the initial state $|\psi_i^m\rangle$ going to the intermediate state $\left|\psi_b^k\right\rangle$ without any further conditions on final state. Similarly, $P_{\left(k_b \to n_f\right)}$ describes the probability that the intermediate state $\left|\psi_b^k\right\rangle$ ends up in $\left|\psi_f^n\right\rangle$ irrespective of what the initial state is. The two probabilities are multiplied as the ensemble satisfying all the three conditions is a subset of the ensemble satisfying two conditions pairwise. The denominator describes the sum of all such probabilities, where the intermediate state can be any eigenstate $\left|\psi_b^j\right\rangle$ of A_b. In simpler terms, the above equation computes what is the probability of obtaining $j = k$ given that there are $j = 1, 2, \ldots d$ possibility of intermediate states.

Now, using retrodiction, let's compute $P_{\left(m_i \leftarrow k_b \leftarrow n_f\right)}$.

$$P_{\left(m_i \leftarrow k_b \leftarrow n_f\right)} = \frac{P_{(m_i \leftarrow k_b)} P_{\left(k_b \leftarrow n_f\right)}}{\sum_j \left(P_{(m_i \leftarrow j_b)} P_{\left(j_b \leftarrow n_f\right)} \right)} \tag{1.9}$$

$$= \frac{\left|\left\langle \psi_b^k \middle| \psi_i^m \right\rangle\right|^2 \left|\left\langle \psi_f^n \middle| \psi_b^k \right\rangle\right|^2}{\sum_j \left(\left|\left\langle \psi_b^j \middle| \psi_i^m \right\rangle\right|^2 \left|\left\langle \psi_f^n \middle| \psi_b^j \right\rangle\right|^2 \right)} \tag{1.10}$$

$$= P_{\left(m_i \to k_b \to n_f\right)} \tag{1.11}$$

Thus, when we consider both ensemble $E_{i,f|b}$, the probability of the intermediate state being $\left|\psi_b^k\right\rangle$ is the same no matter whether we use forward time evolution or backward time evolution. Note that we can insert multiple intermediate measurements and as long as we sum over all the possible sets of intermediate values, we can compute the probability of obtaining a sequence of intermediate states.

Now that we have a time symmetric way of computing probabilities, let us return to the question of what exactly the intermediate state is? Is it $|\psi_i^m\rangle$ or $\left|\psi_f^n\right\rangle$ or some other

state? From the Eqn. 1.10 we can see that whenever $\left|\psi_b^k\right\rangle = |\psi_i^m\rangle$ or $\left|\psi_b^k\right\rangle = \left|\psi_f^n\right\rangle$, the conditional probability of obtaining the same becomes 1. If it is the initial state, the numerator simplifies to $|\langle\psi_i^m|\psi_i^m\rangle|^2 \left|\left\langle \psi_f^n \middle| \psi_i^m \right\rangle\right|^2 = \left|\left\langle \psi_f^n \middle| \psi_i^m \right\rangle\right|^2$. In the denominator only the term $j = m$ survives as all other states would be orthogonal and thus the denominator also reduces to $\left|\left\langle \psi_f^n \middle| \psi_i^m \right\rangle\right|^2$. Since the probability of the intermediate state being $|\psi_i^m\rangle$ is 1 (among a set possible intermediate states forming orthogonal basis), we can claim certainty. It is compelling to consider the state to be an element of physical reality for the ensemble[1.2] based on the probability of obtaining it being 1, but that needs additional assumptions on the ontology of quantum state. The complications further arise because $\left|\psi_f^n\right\rangle$ can be also claimed to be the state at time t_b since the probability of obtaining the same from Eqn. 1.10 is also 1. At this stage we can ignore the time symmetric approach since it seems to give two different solutions with certainty. Or we can seriously consider both the answers to be a part of the complete description of quantum state.

The description of the quantum state only involving ψ has been argued to be incomplete because using this state description in an EPR like setup we cannot predict the values of certain observables with certainty when we can show that they are uniquely determined from the conservation laws. With an incomplete description of state, we can at best give probabilistic prediction. In classical mechanics, if we had started with an incomplete description of state, we would also have had to resort to probabilistic description. For example, consider placing a ball on top of an inverted bowl. It is practically impossible to place it exactly at the centre and hence eventually the ball would roll down. To which direction it rolls down would seem random and indeterministic. But as we understand, the apparent randomness arises from the lack of precise knowledge of the initial state (position and momentum). Quantum mechanics could fundamentally be similar. The outcomes may seem random due to lack of precise initial state which may include other hidden variables along with the ψ. But if the underlying theory is time-symmetric then we may be able to apply the two-state description to obtain complete information about the system. For instance, if we are provided with the information that at time t_i the ball is on top of the bowl i.e. the partial state is described by x_i and at a later time t_f the ball has rolled down to a state x_f, we can use these two boundary conditions in time to calculate the state of the ball at any intermediate time $t_i < t < t_f$. This is in essence what we do in Lagrangian

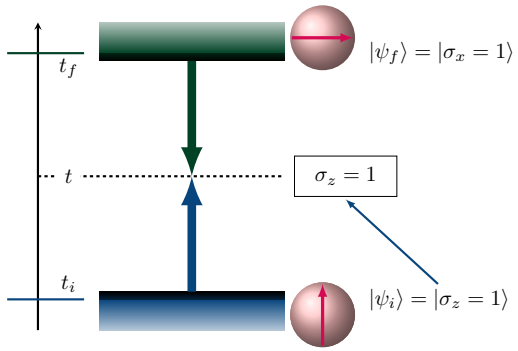
formulation: convert the initial value problem to the boundary value problem.

We may hope to achieve something similar in quantum mechanics. The knowledge of the two state vectors at two times $\{|\psi_i\rangle(t_i)\ ,\ |\psi_f\rangle(t_f)\}$ can be used as two boundary conditions using which we aspire to assign a unique value to any observable of interest[1.23, 1.24].

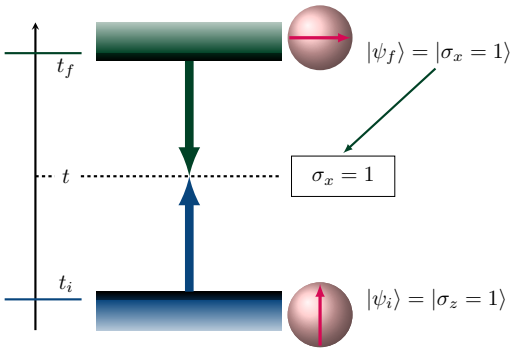
1.3 Properties of pre- and post-selected ensembles - weak value

The prepared state which often is the selection from the results of a measurement at time t_i on the quantum particles coming from some source is formally called as the *pre-selected* state $|\psi_i\rangle$ at time t_i. The process of creating the pre-selceted state is called pre-selection. The state $|\psi_f\rangle$ is the state at time t_f which guarantees the certainty of detection at a later time t_d. This also means any state orthogonal to $|\psi_f\rangle$ have zero probability of detection. With the two state vectors as the pre and post-selcted state, let us try to identify the properties of the particle[1.25] in between the time t_i and t_f.

Let's return to the example [1.26, 1.27] of a spin-1/2 system for simplicity because any spin operator commutes with the free particle Hamiltonian. Let the pre-selected state be the spin up state which is the eigenstate of the σ_z operator corresponding to the eigenvalue $+1$ i.e., $|\psi_i\rangle = |\uparrow\rangle = |\sigma_z = +1\rangle$. Let the post-selected state be the spin right $|\rightarrow\rangle$ which is the eigenstate of the σ_x observable with eigen value $+1$. What is the value of the observable σ_z at any time $t : t_i < t < t_f$? Since the preselected state had the value $+1$ and there is no interaction, one can argue that the value of σ_z would be $+1$ any time in between. And surely, if we make an intermediate σ_z measurement, we shall obtain the eigenvalue $+1$. This observation corresponds to the prediction made from the pre-selected state with forward time evolution, as illustrated in the Fig. 1.2a.



(a) The state of the spin is prepared as $|\uparrow\rangle$ which is the eigenstate of the observable σ_z corresponding to the eigenvalue of $+1$. If we make a measurement of the observable σ_z, we would find it to have the same value $+1$. Some of these particles would end up in the final state $|\psi_f\rangle$. Thus, trivially the only outcome for a σ_z measurement at time t that is consistent the state to be $|\psi_i\rangle$ at t_i and $|\psi_f\rangle$ at t_f is 1.



(b) The result of a measurement of σ_x on an eigenstate of σ_z can either be $+1$ or -1. Thus we would have two outcomes in the intermediate measurement of σ_x on the pre-selected state. But only the states with the outcome of $+1$ are compatible i.e., non-orthogonal with the final state $|\psi_f\rangle$

Figure 1.2: Intermediate property of a particle between two states of the particle in different times

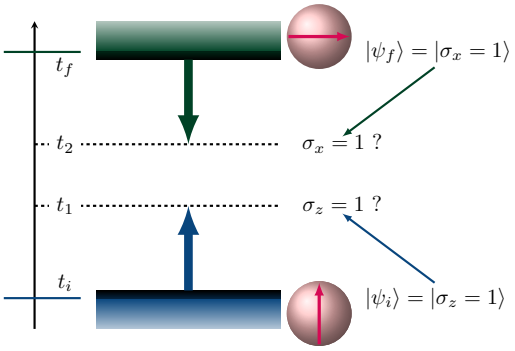
Similarly, if we ask what is the value of the observable σ_x in between pre- and post-selection, we may retrodict that the value of σ_x is also $+1$. Using forward time evolution, one may tend to infer that there should be 50% chance of getting the value to be $+1$ and 50% chance to get the value -1. However, note that the post-selection does not filter out all the particles with $\sigma_x = -1$. Hence, for the pre-and post-selected ensemble, the retrodicted value of $+1$ seems to be the only solution as illustrated in Fig. 1.2a.

1.3.1 Simultaneous Measurements

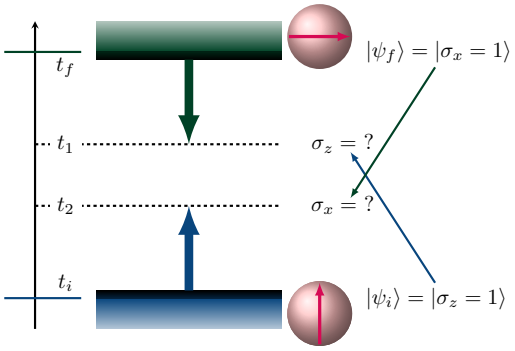
But how can we know the value of both σ_z and σ_x simultaneously ? Does it not violate the uncertainty relation ? We have $[\sigma_x, \sigma_z] = 2i\sigma_y$. So the only way consistency can be achieved if we know that the value of σ_y is 0. But it has only two eigenstates which can take the value 1 or -1. Therefore, the values of the operators are not the standard eigenvalues. The properties of a pre- and post-selected system must be therefore describing something different. So far we have considered σ_z measurement at time $t : t_i < t < t_f$ and σ_x measurement $t : t_i < t < t_f$ in separate runs of the experiment. May be the values obtained are regular eigenvalues, just that the result of both these cases do not hold simultaneously

in the same run of the experiment. Let's now consider measuring them in the same run of the experiment.

Let's have σ_z and σ_x measurement at times $t_1 : t_i < t_1 < t_f$ and $t_2 : t_i < t_2 < t_f$ respectively. When $t_2 > t_1$, we first measure σ_z as shown in Fig. 1.3a. Since the pre-selected state $|\uparrow\rangle$ is an eigenstate of the σ_z operator we get corresponding eigenvalue $+1$ as the result of the measurement. Then we make the measurement on σ_x at time t_2. There are two possible outcomes, but the state corresponding to the eigenvalue $+1$ is compatible with the post-selected state $|\rightarrow\rangle$. Thus, we can get the value of $\sigma_z = 1$ and the value of $\sigma_x = 1$ for the same particle subjected to pre- and post-selection.



(a) In the first measurement of σ_z the only possible outcome compatible with $|\psi_i\rangle$ is $\sigma_z = 1$. In the subsequent σ_x measurement at time t_2, we have have two possible outcomes for the measurement of σ_x. However, $\sigma_x = -1$ is orthogonal to the finals state $|\psi_f\rangle$. Thus the only possible outcome of σ_z followed by the measurement of σ_x which satisfies both the initial and final state is that both have values $+1$.



(b) The temporal order of the two measurements are interchanged. σ_x measurement at t_2 now precedes the measurement of σ_z at t_1. Both the outcomes $+1$ and -1 for σ_x are compatible with the initial state $|\psi_i\rangle$. In the subsequent measurement, σ_z can have outcomes $+1$ or -1 compatible with either outcome of the previous σ_x measurement. Again none of the outcomes of σ_z measurement are incompatible with the final state $|\psi_f\rangle$. Thus, we cannot disregard any of the outcomes.

Figure 1.3: Both the measurement are made on the same particle. In (a), we first make the measurement of σ_z of which the initial state is an eigenstate. The second measurement is of σ_x and the final state is the eigenstate of it. In (b), the order of measurements are reversed. Neither σ_x nor σ_z could be ascertained in this case. Note that the reason behind this is the fact that σ_x does not commute with σ_z and hence they do not share the eigenstates for which the eigenvalues of both the observables can be defined simultaneously.

But what happens if $t_2 < t_1$ as is the case in Fig. 1.3b? The outcome of σ_x measurement at time t_1 can have both values $+1$ and -1. Neither can be discarded claiming that they would later not form the part of post-selected ensemble because there exists a measurement with σ_z at time t_2. The result of which again can be both $+1$ and -1 because the state after

the first measurement is no more an eigenstate of σ_z. Each outcome has a probability of 50% to be post-selected with state $|\rightarrow\rangle$. Thus, simultaneous measurement of σ_x and σ_z can give us all possible eigenvalues. Surprisingly, this particular choice of measurement operators also do not seem to change probability of successful post-selection. Without intermediate measurements, half of the pre-selected particles would have satisfied post-selection. When we make two intermediate measurements, first σ_x and then σ_z, the state results in a mixture of 50% particles being $|\uparrow\rangle$ and 50% particles being in $|\downarrow\rangle$. Each of them have 50% chance of getting post-selected. Thus, cannot claim that the two measurements disturbed the ensemble by changing the probability of post-selection.

1.3.2 Imposing Time Symmetry on the value of Observables

Thus, if not the ensemble, the process of measuring two observables must be disturbing time symmetry itself. Hence, let us have a revisit at measuring one observable which can give some information about σ_x and σ_z. What should be the outcome if we make an intermediate measurement of the observable of spin along the diagonal represented by the Hermitian unitary operator $\sigma_d = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$?

If the measurement respects time symmetry as shown in the Fig. 1.4, then the value of σ_x should be 1 to be compatible with post-selection and the value of σ_z has to be also 1 because of pre-selection. Thus, the value of the observable σ_d must be $\frac{1}{\sqrt{2}}(1+1) = \sqrt{2}$. Absolute non-sense !!! Or so it appears. How can the result of measurement of σ_d can be anything except being one of its eigenvalues of $+1$ and -1. It cannot be also a describtion of being statistical average over a number of runs of the experiment because no matter what the statistics are the average of the eigenvalues is bounded by the extreme eigenvalues themselves. Hence, the answer of $\sqrt{2}$ can never be obtained as a result of either single measurement or as a statistical average over many measurements.

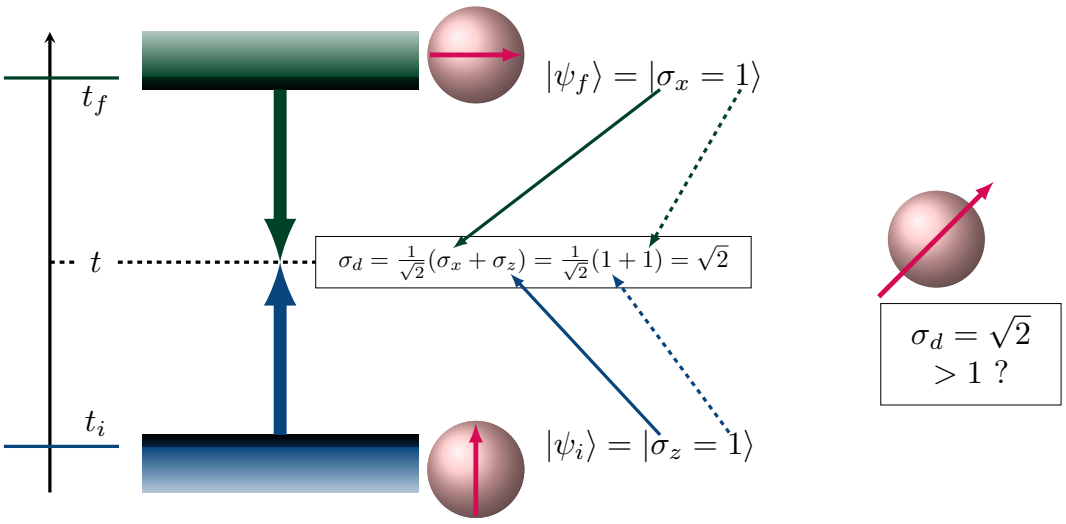


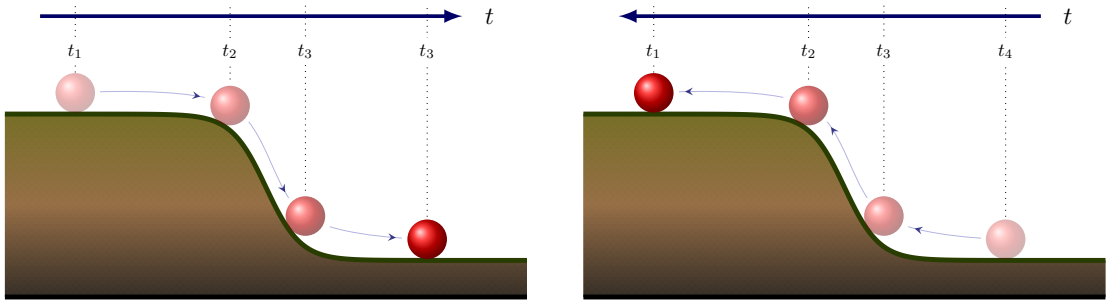
Figure 1.4: We consider the measurement of the Hermitian Unitary operator σ_d at an intermediate time t. Neither the initial state $|\psi_i\rangle$ nor the final state $|\psi_f\rangle$ are the eigenstates of this observable. However, the result of the measurement will be one of its eigenstates $+1$ or -1. However, any of the above outcome would change the probability that a particle is subjected to both being in the state $|\psi_i\rangle$ at time t_i and $|\psi_f\rangle$ at t_f. But if in a leap of imagination, we consider a measurement that would respect time symmetry, we would expect the outcome to be such that the projections on the initial and final state are maintained. In that case what would be the value of the observable ? Since both σ_z and σ_x projections has to be $+1$ to satisfy $|\psi_i\rangle$ and $|\psi_f\rangle$, the value of σ_d seems to be $\sqrt{2}$. This is beyond the eigenvalues and hence cannot be measured in practice with the kind of measurement that yields only the eigenvalues.

Perhaps, the above should have been enough to ditch the idea of time-symmetric formulation of quantum mechanics. The nature seems to not respect time symmetry when we make a measurement and possibly the fundamental dynamical laws do not respect time-reversal symmetry. But there is a last hook to cling on. Even if the fundamental laws of dynamics are time-symmetric in classical mechanics, the solution can be time-asymmetric.

1.3.3 Discontinuity in potential breaking Time Reversal Symmetry : Classical analogue

Imagine a ball rolling due to gravity on a landscape with a smooth slope as shown in Fig. 1.5a. If the momentum at any point in time is reversed, the ball would retrace the trajectory which would be indistinguishable from ball going backward in time. If the event is recorded as a video, it is difficult to tell whether the trajectory of the ball shown in Fig.

1.5b represents the ball climbing the slope when thrown from right towards left or is the video of what is represented in Fig. 1.5a just played backwards in time.



(a) Snapshots of a ball rolling down a smooth slope from left to right arranged in the order $t_1 < t_2 < t_3 < t_4$. The past snapshots are represented in decreasing order of opacity.

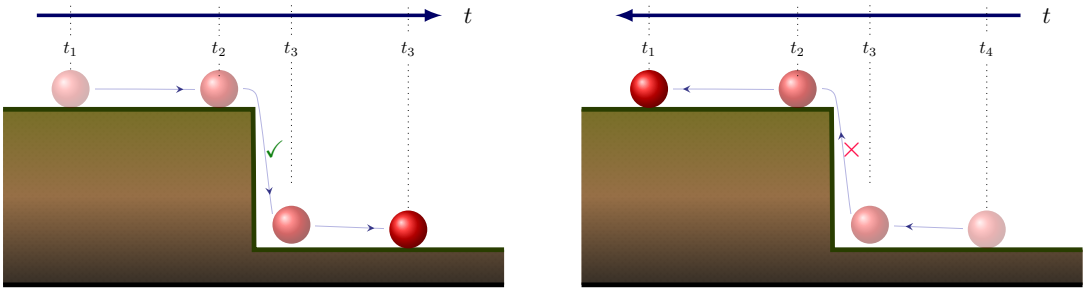
(b) Snapshots of the ball rolling down the smooth slope from left to right but arranged in the order $t_1 > t_2 > t_3 > t_4$ to indicate the reversed direction of time.

Figure 1.5: If the potential is smooth, the time reversed arrangement of snapshots seems physically plausible. In fact, the time reversed arrangement of the ball rolling down from left to right and speeding up as it moves down the slope is indistinguishable from a ball rolling up from left to right and slowing down as it moves up the slope.

However, this aspect of time reversal symmetry do not get manifested if instead of a smooth slope we have a step potential. The ball can easily go from left to right as shown in Fig. 1.6a. But in practice, when we throw the ball from right to left, the ball would hit the wall of the step and then bounce back towards right. It would not jump up the step as depicted in Fig. 1.6b. Of course, the laws of classical dynamics are still time reversible. Further, one can argue that, if all the interactions are conservative then the ball would never simply roll after getting dropped at the step as shown in Fig. 1.6a. Instead, it would keep on bouncing. Thus, the initial condition shown in Fig. 1.6b is never a momentum reversed solution of 1.6a. The time reversal symmetry is broken not only due to the step potential but also, in practice some energy eventually being lost due to non-conservative interaction such as friction. But during the process of quantum measurement a similar interaction occurs between the quantum system and the environment. In absence of measurement, one is content with the unitary evolution which is time reversible [4]. Thus, if we draw the analogy it is the discontinuity in the potential that breaks the time reversibility of

[4]Strictly speaking the expectation value of the observable would be time reversible. The uncertainty would increase both forward and backward in time

the solution. Otherwise, if the potential is smooth, the solution is approximately [5] time reversible.



(a) Snapshots of the ball moving from left to right through the step arranged in the time order $t_1 < t_2 < t_3 < t_4$.

(b) Snapshots of the ball moving from left to right through the step arranged in the reverse time order $t_1 > t_2 > t_3 > t_4$.

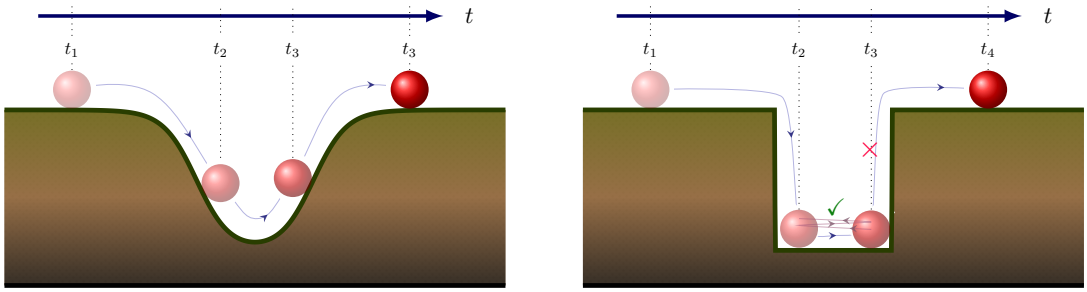
Figure 1.6: When we have a step potential (large enough discontinuity) the time reversed arrangement of snapshots does not appeal as physically plausible. If a ball is coming from right to left, it would reflect from the wall formed by the step and then move from left to right. It would not be able to jump up the step. Thus, the time reversed solution is not equivalent to starting with a different initial condition with momentum reversed and then evolving forward in time. One may argue that the above essentially involves loss of energy of the ball to friction and damping as it falls and rolls. Otherwise, the ball would keep bouncing. In that case if the momentum is reversed it would have a vertical component as well which would help it bounce up to the height of step. But interaction, even in quantum world involves exchange of energy. Therefore, the analogy survives in the context of interaction between a quantum particle subjected to large change in potential, which often is the case during the process of quantum measurement.

However, now consider a square well shown in 1.7b instead of a step potential. Once the ball rolls down the well, it cannot jump up [6]. All the energy of the ball can only help it bounce back and forth. Even if we reverse the momentum, the ball would not spontaneously climb up the well. The sequence of snapshots played backwards in time would appear ridiculously gravity defying fiction compared to the motion forward in time.

[5] One may attempt to be quantitative here. Basically, given a step potential even a very small amount of energy lost to friction or interaction with environment is enough to give an entirely different solution. How different the solutions are can be quantified by the integral $\int_{t_i}^{t_f} |x_1(t) - x_2(t_f - t)| dt$. Here, $x_1(t)$ is the trajectory of particle going forward in time when the initial condition represents the ball being thrown from left. $x_2(t)$ represents trajectory of the ball thrown from right to left. However, the trajectory is evolved considering time t_f as the origin. Then one may argue that the time reversed solution, in case of smooth potential, would not be different as we take the limit when energy lost to environment can be neglected. In a step potential, slight energy loss would mean that the ball would hit the step and thus evolve into a completely different trajectory.

[6] Note that we are considering negligible rolling friction but considerable coefficient of restitution as often is the case in practice

But if the well is smooth as shown in Fig. 1.7a, there is nothing preventing in realizing the approximately time reversible solution in practice. The difference between the two cases is that in one of them the potential surface was smooth where as in case of the well there was discontinuity.



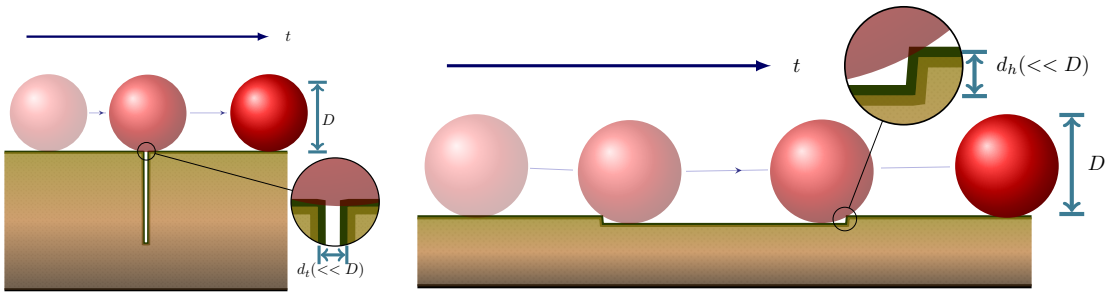
(a) Given a smooth furrow, the particle moving from left to right can come up after rolling down. If the snapshots are played back in time, this would appear indistinguishable from the particle moving from right to left.

(b) In a square well, the ball would fall in the well and, given some loss of energy in the bounce, would get trapped in the well. It is unlikely to jump up the well.

Figure 1.7: When the interaction of a particle with a potential has a compact support i.e., it is short-lived, it makes the particle get back to the same (kinetic) energy prior to the interaction. Thus, on average, there is no net exchange of energy. In the potential is smooth, time reversal symmetry is maintained. But in a square well, where the potential is discontinuous, the time reversal symmetry of the solution seems not be respected.

If we make the slopes of the walls of the well slanted instead of vertical and create a smooth[7] surface, we can make the trajectory of the ball respect time symmetry. But what determines whether the surface is smooth or discontinuous ? Practically, every surface cannot be smooth down to the infinitesimal scale. The scale determining factor is the size of the ball. If the depth of the well was much smaller than the radius of the ball, then it could easily roll upwards. So even if there are discontinuities, they affect the time symmetry of trajectory of the ball only when the amount of discontinuity (depth of the well) is comparable to size of the ball or larger.

[7]Here, smoothness demands at continuity of the function and all its derivative upto at least 2nd order



(a) The ball may roll over a very square well if it has to interact with that for a very small amount of time. In other words, if the width of the well given by d_t is much much smaller compared to the diameter D of the ball, the ball just keeps rolling from left to right almost unaffected.

(b) If the height of the step d_h is very small compared to the diameter D of the ball, the ball would keep rolling. Its time reversed solution would approximately be the same as that of a ball moving from right to left.

Figure 1.8: Even if there are discontinuities, if the support (interaction time) or value (strength/height) of the potential or both are much smaller compared to the size of the ball, the ball would not get much affected by the presence of the discontinuity. We can draw the analogy that if either the strength of the measurement interaction can be reduced or the the interaction last for very small time, the solution may approximately be time symmetric.

The process of quantum measurement could be introducing such discontinuity in the overall Hamiltonian. Therefore, even if the laws would be fundamentally respecting time-reversal symmetry, the solutions would not. Once we identify what is the analogue of the well and the size of the ball, we may then attempt to reduce its depth to restore time symmetry of the solutions. Thus, we need to introduce a new kind of measurement protocol which respects time-symmetry [1.28, 1.29]. And if it is achieved, we may get the value of $\sqrt{2}$ for the spin along the diagonal direction [1.30]. As we would explore in the next section, such measurements are possible and are known as weak measurements (they weaken or reduce the depth of the well to restore time symmetry). The values which are the result of the measurement are therefore called as weak values [1.31, 1.32]. Thus, the value of the spin along the diagonal for the mentioned pre- and post-selected system which is $\sqrt{2}$ is neither the eigen-value nor the expectation value of the observable, but rather its weak-value for the given pre- and post-selected states.

1.4 Weak Measurements

What exactly is the process of quantum measurement ? And how it fundamentally differs from when we measure anything in the classical world. In simplified terms, a measurement is a process that links an inaccessible quantity of a system to some property of the apparatus which can be read off easily and can be compared to pre-defined standards. As an example, consider measurement of temperature of human body. We place the thermometer to record the temperature of our body. We link the otherwise inaccessible quantity of temperature of the system (human body) to the property of expansion of the mercury in the apparatus(thermometer). Once, we manage to link, then since we know how to measure change in lengths of objects we can read off how much the mercury expanded and therefore can assign values to the temperature. Another example is the measurement of current in an electrical wire. We place an ammeter in series. There is a coupling of the ammeter's pointer needle to the current in the wire by virtue of electromagnetic induction. The current in the wire which was otherwise inaccessible now can be quantified by the shift in the pointer needle of the ammeter. But do they change the state of the system? Well, they are designed not to change much, but the very fact that there is an interaction between the system and the apparatus and the state of the apparatus changed due to interaction (otherwise we would not be able to read anything), by conservation laws the system must have undergone some changes as well. The thermometer takes away some heat if mercury expands or gives away some heat if the body is cold. Deflecting the pointer's needle in an ammeter, and that must come at the expense of a slight decrease in the current in the wire. However, one keeps the resistance of the ammeter as low as possible and the specific heat capacity of the liquid inside the thermometer to be low so that they do not significantly affect the property of the system that is being measured.

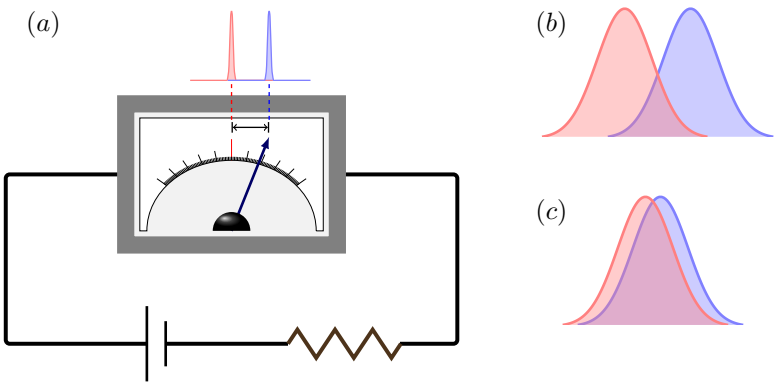


Figure 1.9: The ammeter connected to a circuit to measure the current is an illustrative example of a pointer. The pointer is calibrated to the red mark when there is no current flowing in the circuit or when it is not connected. When the circuit is connected, the shift of the pointer in the measuring device(ammeter) gets coupled to the amount and direction of current in the circuit. The more the current in the circuit, the more deflection of the pointer needle. (a) For a precise measurement, the pointer needle must have minimal fluctuation while calibrating as well as while reading a steady current. The signal-to-noise ratio is better if the corresponding distribution represented by the red and blue Gaussians respectively have well separated means compared to their standard deviation. (b) When the fluctuation in the pointer is more and the distribution of pointer positions for the current reading overlaps with the distribution of the fluctuating pointer position during calibration, the measurement is imprecise. If we take enough readings, we can determine the mean value of the current. However, it is difficult to attribute a single reading to the current of the circuit. (c) In the limit where the two distributions overlap, it is impossible to comment on whether there is current in the circuit or not. But if there exist a small difference, we have to take very large number of readings so that the uncertainty in determining the mean current is much less than the slight separation of the distribution mean from the reference.

We do not have the luxury in the quantum world. Imagine measuring the current due to motion of a single electron using a galvanometer hoping that the electromagnetic induction would cause its pointer to deflect. If the electron is moving slowly, most likely it would not have enough energy to deflect the needle and hence no measurement would be possible. If it were accelerated to very high speeds and it manages to move the pointer needle a bit, the energy spend would decrease the speed of the electron significantly and hence the property of the system would be significantly affected by virtue of the measurement. Let's now formalize the above concepts in the mathematical language used in quantum mechanics and discuss the von-Neummann measurement scheme [1.6]. The description to follow is certainly inspired from the go to example of measuring the spin of a quantum particle [1.33, 1.34], but would be applicable in general.

1.4.1 von Neumann Measurement Interaction

Let the observable of the system which is to be measured be denoted as $\hat{A}$. We aim to interact the system with the pointer so that the pointer variable that can be read off is changed by virtue of the interaction. If nature is fundamentally described by quantum mechanics, we should also have a quantum description of the pointer. Let the pointer variable be denoted by the operator $\hat{X}$ as often measurement is finally related to reading the position of some component in the apparatus, be it the in case of the thermometer or ammeter. To be able to measure the value of the observable $\hat{A}$, the pointer variable must also change its value depending on the value of the observable $\hat{A}$. Thus, $\hat{A}$ couples to the generalized momentum $\hat{P} = \frac{\hbar}{i}\frac{\partial}{\partial X}$ of the measuring apparatus, which is the generator of translation of the pointer variable $\hat{X}$. To have a finite translation of the pointer variable, the interaction must last for a finite time as well. Thus, we can have a weight associated with the interaction strength which can be described by an impulse function having compact support in time. Let this weight be given by $g(t)$ satisfying $\int_{t_-}^{t_+} g(t)dt = 1$. Here, t_- denotes the time at which the interaction can be considered to have an onset. $t_+ > t_-$ is the time at which the interaction can be considered over. Most interactions do not fall in this category since the strength of fundamental interactions like gravity and electromagnetism falls off smoothly. So two charged particles which are set to approach each other do have interaction no matter how far they are. But for all practical purposes, when the particles are far apart the force would be so small that the effect on the trajectories can be neglected. We can therefore choose t_- and t_+ to be finite and small based on the precision with which we would want the result of the measurement. Thus, now we can define the interaction between the system and the pointer with the following interaction Hamiltonian.

$$H_I = g(t)\gamma\hat{A} \otimes \hat{P} \tag{1.12}$$

Here, γ is the scaling factor governing the maximum strength of the interaction. This Hamiltonian acts on the joint state of the system and the pointer and evolves them together. The initial joint state comprising of the system and the pointer can be described as a separable state at time t_i because of lack of any interaction in the past. Hence, we can write the joint initial state $|\Psi_i\rangle$ as

$$|\Psi_i\rangle = |\psi_i\rangle \otimes |\chi_i\rangle \tag{1.13}$$

Here, the $|\psi_i\rangle$ is the initial state of the system. For simplicity, we assume that the the operator $\left[\hat{A}, H_s\right] = 0$ commutes with the Hamiltonian H_s of the system so that the state of the system changes due to the interaction. However, the same may not be usually demanded for the pointer. If we are interested in position of the measuring apparatus as the pointer, its free Hamiltonian H_p would keep evolving it. If we start with a position eigenstate of the pointer, its momentum would be uncertain. Further the Hamiltonian H_p would introduce uncertainty in position as time evolves. To be able to read off change in the value of position, we must have changes due to the Interaction Hamiltonian much much larger than the uncertainty introduced by the free Hamiltonian. Otherwise, we can start with some uncertainty in position so that the uncertainty in position do not grow rapidly with time. In either case the interaction strength γ has to be sufficiently large enough to overcome this uncertainty. This, unfortunately also means that the value of the observable cannot be read off precisely in a single run of the experiment without affecting the property itself. With this approximation, the total Hamiltonian can be effectively approximated as the interaction Hamiltonian, i.e, $H = H_s + H_p + H_I \approx H_I$.

The, unitary evolution operator can therefore be written as,

$$U = \exp\left(-\frac{i}{\hbar}\int_{t_-}^{t_+} H_I dt\right) = \exp\left(-\frac{i}{\hbar}\int_{t_-}^{t_+} g(t)\gamma\hat{A}\otimes\hat{P}dt\right) \tag{1.14}$$

Since the operators on the system and the pointer state i.e, $\hat{A}$ and $\hat{P}$ do not have any explicit time dependence, we can simplify the above operator by performing the time integration with the understanding that the operator would be acting on state with no implicit time dependence under the approximation that H_p is negligible. Thus, we may simplify the evolution by performing the integral $\int_{t_-}^{t_+} g(t)dt = 1$. Thus, we have $U = \exp\left(-\frac{i}{\hbar}\gamma\hat{A}\otimes\hat{P}\right)$

To be able to operate U on $|\Psi_i\rangle$, we first need to express $|\psi_i\rangle$ and $|\chi_i\rangle$ in terms of linear combination of the eigenstates of the operator $\hat{A}$ and $\hat{P}$ respectively. Thus, we express,

$$|\psi_i\rangle = \sum_{m=0}^{d} c_m |a_m\rangle \tag{1.15}$$

Here, $|a_m\rangle$ is the m-th eigenstate of the operator $\hat{A}$ out of the d eigenstates arranged in ascending order of the eigenvalues a_m. The c_m are the coefficient of the $|\psi_i\rangle$ along the eigenstate $|a_m\rangle$

The initial pointer state is preferred to have $\langle \hat{P} \rangle = 0$, so that in absence of the interaction, the value of the pointer variable $\hat{X}$ do not evolve in time. A Gaussian distribution in x satisfies the above preference. Hence, we can have the initial pointer state as follows:

$$|\chi_i\rangle = \int \sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right) |x\rangle \tag{1.16}$$

In the above expression, the integral is over all the position eigenstates $|x\rangle$ which forms the continuous variable basis. Here, μ is the mean of the Gaussian distribution and σ is the standard deviation of the distribution. The σ is assumed to be sufficiently large enough, so that the momentum uncertainty can be made lower. This allows us to ignore the increase in σ over time due to the free Hamiltonian H_p of the pointer. We had already ignored H_p and hence for consistency assume that σ does not change appreciably during the timescale in consideration. Note that since the probability distribution of the pointer should be a Gaussian, the probability amplitude $|\chi_i\rangle$ is written as the square root of the Gaussian such that $\int_{-\infty}^{\infty} |\chi_i|^2 dx = 1$. Also, all the coefficients of $|x\rangle$ are considered to be real so that during the evolution the plane of the wave-front maintains a constant phase.

The evolution operator now can be applied to find the final evolved state after the interaction $|\Psi_e\rangle$

$$|\Psi_e\rangle = U |\Psi_i\rangle \tag{1.17}$$

$$= \exp\left(-\frac{i}{\hbar}\gamma \hat{A} \otimes \hat{P}\right) |\psi_i\rangle \otimes |\chi_i\rangle \tag{1.18}$$

$$= \exp\left(-\frac{i}{\hbar}\gamma \hat{A} \otimes \frac{\hbar}{i}\frac{\partial}{\partial x}\right) \left(\sum_{m=0}^{d} c_m |a_m\rangle\right) \otimes \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.19}$$

In the last line, we have dropped the $|x\rangle$ terms in the representation of the pointer state because the operator $\hat{P}$ is also described in position representation. The resultant state after the operation is the entanglement between the system's degree of freedom and the pointer.

$$|\Psi_e\rangle = \exp\left(-\hat{A}\otimes\gamma\frac{\partial}{\partial x}\right)\left(\sum_{m=0}^{d} c_m|a_m\rangle\right)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.20}$$

$$= \sum_{m=0}^{d}\left(c_m\exp(A)|a_m\rangle\right)\otimes\left(\exp\left(-\gamma\frac{\partial}{\partial x}\right)\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.21}$$

The first term in the tensor product that consists of the system degree of freedom can be evaluated because $|a_m\rangle$ are the eigenstates of $\hat{A}$.

$$\exp(A) = \sum_{i=0}^{\infty}\frac{A^i}{i!} = \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots \tag{1.22}$$

$$\Rightarrow \exp(A)|a_m\rangle = \left(\mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots\right)|a_m\rangle \tag{1.23}$$

$$= \left(1 + a_m + \frac{a_m^2}{2!} + \frac{a_m^3}{3!} + \ldots\right)|a_m\rangle \tag{1.24}$$

$$= \sum_{i=0}^{\infty}\frac{a_m^i}{i!}|a_m\rangle = \exp(a_m)|a_m\rangle \tag{1.25}$$

Using the above, now we can rewrite Eqn. (1.21) such that the term $\exp(a_m)$ now couples with γ.

$$|\Psi_e\rangle = \sum_{m=0}^{d}\left(c_m\exp(a_m)|a_m\rangle\right)\otimes\left(\exp\left(-\gamma\frac{\partial}{\partial x}\right)\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.26}$$

$$= \sum_{m=0}^{d}\left(c_m|a_m\rangle\right)\otimes\left(\exp\left(-\gamma a_m\frac{\partial}{\partial x}\right)\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.27}$$

To simplify the above, we recall the Taylor series expansion of a real function $f(x-a)$ about x.

$$f(x-a) = f(x) + \frac{\partial f(x)}{\partial x}(-a) + \frac{1}{2!}\frac{\partial^2 f(x)}{\partial x^2}(-a)^2 + \frac{1}{3!}\frac{\partial^3 f(x)}{\partial x^2}(-a)^3 + \ldots \tag{1.28}$$

$$= \left(\sum_{i=0}^{\infty}\left(\frac{\partial}{\partial x}\right)^i \frac{(-a)^i}{i!}\right) f(x) \tag{1.29}$$

$$= \exp\left(-a\frac{\partial}{\partial x}\right) f(x) \tag{1.30}$$

By comparing, Eqn. (1.30) with the second term acting on the pointer state in Eqn. (1.27), we can evaluate the final state as follows:

$$|\Psi_e\rangle = \sum_{m=0}^{d} (c_m |a_m\rangle) \otimes \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{((x-a_m\gamma)-\mu)^2}{4\sigma^2}\right)\right) \tag{1.31}$$

The above state represents the entanglement between the spin eigenstates $|a_m\rangle$ and the Gaussian pointers which are shifted by the corresponding eigenvalue multiplied with the interaction strength γ. In general, one needs to take care that the given pointer function has convergence when expanded using Taylor series.

Let's denote a_{min} to be smallest separation between any pair of eigenvalues a_m. If $\gamma\, a_{min}/\sigma >> 1$, then different eigenstates $|a_m\rangle$ can be distinguished by the shift of the pointer, i.e., whether the initial state was $|a_1\rangle$ can be distinguished from the state $|a_2\rangle$ for a single particle as the corresponding shifts in the pointer variable $a_1\gamma$ and $a_2\gamma$ can be distinguished because $(a_2 - a_1)\gamma >> \sigma$. In general, when the initial state is in a superposition of eigenstates of A, the pointer degree of freedom remains entangled. However, in each run of the event, we get one distinguishable outcome of the pointer. How this entangled state reduces to one of the outcome is still one of the unresolved problem in the description of quantum mechanics. However, in the context of statistical ensemble, tracing out the system degree of freedom results in a mixed state of the pointer variable. Many attempts to provide rigour to this description do exist, like the Decoherence models and spontaneous collapse models, but they require extra assumptions or preferences to be invoked. However, in this section, the emphasis is to comment on the property of the system from knowing the pointer variable. Let us return to the spin half particle subjected to the Stern Gerlach as our go to example.

If the particle is prepared in the state spin $|\uparrow\rangle$ and it has a transverse spatial distribution which is a Gaussian probability amplitude centred about zero given by $|g(0)\rangle$, what is the state after the interaction with the Stern Gerlach apparatus oriented along the diagonal i.e. the operator being $\sigma_d = \sigma_{\pi/4}$?. There are two eigenstates as $|\nearrow\rangle$ and $|\swarrow\rangle$. Given the interaction strength is γ, the joint state of the system and the pointer after the interaction is given as

$$c_{\nearrow}|\nearrow\rangle|g(\gamma)\rangle + c_{\swarrow}|\swarrow\rangle|g(-\gamma)\rangle \tag{1.32}$$

Here, the coefficients are determined as $c_{\nearrow} = \langle\nearrow|\uparrow\rangle$ and $c_{\swarrow} = \langle\swarrow|\uparrow\rangle$ respectively. In each outcome, either the particle lands at the detector plane with the pointer variable shifted by γ or $-\gamma$. When it is γ, we infer that the system has reduced to the state $|\nearrow\rangle$ and when it is $-\gamma$, we infer the state to be $|\swarrow\rangle$. However, when $\gamma << \sigma$, we cannot distinguish the two pointer states. That also means that we cannot infer which state the system has collapsed to after the interaction. Or does the system collapse at all ? Let's investigate by taking the limit $\frac{\gamma}{\sigma} \to 0$. In this case, we have $\lim_{\gamma\to 0}\langle g(\gamma)|g(-\gamma)\rangle \approx 1$. Thus the two gaussians are indistinguishable. Then the state after the interaction is given by

$$\lim_{\gamma\to 0} c_{\nearrow}|\nearrow\rangle|g(\gamma)\rangle + c_{\swarrow}|\swarrow\rangle|g(-\gamma)\rangle = (c_{\nearrow}|\nearrow\rangle + c_{\swarrow}|\swarrow\rangle)|g(0)\rangle = |\uparrow\rangle|g(0)\rangle \tag{1.33}$$

The system state is now separable in this limit and is still $|\uparrow\rangle$. However, if we repeat this experiment with an ensemble of N particles, in the limit $N \to \infty$, we can infer the mean of the resultant pointer. Since $|c_{\nearrow}|^2 > |c_{\swarrow}|^2$, the average of the distribution of the pointer variable would shift towards positive. Of course, it is impossible to say from the distribution of the pointer what exactly the initial state is. Any relative phase between the two eigenstates of the observable would result in the same distribution and hence the same mean. However, if the initial state is the eigenstate $|\nearrow\rangle$ of the observable, the mean of the distribution of the pointer state would converge $\lim_{N\to\infty}$, the initial state can be inferred to be the eigenstate $|\nearrow\rangle$. Note that this is not generically true and holds only when the eigenvalue in consideration is extreme. For instance, consider a spin-1 particle. If the distribution of the pointer converges to zero, whether the state is an equal weighted

superposition of $|+1\rangle$ and $|-1\rangle$ or the state is actually $|0\rangle$, cannot be uniquely determined just by knowing the mean value of the pointer variable distribution.

1.4.2 Weak Interaction

Formally, when $\gamma\, a_{max} << \sigma$, we call the interaction to be *weak*. And in contrast, when $\gamma\, a_{min} >> \sigma$, the interaction is *strong*. When $\gamma\, a_{max/min} \approx \sigma$, we have the *intermediate* interaction strength representing the transition from *strong* interaction to *weak* interaction. The weak interaction limit if often described by truncating the power series expansions in Eqn. (1.22) and Eqn. (1.28) after the first order terms. This implies that the evolution operator is approximated and only linear order terms of the operator $\hat{A}$ are considered.

$$\exp\left(-\frac{i}{\hbar}\gamma\hat{A}\otimes\hat{P}\right) \approx \mathbb{1} - \frac{i}{\hbar}\gamma\hat{A}\otimes\hat{P} \tag{1.34}$$

The above expression should only be considered in the limit $\gamma \to 0$, else the truncated evolution operator is not guaranteed to be unitary. Under this limit, the evolved state in Eqn. (1.19) can be approximated.

$$\lim_{\gamma\to 0}|\Psi_e\rangle = \left(\mathbb{1} - \frac{i}{\hbar}\gamma\hat{A}\otimes\hat{P}\right)\left(\sum_{m=0}^{d} c_m|a_m\rangle\right)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.35}$$

$$= \left(\sum_{m=0}^{d} c_m|a_m\rangle\right)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) - \left(\sum_{m=0}^{d} c_m A|a_m\rangle\right) \otimes\gamma\frac{\partial}{\partial x}\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) \tag{1.36}$$

$$= \left(\sum_{m=0}^{d} c_m|a_m\rangle\right)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right) - \left(\sum_{m=0}^{d} c_m a_m|a_m\rangle\right)\otimes\gamma\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right)\left(-2\frac{(x-\mu)}{4\sigma^2}\right) \tag{1.37}$$

$$= \sum_{m=0}^{d}\left[(c_m|a_m\rangle)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right)\left(1+\gamma a_m\frac{2(x-\mu)}{4\sigma^2}\right)\right] \tag{1.38}$$

Since $\gamma << \sigma$, we can treat the term $\gamma a_m\frac{2(x-\mu)}{4\sigma^2} \to 0$. Therefore we can approximate the last term in the expression 1.38 to an exponential.

$$1+\gamma a_m\frac{2(x-\mu)}{4\sigma^2} \approx \exp\left(\gamma a_m\frac{2(x-\mu)}{4\sigma^2}\right) \tag{1.39}$$

Substituting the exponential approximation in the expression 1.38, we can simplify $|\Psi_e\rangle$.

$$\lim_{\gamma\to 0}|\Psi_e\rangle = \sum_{m=0}^{d}\left[(c_m|a_m\rangle)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)\right)\exp\left(\gamma a_m\frac{2(x-\mu)}{4\sigma^2}\right)\right] \tag{1.40}$$

$$= \sum_{m=0}^{d}\left[(c_m|a_m\rangle)\otimes\left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2-2\gamma a_m(x-\mu)}{4\sigma^2}\right)\right)\right] \tag{1.41}$$

We employ the familiar trick of squaring the exponent and find the following appropriate form of the expression.

$$((x-\mu)-\gamma a_m)^2 = (x-\mu)^2 - 2\gamma a_m(x-\mu) + \gamma^2 a_m^2 \tag{1.42}$$

$$\approx (x-\mu)^2 - 2\gamma a_m(x-\mu) \tag{1.43}$$

The last term in the above Eqn. 1.42 can be neglected in the limit $\gamma \to 0$. Now, the pointer state in the Eqn. 1.41 can be expressed as a displaced Gaussian.

$$\lim_{\gamma\to 0}|\Psi_e\rangle = \sum_{m=0}^{d}\left[c_m|a_m\rangle \otimes \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{((x-\gamma a_m)-\mu)^2}{4\sigma^2}\right)\right)\right] \tag{1.44}$$

Note that the above expression is essentially the same as 1.31. The pointer state is entangled state of various eigenstates$|a_m\rangle$ of the observable A and the various Gaussians shifted by γa_m where a_m is the corresponding eigenvalue. Irrespective of the strength of interaction, we end up with the same expression. This is not surprising since the Eqn. 1.31 was obtained without any restriction on strength of interaction. So far, the only difference is that all the different Gaussian pointer states corresponding to the different eigenstates of a_m would overlap when $\gamma \to 0$. Thus, the exact eigenstate cannot be inferred upon a single detection.

1.4.3 Pointer shift after Post-selection

After the interaction, let the system be subjected to a measurement corresponding to the observable B. Out of the many possible outcomes, we only consider the ensemble in which the outcome was the particular eigenstate $|b_n\rangle$. The resultant state of the system and the pointer can be obtained by applying the corresponding projection operator.

$$|b_n\rangle\langle b_n|\Psi_e\rangle = \sum_{m=0}^{d}\left[c_m|b_n\rangle\langle b_n|a_m\rangle \otimes \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{((x-\gamma a_m)-\mu)^2}{4\sigma^2}\right)\right)\right] \tag{1.45}$$

Since, we are interested in the intermediate measurement with observable A, we now represent the state $|b_n\rangle$ as a linear combination of eigenstates of A, i.e., $|b_n\rangle = \sum_{k=0}^{d} c'_k|a_k\rangle$.

$$|b_n\rangle \langle b_n|\Psi_e\rangle \tag{1.46}$$

$$= \sum_{m=0}^{d} \left[c_m |b_n\rangle \left(\sum_{k=0}^{d} \bar{c'_k} \langle a_k| \right) |a_m\rangle \otimes \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{((x-\gamma a_m)-\mu)^2}{4\sigma^2} \right) \right) \right] \tag{1.47}$$

$$= |b_n\rangle \sum_{m=0}^{d} \left[c_m \bar{c'_m} \left(\sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{((x-\gamma a_m)-\mu)^2}{4\sigma^2} \right) \right) \right] \tag{1.48}$$

Therefore, the final pointer state is described as a superposition of displaced Gaussians.

$$\left|\chi_f\right\rangle = \langle b_n|\Psi_e\rangle = \sum_{m=0}^{d} c_m \bar{c'_m} \sqrt[4]{\frac{1}{2\pi\sigma^2}} \exp\left(-\frac{((x-\gamma a_m)-\mu)^2}{4\sigma^2} \right) \tag{1.49}$$

In the above Eqn. 1.49 , the coefficients c_m and c'_m describes the component of the initial state ψ_i and final state $\psi_f = |b_n\rangle$ along the eigenstate $|a_m\rangle$ of the observable A. They are on the same footing and hence the pointer shift symmetrically depends upon both the pre-selected state and the post-selected state.

Now let us revisit the entire calculation of deriving the pointer shift, but this time without any assumption regarding the pointer state. The initial and final joint state of the system and pointer are described as $|\psi_i\rangle |\chi_i\rangle$ and $\left|\psi_f\right\rangle \left|\chi_f\right\rangle$ respectively. The evolution operator is given by $\exp\left(-\frac{i}{\hbar}\gamma A \otimes P\right)$. The final pointer state can then be computed as follows:

$$|\psi_f\rangle|\chi_f\rangle = \exp\left(-\frac{\imath}{\hbar}\gamma A\otimes P\right)|\psi_i\rangle|\chi_i\rangle \tag{1.50}$$

$$\Rightarrow |\chi_f\rangle = \langle\psi_f|\exp\left(-\frac{i}{\hbar}\gamma A\otimes P\right)|\psi_i\rangle|\chi_i\rangle \tag{1.51}$$

$$= \langle\psi_f|\left[1-\gamma\frac{i}{\hbar}A\otimes P+\frac{1}{2}\left(\frac{i}{\hbar}A\otimes P\right)^2+\ldots\right]|\psi_i\rangle|\chi_i\rangle \tag{1.52}$$

$$= \left[\langle\psi_f|\psi_i\rangle+\langle\psi_f|A|\psi_i\rangle\left(-\gamma\frac{i}{\hbar}P\right)+\frac{1}{2}\langle\psi_f|A^2|\psi_i\rangle\left(-\gamma\frac{i}{\hbar}P\right)^2+\ldots\right]|\chi_i\rangle \tag{1.53}$$

$$= \langle\psi_f|\psi_i\rangle\left[1+\frac{\langle\psi_f|A|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}\left(-\gamma\frac{i}{\hbar}P\right)+\frac{1}{2}\frac{\langle\psi_f|A^2|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}\left(-\gamma\frac{i}{\hbar}P\right)^2+\ldots\right]|\chi_i\rangle \tag{1.54}$$

To be able to approximate the series in the above expression to first order in γ, we require that

$$|\gamma^n\langle\psi_f|A^n|\psi_i\rangle| << |\gamma\langle\psi_f|A|\psi_i\rangle| << 1 \qquad \text{and} \tag{1.55}$$

$$|\gamma^n\langle\psi_f|A^n|\psi_i\rangle| << |\langle\psi_f|\psi_i\rangle| \qquad \forall\, n>1 \tag{1.56}$$

We further assume that $|\chi_i\rangle$ is well behaved such that $\gamma\langle\chi_i|P^n|\chi_i\rangle << 1$. Under the above assumptions, we can approximate the pointer state in Eqn. 1.54 as below.

$$|\chi_f\rangle \approx \langle\psi_f|\psi_i\rangle\left[1+\frac{\langle\psi_f|A|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}\left(-\gamma\frac{i}{\hbar}P\right)\right]|\chi_i\rangle \tag{1.57}$$

$$\approx \langle\psi_f|\psi_i\rangle\exp\left(-\frac{i}{\hbar}\gamma\frac{\langle\psi_f|A|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}P\right)|\chi_i\rangle \tag{1.58}$$

The exponential term in the above expression can be identified as the position translation operator in which momentum P acts as the generator. The amount of translation is determined by the coefficient $\gamma\frac{\langle\psi_f|A|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}$. Thus, the pointer shifts by an amount scaled by a value which given $\gamma << 1$ is dependent upon the states and the observable A. This value that results in the shift of the pointer in a measurement where the interaction is weak and the system is subjected to both pre-selection and post-selection is known as the *weak value*.

$$ {}_{\psi_f}\langle A\rangle^{\psi_i}_{(w)} = \frac{\langle \psi_f | A | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \tag{1.59} $$

1.5 Weak value

The weak value in general is a complex quantity. Since the initial and final states are allowed to be nearly orthogonal the weak value is unbounded as $\langle \psi_f | \psi_i \rangle \to 0$ provided A is not the eigen operator for $|\psi_i\rangle$ or $|\psi_f\rangle$. This implies that, when weak value is real, weak measurement can make the pointer variable shift by a distance much larger than what the strong measurement could. This can help to amplify small signals . However, the cost is that success rate of pre- and post-selection would decrease dramatically. However, we have to take care of the assumptions here. Since we would require $\gamma \langle A\rangle_w << 1$, when $|\langle A\rangle_w| \to \infty$ we need to establish the bound for γ. In particular, when the initial pointer state is Gaussian $g_x(\mu,\sigma)$ with mean μ and standard deviation σ, the expression 1.58 can simplified as follows:

$$ |\chi_f\rangle = \langle \psi_f | \psi_i \rangle \exp\left(-\frac{i}{\hbar}\gamma \langle A\rangle_w P\right) g_x(\mu,\sigma) \tag{1.60} $$

$$ \overset{assumption}{\approx} \langle \psi_f | \psi_i \rangle \, g_x(\mu + \gamma \langle A\rangle_w, \sigma) \tag{1.61} $$

The assumption in the above expression is that the final pointer state still needs to be approximately described by a Gaussian which is shifted. For this, we require $\gamma \langle A\rangle_w << \sigma$. As pre- and post-selection approach orthogonality, for a given strength of interaction γ, we need to have a larger initial pointer uncertainty σ to be able to obtain the large weak value $\langle A\rangle_w$. The fact that we can obtain $|\langle A\rangle_w| > a_{max}$ at first may seem non-intuitive and almost may feel wrong. How can a linear combination of eigenstates with eigen values a_m give rise to a resultant state with shift larger than a_{max}. The reason this is possible is because the coefficients are in general complex and hence the various pointer states corresponding to different orthogonal eigenstates $|a_m\rangle$ can interfere once the system is projected to the final state $|\psi_f\rangle$. Thus, the pointer shifts proportional to the weak value as a result of *interference*. And as with any interference scenario, to have good visibility in the

resultant pattern we need minimal distinguishability of the superposing states. Therefore, to obtain weak value, the pointer states corresponding to different eigenstates must have maximal overlap.

In general, we can have the pre-selected state[8] as a mixed state described by the density matrix ρ_i. Then, the weak value of the observable A is obtained as

$$\langle A \rangle_w = \frac{\langle \psi_f | A \rho_i | \psi_f \rangle}{\langle \psi_f | \rho_i | \psi_i \rangle} \tag{1.62}$$

Note that if $\rho_i = |\psi_i\rangle \langle \psi_i|$, i.e, a pure state the above expression reduces to the expression for weak value for pure pre-selected state given in Eqn. 1.59.

1.5.1 ABL rule

So far we had only the interaction Hamiltonian that couples the system and the pointer giving rise to the unitary time evolution. However, in general, there is time evolution due to the free Hamiltonian. If a system is prepared in state $|\psi_i\rangle$ at time t_i and finally measured at time t_f with the operator A_f and one of the eigenstate $|\psi_f\rangle$ is post-selected, the weak value of the observable A must account for the evolution of state. In particular, consider the simpler case when the state at time $t : t_i < t < t_f$ is in one of the eignestate $|a_m\rangle$ of the observable A. The transition amplitude from t_i to t is given by $c_{t_i \to t} = \langle a_m | U(t_i \to t) | \psi_i \rangle$. Similarly, the transition amplitude from the intermediate state $|a_m\rangle$ to the final state $|\psi_f\rangle$ at time t_f, is given by $c_{t \to t_f} = \langle \psi_f | U(t \to t_f) | a_m \rangle$. The overall amplitude for the entire process of the particle being subjected to pre- and post-selection and being in state $|a_m\rangle$ at an intermediate time t is given by the amplitude

$$c_{t_i \to t \to t_f} = c_{t \to t_f} c_{t_i \to t} \tag{1.63}$$

$$= \langle \psi_f | U(t \to t_f) | a_m \rangle \langle a_m | U(t_i \to t) | \psi_i \rangle \tag{1.64}$$

Therefore, the conditional probability of the particle being in state $|a_m\rangle$ at time t while

[8] Although the post-selected state could also be considered as a mixed state in general, there is little literature on that. In practice, both pre-selected and post-selected states obtained in an experiment would be mixed. Therefore, a generic formula for weak value where both states can be in general mixed might be useful.

subjected to both pre- and post-selection is given by the following, which is known as the ABL rule [1.35, 1.5].

$$P(|a_m\rangle(t)\Big||\psi_i\rangle(t_i),|\psi_f\rangle(t_f)) = \frac{|\langle\psi_f|U(t\to t_f)|a_m\rangle\langle a_m|U(t_i\to t)|\psi_i\rangle|^2}{\sum_k|\langle\psi_f|U(t\to t_f)|a_k\rangle\langle a_k|U(t_i\to t)|\psi_i\rangle|^2} \tag{1.65}$$

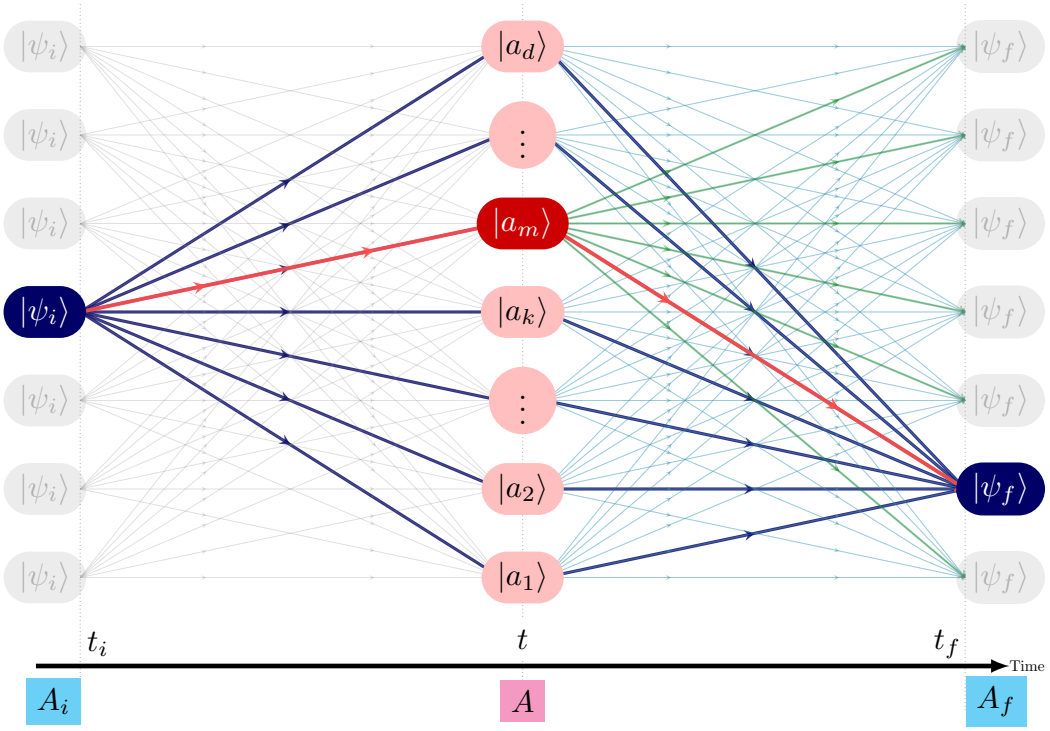


Figure 1.10: At time t_i, one of the many possible eigenstates of the observable A_i was choosen as the initial state $|\psi_i\rangle$. The state of the system at time t is enquired in terms of the eigenstate of the observable A. For a d- dimensional Hilbert space all d eigenstates are in general possible. All the possible outcomes may evolve into any of the eigenstate of A_f when we make the final measurement with the observable A_f at time t_f. However, we post-select only one of the eigenstate $|\psi_f\rangle$. All such paths between the pre-selected and post-selected state are shown with blue arrows. In one run of the experiment, when we make a strong measurement corresponding to the observable A, we obtain one of the eigenstate, say $|a_m\rangle$. The probability of obtaining the same conditioned to pre- an post-selection is the ratio of probability of the particle taking the path as shown in red to the sum of all the probabilities of the particle taking the paths shown in blue.

The numerator in the ABL rule Eqn. 1.65 is the unconditional probability of the transition from the initial state $|\psi_i\rangle$ to the final state $|\psi_f\rangle$ via the state $|a_m\rangle$ at an intermediate time. The denominator gives us the sum of probabilities for all such paths via the state

$|a_k\rangle$. The summation is carried over all the eigenstates of the Hermitian observable A that forms a orthogonal basis. Therefore in the frequentist sense of probability, the ratio of the two gives us the probability that the intermediate state is $|a_m\rangle$ conditioned to both pre- and post-selection.

As a special case, $P(|a_m\rangle(t)||\psi_i\rangle(t_i),|\psi_f\rangle(t_f))=0$, the intermediate state $|a_m\rangle$ would never occur in the given pre- and post-selected ensemble. Asking any questions in the context of such states would be deemed as *counterfactual*. On the other hand, when $P(|a_m\rangle(t)||\psi_i\rangle(t_i),|\psi_f\rangle(t_f))=1$, the intermediate state $|a_m\rangle$ is guaranteed to occur in every run of the experiment for the pre- and post-selected ensemble. Such states and its associated properties like the eigenvalue a_m associated with the state would be regarded as *elements of physical reality*. When we have intermediate strong measurement satisfying either $[A,A_i]=0$ or $[A,A_f]=0$ (or both) , the shared eigenstate would trivially correspond to as an element of physical reality. In all such cases, if we perform a weak measurement and obtain the weak value to be the same as the eigenvalue a_m, we can claim that we would indeed have obtained the same result had we performed a strong measurement instead.

Can we obtain the weak value equal to one of the eigenvalue of A non-trivially i.e., when $[A,A_i]\neq 0$ and $[A,A_f]\neq 0$. As an example, consider a qutrit with the initial state as one of the eigenstate of the $\sigma_x^{(3)}$ operator. We also choose the final state to be one of the eigenstate of the $\sigma_y^{(3)}$ observable. The particular choice of eigenstates given below in Eqn. 1.66 are in represented in the computational basis $\{|0\rangle,|1\rangle,|2\rangle\}$.

$$|\psi_i\rangle=\frac{1}{2}\begin{pmatrix}1\\ \sqrt{2}\\ 1\end{pmatrix}\qquad |\psi_f\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}i\\ 0\\ 1\end{pmatrix} \tag{1.66}$$

Let's consider the weak value of the projector $\Pi_1=|1\rangle\langle 1|$. We find that

$$\langle\Pi_1\rangle^{(w)}=\frac{\langle\psi_f|\Pi_1|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}=0 \tag{1.67}$$

The weak value obtained is one of the eigenvalue of the observable Π_1. The observable does not commute with $\sigma_x^{(3)}$ or $\sigma_y^{(3)}$. Thus, can we ascribe any sense of physical realism to the state of the system at the intermediate time? As a caution, let's note that there are

two eigenstates corresponding to the eigenvalue of 0 for the observable Π_1. When there is degeneracy, we cannot ascribe much significance to the eigenstates themselves. But is the weak value of 0 signifies an element of reality ? Let's go to the ABL rule and consider all the possible outcomes of the hypothetical measurement of the observable Π_1 at an intermediate time. When the outcome of the intermediate measurement yields the value 1, we know that the corresponding state is given by $(0,1,0)$ as it is the only eigenstate with that eigenvalue. This state is orthogonal to $\left|\psi_f\right\rangle$. Thus, the probability of transition from pre-selected state to the post-selected with a state in between with value of observable Π_1 being 1 is 0.

$$
\begin{aligned}
& P(\Pi_1 = 1 \big| \left|\psi_i\right\rangle(t_i), \left|\psi_f\right\rangle(t_f)) \\
& = \frac{|\left\langle\psi_f\right| (\Pi_1 = 1) \left|\psi_i\right\rangle|^2}{|\left\langle\psi_f\right| (\Pi_1 = 1) \ \left|\psi_i\right\rangle|^2 + |\left\langle\psi_f\right| (\Pi_1 \neq 1) \ \left|\psi_i\right\rangle|^2} = 0 \qquad (1.68) \\
& \Rightarrow P(\Pi_1 \neq 1 \big| \left|\psi_i\right\rangle(t_i), \left|\psi_f\right\rangle(t_f)) = 1 \qquad (1.69) \\
& \Rightarrow P(\Pi_1 = 0 \big| \left|\psi_i\right\rangle(t_i), \left|\psi_f\right\rangle(t_f)) = 1 \qquad (1.70)
\end{aligned}
$$

Since the only other possible eigenvalue is 0, it must occur with certainty for the pre- and post-selected ensemble. Therefore, when the weakvalue becomes the eigenvalue of the observable, the ABL rule suggest that in the hypothetical intermediate strong measurement we would have obtained the same value as well.

However, it is possible that for a given pre(or post)-selected state, for two different post(or pre)-selected states the weak value of the observable may be the same eigenvalue. Let's explore whether in both case the ABL rule indicates the weak value to be an element of physical reality. Let's have $\left|\psi_i\right\rangle = (1/2)(1,\sqrt{2},1)$ as in the previous example. Now let's consider the final state to be $\left|\psi_f\right\rangle = (0,1,0) = |1\rangle$. The observable under consideration is $\sigma_z^{(3)}$. We find the weak value ${}^{\psi_f}\left\langle\sigma_z^{(3)}\right\rangle_w^{\psi_i} = 0$. Since $\left|\psi_f\right\rangle$ is one of the eigenstate of the observable $\sigma_z^{(3)}$, we are not surprised to be find the corresponding eigenvalue as the weak value. There are three eigenstates of σ_z and hence let's explicitly compute what probability the ABL rule ascribes to the state $|1\rangle$.

$$
\begin{aligned}
&P(|1\rangle \,\big|\, |\psi_i\rangle, |\psi_f\rangle = (0,1,0)) \\
&= \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2}{|\langle\psi_f|0\rangle\langle 0|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\ |\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\ |\psi_i\rangle|^2} = 1
\end{aligned}
\tag{1.71}
$$

Since the post-selected state is $|1\rangle$, the first and last terms in the denominator of the above expression goes to zero due to orthogonality. Thus, the weak value of 0 indicates that the corresponding eigenstate $|1\rangle$ occurs with certainty as per the ABL rule. Now for the same initial state, we $|\psi_i\rangle = (1/2)(1,\sqrt{2},1)$, we change the final state to $|\psi_f\rangle = (1/\sqrt{2})(1,0,1)$. This state is no longer the eigenstate of the observable $\sigma_z^{(3)}$. Yet, if we compute the weak value, we obtain it to be 0, one of the eigenvalues of the observable $\sigma_z^{(3)}$. Can we conclude that the intermediate state would be the corresponding eigenstate $|1\rangle$? Let's compute the probability using the ABL rule.

$$
\begin{aligned}
&P(|1\rangle \,\big|\, |\psi_i\rangle, |\psi_f\rangle = \frac{1}{2}(1,0,1)) \\
&= \frac{|\langle\psi_f|1\rangle\langle 1|\psi_i\rangle|^2}{|\langle\psi_f|0\rangle\langle 0|\psi_i\rangle|^2 + |\langle\psi_f|1\rangle\langle 1|\ |\psi_i\rangle|^2 + |\langle\psi_f|2\rangle\langle 2|\ |\psi_i\rangle|^2} = 0
\end{aligned}
\tag{1.72}
$$

We obtain the above by explicitly computing the probability amplitude in the numerator.

$$
\langle\psi_f|1\rangle\langle 1|\psi_i\rangle = \frac{1}{2}\begin{pmatrix}1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix}1\\ \sqrt{2}\\ 1\end{pmatrix} = 0 \tag{1.73}
$$

However, if we compute the probability of obtaining the state $|0\rangle$ or $|2\rangle$ we obtain both of them to be $1/2$.

$$
P(|0\rangle \,\big|\, |\psi_i\rangle, |\psi_f\rangle) = \frac{1}{2} \qquad P(|2\rangle \,\big|\, |\psi_i\rangle, |\psi_f\rangle) = \frac{1}{2} \tag{1.74}
$$

Thus, as per the ABL rule, none of the eigenstates of the observable σ_z occurs with certainty. Yet, we had obtained the weak value to be equal to the eigenvalue. This occurred because $(p = 1/2) \times 1 + (p = 0) \times 0 + (p = 1/2) \times -1 = 0$. The average of the eigenvalues weighed with the probabilities computed with the ABL rule happened to be another eigenvalue.

Also, if we compute the weak value of the various projectors we obtain the following:

$$\langle \Pi_0 \rangle_w = \frac{1}{2} \qquad \langle \Pi_1 \rangle_w = 0 \qquad \langle \Pi_2 \rangle_w = \frac{1}{2} \tag{1.75}$$

The weak values of observables are additive i.e, for the same pre- and post-selection, we have $\langle A+B \rangle_w = \langle A \rangle_w + \langle B \rangle_w$. Therefore, we can decompose $\sigma_z^{(3)}$ in terms of the projectors to compute the weak value.

$$\sigma_z^{(3)} = \Pi_0 - \Pi_2 \tag{1.76}$$

$$\Rightarrow \left\langle \sigma_z^{(3)} \right\rangle_w = \langle \Pi_0 \rangle_w - \langle \Pi_2 \rangle_w \tag{1.77}$$

$$= \frac{1}{2} - \frac{1}{2} = 0 \tag{1.78}$$

As a side note, we note that the product rule does not apply to the weak values i.e. $\langle AB \rangle_w \neq \langle A \rangle_w \times \langle B \rangle_w$. This can be easily observed in the above example as $\langle \Pi_0 \Pi_2 \rangle_w = 0$ where as $\langle \Pi_0 \rangle_w \times \langle \Pi_2 \rangle_w = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \neq 0$.

One needs to be careful in the interpretation of the weak value. Merely, it being same as the eigenvalue do not make the corresponding eigenstate an element of physical reality in the pre- and post-selected system.

However, if the eigenvalue turns out to be maximum or minimum of the possible set of eigenvalues, then it is impossible to obtain the same eigenvalue by weighted average of other eigenvalues. Therefore, if the eigenvalue is extreme and we obtain the weak value as that eigenvalue ABL rule would suggest that the corresponding eigenstate occurs with certainty and the weak value would represent an element of physical reality. As a corollary, since both the eigenvalues of a dichotomic observable would be extreme, obtaining one of them as the weak value would ascribe certainty of obtaining that value in the ABL rule [1.35]. The other eigenvalue would never occur. Therefore, projectors (in any dimension) are often used in evaluating the weak value so that when the weak value is equal to 0 or 1 it immediately suggests what the probability computed using ABL rule would be. In 2-dimensional Hilbert space, observables do not have degenerate eigenstates and hence the connection between ABL rule and the weak value is trivial, i.e., if the weak value is the

same as the eigenvalue it is a certainty.

1.5.2 Two-state vector formalism

We would now like to include the evolution operators in Eqn. 1.65 directly in the description of state. By convention, we shall call $\left|\psi_{pre}\right\rangle = U(t_i \to t)\left|\psi_i\right\rangle$ as the *pre-selected state*. This is the prepared state at time t_i evolved forward in time till t. It is assumed that any interaction or measurement of A at time t takes negligible time. If interaction is over a period $2t_\Delta$, then we evolve the state till $t - t_\Delta$. In particular, we would consider $t_\Delta \to 0$ if the interaction is impulsive and the interaction strength has a compact support. We would now attempt to apply the unitary to the final state $\left|\psi_f\right\rangle$ instead of the eigenstate $|a_m\rangle$ as per Eqn. 1.65.

$$\left\langle\psi_f\right| U(t \to t_f) = \left\langle U^\dagger_{t\to t_f}\psi_f\right| = \left\langle U_{t_f\to t}\psi_f\right| \overset{def}{=} \left\langle\psi_{post}\right| \tag{1.79}$$

We could perform the last step because of the following equivalence for the given Hamiltonian H.

$$U^\dagger_{t\to t_f} = \left[\exp\left(-iH(t_f - t)\right)/\hbar\right]^\dagger = \exp\left(iH(t_f - t)/\hbar\right) = \exp\left(-iH(t - t_f)/\hbar\right) = U_{t_f\to t} \tag{1.80}$$

However, $\left\langle U_{t_f\to t}\psi_f\right|$ has a significantly different interpretation. It is the final state evolved backward in time from t_f to t. We shall call this backward evolved state as the *post-selected state* $\left|\psi_{post}\right\rangle$, which guarantees the detection of the selected state $\left|\psi_f\right\rangle$. Note that any state orthogonal to $\left|\psi_{post}\right\rangle$ would also remain orthogonal to $\left|\psi_f\right\rangle$ and hence would not be successfully detected. In some cases, the unitary evolution would be inconsequential as it would evolve the state by adding a global phase. In such cases, we would loosely treat $\left|\psi_f\right\rangle$ and $\left|\psi_{post}\right\rangle$ as equivalent and use them interchangeably. Similarly, unless otherwise the unitary evolution is explicitly mentioned we would treat $\left|\psi_{pre}\right\rangle$ and $|\psi_i\rangle$ interchangeably. Now, the ABL rule can be written as in terms of the initial state evolving forward in time $|\overrightarrow{\psi_i}\rangle$ and the final state evolving backward in time $\left|\overleftarrow{\psi_f}\right\rangle$.

$$P(|a_m\rangle\,|(\langle\overleftarrow{\psi_f}||\overrightarrow{\psi_i}\rangle)) = \frac{|\langle\overleftarrow{\psi_f}|\Pi_m|\overrightarrow{\psi_i}\rangle|^2}{\sum_k |\langle\overleftarrow{\psi_f}|\Pi_k|\overrightarrow{\psi_i}\rangle|^2} \tag{1.81}$$

The probability assignment in the ABL rule depends symmetrically on the pre-selected and the post-selected state. The LHS of the above euqation Eqn. 1.81 remains invariant on exchange of $|\overrightarrow{\psi_i}\rangle$ and $|\overleftarrow{\psi_f}\rangle$. Therefore, one may interpret that the value of an observable at time t in a pre- and post-selected system not only depends on the state in its past but also on the state that it is going to be in the future. The above do not necessarily imply retro-causality i.e., future affecting the present. This two states must be treated as boundary conditions in time. Just like in Lagrangian formalism we compute the path after fixing the state at two different times, these two states do the same purpose.

Thus, we can attempt to have the complete description of quantum mechanics [1.32, 1.29] by having the quantum state as a set of two state vectors: $\langle\overleftarrow{\psi_f}||\overrightarrow{\psi_i}\rangle$. In general, we can describe the quantum state as the superposition of many two-state vectors as $\sum_k \alpha_k \langle\overleftarrow{\psi_f}|_k|\overrightarrow{\psi_i}\rangle_k$. Since the two-vector quantum state describes a pre- and post-selected system uniquely, it also allow us to attribute a definite weak value of an observable for a given quantum two-state. Since, weak value would be experimentally obtained using weak measurement, the two-state vector allows us to predict the result of a weak measurement with certainty. In this sense, we ascribe the element of physical reality to the weak value $\langle A\rangle_w = \frac{\langle\overleftarrow{\psi_f}|A|\overrightarrow{\psi_i}\rangle}{\langle\overleftarrow{\psi_f}||\overrightarrow{\psi_i}\rangle}$.

The two-state vector formalism provides insightful framework towards a time-symmetric formulation of quantum mechanics. However, most of questions arising in a pre- and post-selected quantum system can be addressed in the standard quantum mechanics. Also, the operational results are so far equivalent with few controversial exceptions. Therefore, we would use the one-state formalism in the context of pre- and post selection wherever possible to compute the weak values. New interpretation in terms of the two-state can be thought of as reformulation of the quantum mechanics giving (almost) the same answers, just like Feynman's path integral was a reformulation of quantum dynamics governed by Schrodinger's equation.

1.5.3 Complex weak values and Pointer Shift - Josza theorem

So far we have compared the weak value with eigenvalues of observable and in doing so implied that we are considering the weak value to be a real number or are considering only the real part of it. For example, while describing the physical effect of weak value, we stated it causes proportional translation of the pointer state as was evident from Eqn. 1.60 and Eqn. 1.61. But strictly speaking, that is true only when the weak value is real. When the weak value is in general complex, it is not simple to comprehend the action of the term $\exp\left(-\frac{i}{\hbar}\gamma\langle A\rangle_w P\right)$ on the pointer state [1.36]. First of all, the pointer state may not always be Gaussian. In general, the requirement for pointer state to be suitable for carrying out weak measurement is that it should have a vanishing[9] current density[1.37]. Now assuming that the pointer state is suitable, we aim to comprehend the effect of the complex weak value on the mean and uncertainty(variance) of the pointer state distribution.

Let the initial pointer state be given as $|\chi_i\rangle$ and after the weak interaction and post-selection it becomes $|\chi_f\rangle$. Using the Eqn. 1.57, we have the relation between the initial and final pointer state as

$$|\chi_f\rangle \approx \langle\psi_f|\psi_i\rangle\,(1-\frac{i}{\hbar}\gamma\langle A\rangle_w P)\,|\chi_i\rangle \tag{1.82}$$

Note that when we truncate the evolution operator to the first order term only, it is no longer unitary. Therefore, the pointer state needs to be appropriately normalized while considering mean and variance. The mean of position and momentum in the initial and final pointer state are therefore given as below:

$$\langle X\rangle_i = \langle\chi_i|X|\chi_i\rangle \qquad \langle X\rangle_f = \frac{\langle\chi_f|X|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle} \tag{1.83}$$

$$\langle P\rangle_i = \langle\chi_i|P|\chi_i\rangle \qquad \langle P\rangle_f = \frac{\langle\chi_f|P|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle} \tag{1.84}$$

We aim to compute the shift in position and momentum of the pointer state i.e., we aim to compute $\Delta X = \langle X\rangle_f - \langle X\rangle_i$ and $\Delta P = \langle P\rangle_f - \langle P\rangle_i$. The weakness of interaction is based on comparison of the shift in the mean position with the uncertainty of the pointer

[9] In general, if the initial pointer state is a mixed state ρ_χ, the vanishing current density implies that $\langle x|P\rho_\chi|x\rangle + \langle x|\rho_\chi P|x\rangle = 0$.

state distribution. Hence, it would be useful to define the variance of the initial pointer state distribution.

$$\sigma_X^2 = \langle\chi_i|X^2|\chi_i\rangle - \langle X\rangle_i^2 \qquad \sigma_P^2 = \langle\chi_i|P^2|\chi_i\rangle - \langle P\rangle_i^2 \tag{1.85}$$

It is assumed that the weak interaction do not change the variance of the pointer distribution.

It is implicitly assumed that the pointer state is a wavefunction of a continuous variable. Thus, we can write the wavefunctions and operators in position domain. Hence Eqn. 1.82 can be written as below.

$$\chi_f(x) = \langle\psi_f|\psi_i\rangle\left(1 - \frac{i}{\hbar}\gamma\langle A\rangle_w(-i\hbar)\frac{\partial}{\partial x}\right)\chi_i(x) \tag{1.86}$$

$$= \langle\psi_f|\psi_i\rangle\left(\chi_i(x) - \gamma\langle A\rangle_w\frac{\partial\chi_i(x)}{\partial x}\right) \tag{1.87}$$

$$\Rightarrow \overline{\chi_f(x)} = \overline{\langle\psi_f|\psi_i\rangle}\left(\overline{\chi_i(x)} - \gamma\overline{\langle A\rangle_w}\frac{\partial\overline{\chi_i(x)}}{\partial x}\right) \tag{1.88}$$

in the last expression, we have taken the complex conjugate of the final pointer state. Now multiplying Eqn. 1.87 and 1.88, we have

$$\overline{\chi_f(x)}\chi_f(x) \approx \overline{\langle\psi_f|\psi_i\rangle}\langle\psi_f|\psi_i\rangle$$

$$\left(\overline{\chi_i(x)}\chi_i(x) - \gamma\langle A\rangle_w\overline{\chi_i(x)}\frac{\partial\chi_i(x)}{\partial x} - \gamma\overline{\langle A\rangle_w}\frac{\partial\overline{\chi_i(x)}}{\partial x}\chi_i(x) + \mathcal{O}(\gamma^2)\right) \tag{1.89}$$

$$= |\langle\psi_f|\psi_i\rangle|^2\left(\overline{\chi_i(x)}\chi_i(x) - \gamma\mathrm{Re}\{\langle A\rangle_w\}\left[\overline{\chi_i(x)}\frac{\partial\chi_i(x)}{\partial x} + \frac{\partial\overline{\chi_i(x)}}{\partial x}\chi_i(x)\right]\right. \tag{1.90}$$

$$\left. - i\gamma\mathrm{Im}\{\langle A\rangle_w\}\left[\frac{\partial\overline{\chi_i(x)}}{\partial x}\chi_i(x) - \overline{\chi_i(x)}\frac{\partial\chi_i(x)}{\partial x}\right]\right)$$

$$= |\langle\psi_f|\psi_i\rangle|^2\left(\rho_{\chi_i} - \gamma\mathrm{Re}\{\langle A\rangle_w\}\frac{\partial\rho_{\chi_i}}{\partial x} - \gamma\mathrm{Im}\{\langle A\rangle_w\}\frac{2m}{\hbar}J_{\chi_i}\right) \tag{1.91}$$

Here, ρ_{χ_i} and J_{χ_i} are the probability density and the probability current density respectively. We have assumed that the particle bearing the pointer state has mass[10] m and is governed in Schrodinger's equation.

Next, we compute the expectation value of position in the final pointer state.

[10]For photons, we may need to take an alternative approach and account for the Maxwell's equation instead of the Schrodinger's equation

$$\langle X\rangle_f = \frac{\int \overline{\chi_f(x)} X \chi_f(x) dx}{\int \overline{\chi_f(x)} \chi_f(x) dx} \tag{1.92}$$

$$= \frac{\int \left(x\rho_{\chi_i} - \gamma \mathrm{Re}\{\langle A\rangle_w\} x \frac{\partial \rho_{\chi_i}}{\partial x} - \gamma \mathrm{Im}\{\langle A\rangle_w\} \frac{2m}{\hbar} x J_{\chi_i} \right) dx}{\int \left(\rho_{\chi_i} - \gamma \mathrm{Re}\{\langle A\rangle_w\} \frac{\partial \rho_{\chi_i}}{\partial x} - \gamma \mathrm{Im}\{\langle A\rangle_w\} \frac{2m}{\hbar} J_{\chi_i} \right) dx} \tag{1.93}$$

$$\approx \int \left(x\rho_{\chi_i} - \gamma \mathrm{Re}\{\langle A\rangle_w\} x \frac{\partial \rho_{\chi_i}}{\partial x} - \gamma \mathrm{Im}\{\langle A\rangle_w\} \frac{2m}{\hbar} x J_{\chi_i} \right) dx \tag{1.94}$$

$$\times \left(1 + \gamma \mathrm{Re}\{\langle A\rangle_w\} \int \frac{\partial \rho_{\chi_i}}{\partial x} dx + \gamma \mathrm{Im}\{\langle A\rangle_w\} \frac{2m}{\hbar} \int J_{\chi_i} dx \right) + \mathcal{O}(\gamma^2)$$

In the denominator, we have evaluated the first term as $\int \rho_{\chi_i} dx = 1$ as the initial pointer state is normalized. Then we have used the Maclaurin series expansion of $1/(1-\gamma) \approx 1 + \gamma + \mathcal{O}(\gamma^2)$ since $\gamma \to 0$. We further simplify by evaluating $\int_{-\infty}^{\infty} \frac{\partial \rho_{\chi_i}}{\partial x} dx = \rho|_{-\infty}^{\infty} = 0$. In the Eqn. 1.94, we identify the first term as $\langle X\rangle_i = \int x\rho_{\chi_i} dx$. Then we simplify the expression by multiplying the two terms and keeping the terms upto first oreder in γ.

$$\langle X\rangle_f = \langle X\rangle_i - \gamma \mathrm{Re}\{\langle A\rangle_w\} \int x \frac{\partial \rho_{\chi_i}}{\partial x} dx - \gamma \mathrm{Im}\{\langle A\rangle_w\} \frac{2m}{\hbar} \int (x - \langle X\rangle_i) J_{\chi_i} dx + \mathcal{O}(\gamma^2) \tag{1.95}$$

We evaluate the second term in the RHS of the above Eqn. 1.95 using integration by parts.

$$\int_{-\infty}^{\infty} x \frac{\partial_t \rho_{\chi_i}}{\partial_x} = -\int_{-\infty}^{\infty} \frac{dx}{dx} \rho_{\chi_i} dx + x\rho_{\chi_i} \Big|_{-\infty}^{\infty} \tag{1.96}$$

$$= -1 + 0 = -1 \tag{1.97}$$

The last term evaluates to 0 because the pointer state is expected to go to zero as x tends to infinity so that the probability can be normalized. For normalization in Hilbert space, it is needed that ψ falls faster [1.1] as $x \to \infty$ compared to the growth of x^2. The third term is also evaluated with integration by parts as follows:

$$\int_{-\infty}^{\infty} J_{\chi_i}(x-\langle X\rangle_i)dx = \int_{-\infty}^{\infty} J_{\chi_i}\frac{1}{2}\frac{\partial(x-\langle X\rangle_i)^2}{\partial x}dx \tag{1.98}$$

$$= -\frac{1}{2}\int_{-\infty}^{\infty}\frac{dJ_{\chi_i}}{dx}(x-\langle X\rangle_i)^2 + \frac{1}{2}J_{\chi_i}(x-\langle X\rangle_i)^2\Big|_{-\infty}^{\infty} \tag{1.99}$$

$$= \frac{1}{2}\frac{d}{dt}\int(\rho_{\chi_i}(x-\langle X\rangle_i)^2)dx \tag{1.100}$$

In the last step, we have used the continuity equation in quantum mechanics that relates probability density with the gradient of current density i.e. $\frac{\partial}{\partial x}J + \frac{d}{dt}\rho = 0$. Further, we have $\int(\rho_{\chi_i}(x-\langle X\rangle_i)^2)dx = \sigma_X^2$.

Therefore, we have

$$\Delta X = \gamma \mathrm{Re}\{\langle A\rangle_w\} - \gamma \mathrm{Im}\{\langle A\rangle_w\}\frac{m}{\hbar}\frac{d}{dt}\sigma_x^2 \tag{1.101}$$

In particular, when the weak value is real, we see that the pointer variable is shifted proportionally. the imaginary part of the weak value affects the pointer shift only when the variance and hence the uncertainty [11] in the pointer state evolves with time. Note that the above is for a particle with mass. For massless particles like photons, we shall reconsider the above expression 1.101 in terms of rate of spread of momentum.

We could compute ΔP similarly. But it is illustrative to compute the expectation value of any pointer observable M in the final pointer state in the Heisenberg picture.

[11] Sometimes it is argued that the wavefunction is real, its variance remains unchanged in time. However, one needs to be careful. Even if the wavefunction can be written as a real function at an initial time, the Schrödinger equation evolves it to be a complex function in general.

$$\langle M\rangle_f = \frac{\langle\chi_f|M|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle} \tag{1.102}$$

$$= \frac{\langle\chi_i|\,(1+(i/\hbar)\gamma\overline{\langle A\rangle_w}P)M(1-(i/\hbar)\gamma\langle A\rangle_w P)\,|\chi_i\rangle}{\langle\chi_i|\,(1+(i/\hbar)\gamma\overline{\langle A\rangle_w}P)(1-(i/\hbar)\gamma\langle A\rangle_w P)\,|\chi_i\rangle} \tag{1.103}$$

$$= \frac{\langle\chi_i|\,(M-(i/\hbar)\gamma\langle A\rangle_w MP+(i/\hbar)\gamma\overline{\langle A\rangle_w}PM)\,|\chi_i\rangle+\mathcal{O}(\gamma^2)}{\langle\chi_i|\,(1-(i/\hbar)\gamma\langle A\rangle_w P+(i/\hbar)\gamma\overline{\langle A\rangle_w}P)\,|\chi_i\rangle+\mathcal{O}(\gamma^2)} \tag{1.104}$$

$$\approx \frac{\langle M\rangle_i+(i/\hbar)\gamma\mathrm{Re}\{\langle A\rangle_w\}\langle[P,M]\rangle_i+(1/\hbar)\gamma\mathrm{Im}\{\langle A\rangle_w\}\langle\{P,M\}\rangle_i}{1+(1/\hbar)\gamma\,2\,\mathrm{Im}\{\langle A\rangle_w\}\langle P\rangle_i} \tag{1.105}$$

$$\approx (\langle M\rangle_i+(i/\hbar)\gamma\mathrm{Re}\{\langle A\rangle_w\}\langle[P,M]\rangle_i + (1/\hbar)\gamma\mathrm{Im}\{\langle A\rangle_w\}\langle\{P,M\}\rangle_i)(1-(1/\hbar)\gamma\,2\,\mathrm{Im}\{\langle A\rangle_w\}\langle P\rangle_i) \tag{1.106}$$

$$= \langle M\rangle_i+(i/\hbar)\gamma\mathrm{Re}\{\langle A\rangle_w\}\langle[P,M]\rangle_i + (1/\hbar)\gamma\mathrm{Im}\{\langle A\rangle_w\}(\langle\{P,M\}\rangle_i-2\langle M\rangle_i\langle P\rangle_i)+\mathcal{O}(\gamma^2) \tag{1.107}$$

Here, $[P,M]=PM-MP$ is the commutator of the two observables and $\{P,M\}=PM+MP$ is the anti-commutator. When $M=P$, the commutator disappears and only the imaginary part of the weak value survives.

$$\Delta P = \frac{1}{\hbar}\gamma\mathrm{Im}\{\langle A\rangle_w\}2(\langle P^2\rangle-\langle P\rangle^2) \tag{1.108}$$

$$= \frac{1}{\hbar}\gamma\mathrm{Im}\{\langle A\rangle_w\}2\sigma_P^2 \tag{1.109}$$

Thus, the imaginary part of the weak value changes the momentum of the pointer proportional to the momentum variance of the pointer state.

References

[1.1] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. 3rd ed. Cambridge University Press, 2018. DOI: 10.1017/9781316995433.

[1.2] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[1.3] N. Bohr. "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" In: *Phys. Rev.* 48 (8 1935), pp. 696–702. DOI: 10.1103/PhysRev.48.696.

[1.4] Gary Bowman. *Essential Quantum Mechanics*. English (US). Oxford University Press, 2008. ISBN: 9780199228928. DOI: 10.1093/acprof:oso/9780199228928.001.0001.

[1.5] Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz. "Time Symmetry in the Quantum Process of Measurement". In: *Phys. Rev.* 134 (6B 1964), B1410–B1416. DOI: 10.1103/PhysRev.134.B1410.

[1.6] J. von Neumann, R.T. Beyer, and N.A. Wheeler. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 2018. ISBN: 9781400889921.

[1.7] Wojciech Hubert Zurek. "Decoherence, einselection, and the quantum origins of the classical". In: *Rev. Mod. Phys.* 75 (3 2003), pp. 715–775. DOI: 10.1103/RevModPhys.75.715.

[1.8] Justin Dressel et al. "Colloquium: Understanding quantum weak values: Basics and applications". In: *Rev. Mod. Phys.* 86 (1 2014), pp. 307–316. DOI: 10.1103/RevModPhys.86.307.

[1.9] Boaz Tamir and Eliahu Cohen. "Introduction to Weak Measurements and Weak Values". In: *Quanta* 2.1 (2013), p. 7. DOI: 10.12743/quanta.v2i1.14.

[1.10] Bengt E. Y. Svensson. "Pedagogical Review of Quantum Measurement Theory with an Emphasis on Weak Measurements". In: *Quanta* 2.1 (2013), p. 18. DOI: 10.12743/quanta.v2i1.12.

[1.11] Lev Vaidman. "Weak-measurement elements of reality". In: *Foundations of Physics* 26 (1996), pp. 895–906. DOI: doi.org/10.1007/BF02148832.

[1.12] Yakir Aharonov et al. "The Weak Reality That Makes Quantum Phenomena More Natural: Novel Insights and Experiments". In: *Entropy* 20.11 (2018). ISSN: 1099-4300. DOI: 10.3390/e20110854.

[1.13] R. Feynman and F. Wilczek. *The Character of Physical Law, with new foreword*. The MIT Press. MIT Press, 2017. ISBN: 9780262341738.

[1.14] Leonard Susskind. *Classical mechanics : the theoretical minimum*. London: Penguin Books, 2014. ISBN: 978-0141976228.

[1.15] Walter Grandy. *Entropy and the time evolution of macroscopic systems*. Oxford New York: Oxford University Press, 2008. ISBN: 9780199546176.

[1.16] G. M. Wang et al. "Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales". In: *Phys. Rev. Lett.* 89 (5 2002), p. 050601. DOI: 10.1103/PhysRevLett.89.050601.

[1.17] Louis Weisner. "Group Theory and its Application to the Quantum-Mechanics of Atomic Spectra, by Eugene P. Wigner, translated from the German by J. J. Griffin, Academic Press, New York and London, 1959. 372 pages". In: *Canadian Mathematical Bulletin* 3.2 (1960), pp. 95–95. DOI: 10.1017/S0008439500025364.

[1.18] J.M. Domingos. "Time reversal in classical and quantum mechanics". In: *International Journal of Theoretical Physics volume* 18 (1979), pp. 213–230. DOI: doi.org/10.1007/BF00670397.

[1.19] Yakir Aharonov and David Z. Albert. "Is the usual notion of time evolution adequate for quantum-mechanical systems? I". In: *Phys. Rev. D* 29 (2 1984), pp. 223–227. DOI: 10.1103/PhysRevD.29.223.

[1.20] David Z. Albert, Yakir Aharonov, and Susan D'Amato. "Curious New Statistical Prediction of Quantum Mechanics". In: *Phys. Rev. Lett.* 54 (1 1985), pp. 5–7. DOI: 10.1103/PhysRevLett.54.5.

[1.21] Yakir Aharonov, David Z. Albert, and Susan S. D'Amato. "Multiple-time properties of quantum-mechanical systems". In: *Phys. Rev. D* 32 (8 1985), pp. 1975–1984. DOI: 10.1103/PhysRevD.32.1975.

[1.22] Y Aharonov and D Rohrlich. "Towards a Two Vector Formulation of Quantum Mechanics". In: *Quantum Coherence*. United States: World Scientific, 1990, pp. 221–231.

[1.23] Y. Aharonov and L. Vaidman. "A New Characteristic of a Quantum System between Two Measurements — A "Weak Value"". In: *Bell's Theorem, Quantum Theory and Conceptions of the Universe*. Ed. by Menas Kafatos. Dordrecht: Springer Netherlands, 1989, pp. 17–22. ISBN: 978-94-017-0849-4. DOI: 10.1007/978-94-017-0849-4_3.

[1.24] Yakir Aharonov and Lev Vaidman. "Properties of a quantum system during the time interval between two measurements". In: *Phys. Rev. A* 41 (1 1990), pp. 11–20. DOI: 10.1103/PhysRevA.41.11.

[1.25] Y. Aharanov et al. "Novel Properties of Preselected and Postselected Ensembles". In: *Annals of the New York Academy of Sciences* 480.1 New Technique (1986), pp. 417–421. DOI: 10.1111/j.1749-6632.1986.tb12443.x.

[1.26] Lev Vaidman and Yakir Aharanov. "The Predictability of the Results of Measurements of Noncommuting Variables". In: *Annals of the New York Academy of Sciences* 480.1 New Technique (1986), pp. 620–621. DOI: 10.1111/j.1749-6632.1986.tb12480.x.

[1.27] Lev Vaidman, Yakir Aharonov, and David Z. Albert. "How to ascertain the values of σ_x, σ_y, and σ_z of a spin-1/2 particle". In: *Phys. Rev. Lett.* 58 (14 1987), pp. 1385–1387. DOI: 10.1103/PhysRevLett.58.1385.

[1.28] B. Reznik and Y. Aharonov. "Time-symmetric formulation of quantum mechanics". In: *Phys. Rev. A* 52 (4 1995), pp. 2538–2550. DOI: 10.1103/PhysRevA.52.2538.

[1.29] Daniele C. Struppa and Jeffrey M. Tollaksen, eds. *Quantum Theory: A Two-Time Success Story*. Springer Milan, 2014. DOI: 10.1007/978-88-470-5217-8.

[1.30] Yakir Aharonov, Sandu Popescu, and Jeff Tollaksen. "A time-symmetric formulation of quantum mechanics". In: *Physics Today* 63.11 (2010), pp. 27–32. DOI: 10.1063/1.3518209.

[1.31] Daniel Greenberger, Klaus Hentschel, and Friedel Weinert, eds. *Compendium of Quantum Physics*. Springer Berlin Heidelberg, 2009. DOI: 10.1007/978-3-540-70626-7.

[1.32] Yakir Aharonov and Lev Vaidman. “The Two-State Vector Formalism: An Updated Review”. In: *Time in Quantum Mechanics*. Springer Berlin Heidelberg, pp. 399–447. DOI: 10.1007/978-3-540-73473-4_13.

[1.33] Yakir Aharonov, David Z. Albert, and Lev Vaidman. “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100”. In: *Phys. Rev. Lett.* 60 (14 1988), pp. 1351–1354. DOI: 10.1103/PhysRevLett.60.1351.

[1.34] 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”. In: *Phys. Rev. D* 40 (6 1989), pp. 2112–2117. DOI: 10.1103/PhysRevD.40.2112.

[1.35] Y Aharonov and L Vaidman. “Complete description of a quantum system at a given time”. In: *Journal of Physics A: Mathematical and General* 24.10 (1991), pp. 2315–2328. DOI: 10.1088/0305-4470/24/10/018.

[1.36] Richard Jozsa. “Complex weak values in quantum measurement”. In: *Phys. Rev. A* 76 (4 2007), p. 044103. DOI: 10.1103/PhysRevA.76.044103.

[1.37] Lars M. Johansen. “Weak Measurements with Arbitrary Probe States”. In: *Phys. Rev. Lett.* 93 (12 2004), p. 120402. DOI: 10.1103/PhysRevLett.93.120402.

Chapter 2

Optical Implementation of Weak Measurement in Polarization degree of freedom

Contents

2.1 Measurement in Polarization degree of freedom 80

2.2 Transverse spatial pointer shift in weak measurement of polarization components 82

2.3 Experimental Setup for weak measurement 102

2.3.1 Experimentally measuring the beam displacement 110

2.3.2 σ_z, Π_H and Π_V measurement with the beam displacer 112

2.3.3 Measuring the expectation value of σ_z observable 113

2.3.4 Weak Measurement of σ_z observable with pre-selected state is the eigenstate 117

2.4 Weak Measurement of the σ_x observable 118

2.4.1 Using waveplates to transform σ_z interaction to σ_x interaction 119

2.4.2 Phase shift in the birefringent crystal 122

2.4.3 Experimental result : Weak value of σ_x observable 127

2.4.4 Effect of tilt on beam displacement 131

2.5 Modifications to σ_x setup for single photon experiments 133

2.5.1 Effects of longitudinal delay and coherence length 133

2.5.2 Compensating the loss of longitudinal coherence 136

The advantages of performing experiments in polarization degree of freedom in optical domain over experiments with spin systems is that photons do not decohere over time due to minimal interaction with environment[1] during their propagation. This enables us observe interference effects relatively easily. However, unlike spin 1/2 particles, we may not presume that an optical source is a stream of single particles. Some of the experiments in this thesis require us to establish the particle nature of light (Ch 4) and there we perform g_2 characterization of the source. In this chapter, we shall discuss the theoretical and experimental aspects of implementation of weak measurement in optical domain with a coherent laser source. The use of birefringent crystals to perform weak measurements was proposed [2.1, 2.2] and experimentally implemented [2.3] soon after the theory of weak measurement was developed [2]. Since then, weak measurements have been implemented in optics [2.5] for various precision measurements in foundational experiments [2.6, 2.7, 2.8]. Two photon interferometry also provides another method for implementation of weak measurement in polarization degree of freedom. This is discussed in the appendix A as the experiments performed in this do not use the technique but it helps us compare and contrast other experiments (See Appendix C) on the Quantum Cheshire Cat which implemented the technique. Single photon source generated from methods like spontaneous parametric down conversion (SPDC) do have smaller coherence length ($\approx$ 0.1 mm) due to large bandwidth (*nm*). The setup for weak measurement need to account for the finite coherence length and any path delay caused by the birefringent crystal need to be compensated. This is discussed later part in this chapter.

2.1 Measurement in Polarization degree of freedom

We, typically, analyze the polarization of light using a polarizer [3] which allows a certain polarization to pass through. This gives us some information about the polarization state

[1] most of the media is air or glass which are almost transparent and dispersion is less.

[2] The scheme in **??** uses birefringent prism to cause angular deviation(momentum shift) which is very similar to how spin -1/2 particles evolve in a Stern–Gerlach setup which was originally used to conceptualize weak measurements in **??**. The scheme described in **??** uses transverse displacement. For some experimental insights on this, see [2.4]

[3] A linear polarizer typically absorbs the light orthogonal to the polarization component which gets transmitted. Glan Taylor or Glan Thompson polarizers can reject the orthogonal polarization in a separate port. But here, we are presuming that the polarizer is a non-unitary operator.

of light. Given an ensemble of identically prepared photons in the state $|\psi\rangle$, we can obtain the expectation value of an observable in the polarization degree of freedom. With an ideal linear polarizer oriented at an angle θ with respect to the horizontal, the state after transmission reduces to a linear polarization oriented at the same angle θ and the state can be labeled $|\theta\rangle$. The transmission probability[4] is given by $|\langle\theta|\psi\rangle|^2$. The orthogonal polarization component $|\theta^T\rangle$ is absorbed, the probability of which is given by $|\langle\theta^T|\psi\rangle|^2$. The expectation value of the unitary observable[5] for the polarization component along θ is then given by $|\langle\theta|\psi\rangle|^2 - |\langle\theta^T|\psi\rangle|^2$. However, practically, the material of the polarizer also absorbs light and thus there is no easy way to determine $|\langle\theta^T|\psi\rangle|^2$ unless we measure the transmission with the polarizer rotated to the angle θ^T. Even if material absorption could be low enough to be ignored, often in single photon experiments, we may not know the exact incident flux. Of course, we can remove the polarizer to measure the incident number of photons and then place it to measure the transmitted number and thereby deducing the number of absorbed photons, using which we can compute the expectation value. But, we may not be able to use such a method to find the information about the intermediate state of polarization of light given pre-selected and post-selected polarization states.

To have the pre- and post-selected ensemble undisturbed, the polarizer should not absorb any polarization component. This leaves the allowed orientation of the polarizer to be along [6] the pre-selected polarization or the post-selected polarization state. Thus, we cannot use a polarizer to ask about the intermediate polarization property. Even if we orient the polarizer along θ and θ^T in two different runs of the experiment, we may not combine the result to comment on the intermediate polarization because the pre- and post-selected ensembles would get disturbed. For example, consider the pre-selected state to be horizontally polarized $|H\rangle$ and the post-selected state to have diagonal polarization to be $|D\rangle$. If we are interested in knowing what is the value of the σ_z observable, we essentially need to determine the polarization component along $|H\rangle$ and $|V\rangle$. If we place the intermediate ideal polarizer along horizontal ($\theta = 0$), we will ensure that the pre- and post-selected state is intact. The same would also happen if we make the polarizer oriented

[4] This is in accordance with the Malus' law

[5] with eigenvalues 1 for eigenvector $|\theta\rangle$ and -1 for the eigenvector $|\theta^T\rangle$

[6] Here, we take the assumption that pre- and post-selected state being linear polarization

along the diagonal $\theta = \pi/4$. Thus, if we need to treat both the results on par, then the intermediate polarization should be along $\theta = \pi/8$ but with a higher magnitude so that its components can be same along horizontal and diagonal. Indeed, the overall transmission increases, when we orient the intermediate polarizer along $\pi/8$. This may tempt us to conclude that the intermediate polarization is along $\pi/8$. But, if it is so, then the orthogonal polarization should not give any transmission. However, if we orient the polarizer along $\theta = \pi/2 + \pi/8 = 5\pi/8$, we see that still there is some transmission. If we treat the transmission as a signature of intermediate polarization, then clearly the polarization do not seem to be orthogonal $\theta = \pi/8$. This is because we are no longer dealing with the same ensemble, as the number of incident photons which gets successfully pre- and post-selected have now changed. Therefore, it is imperative to ensure that the evolution operator associated with the intermediate measurement is unitary. Hence, we would require that the polarization orthogonal to what we intend to measure do not get absorbed. Since, transmission is not an indicator, we now need the polarization to couple to another pointer state from which we can extract the information. Polarizing beam displacers, which are birefringent crystals that couples the polarization component with transverse spatial momentum, becomes candidates for implementation of the unitary evolution.

In this chapter, we shall study how the birefringent crystal can be used to measure pointer shifts corresponding to weak values of Pauli observables in polarization degree of freedom. Such coupling would be used in experiments described in Ch. 3 and Ch 4. In the method we use another degree of freedom of the same photon itself is coupled to the observable. Another method to determine the weak value does exist which uses two-photon interference to couple average polarization of one photon to observable of another photon. This is discussed in the Appendix A.

2.2 Transverse spatial pointer shift in weak measurement of polarization components

The birefringent crystal couples the transverse spatial degree of freedom of the beam to the polarization degree of freedom. In this section, we shall use this coupling to perform weak interaction in a pre- and post selected system and then calculate the shift in the center of

the transverse profile which we shall consider as the pointer for measuring the weak value. We shall assume that the beam is travelling along the $\hat{z}$ direction. We shall assume that the beam displacer is oriented in such a way that the extra-ordinary ray gets a shift along the direction $\hat{x}$. The initial transverse spatial intensity distribution along $\hat{x}$ shall be considered to be a gausssian centred at $x = 0$. The centre of this gaussian shall serve as the mean of the pointer variable. The uncertainty in the pointer state is given by the beam width which is given by $w_0 \approx 4\sigma_0$, where σ_0 is the standard deviation of the gaussian distribution. We shall also assume that the divergence of the beam is zero, i.e., its σ_0 is constant as a function of beam propagation along $\hat{z}$. This approximation is reasonable when σ_0 is large enough to begin with (orders of magnitude larger than the wavelength) and the beam displacement caused by the polarizing beam displacer is comparatively small.

We allow the initial polarization, the system, to be any general elliptic polarization state $|\psi\rangle_i$ and the initial transverse beam profile to be given by the gaussian amplitude $g_i(x)$.

$$|\psi_i\rangle\,|g_i(x)\rangle = (\cos(\theta_i/2)\,|H\rangle + e^{i\phi_i}\sin(\theta_i/2)\,|V\rangle) \otimes \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right) \exp\left(-\frac{x^2}{4\sigma_0^2}\right) \tag{2.1}$$

Here, $0 \leq \theta_i \leq \pi$ and $0 < \phi_i <\leq 2\pi$ are the angles corresponding to the spherical polar coordinates for the representation of the polarization in the Poincare sphere. We have considered a pure state for simplicity as well as because that is what the experiments would be concerned with. Note that the transverse spatial wavefunction is the square root of the gaussian intensity profile as it represents the probability amplitude.

The observable of interest would be a polarization component along the direction denoted by the angles $0 \leq \theta_m \leq \pi$ and $0 < \phi_m \leq 2\pi$ as represented in the Poincare sphere. Upon a strong measurement, we expect the polarization state to reduce to one of the following two state vectors which are orthogonal t each other.

$$|\psi(\theta_m/2,\phi_m)\rangle = \begin{pmatrix} \cos(\theta_m/2) \\ e^{i\phi_m}\sin(\theta_m/2) \end{pmatrix} \tag{2.2}$$

$$|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle = \begin{pmatrix} \sin(\theta_m/2) \\ -e^{i\phi_m}\cos(\theta_m/2) \end{pmatrix} \tag{2.3}$$

The operator, say A, which would correspond to the observable for the above measurement can be described as follows:

$$\begin{aligned} A(\theta_m,\phi_m) = & |\psi(\theta_m/2,\phi_m)\rangle\langle\psi(\theta_m/2,\phi_m)| \\ & -|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle\langle\psi((\pi-\theta_m)/2,\pi+\phi_m)| \end{aligned} \tag{2.4}$$

$$= \begin{pmatrix} \cos(\theta_m/2) & e^{-i\phi_m}\sin(\theta_m/2) \\ e^{i\phi_m}\sin(\theta_m/2) & -\cos(\theta_m/2) \end{pmatrix} \tag{2.5}$$

The operator A is hermitian (and unitary as well) and has eigenvalues $\{1,-1\}$, corresponding to the eigenstates $|\psi(\theta_m/2,\phi_m)\rangle$ and $|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle$ respectively.

Ideally, we wish to couple the observable A with the momentum[7] conjugate to our pointer variable x as is usually done in von Neumann's description of measurement procedure [2.9]. Next we write the interaction Hamiltonian that couples transverse spatial profile of the beam with the polarization observable by drawing analogy to how one would treat a spin-1/2 particle in a Stern-Gerlach apparatus [2.10, 2.1, 2.11].

$$H_I = g(t)\gamma A(\theta_m,\phi_m)\otimes P_x \qquad \text{with} \int_{t-t_\varepsilon}^{t+t_\varepsilon} g(t) = 1 \tag{2.6}$$

Here $g(t)$ is a compact function which is non-zero during the interaction and whose time integral turns out to be 1. Here, t and $2t_\varepsilon$ represents the mean time and duration of interaction respectively. The variable γ represents the strength of interaction. In practice, when a beam is incident on a birefringent, the interaction time depends on the length of the crystal and the interaction strength depends on the angle of refraction. But the result after

[7]The description of polarization measurement in optics is analogous or isomorphic to the measurement of a spin half particle in a Stern-Gerlach setup. In the stern gerlach set up, the transverse position gets coupled and the spin gets a momentum kick. In optics, the refraction occurs within the media, i.e., the momentum is coupled, but upon emergence the beam travels along the same direction but with displacement.

the interaction is over is a relative beam displacement which we shall be related to γ (See the section 2.3 for details). Here, we have the $P_x = -i\hbar\partial_x$ is the momentum operator (the position translation operator) for the transverse spatial degree of freedom along $\hat{x}$.

After the interaction, the joint system pointer state is evolved by the unitary evolution operator $U = \exp\left(\int_{t-t_\varepsilon}^{t+t_\varepsilon} -\frac{i}{\hbar} H_I dt\right)$.

$$|\Psi_e\rangle_{(t+t_\varepsilon)} = \exp\left(\int_{t-t_\varepsilon}^{t+t_\varepsilon} -\frac{i}{\hbar} H_I dt\right) |\psi_i\rangle\,|g_i(x)\rangle \tag{2.7}$$

$$= \exp\left(\int_{t-t_\varepsilon}^{t+t_\varepsilon} -\frac{i}{\hbar}\frac{\hbar}{i}\gamma A(\theta_m,\phi_m)\otimes\frac{\partial}{\partial x}dt\right) |\psi_i\rangle\,|g_i(x)\rangle \tag{2.8}$$

$$= \exp\left(-\gamma A(\theta_m,\phi_m)\partial_x\right) |\psi_i\rangle\,|g_i(x)\rangle \tag{2.9}$$

Next, we can write $|\psi_i\rangle$ as superposition of the eigenstates of $A(\theta_m,\phi_m)$ so that we can act the evolution operator [8].

$$|\psi_i\rangle = c_1\,|\psi(\theta_m/2,\phi_m)\rangle + c_2\,|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle \qquad (\text{, where, }) \tag{2.10}$$

$$c_1 = \cos(\theta_i/2)\cos(\theta_m/2) + e^{i(\phi_i-\phi_m)}\sin(\theta_i/2)\sin(\theta_m/2) \tag{2.11}$$

$$c_2 = \cos(\theta_i/2)\sin(\theta_m/2) - e^{i(\phi_i-\phi_m)}\sin(\theta_i/2)\cos(\theta_m/2) \tag{2.12}$$

We then proceed towards evaluation of the expression 2.9 by substituting $|\psi\rangle$ from the expression 2.10.

$$|\Psi_e\rangle = \exp\left(-\gamma A(\theta_m,\phi_m)\partial_x\right) \left(c_1\,|\psi(\theta_m/2,\phi_m)\rangle + c_2\,|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle\right)|g_i(x)\rangle \tag{2.13}$$

$$= \begin{aligned} &c_1\,|\psi(\theta_m/2,\phi_m)\rangle\, e^{(-1)\gamma\partial_x}\,|g_i(x)\rangle \\ &+ c_2\,|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle\, e^{(-(-1)\gamma)\partial_x}\,|g_i(x)\rangle \end{aligned} \tag{2.14}$$

In the last step, we replaced the operator A with the eigenvalue. Hence, after the interaction, the joint system-pointer state evolved into an entangled state of the system being in either of the eigenstate which are coupled to appropriate translated pointer state. To evaluate the action of $e^{\gamma\partial_x}$ on $|g_i(x)\rangle$, we expand the exponential of the operator as the series.

Expanding the exponential we have

[8] We shall use the trick that $e^A\,|\psi\rangle = e^a\,|\psi\rangle$, given $A\,|\psi\rangle = a\,|\psi\rangle$

$$e^{\gamma\partial_x} = \mathbb{1} + \gamma\partial_x + \frac{\gamma^2\partial_x^2}{2!} + \frac{\gamma^3\partial_x^3}{3!}\ldots \tag{2.15}$$

Let's evaluate the action of first two terms in the above expression on the gaussian probability amplitude.

$$\mathbb{1}\left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right) = \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right) \tag{2.16}$$

$$\gamma\,\partial_x\left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right) = \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right)\frac{-2}{4\sigma_0^2}\,\gamma\,x \tag{2.17}$$

The sum of the above two terms is a truncation of $e^{\gamma\partial_x}$ acting on the initial gaussian pointer state up to the first order in γ.

$$e^{\gamma\partial_x}|g_i(x)\rangle = \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right)\left(1-\frac{2}{4\sigma_0^2}\gamma x\right) + \mathcal{O}(\gamma^2) \tag{2.18}$$

The above term does not look like a gaussian unless we make the further assumption that $\gamma <<< \sigma_0$, corresponding to the weak interaction. Under, this approximation, we have

$$\exp\left(-\frac{(x+\gamma)^2}{4\sigma_0^2}\right) = \exp\left(-\frac{x^2+2\gamma x+\gamma^2}{4\sigma_0^2}\right) \tag{2.19}$$

$$= \exp\left(-\frac{x^2}{4\sigma_0^2}\right)\exp\left(-\frac{2\gamma x}{4\sigma_0^2}\right)\exp\left(-\frac{\gamma^2}{4\sigma_0^2}\right) \tag{2.20}$$

We can ignore the third term as its series expansion would have the leading term, after 1, as γ^2. The second term can be expanded up to the γ. Then we have the following approximate result.

$$\exp\left(-\frac{(x+\gamma)^2}{4\sigma_0^2}\right) \approx \exp\left(-\frac{x^2}{4\sigma_0^2}\right)\left(1-\frac{2}{4\sigma_0^2}\gamma x\right) \tag{2.21}$$

$$\Rightarrow e^{\gamma\partial_x}|g_i(x)\rangle \approx \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\exp\left(-\frac{x^2}{4\sigma_0^2}\right)\left(1-\frac{2}{4\sigma_0^2}\gamma x\right) \tag{2.22}$$

$$= \exp\left(-\frac{(x+\gamma)^2}{4\sigma_0^2}\right) = |g(x+\gamma)\rangle \tag{2.23}$$

Thus, in the weak interaction limit i.e., $\gamma << \sigma_0$, we can truncate the infinite exponential series to the first term with reasonable approximation which shows that the gaussian shifts by an amount γ. The center of the gaussian after the interaction becomes $-\gamma$. Similarly, we have $e^{-\gamma\partial_x}|g_i(x)\rangle \approx |g(x-\gamma)\rangle$, i.e., the gaussian centred about γ.

Next, we can compute the finite sum $\sum_{k=0}^{n}\frac{\gamma^k}{k!}\partial_x^k|g_i(x)\rangle$ and compare it with $g(x+\gamma)$.

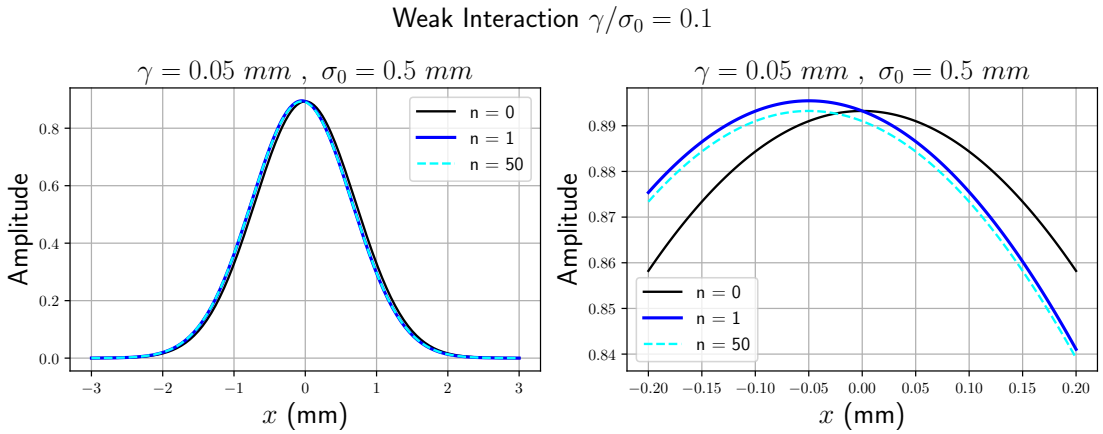


Figure 2.1: After weak interaction, the pointer state shifts by γ. If we sum the first two terms up to first power of γ i.e., $n = 1$, we get a solution which is very close to the case when we sum all the terms up to $n = 50$. The left plot shows that the shift is tiny compared to the width of the gaussian. The plot on right gives a zoomed in version of the same. The shift are with respected to $g_i(x)$ shown with the label $n = 0$

.

From, the above it may seem as of now that, the interaction resulting in shifting of the gaussian is valid only as an approximation of the weak interaction limit. However, when we add enough terms, we observe that even if $\gamma > \sigma_0$, the pointer state shifts by γ.

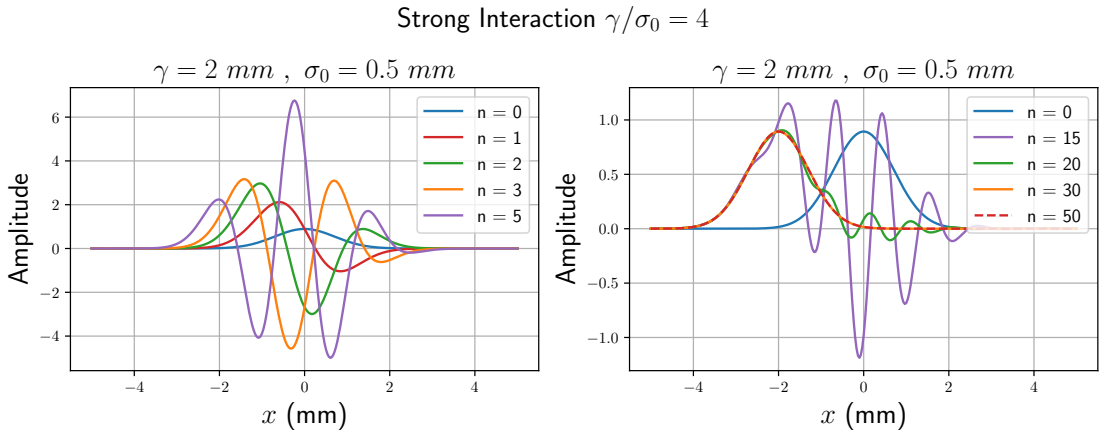


Figure 2.2: In the strong interaction limit, i.e., when $\gamma >> \sigma_0$, the first order approximation is no where close to being a gaussian. All higher order terms becomes are oscillatory. Once we add enough terms (n=20) we start to observe that the pointer state tends to become a gaussian with an oscillatory tail. Upon adding more terms the solution tends to be $|g(x+\gamma)\rangle$

.

Therefore, irrespective of whether the interaction strength is strong or weak, it seems that $e^{\pm\gamma\partial_x}|g_i(x)\rangle = |g(x\pm\gamma)\rangle$. To prove the above result, the evolution operator $e^{\pm\gamma\partial_x} = \exp\left(\mp\frac{i}{\hbar}P_x\right)$ is expressed as an exponential of the momentum operator P_x. Next, if we express $|g_i(x)\rangle$ in momentum representation, we can proceed by replacing the operator with the eigenmomentum. Then, we can revert to the position representation.

In this exercise, the following expressions relating position and momentum representation shall be useful.

$$\hat{X} = i\hbar\partial_p \qquad \hat{P} = -i\hbar\partial_x \tag{2.24}$$

$$|x\rangle = \delta(x-x') \qquad |p\rangle = \delta(p-p') \tag{2.25}$$

The first tells us the representation of position operator in momentum representation and vice versa. The latter tells us that the eigenstate of position and momentum operator in their respective representation are Dirac delta distributions. Next, we need to evaluate the inner product between the position and momentum eigenstates so that we can use the same while transforming the initial pointer state in momentum representation.

$$\hat{P}\,|p\rangle = p\,|p\rangle \qquad \text{(momentum eigenstate)} \tag{2.26}$$

$$\Rightarrow -i\hbar\partial_x\,|p\rangle = p\,|p\rangle \qquad \text{(operator in position representation)} \tag{2.27}$$

$$\Rightarrow |p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{i}{\hbar}px\right)|x\rangle \qquad \text{(Solving the differential equation)} \tag{2.28}$$

$$\Rightarrow \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{i}{\hbar}px\right) \tag{2.29}$$

Here, the term $\frac{1}{\sqrt{2\pi\hbar}}$ does not actually normalize the state but rather is used as a useful constant which preserves the state while going from position to momentum representation and then returning to the position representation.

We express the initial pointer state in the momentum representation, and obtain the wavefunction as a gaussian in p with standard deviation given by $\frac{\hbar}{2\sigma_0}$.

$$g_i(x)\,|x\rangle \to \int |p\rangle\,\langle p|x\rangle\,g_i(x)dx \tag{2.30}$$

$$= |p\rangle\,\frac{1}{\sqrt{2\pi\hbar}}\int_{\infty}^{\infty} g_i(x)\exp\left(\frac{i}{\hbar}xp\right)dx \tag{2.31}$$

$$= |p\rangle\,\frac{1}{\sqrt{2\pi\hbar}}\left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{4\sigma_0^2}\right)\exp\left(\frac{i}{\hbar}xp\right)dx \tag{2.32}$$

$$= |p\rangle\,\frac{1}{\sqrt{2\pi\hbar}}\left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\sqrt{4\pi\sigma_0^2}\exp\left(-\frac{p^2}{(\hbar/\sigma_0)^2}\right) \tag{2.33}$$

Next, we evaluate $\exp\left(\mp\frac{i}{\hbar}\gamma P_x\right)|p\rangle = \exp\left(\mp\frac{i}{\hbar}\gamma p\right)|p\rangle$ since $\hat{P}_x\,|p\rangle = p\,|p\rangle$. Thus, after the interaction, the evolved pointer state $\xi(p)$ can be written in the momentum representation.

$$|\xi(p)\rangle = \exp\left(\mp\frac{i}{\hbar}\gamma p\right)|p\rangle\,\frac{1}{\sqrt{2\pi\hbar}}\left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right)\sqrt{4\pi\sigma_0^2}\exp\left(-\frac{p^2}{(\hbar/\sigma_0)^2}\right) \tag{2.34}$$

Next, we convert the description of the state in momentum representation to position representation $\xi(p) \to \xi(x)$.

$$|\xi(p)\rangle \rightarrow \int |x\rangle \langle x|p\rangle \xi(p) dp \tag{2.35}$$

$$= |x\rangle \frac{1}{\sqrt{2\pi\hbar}} \int_{\infty}^{\infty} \xi(p) \exp\left(\frac{i}{\hbar} xp\right) dp \tag{2.36}$$

$$= |x\rangle \frac{1}{2\pi\hbar} \left(\frac{\sqrt{4\pi\sigma_0^2}}{\sqrt[4]{2\pi\sigma_0^2}}\right) \int_{\infty}^{\infty} \exp\left(\frac{i}{\hbar}(x \mp \gamma)p\right) \exp\left(-\frac{p^2}{(\hbar/\sigma_0)^2}\right) dp \tag{2.37}$$

$$= |x\rangle \frac{1}{2\pi\hbar} \left(\frac{\sqrt{4\pi\sigma_0^2}}{\sqrt[4]{2\pi\sigma_0^2}}\right) \exp\left(-\frac{(x \mp \gamma)^2}{4\sigma_0^2}\right) \frac{\sqrt{\pi}\hbar}{\sigma_0} \tag{2.38}$$

$$= \left(\frac{1}{\sqrt[4]{2\pi\sigma_0^2}}\right) \exp\left(-\frac{(x \mp \gamma)^2}{4\sigma_0^2}\right) |x\rangle \tag{2.39}$$

$$= |g(x \mp \gamma)\rangle \tag{2.40}$$

Therefore, we conclude that irrespective of the interaction strength, the center of gaussian pointer gets shifted by γ, i.e., $e^{\pm\gamma\partial_x} |g_i(x)\rangle = |g(x \pm \gamma)\rangle$. The results does hold without any approximation and is also applicable to intermediate interaction strength as shown below.

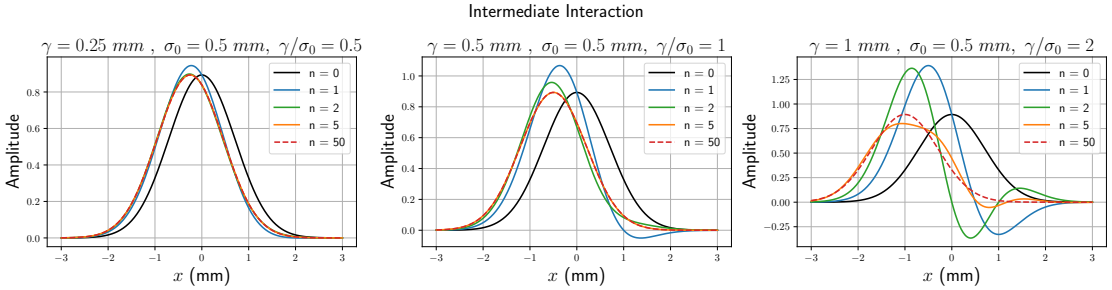


Figure 2.3: For intermediate interaction strength, as we consider more terms in the sum, the result converges to a gaussian whose center is shifted by γ.

Now, we can simplify the expression of $|\Psi_e\rangle$ described in Eqn. 2.14 and write it in terms of the shifted gaussians.

$$|\Psi_e\rangle = \begin{aligned} & c_1 |\psi(\theta_m/2, \phi_m)\rangle |g(x-\gamma)\rangle \\ & + c_2 |\psi((\pi-\theta_m)/2, \pi+\phi_m)\rangle |g(x+\gamma)\rangle \end{aligned} \tag{2.41}$$

Therefore, after the interaction, we have an entangled state. If we make a detection at this point, we have find the particle being a part of the gaussian $|g(x-\gamma)\rangle$, which is a gaussian with centre at γ, we would conclude that the polarization state is $|\psi(\theta_m/2,\phi_m)\rangle$. This would occur with probability $|c_1|^2$. Similary, if we detect the particle to be a part of $|g(x+\gamma)\rangle$, the gaussian with centre at $-\gamma$, we would conclude the polarization state to be $|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle$. Of course, to be able to distinguish whether one detected particle belongs to $|g(x-\gamma)\rangle$ or $|g(x+\gamma)\rangle$, we need $\gamma >> \sigma_0$. In such a case, we have $\langle g(x-\gamma)|g(x+\gamma)\rangle \approx 0$. If we take the weak interaction limit, since $\gamma << \sigma_0$, we would have $\langle g(x-\gamma)|g(x+\gamma)\rangle \approx 1$. Thus, given a detection, we would not be able to comment on whether the polarization of the particle is $|\psi(\theta_m/2,\phi_m)\rangle$ or $|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle$.

However, instead of making the detection, we would post-select the system and select a particular polarization component. Let the polarization state being post-selected is given by $\langle\phi| = \left(\cos\left(\theta_f/2\right), e^{-i\phi_f}\sin\left(\theta_f/2\right)\right)$. Then, upon, successful post-selection, the final pointer state [9], would become, $\left|\chi_f\right\rangle$ as given in the following.

$$\left|\chi_f\right\rangle = \begin{array}{l} c_1\langle\phi|\psi(\theta_m/2,\phi_m)\rangle\,|g(x-\gamma)\rangle \\ \quad + c_2\langle\phi|\psi((\pi-\theta_m)/2,\pi+\phi_m)\rangle\,|g(x+\gamma)\rangle \end{array} \tag{2.42}$$

$$\Rightarrow \left|\chi_f\right\rangle = C_1\,|g(x-\gamma)\rangle + C_2\,|g(x+\gamma)\rangle \tag{2.43}$$

The pointer state has two components, the gaussians shifted by γ multiplied by the eigenvalues ± 1. Here, we have termed the complex coefficients of the gaussians as C_1 and C_2 respectively,

$$C_1 = c_1\langle\phi|\psi(\theta_m/2,\phi_m)\rangle \tag{2.44}$$

$$= \begin{array}{l} (\cos(\theta_i/2)\cos(\theta_m/2) + e^{i(\phi_i-\phi_m)}\sin(\theta_i/2)\sin(\theta_m/2)) \\ \quad (\cos(\theta_m/2)\cos\left(\theta_f/2\right) + e^{i(\phi_m-\phi_f)}\sin(\theta_m/2)\sin\left(\theta_f/2\right)) \end{array} \tag{2.45}$$

[9] Note that, the system's state does not necessarily reduce to $|\phi\rangle$ for the post-selected ensemble. Only the component of the evolved state gets projected on to the $|\phi\rangle$. The difference is subtle and shall be clear in an example discussing weak measurement of pauli observables.

$$C_2 = c_2 \langle\phi|\psi((\pi-\theta_m)/2, \pi+\phi_m)\rangle \tag{2.46}$$

$$= \frac{(\cos(\theta_i/2)\sin(\theta_m/2) - e^{i(\phi_i-\phi_m)}\sin(\theta_i/2)\cos(\theta_m/2))}{(\cos(\theta_m/2)\sin(\theta_f/2) - e^{i(\phi_m-\phi_f)}\sin(\theta_m/2)\cos(\theta_f/2))} \tag{2.47}$$

The center of the pointer state distribution is the average of the intensity distribution weighed with the variable x and can be obtained from the expectation value of x in the pointer state.

$$\frac{\langle\chi_f|x|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle} = \frac{(C_1^*\langle g(x-\gamma)| + C_2^*\langle g(x-\gamma)|)\, x\, (C_1|g(x-\gamma)\rangle + C_2|g(x+\gamma)\rangle)}{(C_1^*\langle g(x-\gamma)| + C_2^*\langle g(x-\gamma)|)(C_1|g(x-\gamma)\rangle + C_2|g(x+\gamma)\rangle)} \tag{2.48}$$

$$= \frac{\begin{array}{c}|C_1|^2\langle g(x-\gamma)|x|g(x-\gamma)\rangle + C_1^*C_2\langle g(x-\gamma)|x|g(x+\gamma)\rangle \\ + C_2^*C_1\langle g(x+\gamma)|x|g(x-\gamma)\rangle + |C_2|^2\langle g(x+\gamma)|x|g(x+\gamma)\rangle\end{array}}{|C_1|^2 + C_1^*C_2\langle g(x-\gamma)|g(x+\gamma)\rangle + C_2^*C_1\langle g(x+\gamma)|g(x-\gamma)\rangle + |C_2|^2} \tag{2.49}$$

We needed to perform the normalization while computing the expectation value because $|\chi_f\rangle$ after post-selection need not remain normalized as we are ignoring the component orthogonal to the post-selected state.

Let's understand the case where the intermediate state of the system is one of the eigenstates of $|A\rangle$. In that case one of the coefficient C_1 or C_2 is 0. If $|A_1\rangle$ and $|A_2\rangle$ are the eigenstates, we can write the coefficients as the following. $C_1 = \langle\phi|A_1\rangle\langle A_1|\psi\rangle$ and $C_2 = \langle\phi|A_2\rangle\langle A_2|\psi\rangle$, where $|A_1\rangle = |\psi(\theta_m/2, \phi_m)\rangle$ and $|A_2\rangle = |\psi((\pi-\theta_m)/2, \pi+\phi_m)\rangle$. If the intermediate state is $|A_2\rangle$ then C_1=0. Then, the pointer shift is obtained as $-\gamma$.

$$\left.\frac{\langle\chi_f|x|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle}\right]_{C_1=0} = \frac{|C_2|^2\langle g(x+\gamma)|x|g(x+\gamma)\rangle}{|C_2|^2} = -\gamma \tag{2.50}$$

Similarly, when the intermediate state is $|A_1\rangle$, then C_2 is 0 and the pointer shifts to γ. The intermediate state can be $|A_1\rangle$(or $|A_2\rangle$) if either the pre-selected or post-selected state or

both[10] are $|A_1\rangle$ (or $|A_2\rangle$). Hence, for a dichotomic observable, if the pre- or post-selected state is one of the eigenstates, the pointer shifts indicates[11] the eigenvalue.

But in general, what does the pointer shift indicate ? First, we note that the transition element $\langle g(x-\gamma)|\,x\,|g(x+\gamma)\rangle = 0$.

$$\langle g(x-\gamma)|\,x\,|g(x+\gamma)\rangle = \frac{1}{\sqrt{2\pi\sigma_0^2}}\int_{\infty}^{\infty}\exp\left(-\frac{(x-\gamma)^2}{4\sigma_0^2}\right)x\exp\left(-\frac{(x+\gamma)^2}{4\sigma_0^2}\right)dx \tag{2.51}$$

$$= \frac{1}{\sqrt{2\pi\sigma_0^2}}\int_{\infty}^{\infty}x\exp\left(-\frac{(x+\gamma)^2}{4\sigma_0^2}-\frac{(x-\gamma)^2}{4\sigma_0^2}\right)dx \tag{2.52}$$

$$= \frac{1}{\sqrt{2\pi\sigma_0^2}}\int_{\infty}^{\infty}x\exp\left(-\frac{(x^2+\gamma^2)}{4\sigma_0^2}\right)dx \tag{2.53}$$

$$= 0 \tag{2.54}$$

The integral evaluates to 0 because the integrand is odd function of x. Further, in the strong interaction limit, when $\gamma >> \sigma_0$, we have $\langle g(x+\gamma)|g(x-\gamma)\rangle \approx 0$ Thus, the pointer shift in the strong interaction limit becomes the following.

$$\left.\frac{\langle\chi_f|x|\chi_f\rangle}{\langle\chi_f|\chi_f\rangle}\right]_{\gamma>>\sigma_0} = \frac{|C_1|^2\langle g(x-\gamma)|\,x\,|g(x-\gamma)\rangle + |C_2|^2\langle g(x+\gamma)|\,x\,|g(x+\gamma)\rangle}{|C_1|^2+|C_2|^2} \tag{2.55}$$

$$= \gamma\frac{|C_1|^2-|C_2|^2}{|C_1|^2+|C_2|^2} \tag{2.56}$$

The above can be interpreted as the pointer being shifted by the amount proportional to the conditional[12] expectation value. Note that the quantity is real and is bounded between the two extreme eigenvalues.

When we are in the weak interaction limit i.e., $\gamma << \sigma_0$, we can ignore the slight lack of overlap between the two shifted gaussians and thus approximate,$\langle g(x+\gamma)|g(x-\gamma)\rangle \approx 1$.

[10] If the pre-selected state and post-selected state are $|A_1\rangle$ and $|A_2\rangle$ respectively, then the pointer shift is ill defined as overlap of pre and post-selection tends to vanish. However, they can be posed in a limiting sense which shall be clear when we consider an example.

[11] The pointer shift is eigen value times the interaction strength γ.

[12] conditioned to post-selection. Without post-selection, the expectation value would be have been given by $\frac{|c_1|^2-|c_2|^2}{|c_1|^2+|c_2|^2}$

Hence, we have the following expression for the pointer shift.

$$= \frac{\langle \chi_f | x | \chi_f \rangle}{\langle \chi_f | \chi_f \rangle} \Bigg]_{\gamma << \sigma_0} = \gamma \frac{|C_1|^2 - |C_2|^2}{|C_1|^2 + C_1^* C_2 + C_1 C_2^* + |C_2|^2} \tag{2.57}$$

$$= \gamma \frac{\frac{1}{2} \begin{pmatrix} \cos(\theta_m)\left(\cos(\theta_i) + \cos\left(\theta_f\right)\right) \\ + \sin(\theta_m)\left(\sin(\theta_i)\cos(\phi_i - \phi_m) + \sin\left(\theta_f\right)\cos\left(\phi_f - \phi_m\right)\right) \end{pmatrix}}{\frac{1}{2}\left(1 + \cos(\theta_i)\cos\left(\theta_f\right) + \sin(\theta_i)\sin\left(\theta_f\right)\cos\left(\phi_f - \phi_i\right)\right)} \tag{2.58}$$

The denominator in expression 2.57, although contains complex terms is eventually a real quantity as can be seen in the expression 2.58 where we have substituted the angles for the Bloch sphere representation which are all real numbers. This is because for a complex number z, $z + z* = 2\,\mathrm{Re}\{z\}$. If we consider $z = C_1^* C_2$, we have $z^* = C_1 C_2^*$. Thus, the middle two terms in the denominator of expression 2.57 turns out to be real. The pointer shift, therefore is a real number multiple of γ, as it is supposed to be. However, the denominator can now be arbitrarily close to 0 as $C_1^* C_2 + C_1 C_2^*$ can be negative. The denominator in expression 2.58 is bounded between $(0, 1)$. When the post-selected state is orthogonal to the pre-selected state i.e., $\theta_f \to \pi - \theta_i$, and $\phi_f \to \pi + \phi_i$, the denominator approaches 0. However, in the same limit, the numerator also tends to 0 and hence the pointer shift is indeterminate. This is understandable because when pre- and post-selection are orthogonal we would hardly have any particles in the ensemble that satisfies pre- and post-selection [13].

Since the numerator of the expression 2.58 is bounded between $(-1, 1)$ multiple of γ and the denominator is bounded between $(0, 1)$, the ratio can correspond to a shift of the pointer which is larger than any possible shift without post-selection. In other words, the multiple factor to γ can be larger than the eigenvalue bounds of ± 1. Such a shift is possible because the expression 2.43 describes the resultant pointer state after post-selection as a superposition of two shifted gaussians with complex coefficients [2.1]. The interference of the two gaussian components, in the limit $\gamma << \sigma_0$ can result in another gaussian with a much larger shift compared to γ. In strong interaction regime, due to lack of overlap between the gaussians interference is not possible and the result is an incoherent mixed state upon detection.

[13] Also, we shall see that the pointer state no longer remains gaussian in this limit for a given beam displacement γ

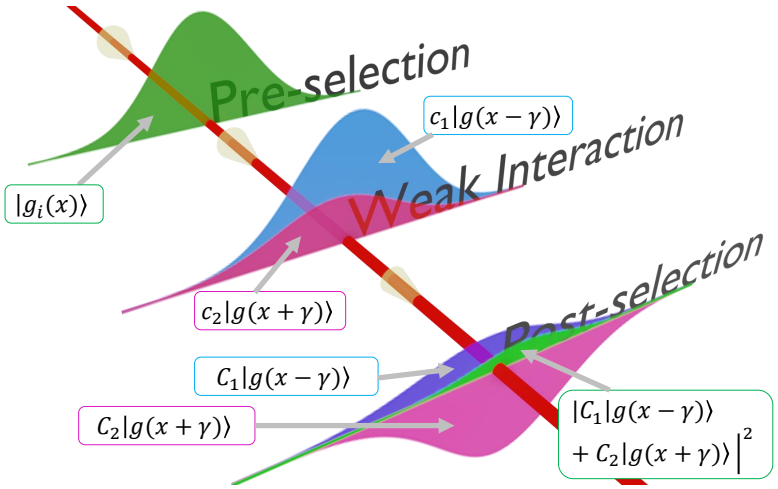


Figure 2.4: Upon weak interaction, the polarization splits into two components which are the eigenstates of the observable. The initial pointer state $g_i(x)$, splits into two shifted gaussians associated with the eigenstates. The gaussian $|g(x-y)\rangle$ associated with eigenstate $|A_1\rangle$ has the complex amplitude $c_1 = \langle\psi|A_1\rangle$ This gaussian is shifted by eigenvalue 1 multiplied with the interaction strength γ. Similarly, the polarization state $|A_2\rangle$ has the amplitude $c_2 = \langle\psi|A_2\rangle$ and is associated with the shifted gaussian $|g(x+\gamma)\rangle$.In weak interaction regime i.e., $\gamma << \sigma_0$, the two shifted gaussians overlap and hence the two polarization components are indistinguishable. This implies the polarization if measured after the weak interaction would behave as like the pre-selected state $|\psi\rangle$, which is the superposition of the eigenstates of A. If we make detection right after the weak interaction, we would get a gaussian shifted by the expectation value. This is because the gaussian and $|g(x-\gamma)\rangle$ and $|g(x+\gamma)\rangle$ would occur with probabilities $|c_1|^2$ and $|c_2|^2$ respectively and there would not be any interference of the two shifted gaussians because the associated polarization eigenstates are orthogonal i.e., $\langle A_1|A_2\rangle = 0$ However, if we post-select with the state ϕ, the gaussian $|g(x-\gamma)\rangle$ gets post-selected with the amplitude $C_1 = \langle\psi|A_1\rangle\langle A_1|\phi\rangle$. Similarly, the gaussian $|g(x+\gamma)\rangle$ gets selected with amplitude $C_2 = \langle\psi|A_2\rangle\langle A_2|\phi\rangle$. These two gaussians after post-selection now have the same polarization and can interfere. In general, C_1 and C_2 are complex and hence the two shifted gaussians superpose with a relative phase which results in another gaussian whose shift can be way higher than γ. Such pointer shifts beyond the shift that would have been caused by extreme eigen values is possible because of coherent combination of post-selected components of the otherwise orthogonal eigenstates of the observable A, for which the weak interaction limit i.e., $\gamma << \sigma_0$ is essential.

But what does the pointer shift correspond to? First of all, the pointer shift can only be reliably obtained if we reconstruct the gaussian with high statistics. The uncertainty in the mean of the reconstructed gaussian i.e, the error in determining the expectation value would be $\sigma_0/\sqrt{N}$ if we reconstruct the gaussian with N particle detections . To measure the pointer shift accurately up to one order lower than γ, we need $\frac{\sigma_0}{\sqrt{N}} << \gamma/10 \Rightarrow N >> 100(\sigma_0/\gamma)^2$.For typical values of $\gamma/\sigma_0 = 0.1$, one would therefore need about 10^4

detections to reconstruct the gaussian[14]. For such a large number of particles it seems that the pointer shift represents some average property of the ensemble conditioned to post-selection. But then if it were a mere statistical incoherent average with each particle having a definite intermediate eigenvalue corresponding to the observable A which is being measured, then one would expect the pointer shift to be bounded by the range of eigenvalue spectrum. Further, each particle in the ensemble is identically prepared and post-selected and given the weak interaction limit we do not know the intermediate state for each particle. Thus, we cannot assign the eigenvalue $+1$ or -1 with a subset of the ensemble. We may assign a value corresponding to the pointer shift for the entire ensemble of particles.

Alternatively, we can consider our system to be a joint state of N particles. We pre and post-select with the states $|\Psi\rangle = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle \ldots$ (N times) and $|\Phi\rangle = |\phi\rangle \otimes |\phi\rangle \otimes |\phi\rangle \ldots$ (N times). First, lets consider the joint observable $\mathbb{A} = A \otimes A \otimes A \ldots$ (N times) which corresponds to the case when polarization of all the photons together affect the pointer. The pointer shift of $\pm\gamma$ can then be associated with the eigenvalues ± 1 of the observable $\mathbb{A}$. The corresponding eigenstates can be $|\mathbb{A}_1\rangle = |A_1\rangle \otimes |A_1\rangle \otimes |A_1\rangle \ldots$ (N times) and $|\mathbb{A}_2\rangle = |A_2\rangle \otimes |A_2\rangle \otimes |A_2\rangle \ldots$ (N times) respectively. But the association of the eigenstates to the eigenvalue is not unique because of degeneracy. For example, if N is odd, then the eigenvalue of -1 for the observable can be associated with the eigenstate $|A_2\rangle \otimes |A_1\rangle \otimes |A_2\rangle \otimes |A_1\rangle \ldots |A_2\rangle$. Thus, the eigenvalue of the joint system of particles may not tell us much about the property of individual particles. Hence, any pointer shift, seemingly, can only be associated with the system of N particles as whole.

But the observable $\mathbb{A}$ does not apply to a scenario were each particle is sent one by one. Here, interaction and hence evolution of each particle needs to be separable. Thus, the operator of interest would be the sum or average of operators acting on individual particles. Hence,the von Neumann measurement procedure described where polarization of each particle couples its own transverse spatial degree of freedom would correspond to the observable $\bar{\mathbb{A}}$ defined below.

[14]Here, the presumption is that the resultant pointer state is approximately a gaussian. Whhen pre and post-selection tends to be orthogonal, we resultant pointer state is not a gaussian and thus we need to be careful in representing the uncertainty in determining the pointer shift.

$$\bar{\mathbb{A}} = \frac{1}{N}\sum_{k=0}^{N} \mathbb{1}^k A \otimes \mathbb{1}^{N-k} \qquad \text{where, } \mathbb{1}^k = \mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1}\ldots \text{k times} \tag{2.59}$$

For such an observable, the eigenvalues 1 and -1, which are at extreme end of the eigenspectrum, are non-degenerate and are associated with eigenstates $|\mathbb{A}_1\rangle$ and $|\mathbb{A}2\rangle$ respectively. Other eigenvalues can be degenerate. Hence, only in the case when pointer shift is ± 1, we can comment on the state of individual particles in the joint system.The result of weak measurement[15] of a system in general is not limited to the eigenspectrum of the observable $\mathbb{A}$ for the joint system of N particles. Thus, from the above considerations alone, we may not consider the pointer shift being reflective of property of individual particles of the system.

However, the requirement that we need a large number of system particles to measure the pointer shift accurately is a consequence of we using the intra-particle degree of freedom as the pointer state. In principle, if we have a system photon, we can couple the polarization of the photon to some degree of freedom of an atom. Under the weak interaction, the photon's polarization would almost remain unaffected and not reduce to an eigenstate with one such interaction. Thus, we can have N such atoms interacting with only one photon considered as the system. All detection shall be on atoms conditioned to the system photon satisfying pre and post-selection. Thus, in principle we can couple N atoms with one system particle and then determine the pointer shift from the detection of N atoms. In such a case[16], the pointer shift would be reflective of intermediate property of the single particle. If the pointer state shifts by $\pm\gamma$, we associate the intermediate state with the eigenstate of the observable and claim to have measured the eigenvalue ± 1. The eigenvalue would be a property of the single particle. Similarly, if we obtain the pointer shift to be $\alpha\gamma$, then we can argue that α represents some value of the observable for the

[15]This may seem quite similar to the strong measurement and the fact that he expectation value is not necessarily one of the eigenvalues can lead us to incorrectly claim that the result of strong measurement of a joint system of N particles is also not an eigenvalue but the expectation value. The key distinction is that the expectation value is an incoherent average and thus is limited within the range of eigenspectrum. The result of weak measurement can shift the pointer beyond it due to coherent combination. Another way of looking at is that if we couple the variable $\mathbb{A}$ with a single pointer particle, in case of strong measurement we would always obtain the pointer shift to be one of the eigenvalues. The result of a weak measurement however could shift the pointer arbitrarily.

[16]How to practically implement such a coupling does not limit us from applying the concept for the case. The theory should be applicable as long as all the limits are applied consistently

single particle. But how can the value of an observable be anything else than its eigenvalues ? Well, the statement on value of an observable being one of its eigen values comes with the concept of strong projective measurements when the system is only subjected state preparation. α could represent the value of an observable for a system subjected to pre and post-selection. Thus, if the pointer shifts by α after coupling to a single particle, we are tempted to associate the value with the system itself. Note that, unlike the case of eigenvalues, we may not be able to associate an intermediate state to the observable corresponding to the pointer shift α and thus we have to be content with associating the value of the observable only. This value α by which the pointer shifts, results from the weak measurement procedure when the system and hence shall be termed as the weak value ${}_{\phi}\langle A\rangle_{\psi}^{(w)}$. of the observable A in the given pre and post-selected states.

So far we have considered a gaussian pointer state. But the results above are generic as long as the current density vanishes [2.12]. Let's consider an arbitrary pointer initial pointer state $|\chi_i(x)\rangle$ which is coupled to the observable A. After unitary evolution and post-selection, the final pointer state is rendered as follows:

$$|\chi_f(x)\rangle = \langle\phi|\exp(-(i/\hbar)\gamma A\otimes P_x)|\psi\rangle|\chi_i(x)\rangle \tag{2.60}$$

$$= \langle\phi|(\mathbb{1} - i\gamma A P_x - \frac{i\gamma^2}{2!}A^2P_x^2\ldots)|\psi\rangle \tag{2.61}$$

$$= \left(\langle\phi|\psi\rangle - \frac{i}{\hbar}\gamma\langle\phi|A|\psi\rangle P_x - \frac{\gamma^2}{\hbar^2 2!}\langle\phi|A^2|\psi\rangle P_x^2\ldots\right)|\chi_i(x)\rangle \tag{2.62}$$

$$= \left(\langle\phi|\psi\rangle + \sum_{k=1}^{\infty}\left(\frac{(-i\gamma)^k}{\hbar^k k!}\langle\phi|A^k|\psi\rangle P_x^k\right)\right)|\chi_i(x)\rangle \tag{2.63}$$

$$= \langle\phi|\psi\rangle|\chi_i(x)\rangle + \langle\phi|\psi\rangle\frac{\sum_{k=1}^{\infty}\left(\frac{(-i\gamma)^k}{\hbar^k k!}\langle\phi|A^k|\psi\rangle P_x^k\right)}{\langle\phi|\psi\rangle}|\chi_i(x)\rangle \tag{2.64}$$

$$= \langle\phi|\psi\rangle|\chi_i(x)\rangle + \langle\phi|\psi\rangle\sum_{k=1}^{\infty}\left(\frac{(-i\gamma)^k}{\hbar^k k!}\frac{\langle\phi|A^k|\psi\rangle}{\langle\phi|\psi\rangle}P_x^k\right)|\chi_i(x)\rangle \tag{2.65}$$

We can truncate the series at the first term if only the following approximation can be made.

$$\left|\frac{\gamma^k}{\hbar^k k!}\frac{\langle\phi|A^{k+1}|\psi\rangle}{\langle\phi|\psi\rangle}P_x^k|\chi_i\rangle\right| << \left|\frac{\gamma}{\hbar}\frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}P_x|\chi_i\rangle\right| \forall\ k\geq 2 \tag{2.66}$$

This approximation implies that pointer shift caused by all higher order terms in the series

expansion is less than the first order term. Hence, we can approximate the final pointer state. This can be ensured if $\gamma/\sigma_0 << 1$ where σ_0 is the position uncertainty of the pointers state. In this approximation,

$$\left|\chi_f(x)\right\rangle \approx \langle\phi|\psi\rangle \left(1 - \frac{i}{\hbar}\gamma\frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}P_x\right)|\chi_i(x)\rangle \tag{2.67}$$

$$\approx \langle\phi|\psi\rangle \exp\left(-\frac{i}{\hbar}\gamma\langle A\rangle^{(w)}P_x\right)|\chi_i(x)\rangle \tag{2.68}$$

In the above expression we have introduced the term $\langle A\rangle^{(w)}$, which is called as the weak value and is given by $\frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}$. The expression 2.68 suggests that after post-selection in weak measurement, we have the pointer shifted by the weak value $\gamma\langle A\rangle^{w}$, the term that becomes the coefficient to P_x in the translation operator (when we compare with $T(x) = \exp\left(-\frac{i}{\hbar}xP_x\right)$ and extract the shift x).

The weak value is in general complex. So only when the weak value is real, it would represent the pointer shift. For real weak values, the expression 2.68 suggests that the post-selection probability is unaffected. The post-selection probability is given by the norm of the final pointer state.

$$\left|\langle\chi_f|\chi_f\rangle\right|^2 = |\langle\phi|\psi\rangle|^2 \exp\left(\frac{i}{\hbar}\gamma\langle A^*\rangle^{(w)}P_x\right)\exp\left(-\frac{i}{\hbar}\gamma\langle A\rangle^{(w)}P_x\right)\langle\chi_i(x)|\chi_i(x)\rangle \tag{2.69}$$

$$= |\langle\phi|\psi\rangle|^2 \langle\chi_i(x)|\exp\left(-2\frac{\gamma}{\hbar}\mathrm{Im}\left\{\langle A\rangle^{(w)}\right\}P_x\right)|\chi_i(x)\rangle \tag{2.70}$$

$$\approx |\langle\phi|\psi\rangle|^2 \left(1 - 2\frac{\gamma}{\hbar}\mathrm{Im}\left\{\langle A\rangle^{(w)}\right\}\langle\chi_i(x)|P_x|\chi_i(x)\rangle\right) \tag{2.71}$$

Thus, if the weak value is real, the post-selection probability remains the same $|\langle\phi|\psi\rangle|$ when compared to the case when we do not perform any weak coupling in the intermediate time. The imaginary part of the weak value appears in the expression 2.71 which decreases the post-selection probability if the expectation value of momentum in the initial pointer state is non-zero.

Since the first order contributions are limited, since most pointers will be initialized with expectation value of momentum to be 0, to get the leading order effect we need to consider the higher order terms. If the approximation given in expression 2.66 is satisfied,

then leading order term for the change in pre and post-selected intensity would come from Eqn. 2.67 as the following:

$$-\frac{\gamma^2}{\hbar^2}|\langle A\rangle^w|^2\langle\chi_i(x)|P_x^2|\chi_i(x)\rangle \approx -\frac{\gamma^2}{4\sigma_0^2\hbar^2}|\langle A\rangle^w|^2 \tag{2.72}$$

Thus, if the pointer state has momentum uncertainty [17] ($\approx 1/(2\sigma_0)$, if σ_0 is the position uncertainty), then the square magnitude of weak value decreases the post-selection probability.

The real weak value shifts the position of the pointer state. What does the imaginary portion do ? If we consider the weak value to be purely imaginary in Eqn. 2.68, we have the pointer state as

$$\big|\chi_f(x)\big\rangle \approx \langle\phi|\psi\rangle\exp\left(\frac{\gamma}{\hbar}\left(\mathrm{Im}\Big\{\langle A\rangle^{(w)}\Big\}P_x\right)\right)|\chi_i(x)\rangle \tag{2.73}$$

$$\Rightarrow \big|\chi_f(p_x)\big\rangle = \langle\phi|\psi\rangle\exp\left(\frac{\gamma}{\hbar}\left(\mathrm{Im}\Big\{\langle A\rangle^{(w)}\Big\}p_x\right)\right)|\chi_i(p_x)\rangle \tag{2.74}$$

The above suggests that the imaginary portion of the weak value adds to the momentum of the pointer state.

If weak value is in general complex, then we can separate it into the real and imaginary parts and write the final pointer state as the following (with $\mathrm{Re}\Big\{\langle A\rangle^{(w)}\Big\} = \langle A_r\rangle^{(w)}$ $\mathrm{Im}\Big\{\langle A\rangle^{(w)}\Big\} = \langle A_i\rangle^{(w)}$):

$$\big|\chi_f(x)\big\rangle \approx \langle\phi|\psi\rangle\, e^{-\frac{i}{\hbar}\gamma\left(\langle A_r\rangle^{(w)}P_x\right)} e^{\frac{1}{\hbar}\gamma\left(\langle A_i\rangle^{(w)}P_x\right)}|\chi_i(x)\rangle \tag{2.75}$$

Next, we compute the expectation value of the position x and momentum p observable when the weak value is real.

$$\langle x\rangle_f\big]_{\langle A_i\rangle^w=0} = \left.\frac{\langle\chi_f(x)|x|\chi_f(x)\rangle}{\langle\chi_f|\chi_f\rangle}\right]_{\langle A_i\rangle^w=0} \tag{2.76}$$

$$\approx \langle\chi_i(x)|\, e^{\frac{i}{\hbar}\gamma\left(\langle A_r\rangle^{(w)}P_x\right)} x\, e^{-\frac{i}{\hbar}\gamma\left(\langle A_r\rangle^{(w)}P_x\right)}|\chi_i(x)\rangle \tag{2.77}$$

$$= \langle\chi_i(x+\langle A_r\rangle^w\gamma)x|\chi_i(x+\langle A_r\rangle^w\gamma)\rangle| \qquad = \langle A_r\rangle^w\gamma \tag{2.78}$$

[17] For a gaussian beam, the momentum expectation value is zero and thus the beam divergence would place the role of momentum uncertainty

$$\langle P_x \rangle_f \Big]_{\langle A_i \rangle^w = 0} = \frac{\langle \chi_f(x) | P_x | \chi_f(x) \rangle}{\langle \chi_f | \chi_f \rangle} \Bigg]_{\langle A_i \rangle^w = 0} \tag{2.79}$$

$$\approx \langle \chi_i(x) | \, e^{\frac{i}{\hbar}\gamma \left(\langle A_r \rangle^{(w)} P_x \right)} P_x \, e^{-\frac{i}{\hbar}\gamma \left(\langle A_r \rangle^{(w)} P_x \right)} \, | \chi_i(x) \rangle \tag{2.80}$$

$$= \langle \chi_i(x + \langle A_r \rangle^w \gamma) P_x \, | \chi_i(x + \langle A_r \rangle^w \gamma) \rangle | \tag{2.81}$$

$$= 0 \qquad \text{(assuming mean momentum of the initial state to be zero)} \tag{2.82}$$

When the weak value is completely imaginary, we compute the expectation value of the position observable turns out to be zero.

$$\langle x \rangle_f \Big]_{\langle A_r \rangle^w = 0} = \frac{\langle \chi_f(x) | x | \chi_f(x) \rangle}{\langle \chi_f | \chi_f \rangle} \Bigg]_{\langle A_r \rangle^w = 0} \tag{2.83}$$

$$\approx \langle \chi_i(p_x) | \, e^{\frac{1}{\hbar}\gamma \left(\langle A_i \rangle^{(w)} P_x \right)} x \, e^{\frac{1}{\hbar}\gamma \left(\langle A_i \rangle^{(w)} P_x \right)} \, | \chi_i(p_x) \rangle \tag{2.84}$$

$$= \langle \chi_i(p_x) | \, e^{\frac{1}{\hbar}\gamma \left(\langle A_i \rangle^{(w)} p_x \right)} x \, e^{\frac{1}{\hbar}\gamma \left(\langle A_i \rangle^{(w)} p_x \right)} \, | \chi_i(p_x) \rangle \tag{2.85}$$

$$\approx \langle \chi_i(p_x) | \left(1 + \frac{1}{\hbar}\gamma \langle A_i \rangle^{(w)} p_x \right) x \left(1 + \frac{1}{\hbar}\gamma \langle A_i \rangle^{(w)} p_x \right) | \chi_i(p_x) \rangle \tag{2.86}$$

$$= \langle \chi_i(p_x) | x | \chi_i(p_x) \rangle + \frac{1}{\hbar}\gamma \langle A_i \rangle^{(w)} \langle \chi_i(p_x) | px + xp | \chi_i(p_x) \rangle \tag{2.87}$$

$$= \frac{1}{\hbar}\gamma \langle A_i \rangle^{(w)} \langle \chi_i(p_x) | px + xp | \chi_i(p_x) \rangle \tag{2.88}$$

In the last step we assumed that the pointer variable x was initialized at 0. Still, a pure imaginary pointer variable do cause translation of position of the pointer when $\langle \chi_i(p_x) | px + xp | \chi_i(p_x) \rangle \neq 0$. This implies that, we could obtain some pointer shift even if the weak value is purely imaginary. However, if our pointer state is given by non-diverging gaussian beam $g(x)$ with plane wavefront, we have $\langle g(x) | px + xp | g(x) \rangle = 0$.

$$\int_{\infty}^{\infty} g(x) \left((-i\hbar)\partial_x (x g(x)) + (-i\hbar) x \partial_x (g(x)) \right) dx \tag{2.89}$$

$$= (-i\hbar) \int_{\infty}^{\infty} g(x) \left(g(x) + x g'(x) + x g'(x) \right) dx \tag{2.90}$$

$$= (-i\hbar)\int_{\infty}^{\infty} g(x)\left(1-\frac{x^2}{\sigma_0^2}\right)g(x)dx \qquad (\text{since } g'(x) = -g(x)/(2\sigma_0^2)) \tag{2.91}$$

$$= 0 \qquad (\text{since } \int x^2 g(x)^2 dx = \sigma_0^2) \tag{2.92}$$

Hence, for well collimated gaussian beam the pointer variable $\langle x\rangle$ would have negligible shift. The imaginary weak value however translates the expectation value of the momentum observable.

$$\langle P_x\rangle_f \big]_{\langle A_r\rangle^w=0} = \frac{\langle \chi_f(x)|P_x|\chi_f(x)\rangle}{\langle \chi_f|\chi_f\rangle}\Bigg]_{\langle A_r\rangle^w=0} \tag{2.93}$$

$$\approx \langle \chi_i(p_x)|\, e^{\frac{1}{\hbar}\gamma\left(\langle A_i\rangle^{(w)}P_x\right)} P_x\, e^{\frac{1}{\hbar}\gamma\left(\langle A_i\rangle^{(w)}P_x\right)}\, |\chi_i(p_x)\rangle \tag{2.94}$$

$$\approx \langle \chi_i(p_x)|\left(1+\frac{1}{\hbar}\gamma\langle A_i\rangle^{(w)}p_x\right)p_x\left(1+\frac{1}{\hbar}\gamma\langle A_i\rangle^{(w)}p_x\right)|\chi_i(p_x)\rangle \tag{2.95}$$

$$\approx \langle \chi_i(p_x)|\, p_x\, |\chi_i(p_x)\rangle + 2\frac{1}{\hbar}\gamma\langle A_i\rangle^{(w)}\langle \chi_i(p_x)|\, p_x^2\, |\chi_i(p_x)\rangle \tag{2.96}$$

$$= 2\frac{1}{\hbar}\gamma\langle A_i\rangle^{(w)}\langle p_x^2\rangle \tag{2.97}$$

Thus if $\langle \chi_i(p_x)|\, p_x\, |\chi_i(p_x)\rangle = 0$, the imaginary part of the weak value couples to the variance of the momentum or divergence of the initial pointer state.

2.3 Experimental Setup for weak measurement

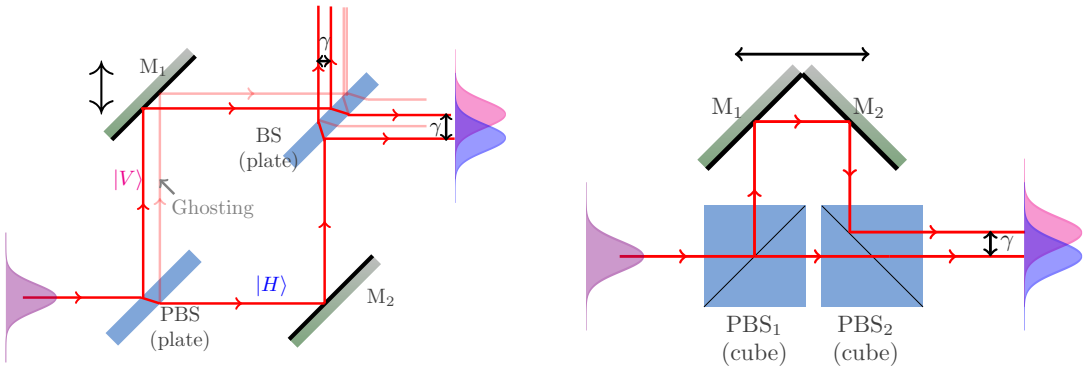
For strong projective measurement of polarization component of photon, we need a linear[18] optical component that behaves differently for a pair of orthogonal polarization. A linear polarizer which typically transmits one optical component and absorbs the component orthogonal to the first is not suitable for weak coupling of the photon's polarization with pointer states. Wire grid polarizers such as WP25M-VIS, Thorlabs and thin film polarizers like LPVIS100, Thorlabs absorb one polarization component which is orthogonal to the orientation of the grid axis. Further, the transmission for the polarization along the orientation of the grid is also not great due to material absorption and reflection. On the

[18]Here linearity is in terms of the electric field, i.e, the behavior of the device is independent of the intensity incident on it.

other hand polarizing plate beamsplitters, like PBSW-808, Thorlabs, separate the incident beam into two polarization components - the 'p'- polarized light which has polarization parallel to the plane of incidence is transmitted and the 's'- polarization with electric field oscillations perpendicular to the plane of incidence is reflected. In principle, we can use two such plate beamsplitters and mirrors to create a Mach-Zehder interferometer (with 50:50 beam splitters replaced with polarizing beam splitters) to combine the two beams and make the two gaussian beams almost overlap as is needed for the weak interaction limit. The translation of the mirrors would allow one to vary the displacement between the two orthogonally polarized beams i.e., vary the parameter γ (See Fig . **??** below). However, polarizing beam splitters would typically suffer from ghosting i.e., the second flat surface of the plate would reflect part of the beam supposed to be transmitted which would lead to complications owing to multiple reflections. Polarizing cube beam splitters, such as PBS252, Thorlabs, do not have ghosting and they can be configured to create a setup in which the overlap between two orthogonal beams can be controlled and hence the interaction strength can be varied.

However, polarizing beam splitters (both plates and cubes) have relatively poor extinction ratio in the reflecting port. If the incident beam itself is not s or p polarized, the reflected beam would be slightly elliptic . This makes the two polarization component separated slightly non-orthogonal which is not ideal for a measuring device if precision is important. The best extinction ratio is provided by birefringent crystals which separate the ordinary and extraordinary polarization component. Among them, the Wollaston prism causes angular deviation of the emergent ordinary and the extraordinary rays. This forms a perfect analogue to the Stern-Gerlach setup for spin-1/2 particles, where one spin-polarization component gets a momentum kick in one direction while the orthogonal component gets the momentum opposite to it. Similar to the Stern Gerlach setup, as the two emergent beams propagate after the interaction is over, the overlap between them decreases with propagation distance. This is indeed desirable for tunability of the interaction strength, but puts a limit on how far we can place the detectors. In the experiment on partial cheshire cat discussed in Ch 3 and 4, it would be evident that it is useful if the overlap between the two states remains the same with propagation distance as that would allow one to modify detection modules without affecting the interaction strength γ/σ_0. The same is of course

achievable by the setups shown in Fig. 2.5.



(a) The plate (or cube) polarizing beam splitter (PBS) separates the two orthogonal polarizations , in this case $|H\rangle$ and $|V\rangle$. For the plate beam splitter, there is typically a ghost reflection from the second surface. For big enough interferometers the ghosting can be avoided by introducing a wedged surface so that the ghost reflected beam is not parallel but at some angle to the reflected beam. The Mach Zehnder interferometer is completed at the plate beam splitter (BS) after the reflections from the two mirror at $\pi/4$ angle of incidence. The translation of either of the mirror (here depicted with M_1) would create a displacement between the horizontally and vertically polarized beams and changing that we can control γ. If the BS is replaced with another PBS, the setup would still work with the two beams going upwards, as per the diagram, after the second PBS. For a similar setup with Sagnac geometry, see [2.13]

(b) The arrangement of two mirrors M_1 and M_2 (or an equivalent retro-reflector) can be translated which causes a displacement between the two beams at the PBS_2, after the two components are separated at the polarizing beam splitter PBS_1. Due to poor extinction ratio of most of the polarizing beam splitters in the reflected port, the refelected beam from PBS_1 would be slightly elliptic instead of ideally considered $|V\rangle$. After couple of reflections from the mirrors and the subsequent reflection from PBS_2, the two emergent beams would not be completely orthogonal. Corner cube retro-reflector can replace the two mirrors for alignment stability, but would need waveplates correction to the polarization. The two emergent beams, however, also have a macroscopic path difference due to which the setup may need alteration for sources with low coherence length. The cube beam splitters typically avoids the ghosting.

Figure 2.5: Setups to create a variable beam displacement between two orthogonal polarizations. Although the above setups provide us with scope to vary the beam displacement γ, we would not be using the above due to the typical poor extinction ratio of the reflected port of beam splitters.

But a monolithic birefringent crystal would allow us not to worry about stabilizing the interferomters shown in 2.5 and give better extinction ratio if we sacrifice the tunability. With Rochon prisms the ordinary ray emerges parallel to the incident ray while the extraordinary ray emerges with an angular deviation. With this, the transverse spatial momentum gets a kick only for the extraordinary component while the ordinary ray is ideally unaffected. With Rochon prisms, we still would have the interaction strength vary

over propagation distance.

The polarizing beam displacer makes both the ordinary and extraordinary polarization component emerge in a direction parallel[19] to the oropagation direction of the incident ray. The beam displacement between the emergent rays depend on the crystal material properties i.e., the refractive indices for the ordinary and emergent rays, the length of the crystal and the wavelength and angle of the incident light. Once these are fixed in the experiment, we cannot vary the beam displacement i.e., γ. But the variation of γ can be traded off for better extinction ratio, since in the experiments to be discussed further, we shall find other parameters which can be varied to study the effect of those parameters on the pointer shift. The typical beam width which will be used in experiments is about 2mm (standard deviation of the gaussian is about $\sigma_0 = 500\ \mu$m). A smaller beam width will lead to beam divergence. A much larger beam width would require the surface of optics to be uniform[20] over a large area. Therefore, we need a much smaller beam displacement. For weak interaction, we need beam displacement as small as possible. But then, we need to be able to observe those small displacements. If the pointing fluctuations of the beam is more than the beam displacement then the detection of displacement would be lost in the noise [21]. Therefore, we choose a beam displacement of about $\approx 50\ \mu$m in our experiments. The polarizing beam displacer used for weak interaction in experiments described in this and subsequent chapters is the calcite crystal PDC12005, Newlight Photonics.

The calcite crystal is a negative uniaxial crystal, i.e., the refractive index n_e of the extraordinary ray is less than the refractive index n_o of the ordinary ray i.e. the birefringence $\Delta n = n_e - n_o < 0$. The optic axis of this polarizing beam displacer crystal is oriented along 45 deg in the y-z plane - as depicted in Fig. 2.6. Upon normal incidence, this is the plane that is spanned by the crystal optic axis and the propagation direction vector k_i

[19] with small angular deviations depending upon how parallel the surfaces of the crystal are cut

[20] One needs to ensure that the wavefront is plane. Therefore, even small dust particles on surface of optics would affect the experiment. Therefore, one needs to ensure cleanliness of optics over the entire beam size. Larger beam size makes this difficult to achieve and hence typically results in poorer transverse coherence.

[21] The pointing fluctuations in the beam are of the order of 1-10μrad, which for a distance of the order of 1m distance results in a transverse beam fluctuation of the order of 1-10μm Thus the beam displacement needs to be larger than this. The pointing fluctuation depends on the opto-mechanics, use of fibers, temperature and vibrations which we later attempt to control or monitor. Using a telescopic arrangement lowers the beam fluctuation and magnifies the beam displacement (See beam pointing fluctuations on RP photonics website). This, in principle, can be used as was done in the first experimental measurement of weak value[2.3]. But, because of the reasons that would be clear in later chapters, we avoid using the telescopic arrangement.

of the incident beam. The ordinary ray, which follows the Snell's law goes undeviated. The polarization of the ordinary ray is perpendicular to plane containing the crystal optic axis and the propagation vector. Its refractive index is independent of the direction of propagation even if incidence were not normal. For the orientation of the crystal depicted in Fig. 2.6, the ordinary ray has horizontal polarization. The velocity of the extraordinary ray, however, depends on the angle it makes with the crystal optic axis. Even upon normal incidence, the extraordinary ray, the polarization of which is in the plane of the optic axis and k_i, can get deviated. This occurs because inside an anisotropic crystal, the direction of flow of energy i.e., the Poynting vector may differ from the direction of wave propagation [2.14].

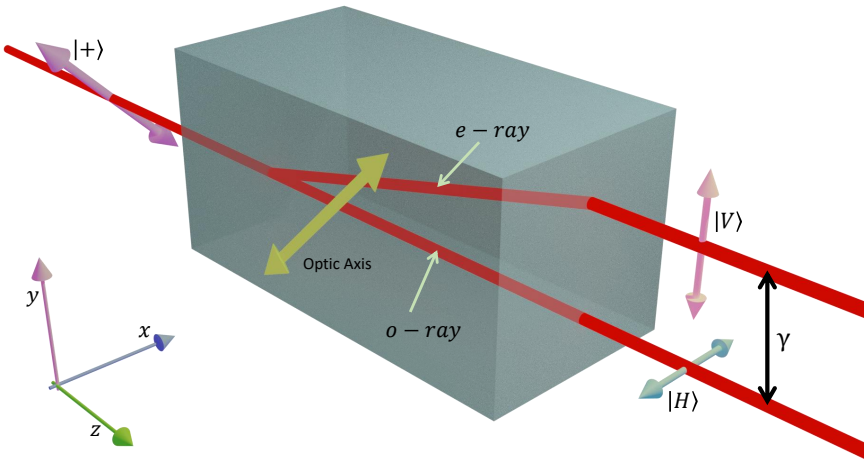


Figure 2.6: Geometry of the polarizing beam displacer: The crystal optic axis is along the diagonal of the y-z plane. If the beam is made incident along z, then the plane formed by the wave propagation vector k_i and the crystal optic axis is the y-z plane. The horizontal polarization (along x axis), is perpendicular to this y-z plane and hence behaves as the ordinary ray. Upon normal incidence on the parallel plate crystal, the ordinary ray emerges undeviated. The polarization component in the y-z plane gets refracted upwards towards y for the negative uniaxial crystal. The extraordinary ray exits the crystal parallel to the ordinary ray but with vertical displacement of γ along the vertical. The polarization of the displaced beam is vertical for the given orientation of the crystal.

In this particular case of negative uniaxial crystal the extraordinary beam gets deviated towards y axis in the plane formed the crystal axis and the wave propagation vector of the incident beam which is normal to the surface. Thus, the extraordinary ray gets a shift along y as shown in Fig. 2.6. The oscillations of the electric field for the extraordinary ray inside the crystal are not perpendicular to the wave-vector. The extraordinary ray upon exit from

the crystal is vertically polarized which is orthogonal to the polarization of the ordinary ray. For the crystal cut as a parallel plate, the ordinary and extraordinary ray emerge parallel to the incident beam but with a relative displacement which would be γ.

If the crystal is rotated by an azimuthal angle θ_z about the z-axis, measured from y-axis (i.e., $\theta_z = 0$ implies emergent extraordinary ray is vertically polarized), then the extraordinary ray, upon exit from the crystal, will have the polarization along the projection of the optic axis along the x-y plane. Thus, the linear polarization of the emergent extraordinary ray would be given by $|P_e\rangle$ and the ordinary ray would have the polarization orthogonal to it as $|P_o\rangle$ as described below [22].

$$|P_e\rangle = \cos(\theta_z)\,|V\rangle - \sin\theta_z\,|H\rangle = \begin{pmatrix} -\sin(\theta_z) \\ \cos(\theta_z) \end{pmatrix} \tag{2.98}$$

$$|P_o\rangle = \cos(\theta_z)\,|H\rangle + \sin\theta_z\,|V\rangle = \begin{pmatrix} \cos(\theta_z) \\ \sin(\theta_z) \end{pmatrix} \tag{2.99}$$

The displacement of the extraordinary ray for negative uniaxial crystal upon normal incidence would also be along a line in the x-y plane making θ_z angle with the y-axis. Thus, if the crystal is rotated by 90 deg from the configuration shown in Fig. 2.6, so that the optic axis is in the x-z plane, then the beam displacement of γ would be along the horizontal i.e., along x. The displaced extraordinary beam would now have polarization $|H\rangle$ and the undeviated ordinary ray would have polarization $|V\rangle$.

For the purpose of creating a beam displacement between two orthogonal polarizations, normal incidence would suffice. But for controlling the phase shift between the emergent rays, the need for which will be later described in section 2.4.2, we may need to tilt the crystal. For oblique incidence, i.e., say the angle of incidence is θ_i, and for the crystal optic axis making an angle θ_c with the incident beam, the angle of refraction θ_o for the ordinary ray is given by formulation [2.15] below which is equivalent to the Snell's law.

$$\tan(\theta_o) = \frac{n\sin\theta_i}{\sqrt{n_o^2 - n^2\sin^2(\theta_i)}} \tag{2.100}$$

[22] This is based on the application of the rotation matrix in two-dimension onto the H-V basis to transform them to ordinary and extraordinary rays for arbritary θ_z

Here, n is the refractive index of air or vacuum in which the crystal is placed. The angle of refraction of the extraordinary ray θ_e described below [2.15, 2.16], on the other hand, depends on θ_c.

$$\tan(\theta_e) = \frac{n\, n_e n_o \sin(\theta_i) + (n_e^2 - n_o^2)\cos(\theta_c)\sin(\theta_c)\sqrt{n_e^2\cos^2(\theta_c) + n_o^2\sin^2(\theta_c) - n^2\sin^2(\theta_i)}}{\left(n_e^2\cos^2(\theta_c) + n_o^2\sin^2(\theta_c)\right)\sqrt{n_e^2\cos^2(\theta_c) + n_o^2\sin^2(\theta_c) - n^2\sin^2(\theta_i)}} \tag{2.101}$$

Due to birefringence, the ordinary ray and extra-ordinary ray would intersect with the exit surface of the crystal at different points. The distance between those points, which is along the edge of the crystal, is however, not the beam displacement γ. The beam displacement is calculated between the parallel wavefronts of the emergent rays which occurs after the beams have crossed the blue line shown in the Fig.2.7.

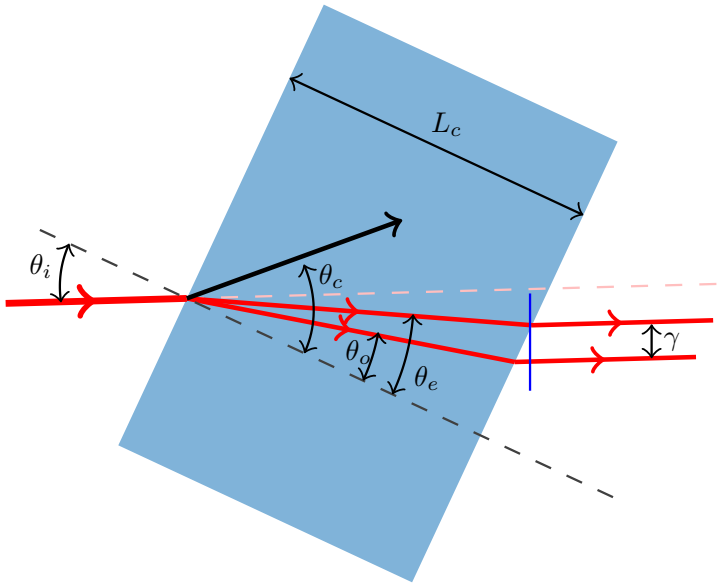


Figure 2.7: Beam deviation with tilted birefringent crystal: When the incident beam is not normal to the surface, the ordinary o-ray also gets parallel displacement given the surface where the beam is incident and emerges are parallel to each other. The extra-ordinary ray also gets a displacement different displacement compared to normal incidence. This implies when the polarizing beam displacer crystal is tilted, we can no longer presume the transverse position of the emergent ordinary ray after refraction to be position of the beam without any interaction. The beam displacement is computed when the wavefronts of the o-ray and e-ray become parallel to each other, which is indicated by the vertical blue line (normal to the emergent beams).

The beam displacement γ is thus obtained as follows.

$$\gamma = L_c(\tan(\theta_e) - \tan(\theta_o))\cos(\theta_i) \tag{2.102}$$

The beam displacement depends on the length of the crystal, the wavelength of light and the material on which n_o and n_e depends, the incident angle θ_i. Note that, when tilted, the ordinary ray also gets a shift. The beam displacement of the extraordinary ray is relative to the shifted ordinary ray and not with the position of the beam for the case when the crystal is removed. At normal incidence, however, they are equivalent. Since the ordinary ray and extra-ordinary ray traverses different optical path length inside the crystal, the emergent beams have a path delay between them. For coherent continuous wave beam, this results in a phase difference between the emergent e-ray and o-ray. For pulsed beams with finite coherence length, this results in temporal delay between the centre of the pulses.

2.3.1 Experimentally measuring the beam displacement

The beam-displacement for the polarizing beam displacer made up of calcite PDC12005, Newlight Photonics is quoted for a 50 μm displacement at 1550 nm wavelength. Using the Cauchy-Sellmeir equations for calcite and using Eqn. 2.102, we calculate the thickness/length L_c of the crystal to be 496.8 μm($\approx 500\mu$m). Using the evaluated L_c, we then calculate the beam displacement for 632.8 nm and 810 nm, the two wavelengths which we are going to use in the chapters ahead. We obtain the beam displacement to be $\gamma_{810\text{ nm}} = 53.2\ \mu m$ and $\gamma_{632.8\text{ nm}} = 54.3\ \mu$m respectively.

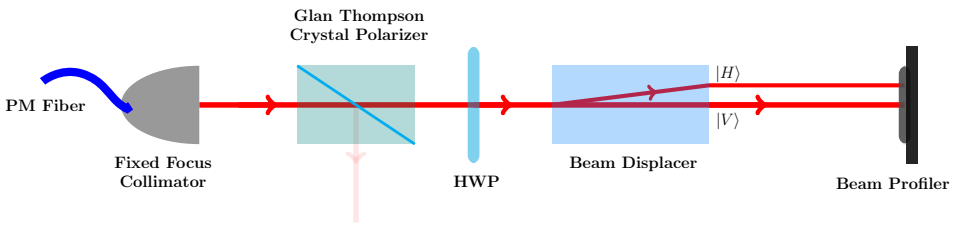


Figure 2.8: Experimental Setup to measure the beam dispacement. Later the HWP was rotated to obtain the expectation value as a function of the prepared linear polarization state.

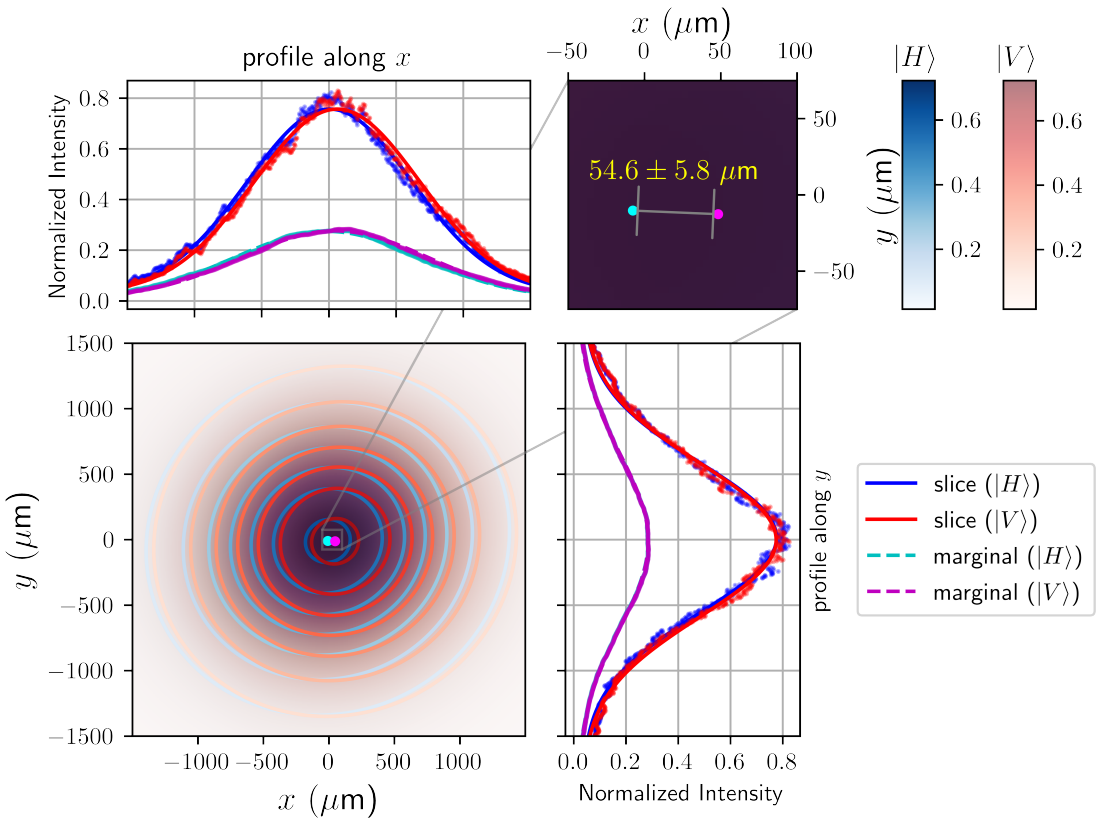


Figure 2.9: Measurement for the beam displacement: The incident polarization was made $|H\rangle$ and then the centre of the beam was measured. Then the incident polarization was changed to $|V\rangle$ and then the center of the beam was measured. The distance between the two centers gives us the beam displacement. The centre was computed using the three methods: First the centroid was calculated from the 2D beam profile and then it was fitted with a 2D gaussian function. Secondly, we slice a row/column from the 2D image with the row/column corresponding to the centroids. We measure the centroid and the centre for the Gaussian fit. Finally, we compress the 2D image into a single row/column i.e., compute the marginal intensity distribution. The centroid and centre obtained from respective gaussian fits for the three profiles (one 2D, and two 1D)slightly differ (about 5-10 μm)from each other as the beam profile was not perfectly a gaussian. Thus, the centroid of profile with horizontal polarization was subtracted from the centroid of the profile with vertical polarization. Similarly, the difference between centers obtained from respective gaussian fits was computed. Then the average and standard deviation of these measures over 3 datasets with 10 images each is reported. The beam profiler was not on a rotation mount and thus its tilt was not exactly matched to the horizontal plane. Nevertheless, the vertical displacement was about 8μm which is already accounted in the net vector distance calculated between the two points.

To experimentally measure the same, we first make the fiber coupled beam from Ti-Sapphire Coherent Mira 900 D laser operated in continuous wave mode at about 810 nm wavelength, horizontally polarized after passing through a Glan Thompson GTH10-

B, Thorlabs polarizer oriented to transmit horizontal polarization. The Glan-Thompson polarizer is chosen over the standard polarizing beam splitters because of couple of orders higher extinction ratio, which would be later important for weak value amplification. Then a half wave plate WPA03-H-810, Newlight Photonics is used to select between horizontal or vertical polarization to be incident on the polarizing beam displacer. The beam displacer was oriented in such a way that the displacement was along x (horizontal plane of the table). The beam profiler WinCamD-UCD15 was used to capture the beam profile, whose centroid and centre from the Gaussian fit was obtained for the two incident polarizations. The distance between the centroids/centres of the beam profile gives the beam displacement γ.

The uncertainty in the beam displacement mainly arises due to beam wander i.e. pointing fluctuations [23]. The uncertainty in the measurement described above turns out to be about 10% of the beam displacement. Note that we are not limited by the pixel size of the camera for the precision on determining the center of the beam as the centroid or the gaussian center can be more precisely determined given enough statistics, in this case, intensity value of pixels on the beam profiler.

2.3.2 σ_z, Π_H and Π_V measurement with the beam displacer

When the optic axis of the birefringent crystal is in the y-z plane, upon normal incidence, the vertical polarization component $|V\rangle$ gets displaced, say by amount γ. Then the coupling interaction can be given by the hamiltonian $H = g(t)\gamma\,\Pi_V \otimes P_y$. Similarly, when the optic axis of the crystal is oriented in the x-z plane, the interaction hamiltonian is given by $H = g(t)\gamma\,\Pi_H \otimes P_x$. Although which eigenstate gets displaced by what amount depends on the observable in the interaction, as far as distinguishability is concerned they are equivalent. In the strong interaction regime, if after the interaction we detect a photon, based on its transverse position, we can identify whether it is $|H\rangle$ or $|V\rangle$ - the eigenstates of σ_z as well. Thus, the measurement can be termed as the measurement of the σ_z observable as well. This is possible because, in two dimensional Hilbert space, we have the following identity, i.e, the sum rule.

[23] Also, slight misalignment in the polarization maintaining fiber caused the beam to oscillate vertically. These oscillations seemed to be correlated with temperature cycle due to air conditioning unit. The cause is not well understood by us, but better alignment of the polarization maintaining fiber reduces the fluctuation

$$\Pi_H + \Pi_V = \mathbb{1} \Rightarrow \Pi_H = \mathbb{1} - \Pi_V \quad \& \quad \Pi_V = \mathbb{1} - \Pi_H \tag{2.103}$$

$$\sigma_z = \Pi_H - \Pi_V = \mathbb{1} - 2\Pi_V \quad \& \quad \sigma_z = 2\Pi_H - \mathbb{1} \tag{2.104}$$

The above relations are between the operators corresponding to the observables which implies that the eigenvalues, expectation values, and eventually the weak values would respect the relations.

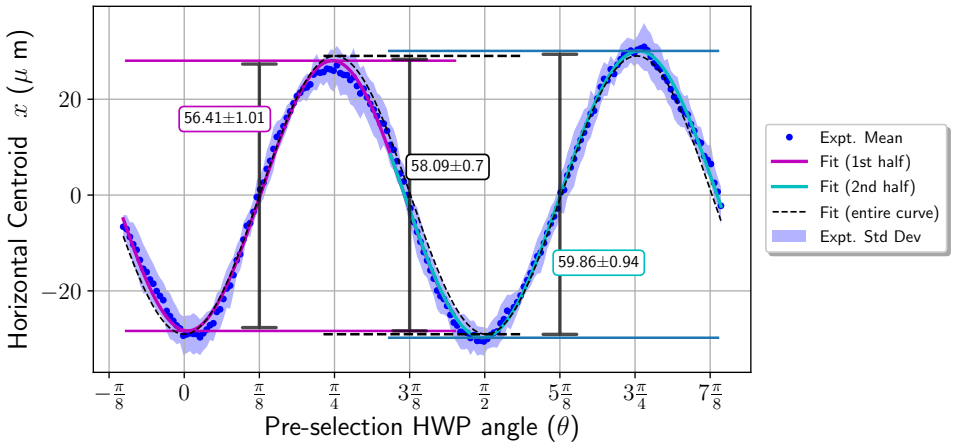
However one needs to interpret the beam displacement correctly. If we wish to treat the coupling with Π_H as σ_z observable, then no beam displacement, which is assigned eigenvalue of 0 for the eigenstate $|V\rangle$ for the observable Π_H, has to be assigned eigenvalue of -1 for the observable σ_z. The expectation value of 0 for the σ_z observable would be a displacement of $\gamma/2$. Thus, we have to interpret the displacement of the pointer as the value of the observable by choosing the following method. In separate runs, make the prepared state (or post-selected state) as the eigenstate of the observable and record the corresponding pointer variables as x_l(for the state with lower eigenvalue) and x_h(for the state with higher eigenvalue) respectively. Then the pointer variable x corresponds to the value $e_l + \frac{x-x_l}{x_h-x_l}(e_h - e_l)$ of the observable, where e_l and e_h are the lower and the higher eigenvalues.

$$x \to e_l + \frac{x - x_l}{x_h - x_l}(e_h - e_l) \tag{2.105}$$

2.3.3 Measuring the expectation value of σ_z observable

The beam displacement obtained through experimental result described in 2.9 may suffer from one systematic error. The beam might have drifted when we take the two separate readings for the two eigenstates. The effect of the drift can be eliminated to some extent if we repeat the observations over many different datasets. However, we can use another method which will be relatively robust to the beam drift. We can measure the expectation value as a function of pre-selected state and fit the expected curve to obtain the beam displacement.

The HWP in Fig 2.8 can be rotated to change the pre-selected state. For HWP angle



H

Figure 2.10: Without any post-selection, we vary the HWP shown in the experimental setup Fig. 2.8 and record the centroid (mean and standard deviation shown over 10 images and 5 repetitions i.e. 50 images total per angle). We fit the function $B + A\cos(4k(\theta - \theta_0))$ to the data. The beam displacement γ is obtained as $2A$. The uncertainty in γ is the error obtained from the covariance matrix while fitting the curve. In this way, the uncertainty in determining the beam displacement is robust against beam wander occurring at a particular pre-selection HWP angle. However, we notice that, both the 90 deg cycles of the HWP in which the centroid shifts to the extreme do not yield the same beam displacement. For the HWP angles ranging $-23°$ to $113°$ the beam displacement turns out to be $56.41 \pm 1.01\mu$m from the fit (shown in magenta).The curve fitting for the second half of the curve gives the beam displacement of 59.86 ± 0.94. This could be due to two reasons: One that the HWP causes some but small shift the centre of the beam, which is later verified to be the case. The other is that the HWP does not act symmetrically about the fast and the slow axis, which later becomes pronounced when we perform weak measurements. For now, the fit of the curve to the different halfs of the data yields slightly different value of k (which ideally should be 1) indicating the asymmetry

θ, the prepared state is given[24] by $(\cos(2\theta), \sin(2\theta))$. The expectation value of the σ_z observable is then computed as a function of θ.

$$\langle \sigma_z \rangle (\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(2\theta) \\ \sin(2\theta) \end{pmatrix} \tag{2.106}$$

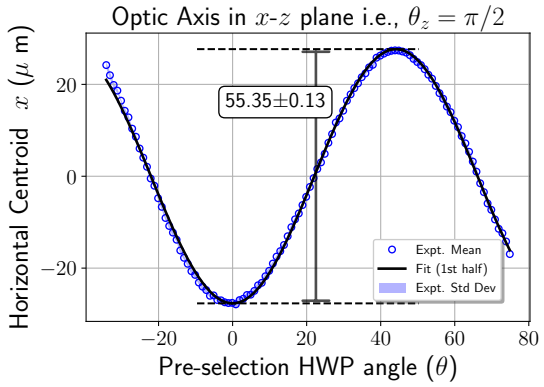
$$= \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \end{pmatrix} \begin{pmatrix} \cos(2\theta) \\ -\sin(2\theta) \end{pmatrix} \tag{2.107}$$

$$= \cos^2(2\theta) - \sin^2(2\theta) = \cos(4\theta) \tag{2.108}$$

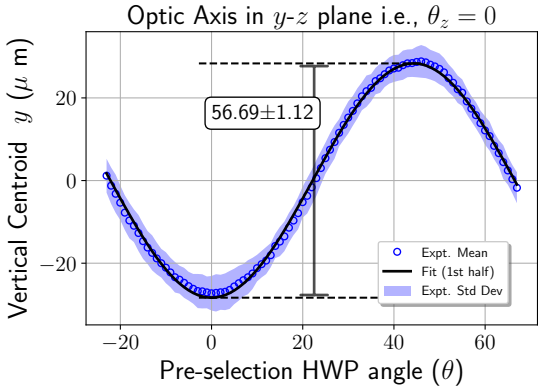
After determining the beam displacement, the next step was to measure the weak value

[24] This is obtained after applying the Jones matrix for a half wave plate on to the horizontally polarized state

of σ_z observable. As a proof of concept, the pointer variable should remain unchanged [25] as a function of pre-selection angle when the post-selected state is $|H\rangle$, one of the eigenstate of σ_z. Experimentally, when we prepared the pre-selected state orthogonal to the post-selected state, introducing the beam displacer slightly increased the post-selection probability. Most likely it was the slight error in setting θ_z which affects the post-selection probability. But there was another visible effect which seemed to some sort of multiple reflection from some component, most likely the beam displacer or the polarizing beam splitter introduced for post-selection. Following [2.4], we then used lens in a confocal configuration as shown in Fig. 2.12a to avoid multiple reflections [26].



(a) The crystal was oriented in such a way that it caused horizontal beam displacement. The use of lens seemed to help in reducing pointing fluctuations (or it could be a coincidence).

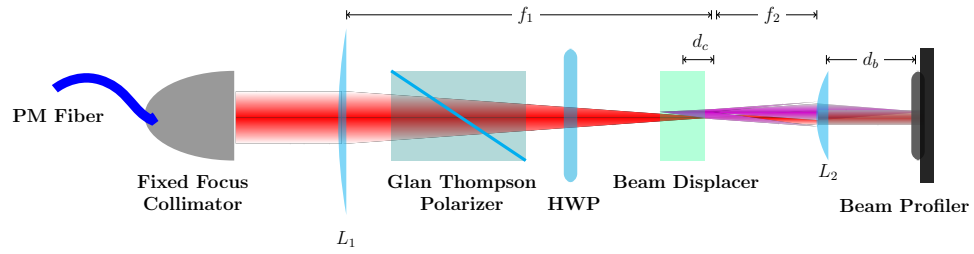


(b) When the crystal was oriented for vertical beam displacement, the beam still drifted up and down. This leads to a larger uncertainty

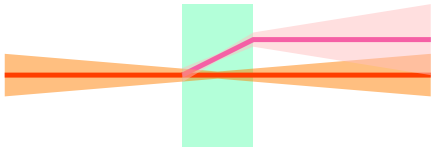
Figure 2.11: Expectation value of σ_z observable (in terms of pointer shift). However, after the lens the beam displacement varies as a function of the distance d_b of the beam profiler camera from the second lens L_2 shown in Fig. 2.12a. The above data shows the case when the beam profiler was placed in a position where the beam displacement was similar without the lens.

[25] except when pre-selection is orthogonal to the post-selected state where it is kind of indeterminate as the error in centre of the beam becomes very large

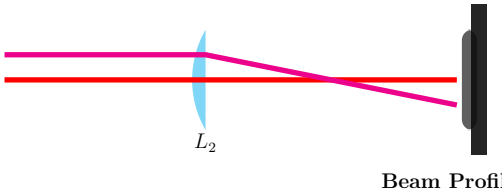
[26] There is an interesting comment in [2.4] on the effect of phase gradient across the wavefront due to the use of lens. The effect apprehended there was not observed in [2.4]. We also did not observe the effects of phase very clearly. But interestingly, when we perform weak measurement often the destructive interference was washed out due improper beam shape arising from dust. Passing the beam through a pinhole aperture often restored the transverse coherence. Instead of closing the aperture, focusing using lens did the same thing. Although we do not have evidence for the phase gradient across the beam, we have signatures of dust and distortion in beam shape washing out the destructive interference pattern



(a) Using 2m polarization maintaining fiber P1-780PM-FC-2, Thorlabs the beam from 810 nm CW Ti-Sapphire laser was coupled and collimated using the available lens 60FC-4-A11-02 as a collimator (not recommended by the vendor because of the longer focal length). Two plano convex lenses, L_1 with focal length $f_1 = 350$mm is arranged confocally with another lens L_2 of focal length $f_2 = 100mm$. When we perform weak measurement, a polarizing beam splitter PBS122, Thorlabs is placed at the focus. The beam displacer is placed slightly before the focus of the first lens with some distance d_c. As this distance d_c is varied, we verify that the beam displacement does not change much.



(b) When the polarizing beam displacer is at the focus, the separation between the beams can be larger than the beam width. So while the interaction occurs the condition $\sigma_0 >> \gamma$ is not satisfied. Often we consider σ_0 to be the uncertainty in the initial pointer state. But evolution of the uncertainty at the time of detection is crucial for the weak interaction limit



(c) The two orthogonally polarized beams emerge from the birefringent crystal parallel to each other. One or both the beams which are off-centred undergoes refraction and the displacement between them becomes a function of the propagation distance. The beam width almost remains the same due to confocal arrangement. Thus the interaction strength can be varied by longitudinal translation of the detector.

Figure 2.12: Confocal lens arrangement to avoid multiple reflections. It also provides a tool to vary the interaction strength.

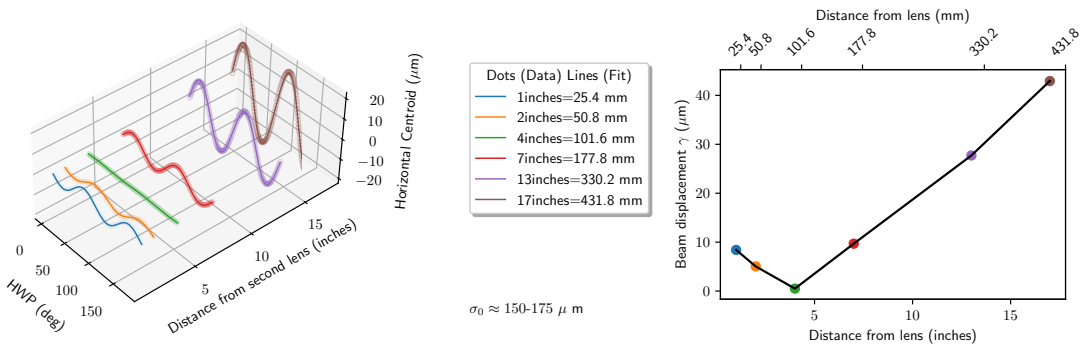


Figure 2.13: The beam displacement changes as a function of distance of the beam profiler d_b from the lens L_2. The beam size almost remains the same(divergence $\approx$ 0.3mrad).

2.3.4 Weak Measurement of σ_z observable with pre-selected state is the eigenstate

The weak measurement of σ_z observable when the post-selected state was kept $|H\rangle$ would yield the weak value 1. Thus, we expect the pointer not to shift at all as we change the pre-selected state. In the setup shown below, the beam diameter at the time of detection was about 700 μm. This corresponds to $\sigma_0 \approx 150\mu$m and describes an interaction strength[27] of $\approx 1/3$.

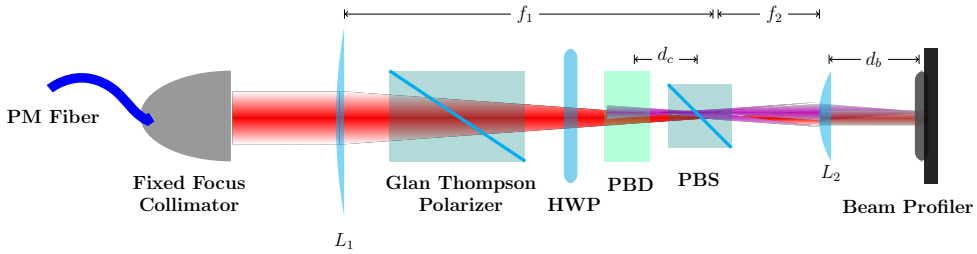


Figure 2.14: For weak measurement a polarizing beam splitter PBS is introduced at the focus of lens L_1. The polarizing beam displacer (PBD) is moved back. But still the beam incident on it is converging and hence multiple reflections are avoided. The centroid of the beam is measured as a function of the pre-selection angle.

.

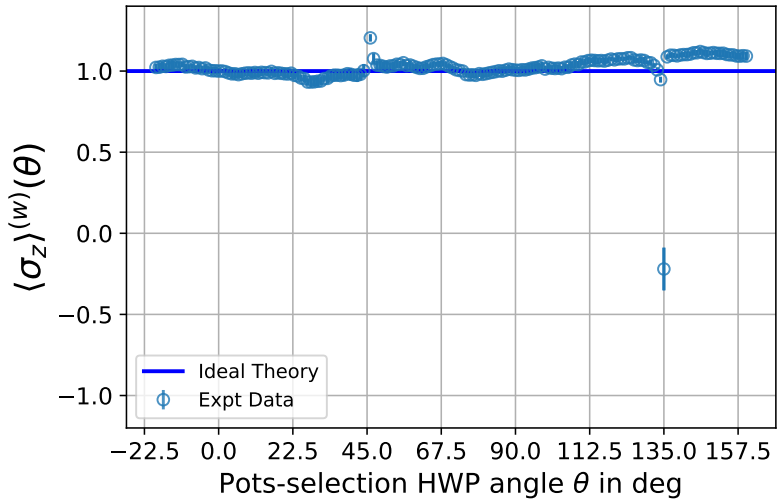


Figure 2.15: When the pre-selected state is $|H\rangle$, we denote the weak value to be 1. The changes and the uncertainty in the pointer are then scaled by $2/\gamma$ to convert the pointer shift to the weak value.

[27] The beam displacement γ corresponds to shift of 2 units in the eigenspectrum for the σ_z observable. Thus the interaction strength can be considered as $1/6$. However, irrespective of the observable description if we resort to distinguishability between the two eigenstates, then we have the interaction strength as $\gamma/\sigma_0 \approx 1/3$

First, we notice two anomalies in the graph 2.15 at $\theta = \pi/4$ (45°) and $\theta = \pi/4 + \pi/2$ (135°). The result when pre-selection becomes orthogonal to the post-selection. The exposure of the beam profile has a limitation and when the intensity is too low, then the measured centroid tends towards the geometric center of the image region which often has noise, i.e., we tend to measure the center of all the pixels with noise rather than the center of the beam. This effect is more prominent at the 135° in the graph. For an ideal wave plate we would have expected the same at 45° as well. But the fast and slow axis of the half-wave plate are not exactly symmetric i.e., in this one causes some of the slight to go through when intending for orthogonal pre-and post-selection . The slight arc that we see over the rotation of the HWP is unlikely to be just by the pointing fluctuation. It is rather caused by the angular deviation of the beam as we rotate[28] the HWP.

2.4 Weak Measurement of the σ_x observable

Since the beam displacer, oriented along $\theta_z = 0, \pi/2$ can be thought of as a means to σ_z observable to the pointer state, as discussed in the subsection 2.3.2, for any other θ_z it can also be thought of as the hermitian unitary observable distinguishing the polarization component along θ_z and $\theta_z \pm \pi/2$. For example, if we orient the beam displacer with $\theta_z = \pi/4$, we have the two eigenstates for the birefringent crystal as $|-\rangle$ and $|+\rangle$ (See expression 2.99). Thus, with is orientation of the beam displacer the observable[29] can be denoted as σ_x (or equivalently Π_+ or Π_-). However, this would produce the beam deviation along the diagonal direction instead of along $\hat{x}$ or $\hat{y}$. This is usually not a problem when we are using a beam profiler / camera as we can measure the pointer shift along the diagonal. But in experiments with single photons discussed later in this chapter and in Ch. 5, we may need to move the coupler/ detector in order reconstruct the beam profile and eventually measure the shift. Doing the scan diagonal to the table is cumbersome, let alone

[28] The effect still occurs even wen the HWP is back aligned. It is much more pronounced with the use of polarizers. This is the reason why we prefer HWP rotation in order to change pre-selection (or later post-selection) compared to directly rotating the polarizer given that we are mostly dealing with linear polarizations for pre- and post selections.

[29] This is analogous to the rotation we have to perform on the Stern Gerlach apparatus to convert from a σ_z measurement to that of a σ_x component. However, due to the mapping of $SU(2)$ to $SO(3)$, the angles by which we need to turn is different. For polarization case, we need rotation of $\pi/4$ for the crystal, where the Stern Gerlach apparatus need to be rotated by 90° so that σ_z gets converted to σ_x.

other complexity arising due to the use of cylindrical lenses, the need for which shall be more clear in Ch. 5. For now, let's note that by rotating the crystal we can transform the σ_z observable to σ_x observable, but we are not going to transform the linearly polarized eigenstates to the circularly polarization just by mere rotation of θ_z. Hence, we cannot measure σ_y polarization component by changing the orientation of the crystal [30].

2.4.1 Using waveplates to transform σ_z interaction to σ_x interaction

The solution is then to transform the observable σ_z into σ_x by using unitary evolution before and after. Since, circular polarization is not involved the unitary transformation needed would be spanned by the use of half wave plates only. For example, when we want to perform σ_x projective measurement say for quantum state tomography , we use a half wave plate oriented at $\pi/8$ before the polarizing beam splitter PBS. The *PBS* projects the polarization into $|H\rangle$ or $|V\rangle$. The HWP before that transforms the state diagonal state $|+\rangle$ into horizontal polarization $|H\rangle$ and similarly transforms the anti-diagonal polarization state $|-\rangle$ to the vertical polarization state $|V\rangle$. Thus, any state at time t_i that will always get transmitted through the PBS is $|+\rangle$. Similary the state $|-\rangle$ is guaranteed to be reflected through the PBS. In this sense, the over all composite device made up on the HWP and the PBS helps us distinguish the two eigenstates of the σ_x observable. However, neither of the two possible projected states $\{|H\rangle, |V\rangle\}$ at time t_d are the eigenstates of the observable σ_x. This is contrary to what we presume from the postulates of quantum mechanics that after the measurement the state would be reduced to one of the eigen states . But from a practical information point of view, this is not catastrophic. We can easily place another two half wave plates after the PBS in the transmitting and reflecting ports so that the state before detection at time t_d becomes $|+\rangle$ and $|-\rangle$ respectively. But that would not bring any change in the probability of detection in the transmitting or reflecting port of the PBS and hence practically not useful. In case of weak measurements, the orthogonal polarization components from the beam displacer undergoes post-selection after which they can interfere and hence we cannot avoid the half-wave plate after the polarizing beam displacer.

[30] This is unlike the Stern Gerlach setup which in principle can be rotated along $\hat{y}$ to measure the σ_y spin component

Here, we should strictly distinguish between the σ_x (or σ_z) observable as operator associated with the measurement and the evolution operator σ_x (or σ_z). The distinction is important and identifying the correct evolution operator helps us in determining the alignment procedure. If we presume the beam displacer to be the σ_z evolution operator, we may attempt to find what operators before and after the birefringent crystal gives us the overall operator to be σ_x. Let us consider the two half waveplates sandwiching the birefringent crystal are oriented at angles θ_1 and θ_2 respectively. Then overall evolution operator becomes as the following:

$$\underbrace{\begin{pmatrix} \cos(2\theta_1) & \sin(2\theta_1) \\ \sin(2\theta_1) & -\cos(2\theta_1) \end{pmatrix}}_{\mathrm{HWP}_1(\theta_1)} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_x} \underbrace{\begin{pmatrix} \cos(2\theta_2) & \sin(2\theta_2) \\ \sin(2\theta_2) & -\cos(2\theta_2) \end{pmatrix}}_{\mathrm{HWP}_2(\theta_2)} = \begin{pmatrix} \cos(2(\theta_1+\theta_2)) & \sin(2(\theta_1+\theta_2)) \\ \sin(2(\theta_1+\theta_2)) & -\cos(2(\theta_1+\theta_2)) \end{pmatrix} \tag{2.109}$$

Thus, if we incorrectly presume to the action of the beam displacer as the evolution operator, then we would incorrectly conclude that any pair of $\theta_1 + \theta_2 = \pi/4$ would transform the σ_z observable to the σ_x observable. This is incorrect as would be shown by the following example. When $\theta_1 = 0$, only the $|H\rangle$ and $|V\rangle$ components gets spatially separated. If $\theta_2 = \pi/4$, then it would only swap $|H\rangle$ and $|V\rangle$. Thus, we would only be able to distinguish $|H\rangle$ and $|V\rangle$ completely and hence the observable would still be σ_z. Thus, the only allowed transformations of σ_z observable to the σ_x observable should happen with[31] $\theta_1 = \theta_2 = \pi/8$. Only at these angles, the $|+\rangle$ and $|-\rangle$ are eigenstates i.e., they get transformed to $|H\rangle$ or $|V\rangle$ and after the emergence from the birefringent crystal get transformed back to $|+\rangle$ and $|-\rangle$.

So what exactly should the evolution operator for the components in the weak measurement limit should be ? This can be answered if we presume the limit $\gamma = 0$. Then the beam displacer does not distinguish between any pair and hence act as the identity evolution operator $\mathbb{1}$. The two half wave plates before and after should also form identity operation so as not to affect the post-selection probability for any given pre-selected state. This is also

[31] Under the transformation, $\theta \to \pi - \theta$ or $\theta \to -\theta$, the Jones matrix of the HWP remains the same upto additional global phase in the evolution operator. Thus the angles of the two half waveplates can be set after these transformations as well when convenient

satisfied with $\theta_1 = \theta_2 = \pi/8$. Thus, while choosing the angle of waveplates, we ensure the overall evolution operator to be unitary and the same time ensure that the transformations (the two waveplates along with the beam displacer)creates a spatial separation between eigenstates of the observable.

We expect that upon insertion of the weak coupling with σ_x, the pre- and post-selection probability would remain unchanged as it should be for any weak interaction. However, it was observed that the post-selected probability was quite sensitive to the alignment of the beam displacer (see Fig. 2.17). The post-selected intensity is indirectly inferred from the inverse of the exposure time needed for a particular image to reach a threshold ADC peak value[32].

First, the two lenses were aligned in confocal configuration. Then the Glan-Thompson (GT) and the polarizing beam splitter (PBS) were placed. Then the two half-wave plates HWP3 and HWP4 were placed to form the identity operation and were kept at $\pi/8$ angle so that they transform the eigenstates of σ_x observable to the eigenstates of σ_z observable and vice-versa. Then the polarizing beam displacer crystal PBD was placed. The post-selected intensity was observed to vary while we were inserting the beam displacer.

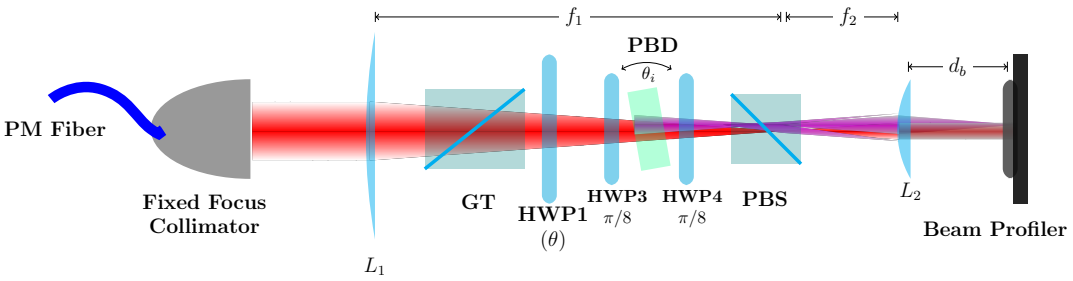


Figure 2.16: The HWP3 and HWP4 oriented at angle $\pi/8$, along with the polarizing beam displacer constitutes the σ_x weak interaction. The polarizing beam displacer was later mounted on a base rotation stage PR01/M, Thorlabs and the pointer shift was observed as a function of the tilt θ_i

Note that the tilt angle of the crystal is in the plane in which the beam displacement occurs (x-z plane in this case). Any tilt in the orthogonal direction (y-z plane) do not seem to affect the post-selection intensity. The pre-selection HWP1 was inserted and as function of the tilt of the crystal, where 0 corresponds to the orientation that matches the back-reflection, the center of the gaussian fit was observed.

[32]For details, please refer to the operation manual of the Dataray beam profiler

At the pre-selection angles $(2n+1)\pi/8$, we observe that the post-selected intensity and the centroid all agree and hence there is no dependence on the tilt of the crystal. At pre-selection angles in multiples of $\pi/4$, the post-selection intensity as well as the center of the beam becomes very sensitive to the tilt of the crystal. This hints that the cause of could be related to interference of the two-eigen polarization components. When the pre-selected state if $|+\rangle$ or $|-\rangle$, the beam is either ordinary or extra-ordinary. When the pre-selection is such that the two components have almost equal amplitude then the effect of interference is maximum. What the tilt of the crystal does is to introduce a relative phase shift between the e-ray and the o-ray as they emerge the crystal. This phase difference propagates to the component which is post-selected thus resulting in interference.

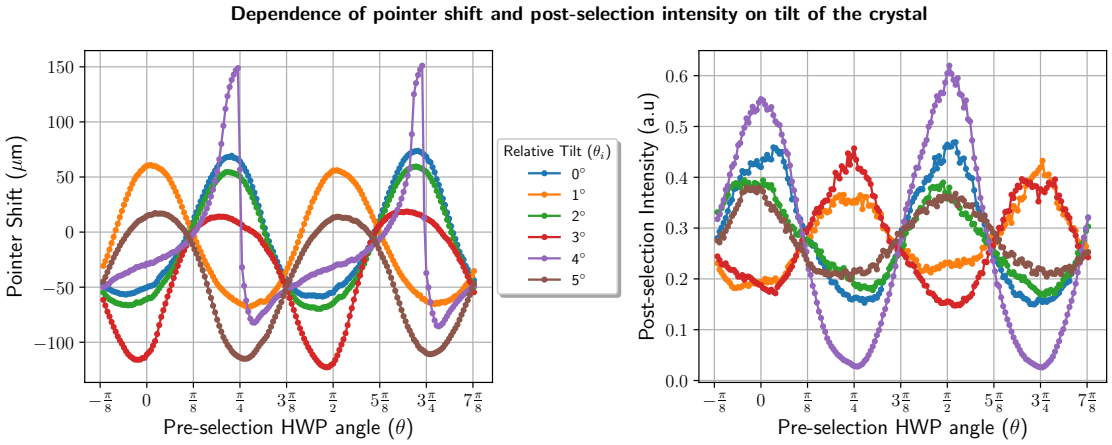


Figure 2.17: [Left] Variation of the center of the beam as the crystal is tilted. [Right] Corresponding variation in relative intensity.

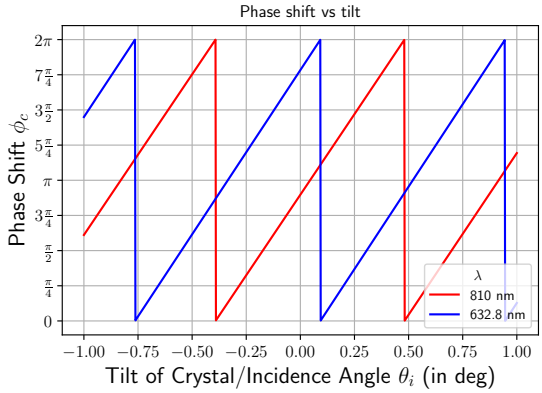
2.4.2 Phase shift in the birefringent crystal

The phase shift between the two polarization components originates because the ordinary and the extra-ordinary ray traverses different longitudinal optical path length to exit the crystal (See 2.7). The longitudinal path length difference is a function of the incident angle, besides wavelength of beam, thickness and optic axis orientation of the crystal. The phase

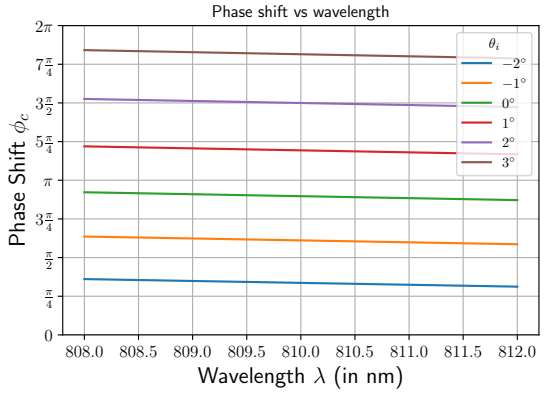
shift arising due to this, is given as follows:

$$\Delta\phi_c = \frac{2\pi L_c}{\lambda}\left(\sqrt{n_o^2 - n^2\sin^2(\theta_i)} + \frac{n(n_o^2-n_e^2)\sin(\theta_c)\cos(\theta_c)\sin(\theta_i)}{n_e^2\sin^2(\theta_c)+n_o^2\cos^2(\theta_c)}\right) + \frac{2\pi L_c}{\lambda}\left(\frac{-n_o\sqrt{n_e^2(n_e^2\sin^2(\theta_c)+n_o^2\cos^2(\theta_c)) - n_e^2 n^2\sin^2(\theta_i)}}{n_e^2\sin^2(\theta_c)+n_o^2\cos^2(\theta_c)}\right) \quad (2.110)$$

Here, n is the refractive index in air, n_o and n_e are the refractive index of the crystal. θ_c is the angle made by the optic axis from the normal. L_c is the length of the crystal which turns out to be about 500 microns. θ_i is the angle of incidence.



(a) The phase shift ϕ_c varies almost linearly with angle of incidence given the range is small (about 4 deg) with a slope of 2.296π/deg. This implies that if the crystal is tilted by $0.871°$, the phase shift will be by 2π.



(b) Dependence on wavelength is linear with a slope of $0.059/nm$. Thus, 2π phase shift occurs if the wavelength is shifted by 16.7nm. For a pulsed laser, use of bandpass filter of 3 nm wide, restricts the uncertainty in phase shift to $\pi/20$

Figure 2.18: For small tilt angles, the phase shift changes almost linearly with angle of incidence and wavelength. But the phase shift is very sensitive to the incidence angle

.

In a scenario, where there is a phase shift between the two orthogonal polarizations displaced after the crystal, what does the pointer shift correspond to ? It is tempting to treat the polarizing beam displacer as the operator $\mathbb{O}(\phi_c)$ given below instead of the σ_z observable.

$$\mathbb{O}(\phi_c) = \begin{pmatrix} 1 & 0 \\ 0 & -1\exp(i\phi_c) \end{pmatrix} \quad (2.111)$$

But $\mathbb{O}(\phi_c)$ is not a class of observables as it is not necessarily hermitian for all ϕ_c. When

sandwiched between the two half wave plates at $\pi/8$, the operator transforms to the following:

$$\begin{pmatrix} \cos(2\pi/8) & \sin(2\pi/8) \\ \sin(2\pi/8) & -\cos(\pi/8) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1\exp(i\phi_c) \end{pmatrix} \begin{pmatrix} \cos(2\pi/8) & \sin(2\pi/8) \\ \sin(2\pi/8) & -\cos(\pi/8) \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} 1-\exp(i\phi_c) & 1+\exp(i\phi_c) \\ 1+\exp(i\phi_c) & 1-\exp(i\phi_c) \end{pmatrix} \quad \text{say, } (= \mathbb{O}(\phi_c, \pi/8)) \tag{2.112}$$

The pre-selected state for angle θ of the HWP1 is given by $(\cos(2\theta), \sin(2\theta))$. When post-selected with the state $|H\rangle$, if we apply the formula to compute the weak value of the operator $\mathbb{O}(\phi_c, \pi/8)$, we obtain the following.

$$\langle \mathbb{O}(\phi_c, \pi/8) \rangle^{(w)} = \frac{\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1-\exp(i\phi_c) & 1+\exp(i\phi_c) \\ 1+\exp(i\phi_c) & 1-\exp(i\phi_c) \end{pmatrix} \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \end{pmatrix}}{\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \end{pmatrix}} \tag{2.113}$$
$$= \frac{1}{2} \frac{(\sin(2\theta) + \cos(2\theta)) + e^{i\phi_c}(\sin(2\theta) - \cos(2\theta))}{\cos(2\theta)} \tag{2.114}$$

$$\langle \sigma_x \rangle^{(w)} = \langle \mathbb{O}(\phi_c, \pi/8) \rangle^{(w)} \big]_{\phi_c=0} = \tan(2\theta) \tag{2.115}$$

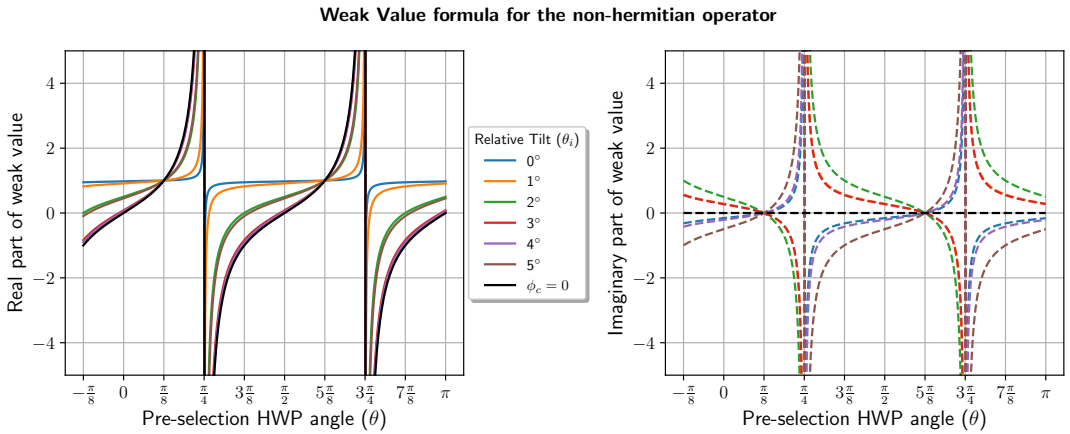


Figure 2.19: If we treat the beam displacer as the operator which introduces a relative phase shift, the real and imaginary part of the weak value varies as a function of the pre-selected state. Some features in the weak value resembles the pointer shift obtained experimentally in the figure 2.17, like at which pre-selection angle we have dependence on the tilt of the crystal and where the pointer shift is unaffected by the tilt. However, other features are not reproduced. The black curves denote the weak value of the σ_x observable was very close the tilt at 6°. This explains why, in the Fig. 2.17, the corresponding curve has the largest amplification and the post-selected intensity goes close to the minimum when pre-selected state is made orthogonal.

The fact that the weak value formula in Fig. 2.19 does not capture all the features of the experimentally obtained curves in Fig. 2.17 is not surprising as $\mathbb{O}(\phi_c, \pi/8)$ is not always an observable to begin with. But we can explicitly compute the pointer shifts by evolving the beam with consideration of transverse spatial and polarization degree of freedom.

$$|\psi_1\rangle = (\cos(2\theta)\,|H\rangle + \sin(2\theta)\,|V\rangle)\,|g(x)\rangle \tag{2.116}$$

(After pre-selection HWP oriented at angle θ)

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}\big((\cos(2\theta)+\sin(2\theta))\,|H\rangle + (\cos(2\theta)-\sin(2\theta))\,|V\rangle\big)\,|g(x)\rangle$$

(After HWP3, the wave plate before the beam displacer oriented at $\pi/8$)

(2.117)

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}\Big(\cos(2\theta)+\sin(2\theta))\,|H\rangle\, g(x-\gamma/2) + (\cos(2\theta)-\sin(2\theta))\,|V\rangle\exp(i\phi_c)\,|g(x+\gamma/2)\rangle\Big)$$
$$\begin{pmatrix}\text{After the polarizing beam displacer, assuming coupling with } \sigma_z\\ \text{observable, with beam separation } \gamma \text{ and relative phase shift } \phi_c\end{pmatrix} \tag{2.118}$$

$$|\psi_4\rangle = \frac{1}{2}\Big(\big((\cos(2\theta)+\sin(2\theta))\,g(x-\gamma/2) + (\cos(2\theta)-\sin(2\theta))\exp(i\phi_c)\,|g(x+\gamma/2)\rangle\big)\,|H\rangle + \big((\cos(2\theta)+\sin(2\theta))\,g(x-\gamma/2) - (\cos(2\theta)-\sin(2\theta))\exp(i\phi_c)\,|g(x+\gamma/2)\rangle\big)\,|V\rangle\Big)$$
$$\text{(After HWP4 oriented at } \pi/8) \tag{2.119}$$

$$|\psi_5\rangle = \frac{1}{2}\Big((\cos(2\theta)+\sin(2\theta))\,g(x-\gamma/2) + (\cos(2\theta)-\sin(2\theta))\exp(i\phi_c)\,|g(x+\gamma/2)\rangle\Big)\,|H\rangle \tag{2.120}$$
(After post-selection, only horizontal component is filtered)

The center of the beam is now given by the expectation value of the position.

$$c_x = \frac{\langle\psi_5|x|\psi_5\rangle}{\langle\psi_5|\psi_5\rangle} \tag{2.121}$$

Further calculation is done by considering the functional form of g as given in the expression 2.1. Performing the integral over x, the denominator of expression 2.121 becomes $\frac{1}{2}\left(1+\exp\left(-\frac{\gamma^2}{2\sigma^2}\right)\cos(\phi_c)\cos(4\theta)\right)$ while the numerator simplifies to $\frac{1}{2}\gamma\sin(4\theta)$. This implies that the transition amplitude is independent of ϕ_c where as the post-selection amplitude (norm of the state) is dependent on that.

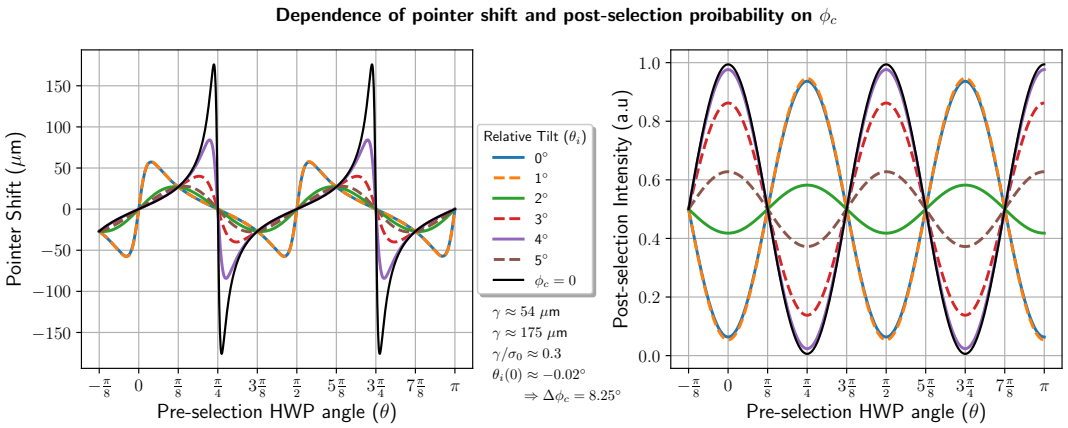


Figure 2.20: Qualitative reproduction of experimental results in Fig 2.17 using expression 2.121. For tilt of 2°, we have pointer shift showing oscillatory behavior. The tilt of 4° produces amplification. Since lens were used and the beam size was less, the pointer shift from σ_x i.e., with $\phi_c = 0$ is also not high. The post-selection intensity also oscillates with pre-selected state and tends towards zero in the region where we obtain amplification. In the above attempt, the 0° tilt was considered to be at $\theta_i = -0.02°$. This zero error in θ_i translate to about $\Delta\phi_c = 8.24°$ phase difference.

From the above, we may conclude that the pointer shift after interaction may not correspond to the weak value of any observable when there is a phase shift between the emergent rays. Unlike the case for strong measurement where the relative phase between the two eigenstates is inconsequential, in weak measurement it is essential for the eigenstates to emerge with zero relative phase shift for the realization of the weak measurement of the observable.

2.4.3 Experimental result : Weak value of σ_x observable

Now that we need to tilt the crystal in order to adjust the phase shift ϕ_c and make it 0, we avoid the multiple back reflections from the crystal. In this way, we avoid the use of lenses and the complication arising from it in terms of setting the detection distance. We follow the following procedure for alignment of the birefringent crystal to realize the weak measurement of σ_x observable.

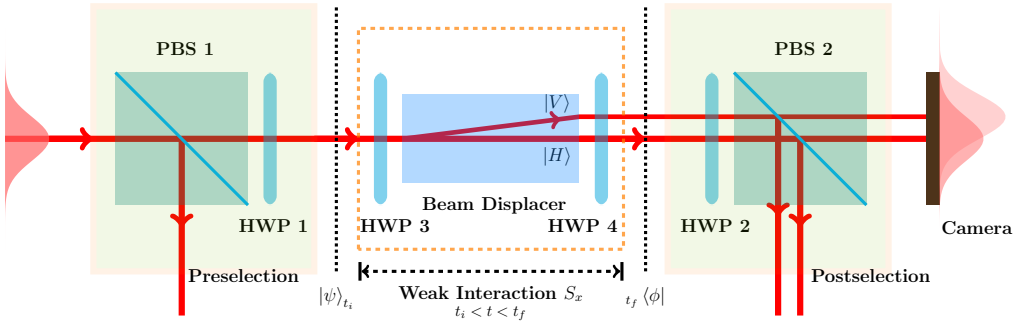


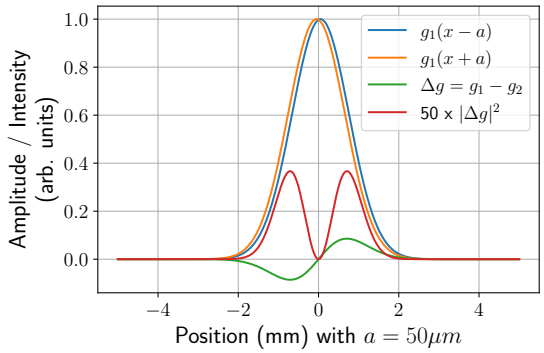
Figure 2.21: Experimental Set up to measure the weak value of σ_x

- We align the PBS 1 (or a Glan Thompson polarizer, which is preferred because of its higher extinction ratio) and the half wave plate HWP 1 to control pre-selected state. The PBS 2 for post-selection is placed along with the HWP 2 if needed. All the components are aligned so that their back reflection matches the incident beam.
- The pre-selection is made orthogonal to the post-selection by ensuring that transmission through PBS 2 is minimized. In this condition, we expect weak value amplification and the centroid is most sensitive to the phase shift between e-ray and o-ray.
- Next, we insert the two half wave plates (HWP3 and HWP4). The first one (HWP3) was inserted such that the fast axis is oriented along $\pi/8(22.5^\circ)$[33]. This is ensured when the reflected and transmitted arm of the PBS2 have same power (in absence of HWP2). The second half wave plate is placed at the same orientation ($\pi/8$) such that minimum transmission occurs through the PBS to ensure that both the waveplates (HWP3 and HWP4) form an identity operation.
- The birefringent crystal is placed and the back-reflection is matched. The azimuthal rotation of the beam displacer θ_z is adjusted along $\hat{x}$ ($\hat{y}$). The crystal is tilted in y-z (x-z) plane i.e., vertically (horizontally) so that we ensure tilt along this direction does not affect the post-selection probability. The vertical (horizontal) tilt is left where the back reflection is matched.
- The birefringent crystal is then tilted in the x-z plane (y-z plane) i.e., horizontally (vertically). This would change the post-selected intensity. We need to ensure the total evolution operator is close to identity. Thus, we adjust the tilt so that the

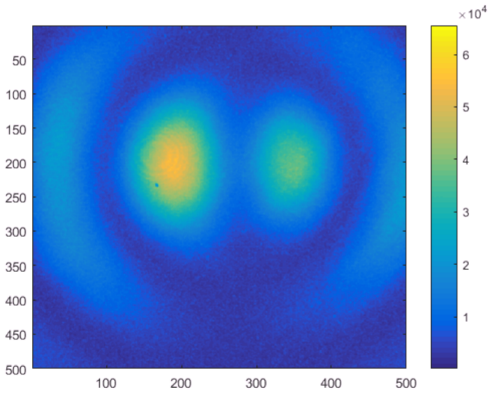
[33] In the experiments discussed in Ch 3 and Ch 4, the reference is always from the horizontal. In some of the experiments the angle was measured from the vertical.

post-selected intensity gets minimized.

- The fine adjustment of the tilt is then done by observing the symmetry of the two lobes in the destructive interference pattern (see Fig. 2.22 below).



(a) [Theory] When the two interfering components are out of phase, the two slightly displaced gaussian amplitudes result in the pattern Δg. The corresponding intensity is what we obtain as the destructive interference pattern.



(b) [Expt] The extra lobes comes from the diffraction pattern of the aperture which was used to spatially filter the beam to increase its transverse coherence.

Figure 2.22: Destructive interference pattern i.e., when pre and post selection are orthogonal, the tilt of the birefringent beam displacer crystal was adjusted so that the minimum power is transmitted as would have been the case without the beam displacer. This results in two gaussians out of phase but slightly displaced which gives rise to the above destructive interference pattern.

When pre and post-selection are orthogonal, the π phase difference between the components after post selection would occur if $\phi_c = 0$. For example, consider the case where pre-selected state is $|H\rangle$ and the post-selected state is $|V\rangle$. The first HWP (before the beam displacer) would change the state to $|D\rangle$. The crystal then separates the H and V components. But the HWP after the crystal converts the two components to $|A\rangle$ and $|A\rangle$ respectively. While projecting $|A\rangle$ to the post-selected state $|V\rangle$, we have the π phase shift with respect to the projection of $|D\rangle$. This results in the interference pattern (2d image of the beam) which appears as two lobes as obtained in the figure below.

Ensuring the symmetry between the two lobes then becomes a signature which can be used to align the birefringent crystal such that $\phi_c = 0$ and we have the desired observable.

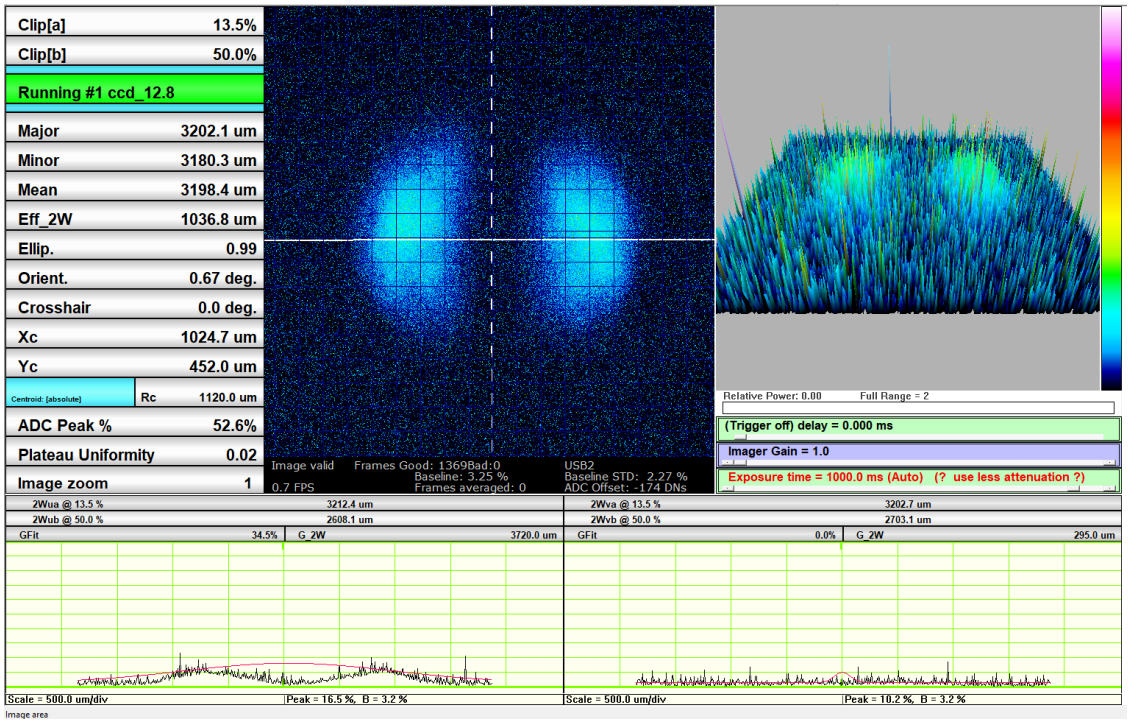
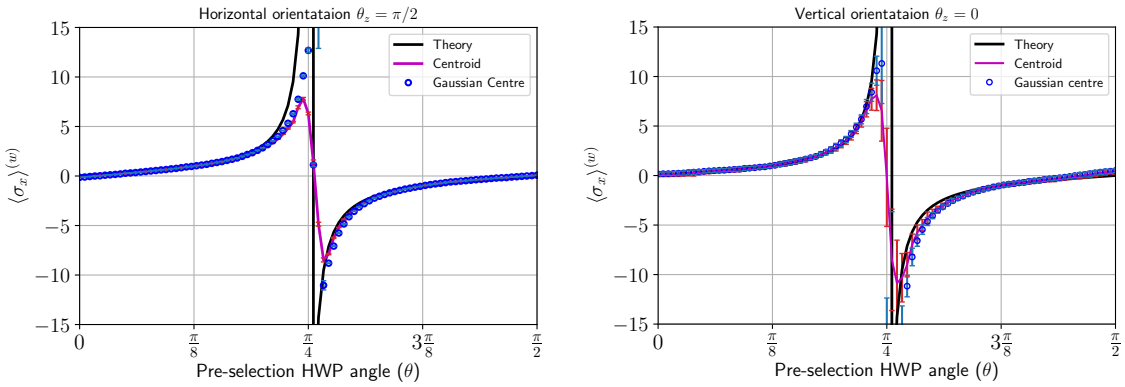


Figure 2.23: Destructive intereference pattern after manage to get better spatial mode of the beam. This is achieved by ensuring that the wavefront does not get distorted by presence of dust and all the beams are well aligned.

Initially, to avoid the regions where camera was not clean we decreased the beam size by closing the aperture. Therefore we see rings at the edges since instead of gaussian the beam is Bessel function. However, we repeated the experiment without the aperture and reproduced all the results. This tells us that the pointer shift would correspond to the weak value even if the transverse profile is not gaussian. For non-gaussian beam shape centroid is a better measure of the pointer variable rather than the centre of the gaussian fit. Later, while performing experiments on quantum cheshire cat we ensure that spatial mode of the beam is such that we do not need any aperture. The signature of spatial coherence is the beam profile as obtained in Fig. 2.23 when pre- and post-selection are orthogonal. In subsequent weak measurement of σ_x observable, apart from what is reported in this subsection, apertures were not used.

The experimental data deviates from the theory when weak value is amplified. This can be due to a finite ratio of beam size to beam displacement and also due to polarization extinction ratio. The beam has periodic walks along the vertical direction after the glan thompson polarizer and therefore the data with vertical shift had more errors.



(a) Crystal oriented to displace the beam along $\hat{x}$ (b) Crystal oriented to displace the beam along $\hat{y}$

Figure 2.24: Weak value of σ_x observable as a function of the pre-selected state for the two crystal orientation. The pointer shift was converted to the scale of weak value by scaling with $1/\gamma$. the value of γ is obtained as the distance of the pointers for the two eigenvalues, which occurs at pre-selected states with $\theta = \pi/8$ and $\theta = 3\pi/8$ respectively. For crystal oriented for horizontal displacement, we have $\gamma = 52.05 \pm 0.64 \mu m$. however, for vertical displacement, this is $44.66 \pm 3.24 \mu m$

Ideally, the zero pointer shift should correspond to the case where the beam displacer is absent. However, insertion of components do cause slight angular deviation of the beam. Even if that is taken care of by placing the components close together, the tilt of the crystal needed to ensure $\phi_c = 0$ would lead to shift of the beam upon insertion of the birefringent crystal. Further, the e-ray and o-ray gets shifted by different amount depending on the tilt. But while describing the observable σ_x, the value 0 should lie symmetrically between the two eigenvalues. For all the above reasons combined, it is not feasible to consider the pointer position without any interaction as the zero value of the observable. Thus, we transform the pointer shift to weak value after scaling by $1/\gamma$ obtained as the distance[34] between the two eigenvalues in the experiment itself.

2.4.4 Effect of tilt on beam displacement

The refraction of the o-ray as such is not an issue while determining Pauli observables in polarization degree of freedom. However, if we expect the beam to overlap with another reference, as would be the case in the quantum cheshire cat experiment discussed in Ch.3 and Ch. 4, it is important to quantify the change. Further,since angle of refraction for

[34]When rotation of HWP causes angular shift of the beam, determination of beam displacement is error prone.

the e-ray and o-ray has different dependence on the incidence angle, in principle we may apprehend the beam displacement γ to be a function of the tilt. In practice, for small angles, γ remains almost constant.

This can be verified by measuring the pointer shift corresponding to the expectation value of the σ_z observable and measuring the distance between the eigenvalues by fitting the curve with sinusoidal function. When we plot the center of e-ray and o-ray components as a function of tilt, we observe that the slope slightly deviates[35] from what is theoretically calculated using the crystal properties. However, the beam displacement between e-ray and o-ray agrees with the theory.

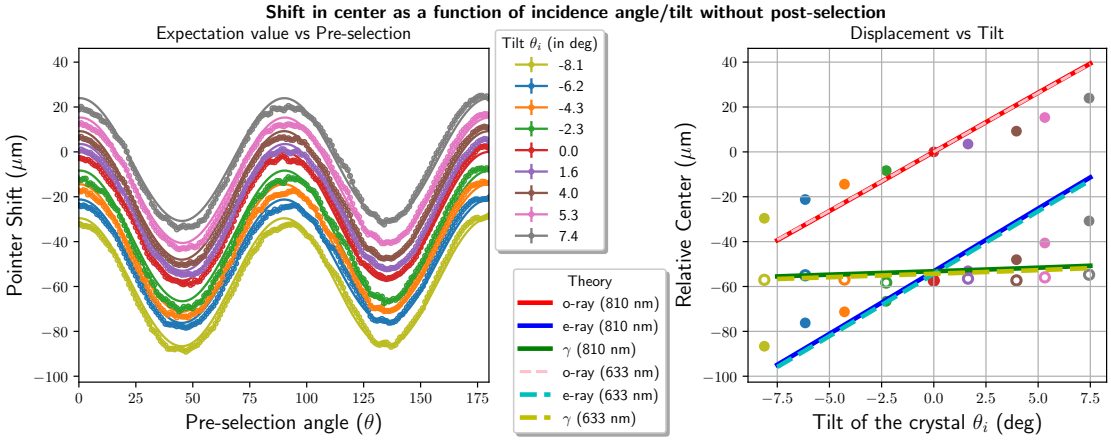


Figure 2.25: Determination of change in center of e-ray and o-ray as a function of tilt of the crystal. This is obtained by plotting and fitting the pointer shift corresponding to the expectation value of σ_z observable as function of pre-selection

We need to tilt the crystal maximum about 0.87° from the back-aligned position as that causes the ϕ_c to change by $2 * np.pi$. But in principle, we can tilt in further and settle at a larger tilt that also gives the destructive interference pattern. The pointer shift for such a case is experimentally analyzed. We measure the pointer shift for the weak measurement of σ_x observable for various tilt angles where we observe the destructive interference pattern.

[35]The tilt was measured based on the readings on base rotation mount. Parallax error may have crept into while taking the readings on the vernier scale.

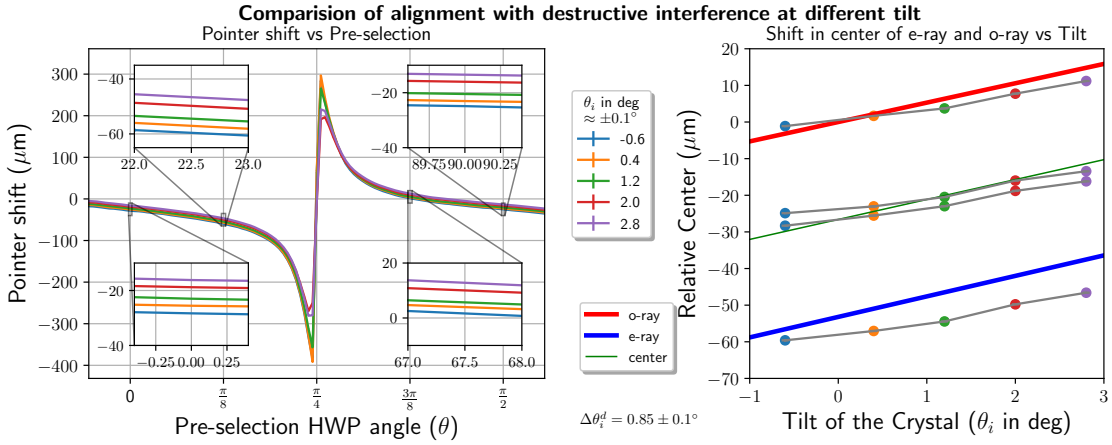


Figure 2.26: Shift of e-ray and o-ray obtained from the pointer shifts corresponding to the case when pre-selected state becomes the eigenstate of the σ_x observable. This is obtained as a function various tilt angles where we obtain destructive interference pattern.

It is determined that about $4-5\mu m$ shift can occur in e-ray or o-ray when we tilt the crystal so that we go from from destructive interference pattern to another.

.

2.5 Modifications to σ_x setup for single photon experiments

2.5.1 Effects of longitudinal delay and coherence length

Unlike a Stern-Gerlach setup where the two-orthogonal components are displaced symmetrically, in a Polarizing Beam Displacer made up of a particular cut of a uniaxial birefringent crystal the two orthogonal polarization components do not get symmetrically displaced. For normal incidence, the o-ray (ordinary polarization) gets transmitted in the same direction where as the e-ray(extraordinary ray) gets displaced. Because of this the e-ray has to travel a longer path before emerging out of the crystal. Therefore in addition to the spatial walk-off (the transverse displacement) there is also a temporal walk-off i.e., the e-ray gets delayed compared to the o-ray as shown in the Fig. 2.28a. The details on how the parameters of the beam affect the weak value are discussed in the appendix G

This is in particular not a problem for the subsequent interference upon post-selection if we are performing the experiment with a coherent source i.e. a laser with large coher-

ence length. However, if we perform the experiment with a pulsed laser source or with single photons which have finite longitudinal coherence length owing to the considerable linewidth in frequency domain, the visibility of the interference would decrease. Therefore, the pointer may not shift as much as it would, had it been a coherent combination of the two components after post-selection. This would therefore decrease the amplification expected to be obtained.

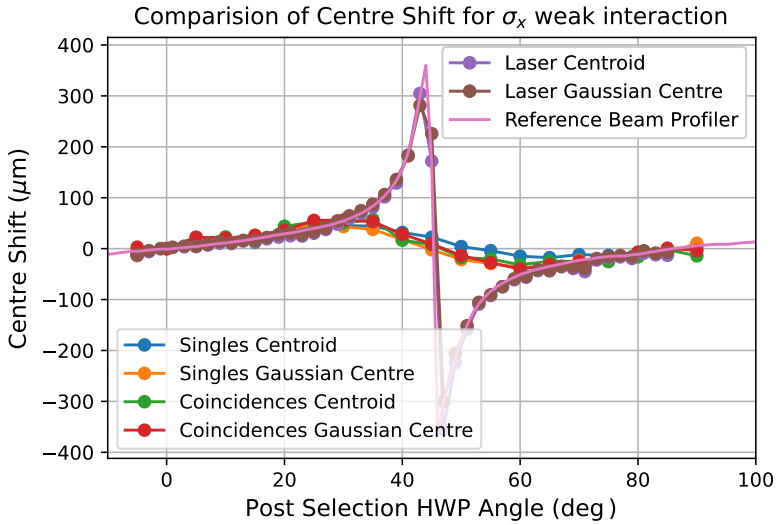
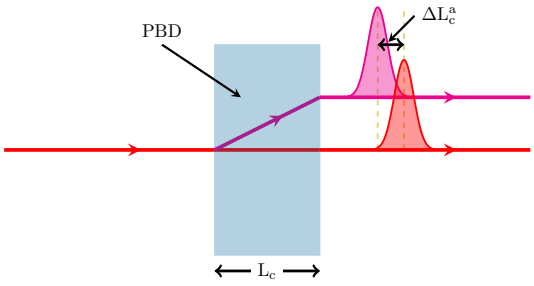
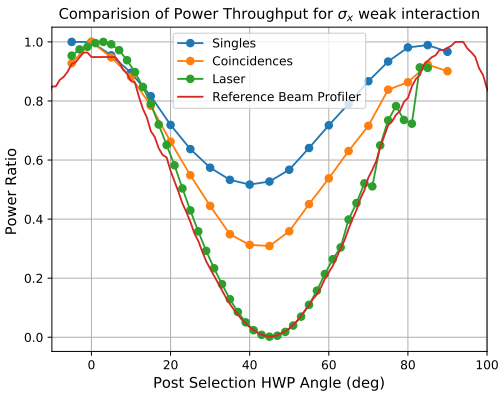


Figure 2.27: Pointer shift obtained for weak measurement of σ_x observable for various sources. With CW laser, we obtain weak value amplification. We measure the centroid as a function of post-selection HWP angle θ_{post}. We then replaced the beam profiler with a multi-mode fiber placed after a cylindrical lens (See Ch 5 for details as this forms the arm B of the interferometer in the Cheshire Cat experiment). As a function of translation of the multimode fiber, we reconstruct the profile from which center from gaussian fit and centroid were computed. They were close to the pointer shift obtained with the reference beam profiler and thus the averaging effect due to the area of the multi-mode fiber is negligible. When we obtain the centre/centroid with the multi-mode fiber but for the single photon source, both the singles and coincidence showed little amplification.



(a) Longitudinal delay between extraordinary and ordinary rays after the birefringent crystal. As the overlap between the two pulses decrease, they do not destructively interfere as much leading to the post-selected power not getting minimized when post-selection is orthogonal to pre-selection.



(b) The relative power obtained for various sources and methods described in Fig. 2.27

Figure 2.28: Effect of longitudinal delay on the post-selection probability

We can rework the expectation by treating the initial state to be a mixed but it is better to compensate the temporal delay so that the weak value obtained would be consistent with pre- and post-selection with pure states.

In the experiment, we use a calcite polarizing beam displacer PDC12005, Newlight Photonics. If the length of the crystal is L_c and the and the displacement between the emergent e-ray and o-ray upon normal incidence is $\gamma(\theta_i = 0) = a$, the physical path traveled by the e-ray inside the crystal is approximately obtained as $\sqrt{L_c^2 + a^2}$. The optical path covered therefore is $n_e(\theta_e)\sqrt{L_c^2 + a^2}$. The optical path traveled by the ordinary ray is $n_o L_c$. Therefore, the longitudinal path difference between the e-ray and o-ray is calculated to be

$$\Delta L_c^a = n_e(\theta_e)\sqrt{L_c^2 + a^2} - n_o L_c \tag{2.122}$$

At 810nm wavelength, the refractive index for the o-ray and e-ray are $n_o = 1.6485$ and $n_e = 1.4818$ respectively . The $ne(\theta_e)$ is calculated using index ellipsoid. Alternatively, we use the formula for the phase shift in expression 2.110 to compute the longitudinal delay $\Delta L_c = \frac{\Delta\phi_c \lambda}{2\pi}$. Thus, the beam displacer would cause a longitudinal delay of about 45μm. The coherence length of a source limited by the bandwidth of the 3 nm bandpass filter has a coherence length of about 110 *mum*. This roughly explains the observed power after post-selection. .

2.5.2 Compensating the loss of longitudinal coherence

To compensate this longitudinal path delay, typically another birefringent crystal is used with optic axis aligned along one of the incident polarizations. Essentially this birefringent crystal works as a long waveplate with the first axis aligned with e-polarization emerging from the beam displacer. The o-ray traveling along the slow axis allows the e-ray to catch up so that both beams emerge together. The length of such a birefringent crystal is chosen so that $(n'_o - n'_e)L'_c = L^a_c$ where, the $'$s refers to the parameters of the compensating crystal i.e, its ordinary refractive index (n'_o), the extraordinary refractive index (n'_e) and its length $L'c$. The advantages of using the birefringent crystal in such a configuration is that the polarization of the emergent beams are not effected. Further, the displacement between the polarization component remains the same even after the longitudinal path compensation.

In our experiment (see Ch 4) however, we have used another polarizing beam displacer identical to the first one to compensate for the longitudinal delay. As shown in the Fig. 2.29, we use a HWP at orientation at 45^{deg} angle from the o-ray, placed after the first beam displacer PBD_1. This changes the o-ray (say with polarization $|H\rangle$) emerging from the first crystal to be $|V\rangle$ and therefore be the e-ray for the next beam displacer PBD_2. Similarly, the e-ray for the first crystal $|V\rangle$ becomes the o-ray for the second crystal.

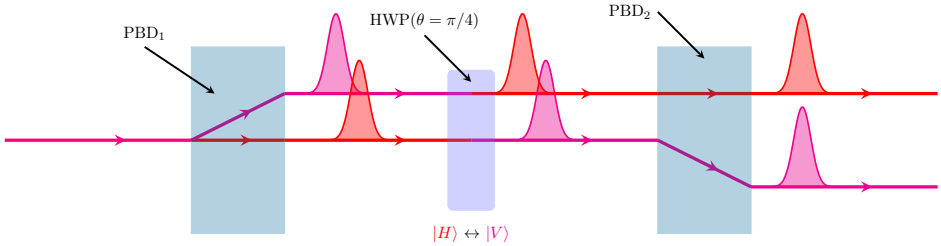


Figure 2.29: Longitudinal path compensation by addition of another crystal. The half-wave plate in the middle ensures that the displacement in the second crystal occurs in the component which was undisplaced in the first crystal. This makes compensation of the longitudinal delay, as well as makes the displacement of the two beams symmetric about the incident beam

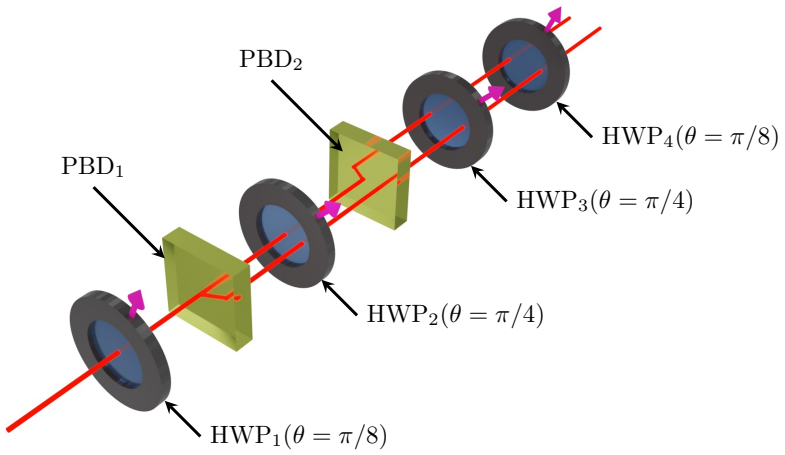


Figure 2.30: Apparatus to measure σ_x polarization component. HWP_3 is added to undo the polarization flip introduced by HWP_2. The two half-wave plates HWP_1 and HWP_4 makes the transformation from σ_z operator to σ_x operator.

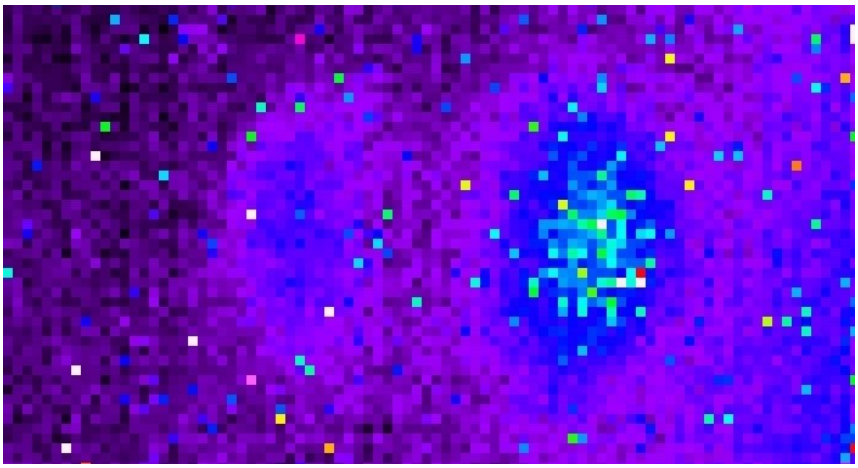


Figure 2.31: The destructive interference pattern obtained with single photon source with the setup described in Fig. 2.30. With the same setup we did obtain the pattern shown in Fig. 2.23 with the pulsed laser. The slight asymmetry in the destructive interference is due slight difference in the spectral distribution of single photon source as compared to the pulsed laser. The frequency bandwidth of femto-second pulsed laser is limited by the 3nm bandpass filter, where as the single photon source has a bandwidth comparable to the bandpass filter (Details in Ch 4)

The slight asymmetry obtained in the destructive interference pattern corresponds to small phase shift between the e-ray and o-ray as they emerge from the polarizing beam displacer. The method used to align the beam displacers is as follows:

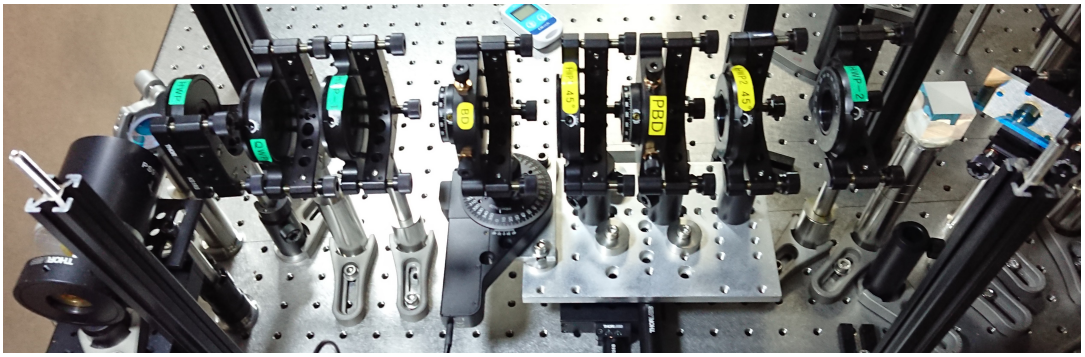
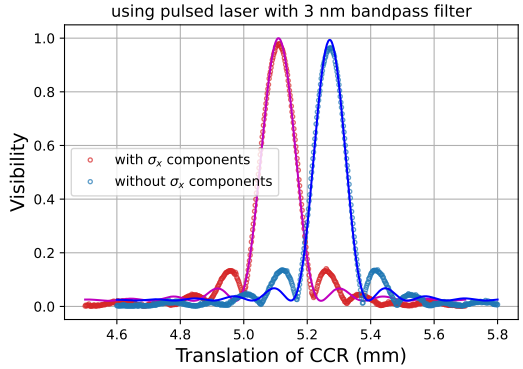
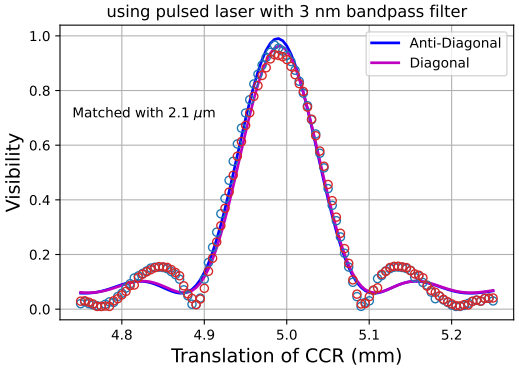


Figure 2.32: Experimental setup for σ_x weak measurement in the reflected arm (arm B) of the partial cheshire cat experimental set up (discussed in Ch. 4). The pre-selected polarization is $|V\rangle$ and the post-selected polarization is $|H\rangle$.



(a) The visibility is plotted as a function of path difference created by translation of one of the components (CCR) in a Mach Zehnder interferometer. Upon insertion of σ_x components, we observe only one peak which is shifted (instead of delay between the two polarization components).



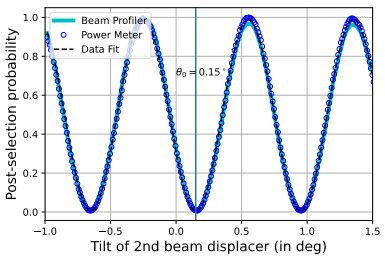
(b) We post-select the two eigenstates of the σ_x. The verify that the delay between those two components is within the uncertainty of 2μm

Figure 2.33: Verification of compensation of longitudinal delay. The emergent beam from the σ_x components is interfered with another reference beam and the visibility as a change in path difference. For details of experimental setup, see Ch 4.

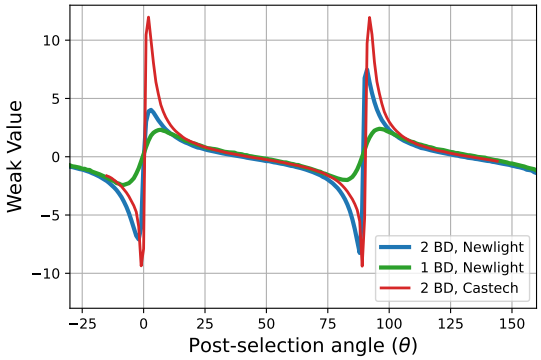
First HWP_1 and HWP_4 are inserted and it was ensured that they form the identity operator i.e., the transmission through the post-selected PBS is minimized. Then HWP_2 and HWP_3 are inserted. Then PBD_2 is mounted and is the faint-back reflection of the pulsed 810 nm laser is matched with the incident beam. Then, the beam displacer is tilted so that the transmission through post-selection PBS is minimized. Then PBD_1 is mounted on a base rotation mount and the transmitted power and centroid are obtained as a function of rotation of it. The rotation of the PBD_1 is then fixed where the transmitted power is

minimized. The power would go very close to zero i.e. upto the extinction ratio if the paths are compensated (else we need to flip the crystal or rotate it axially by 180 degrees.). This method ensures that the overall insertion of components is close to a identity operator upto the separation of beams.

Since the σ_x component is inserted in one arm of a Mach Zehnder interferometer for the partial Cheshire cat experiment with single photons (described in Ch4), we can use the other arm of it as a reference beam to verify longitudinal path delay compensation. With a pulsed source with 3 nm bandpass filter, we can plot the visibility as a function of path difference changed by scanning the CCR (See Ch 4 for detailed experimental set up). We observe a single peak in 2.33a after the σ_x components are inserted (red curve). Upon post-selection of the diagonal and anti-diagonal polarization, we observe that the maximum interference of each of the component occur at the same path delay in the interferometer. This confirms that the alignment of the two beam displacers even after the tilt of the second one ensured that longitudinal path delay between two eigenstates of the σ_x component is compensated.



(a) The first beam displacer PBD1 is aligned with matching the back-reflection. The second beam displacer PBD2 was placed on a motorized rotation mount. After post-selection, we monitor the transmitted power using beam profiler (derived from ADC peak and exposure) and directly using power meter in separate runs. The pre and post-selection were kept orthogonal and the transmitted power oscillated as a function of the tilt. The minimum power really went down to the noise floor thus indicating compensation of longitudinal path delay. The exact angle at which PBD2 was to be kept was determined by fitting the curve and finding the angle for minimum intensity closest to the back-aligned configuration.



(b) Weak value of σ_x observable as a function of post-selected HWP angle with femto-second pulsed laser and 3 nm bandpass filter. The pre-selection was kept $|H\rangle$. When only one beam displacer was there, we do not see much amplification.After the setup was changed to include both the beam displacers, we have weak value amplification. But with the use of Newligh Photonics waveplate, the fast axis and slow axis were not behaving symmetrically. Thus we have more amplification at post-selection angle 90° compared to 0°. This is solved by replacing with the waveplate from Castech, which shows similar performance along both axis.

Figure 2.34: Alignment of the tilt of beam displacer with longitudinal path compensation. This restores weak value amplification as the two beam displacers compensates the path difference created by each other.

For the laser version of the partial cheshire cat experiment discussed in Ch 3, only one beam displacer is sufficient due to the long coherence length of the laser. In Ch 4, the technique mentioned in this section is used.

References

[2.1] 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”. In: *Phys. Rev. D* 40 (6 1989), pp. 2112–2117. DOI: 10.1103/PhysRevD.40.2112.

[2.2] James M Knight and Lev Vaidman. “Weak measurement of photon polarization”. In: *Physics Letters A* 143.8 (1990), pp. 357–361. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(90)90371-T.

[2.3] N. W. M. Ritchie, J. G. Story, and Randall G. Hulet. “Realization of a measurement of a “weak value””. In: *Phys. Rev. Lett.* 66 (9 1991), pp. 1107–1110. DOI: 10.1103/PhysRevLett.66.1107.

[2.4] N.W.M Ritchie. “Measuring the weak value in an optical experiment”. In: *Master’s Thesis, Rice University* (1991). DOI: https://hdl.handle.net/1911/13508..

[2.5] Justin Dressel et al. “Colloquium: Understanding quantum weak values: Basics and applications”. In: *Rev. Mod. Phys.* 86 (1 2014), pp. 307–316. DOI: 10.1103/RevModPhys.86.307.

[2.6] Jeff S. Lundeen et al. “Direct measurement of the quantum wavefunction”. In: *Nature* 474.7350 (2011), pp. 188–191. DOI: 10.1038/nature10120.

[2.7] Onur Hosten and Paul Kwiat. “Observation of the Spin Hall Effect of Light via Weak Measurements”. In: *Science* 319.5864 (2008), pp. 787–790. DOI: 10.1126/science.1152697.

[2.8] Sacha Kocsis et al. “Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer”. In: *Science* 332.6034 (2011), pp. 1170–1173. DOI: 10.1126/science.1202218.

[2.9] John Von Neumann. *Mathematical foundations of quantum mechanics*. en. Ed. by Nicholas A Wheeler. Princeton, NJ: Princeton University Press, 2018.

[2.10] Yakir Aharonov, David Z. Albert, and Lev Vaidman. “How the result of a measurement of a component of the spin of a spin-$1/2$ particle can turn out to be 100”. In: *Phys. Rev. Lett.* 60 (14 1988), pp. 1351–1354. DOI: 10.1103/PhysRevLett.60.1351.

[2.11] Dipankar Home et al. “Aspects of nonideal Stern–Gerlach experiment and testable ramifications”. In: *Journal of Physics A* 40 (2007), pp. 13975–13988.

[2.12] Lars M. Johansen. “Weak Measurements with Arbitrary Probe States”. In: *Phys. Rev. Lett.* 93 (12 2004), p. 120402. DOI: 10.1103/PhysRevLett.93.120402.

[2.13] Luis José Salazar-Serrano et al. “Demonstration of a highly-sensitive tunable beam displacer with no use of beam deflection based on the concept of weak value amplification”. In: *Opt. Express* 23.8 (2015), pp. 10097–10102. DOI: 10.1364/OE.23.010097.

[2.14] Max Born et al. “Optics of crystals”. In: *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. 7th ed. Cambridge University Press, 1999, pp. 790–852. DOI: 10.1017/CBO9781139644181.024.

[2.15] M. Avendaño-Alejo. “Analysis of the refraction of the extraordinary ray in a plane-parallel uniaxial plate with an arbitrary orientation of the optical axis.” In: *Opt. Express* 13.7 (2005), pp. 2549–2555. DOI: 10.1364/OPEX.13.002549.

[2.16] Maximino Avendaño-Alejo and Martha Rosete-Aguilar. “Optical path difference in a plane-parallel uniaxial plate”. In: *J. Opt. Soc. Am. A* 23.4 (2006), pp. 926–932. DOI: 10.1364/JOSAA.23.000926.

Chapter 3

Quantum Cheshire Cat

Contents

3.1 The partial quantum chehsire cat effect 144

3.1.1 A Pre- & post-selection paradox 144

3.1.2 Requirements for experimental demonstration 145

3.2 Optical setup to realize partial cheshire cat 149

3.2.1 Pre- and Post-selection 149

3.2.2 Cheshire Cat Effect & Observables 155

3.3 Weak measurements for the Cheshire cat effect 163

3.3.1 Need for weak interaction 163

3.3.2 Weak measurement of path projectors 164

3.3.3 Null weak values and pointer shift 166

3.3.4 The sum rule for 'which-path' observable 169

3.3.5 Polarization component in a given arm B 170

3.3.6 Polarization component in a given arm A 172

3.3.7 Polarization in the interferometer 173

3.3.8 Joint weak measurements 174

3.4 Has quantum Cheshire cat already been experimentally demonstrated ? . . 176

3.4.1 Literature on demonstration of cheshire cat effect 176

3.4.2 Use of absorbers 177

3.4.3 Use of polarization/spin-rotators 179

3.4.4 Distinction between evolution operators and measurement operators that couples the pointer states to the observable 180

3.4.5 Inference of weak value from intensity 181
3.5 Experimental Implementation with proper weak measurements 185
3.5.1 Experimental Set up 186
3.5.2 Methods 188
3.5.3 Error Analysis 190
3.6 Experimental Results & Analysis 191

3.1 The partial quantum chehsire cat effect

3.1.1 A Pre- & post-selection paradox

Treating photon either as a particle or wave has led to many counter-intuitive manifestations of wave-particle duality such as non-locality and retro-causality in delayed choice which-path experiments [3.1, 3.2, 3.3, 3.4]. When the particles are further subjected to pre-selection and post-selection, the quantum interference effects often lead us to paradoxical conclusions [3.5, 3.6, 3.7, 3.8]. These paradoxes arise as we attempt to impose our understanding of the classical world onto the realm of quantum mechanics. One such notion that we intuitively build from observations in the classical world is that the properties of the particle are localized to the particle object. When we ask what is the color of the ball, the color is describing some property where the ball is located. It seems absurd to ask about if the color of the ball can be present somewhere in a spatial region in which the ball is not physically present. Such concepts, however, where the property of the object is disembodied from the object itself has found its way into the fiction. The Cheshire cat is such a popular fictional character which can have its grin left in a location without the rest of the body [3.9]. A cat may or may not have a grin. But what does it even mean to conceptualize the grin without the cat? Such questions may only serve our imaginations and be irrelevant towards understanding the physical world. Or so does it seem in the classical world.

But quantum mechanics, when subjected to boundary conditions in time, presents us with several pre- and post-selection paradoxes demonstrating either contextuality or the need to consider the effect of measurement disturbance for non-contextual hidden

variable[3.10]. The Quantum Cheshire Cat effect is one such pre- and post-selection paradox where a quantum particle within an interferometer subjected to pre-selection and post-selection exhibits its properties in the spatial region where the particle is absent, i.e, the particle seemingly takes one path of the interferometer while one can have signatures of some of its properties in the other arm of the interferometer. For example, a polarization component of the photon or a spin component of a neutron, which are properties of the particle, can give us a signature of interaction with the measuring apparatus in the arm of an interferometer in which the particle do not seem to affect the measurement. This occurs only within the intermediate time in between which the particle is subjected to pre- and post-selection, and thus the properties are in no way separated from the object permanently. Further, the signature of some component of the polarization or spin can be found in the arm which the particle takes - a portion of the grin stays with the cat. Thus, the phenomenon is called the quantum partial cheshire cat effect, which provokes fundamental questions about fundamental association of properties with particle description in the foundational aspects of quantum mechanics.

3.1.2 Requirements for experimental demonstration

Experimental realization of such pre-and post-selection paradoxes requires that we should be able to reasonably apply all the conclusions drawn from results of different measurements to an individual quantum particle, even though the experiment is often always performed with an ensemble of identically prepared particles. Further, since we are essentially required to probe intermediate properties of a particle subject to state preparation and post-selection, we cannot perform interactions that disturbs the particle. We therefore need to perform the experiment with minimally disturbing[1] probes, i.e, we have to implement weak measurements. Since we have to measure the two observables, 'which-arm' of the interferometer the particle is in and a particular polarization component in the arm where the particle is expected not to be taking, we need at least two measurements. The result of the two measurements need to be simultaneously applied and hence in the due course of measurement we should not create two ensembles which can be distinguished in any way.

[1] Theoretically, we can take the limit of strength of interaction towards zero. But in experiments, for any given interaction strength there would be some disturbance.

For example, consider a diagonally polarized photon incident of a polarizing beam splitter (PBS). We can measure two properties after the interaction. We can ask, 'which-port' of the PBS the particle went to and also what is its polarization. To measure the 'which-port' we can place a detector say in the reflecting arm of the PBS. About half of the time, we would detect the photon there. If we place a polarizer with the transmission axis oriented along the horizontal in the transmitting arm of the PBS, we can claim that half of the time the polarization of the photon after the PBS is along the horizontal. Now, we cannot claim that these two results would simultaneously apply, i.e., it would be wrong to imply that the particle is found in the reflecting port but the polarization of the same photon in the transmitting arm is found to be horizontal. This is because, the measurement reduces the state into two distinguishable, and thus, mutually exclusive ensembles. When we detect the polarization in the transmitting arm of the PBS, we also always find the photon in that path. Measurement of polarization in the given arm also gives information about 'which-port' the particle went to.

Therefore, to realize a partial cheshire cat in a pre and post-selected system, we need intermediate measurements which would not allow us in any way to segregate individual particles into distinguishable ensembles based on the outcome. Hence, we need to perform weak measurements. Additionally, we would require the measurement of one observable should not give us any information about another. This can be ensured only when we have two different pointer variables coupled to each observable and in no way the two pointers are themselves coupled. If the pointer variables belong to the same degree of freedom[2], then they need to be orthogonal.

But is it not easier to have two runs[3] of the experiments with weak measurement of one of the observable, say the 'which-path' observable, in one run and the other observable i.e., polarization in a given path in another run? Yes, but then by the design of the experiment we are having two ensembles. For one of them we know the weak value of the 'which-path' observable. For the other, we know the weak value of the 'polarization component

[2]If the pointer variables belong to two separate degrees of freedom or two different particles entirely, then they would be independent by default unless explicitly coupled otherwise

[3]One run of the experiment implies the case in which no experimental setting in between the pre and post selection changes. Here, we do not imply that the experiment is performed on only one particle. We consider and identically prepared and post-selected ensemble of particles with all the settings in the experiments (between pre and post-selection) remaining the same as one run of the experiment.

in a given path'. Can we apply the conclusions drawn from each run of the experiment simultaneously? The skeptic side in us would definitely argue that strictly speaking, there is no photon common to both the runs and hence on no particular photon both the observations were made. But inspired by the developments in the theory of weak measurements which respects time symmetry we may jump in to argue that implementing weak measurements would minimally disturb the system and hence even if we make both the observations on the same ensemble in a given run, the intermediate properties inferred from the weak value would remain the same as obtained in two separate runs of the experiment. The skeptic would then quickly point that what we have resorted to is counterfactual reasoning. We have not made the joint measurement of the two observables on any single particle but claiming the simultaneous applicability of result of two separate runs of the experiment. That is the recipe to arrive at paradoxical conclusions in quantum mechanics. But the problem arises with combining the results of strong measurements only, we may argue. With weak measurements, it should be fine. Or would it be?

Let's consider a system pre-selected in the state $|\psi\rangle$ and post-selected in the state $|\phi\rangle$. We aim to jointly observe the set of observables $\{A_1, A_2, \ldots A_n\}$ which would coupled to the independent pointer states $\{|\chi_1\rangle, |\chi_2\rangle, \ldots |\chi_n\rangle\}$ respectively. For simplicity, lets consider that the pointers are shifted by s_i for the observables A_i. The displacement s_i in the limit of $s_i << \sigma_i$, where σ_i is the standard deviation representing the uncertainty corresponding to the pointer state $|\chi_i\rangle$, would be proportional[4] to the weak value $a_i^{(w)} = \frac{\langle\phi|A_i|\psi\rangle}{\langle\phi|\psi\rangle}$. Presuming all the pointer states would be orthonormal in the same degree of freedom, the net displacement of the joint pointer state in n dimensions could be $\bar{s} = \sqrt{\sum_{i=1}^{n} s_i^2} \approx \mathcal{O}(\sqrt{n})$. If $\frac{\sigma_i}{s_i} \approx m \ \forall \ i$, then we require $\sqrt{n} << m$ in order to be in the weak interaction regime. Thus, clearly, there is some limit to how many observables can be simultaneously weakly measured given a finite value of m in the experiment. Therefore, it would be non-trivial to claim whether the results of simultaneous multiple weak measurements can be always be counter-factually obtained by combining the result of each weak measurement. Therefore, the demonstration of partial cheshire cat experiment need to explicitly perform the joint measurement and verify that weak values, up to the experimental errors remains the same when the two observables are jointly measured or

[4]For simplicity, we can assume that the weak values turn out to be all real.

measured in two separate runs of the experiment.

We may however perform a quantitative analysis of the pointer shift and determine the upper bound of error in the prediction of the counterfactual pointer shift due to joint weak measurements. But still, we would run into the following conceptual bottleneck: To introduce the coupling of observable A_1 to the pointer $|\chi_1\rangle$ we would have suitably arranged components in the experiment to realize the weak interaction. This defines one context for the experiment. When we wish to determine the weak value of A_2 in a separate run of the experiment, we have to modify the setup to realize the appropriate coupling with $|\chi_2\rangle$. This would imply a different context for the experiment. Whether the weak values merely reveal the physical reality associated with a pre and post-selected particle or it depends on the context of the experiment is debatable. While it has been argued that anomalous weak values or weak values in case of pre- and post-selection paradoxes signify contextuality at some level [5] [3.11, 3.12] , it has also been argued that the weak value is a physical property of a single particle much like the eigenvalue [3.13, 3.14]. Therefore, considering the aspects of counterfactual reasoning, contextuality and ability of combine the results of weak measurements in separate runs, a loophole free demonstration of the quantum cheshire cat would require joint weak measurement of 'which-path' the particle took and 'property in a given path' observables in a single run of the experiment, i.e. without changing the experimental context between pre- and post-selection.

Although Quantum Cheshire Cat is claimed to have been experimentally realized by few groups, unambiguous demonstration of the effect by obtaining weak values of both the observables, in a single run of the experiment with an ensemble of single photons, from the simultaneous shift of the pointers due to joint weak measurement has not been reported [3.15]. In this chapter, we discuss the experimental realization of partial cheshire cat effect by performing weak measurement of the relevant observables one by one. We shall then perform joint weak measurement of a pair of the observables and compare the weak values with what was obtained in the case where one of the observable was measured. For the experiment discussed in this chapter, we have used a laser source. To verify the applicability of the weak values obtained here to an ensemble of single photons,

[5] level of hidden variables or at the level of quantum mechanics

we performed another experiment with single photon source[6] and measured joint weak value for one pair of observables, the details of which are discussed in the next chapter.

3.2 Optical setup to realize partial cheshire cat

For demonstration of Quantum Partial Cheshire Cat effect, the 'which-path' degree of freedom associated with a photon in a Mach-Zehnder interferometer shall serve the purpose of the cat and the σ_x linear component in the polarization degree of freedom shall be the grin. We shall have to obtain signature of σ_x component of the polarization in one of the arm of the Mach-Zehnder interferometer where there is evidence of photon being absent. This can be achieved for particles subjected to pre-selection and post-selection in an interferometer, as described next.

3.2.1 Pre- and Post-selection

The first step towards realizing a quantum cheshire cat is to ensure that the photon can take only one of the two possible path in the interferometer when conditioned to pre-and post-selection. One way to implement the same is to prepare the state such that the polarization of the photon (or any other degree of freedom) in the arms of the interferometer are orthogonal to each other. In that case, we can filter only one of the polarization during post selection, thus ensuring that subjected to pre- and post-selection the photon takes only one path of the interferometer.

Mach-Zehnder Interferometer Consider a beam of horizontally polarized light. This can be achieved by passing the beam through a polarizing beam splitter PBS_{pre}. The polarization of photon which is part of that beam at time t_s, see Fig. 3.1, would be $|H\rangle$. The Mach-Zehnder interferometer consists of the two (non-polarizing) beam splitters BS_1 and BS_2 and the two mirrors M_R and M_T. Let's name the two arms of the interferometer as arm A and arm B. The arm, which is the transmitting port of the first beam splitter BS_1 is labeled as arm A (also aliased as arm R). Similarly, the arm of the Mach-Zehnder interferometer which lies in the reflecting port of the beam splitter BS_1 is named arm B

[6] heralded stream of photon generated via spontaneous parametric down conversion

(also aliased as arm R). When horizontally polarized light is made incident on the beam splitter BS_1, the polarization in the two output ports, right after the beam splitter, is also horizontal. In order to make the polarization in one arm orthogonal to the other, we insert a half-wave plate (HWP) in the arm B in a configuration that transforms the horizontally polarized beam into a vertically polarized beam. This occurs when the fast axis of the HWP makes an angle $\pi/4$ with respect to the horizontal. This makes the polarization in arm A and arm B to be orthogonal. If we filter only horizontally polarized light during post-selection, we can ensure that all the light that are pre- and post-selected indeed could only go through arm A. But what would be the pre-selected state? For that, along with describing 'which-path' and polarization degree of freedom, we also should describe the relative phase shift between the two paths.

The reflection mode[7] in a beam splitter, in general, can have phase shift relative to the transmission mode. Two of the common ways to represent the unitary transformation of the modes is as follows:

$$BS_{sym} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \qquad BS_{asym} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{3.1}$$

BS_{sym} represents the operator transformation in which phase shift upon reflection is the same $\pi/2$ irrespective of which port the light is incident from. For instance, consider the input port of the BS_1 on which light is being incident as $\hat{m}_s$ and the empty incident port as $\hat{m}_{s'}$. With the symmetric beam splitter, these two input modes would get transformed to the two output modes $\hat{m}_a$ and $\hat{m}_b$ as follows:

$$\hat{m}_s \xrightarrow{BS_{sym}} \frac{1}{\sqrt{2}}(\hat{m}_a + i\,\hat{m}_b) \tag{3.2}$$

$$\hat{m}_{s'} \xrightarrow{BS_{sym}} \frac{1}{\sqrt{2}}(i\,\hat{m}_a + \hat{m}_b) \tag{3.3}$$

Here, $\hat{m}_a$ and $\hat{m}_b$ represents the creation operator for the output modes of BS_1 corresponding to arm A and arm B respectively. For the second beam splitter BS_2, these two modes behaves as the two input modes. Let the two output modes of the second beam

[7] The modes in this context would be the ports of the beam splitter. In quantum optics, we represent them creation operators, which when acting on vacuum excites the field.

splitter be labeled as $\hat{m}_d$ and $\hat{m}_{d'}$ representing the ports where we would have the detector and the port which we would ignore during post-selection respectively. The transformation at BS_2 is given by

$$\hat{m}_b \xrightarrow{BS_{sym}} \frac{1}{\sqrt{2}}(\hat{m}_d + i\,\hat{m}_{d'}) \tag{3.4}$$

$$\hat{m}_a \xrightarrow{BS_{sym}} \frac{1}{\sqrt{2}}(i\,\hat{m}_d + \hat{m}_{d'}) \tag{3.5}$$

Ignoring the polarization degree of freedom [8], the light incident on the Mach-Zehnder interferometer from the input port $\hat{s}$ of the first beam splitter, evolves such that all the light exits[9] in the port d i.e, the port of the second beam splitter BS_2 in which the detection is made.

$$\hat{m}_s \xrightarrow[\text{BS}_1]{BS_{sym}} \frac{1}{\sqrt{2}}(\hat{m}_a + i\,\hat{m}_b) \tag{3.6}$$

$$= \frac{1}{\sqrt{2}}\left(\hat{m}_a \xrightarrow[\text{BS}_2]{BS_{sym}} \frac{1}{\sqrt{2}}(i\,\hat{m}_d + \hat{m}_{d'}) + i\left(\hat{m}_b \xrightarrow[\text{BS}_2]{BS_{sym}} \frac{1}{\sqrt{2}}(\hat{m}_d + i\,\hat{m}_{d'})\right)\right) \tag{3.7}$$

$$= \frac{1}{2}(2i\hat{m}_d + (1+i^2)\hat{m}_{d'} \tag{3.8}$$

$$= \hat{m}_d \qquad \text{(ignoring global phase)} \tag{3.9}$$

If we consider an asymmetric beam splitter instead, then all the beam would exit in the port d'.

$$\hat{m}_s \xrightarrow[\text{BS}_1]{BS_{asym}} \frac{1}{\sqrt{2}}(\hat{m}_a + \hat{m}_b) \tag{3.10}$$

$$= \frac{1}{\sqrt{2}}\left(\hat{m}_a \xrightarrow[\text{BS}_2]{BS_{asym}} \frac{1}{\sqrt{2}}(-\hat{m}_d + \hat{m}_{d'}) + \left(\hat{m}_b \xrightarrow[\text{BS}_2]{BS_{asym}} \frac{1}{\sqrt{2}}(\hat{m}_d + \hat{m}_{d'})\right)\right) \tag{3.11}$$

$$= \hat{m}_{d'} \tag{3.12}$$

However, if we introduce a relative phase shift φ between arm A and arm B, we can always adjust that to ensure that all the light exit in port d. In case the two beam splitters

[8] i.e., if the HWP in arm B was not present

[9] Here, we wave ignored any other phase shift due to propagation in the two arms of the interferometer. Any phase shift in the two arms is assumed to be similar such that there is no overall relative phase shift.

are asymmetric, we need to make $\varphi = \pi$ and in case they are symmetric, we need to have $\varphi = 0$. But in general, the beam splitter, does not necessarily introduce a phase shift of π or $\pi/2$ in the reflected port when compared to the beam in the transmitted port [10]. The phase shift φ is usually arbitrary which needs to be adjusted to get the beam in the desired port. Therefore, without the loss of generality, we shall consider the beam splitter BS_1 to perform the following transformation.

$$\hat{m}_s \xrightarrow{BS_1} \frac{1}{\sqrt{2}}(\hat{m}_a + \exp\{i\varphi\}\hat{m}_b) \tag{3.13}$$

In the schematic shown in Fig. 3.1, we have shown that the phase shift is introduced or adjusted by an additional component called the phase shifter, which is essentially a glass block with adjustable thickness. In practice, adjusting effective optical path length would require tilting of the glass block which would introduce transverse beam deviation as well. Therefore, the phase shifter shown in this figure 3.1 is representational only and in the experiment we shall use a retro-reflector mounted on a piezo to achieve the same effect.

Pre-selection We consider the preselected state to be at the time t_{pre} shown in Fig. 3.1 which corresponds to a time after the polarization in arm B has been transformed from $|H\rangle$ to $|V\rangle$ by the action of the half-wave plate (HWP). The pre-selected state is given by

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|A\rangle|H\rangle + \exp(i\varphi)|B\rangle|V\rangle) \tag{3.14}$$

The pre-selected state can be regarded as the state that is prepared at time t_{pre} after the state generated from the source evolves through all the components. The process of transforming the generated state from source at t_s to the pre-selected state at time t_{pre} by the use of unitary evolution through the optical components is referred to as pre-selection. The pre-selected state is written in the polarization degree of freedom and 'which-path' degree of freedom of the light. Therefore, we would also desire to write the post-selected state in the same degree of freedom. As shown in Fig. 3.1, we would consider only those

[10] The phase shift of π upon reflection is often associated with the boundary conditions and half-silvering of beam splitters. The phase shift of $\pi/2$ comes because of frustrated total internal reflection. However, in practice, the uncertainty in size of glass prisms that forms the cube beam splitters is larger than the wavelength of light. Thus, the phase shift between the transmitted and the reflected port becomes arbitrary.

particles which end up in the camera i.e, which are transmitted through the PBS_{post} and have polarization at $|H\rangle$ at time t_d. We need to write the post-selected state in terms of the 'which-path' degree of freedom and therefore consider a time t_{post} before the light interacts with the BS_2. The beam splitter combines some component of the beams from arm A and arm B while the PBS_{post} filters the horizontal polarization component for detection in camera. This entire process is called as post-selection. So what would be the post-selected state at the time t_{post}? The post-selected state is the state which under subsequent evolution is guaranteed to be detected at the camera, i.e., the probability of detection is 1. As a corollary, any state which is orthogonal to the post-selected state is guaranteed to be not-detected. To determine the post-selected state, one easy way is to evolve the detected state at time t_d and then evolve the state back in time to t_{pre}. When the polarization state $|H\rangle$ at time t_d is evolved backwards in time through the PBS_{post} it stays the same[11].

Post-selection Next, we need to evolve the polarization state $|H\rangle$ in port d backwards through the beam splitter BS_2 using the following transformation.

$$\hat{m_d} \rightarrow \frac{1}{\sqrt{2}}\left(\hat{m_b} + \exp\left(-i\varphi'\right)\hat{m_a}\right) \tag{3.15}$$

In the above, not that a beam incident in port d would get transmitted to port b and get reflected to port a. For a phase symmetric beam splitter $\varphi' = \pi/2$ and for an asymmetric beam splitter it can be 0 or $-\pi$ depending on the orientation. However, for the purpose of interference, only the relative phase between arm A and arm B would matter and therefore, one can absorb the effect of φ' into the phase term φ considered in the description of the pre-selected state in Eqn. 3.14. Thus, we can treat $\varphi' = 0$, without any loss of generality. Considering the above the post-selected state at time t_{post} is written as the following.

$$|\phi\rangle = \frac{1}{\sqrt{2}}\left(|A\rangle + |B\rangle\right)|H\rangle \tag{3.16}$$

[11] This is because, any beam from the reflected port of the PBS_{post} is not considered because that does not form the part of the post-selected ensemble.

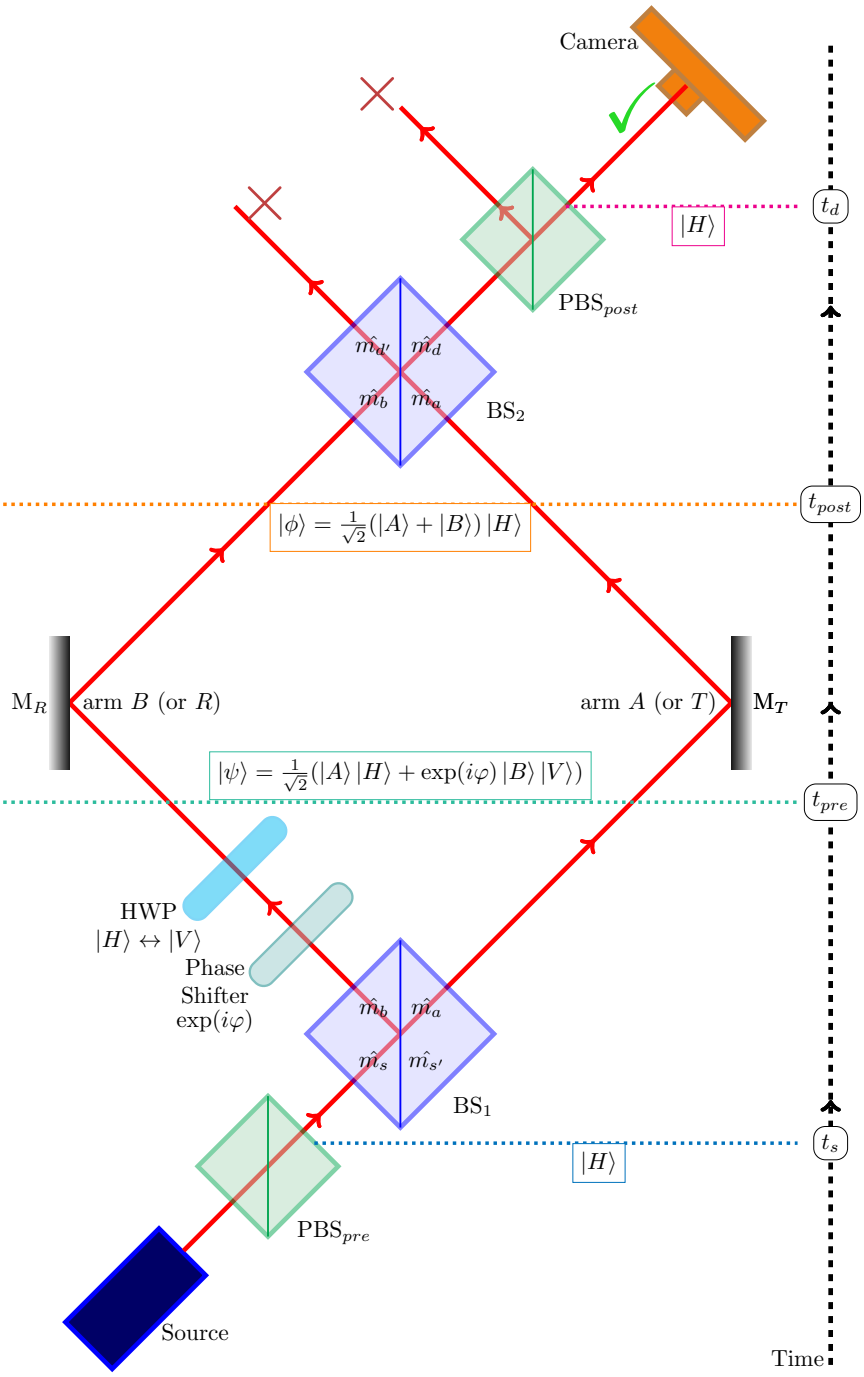


Figure 3.1: Schematic for Partial Cheshire Cat Effect: The Mach-Zehnder interferometer (MZI) constitutes the two beam splitters BS_1 and BS_2 along with the mirrors M_T and M_R in the arm A (corresponding to the transmitting port of the first beam splitter BS_1) and the arm B (corresponding to the reflecting port of the first beam splitter BS_1) respectively. The phase difference between the two arms of the MZI is controlled by the phase shifter. The PBS_{pre} along with the HWP in arm B help prepare the pre-selected state at time t_{pre}. We only consider the light which is post-selected to reach the camera placed after the PBS_{post} placed in one of the output port (port d) of the second beam splitter BS_2. The post-selected state $|\phi\rangle$ guarantees a successful detection at the camera.

Note that an equal superposition of beams from arm A and arm B is required to ensure that if this post-selected state is evolved forward in time, all the light will exit through the port d and none through port d' with the understanding that the phase-shift is taken care of. In case the splitting ratio of the beam splitter BS_2 is not 50:50, then the post-selected state needs to be appropriately modified [12]. As a corollary, we the states which are orthogonal to $|\phi\rangle$, such as $|\phi\rangle_1^{\perp} = \frac{1}{\sqrt{2}}(|A\rangle - |B\rangle)|H\rangle$ or $|\phi\rangle_2^{\perp} = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)|V\rangle$, are guaranteed not to be detected in the camera. The state $|\phi\rangle_1^{\perp}$ would evolve to exit in port d', while the the state $|\phi\rangle_2^{\perp}$ to get reflected by the PBS_{post}.

The post-selected is not necessarily a state that could be obtained by evolution of the pre-selected state. After pre-selection, we have vertically polarized state in arm B. In the post-selected state, we have horizontal polarization in both the arms. Hence, without any additional components, the pre-selected state would never evolve to the post-selected state. Thus, the pre- and post-selected state serve as two independent boundary conditions at time t_{pre} and t_{post} respectively.

3.2.2 Cheshire Cat Effect & Observables

Given a particle is subjected to pre- and post-selected state at times t_{pre} and t_{post} respectively, in which arm of the Mach-Zehnder interferometer the particle could be at any intermediate time $t_{pre} < t_i < t_{post}$? In general, the answer could be that the particle is in arm A or arm B or in a superposition state. But because of the boundary conditions in time, let us evaluate in 'which-path' the particle is ? Here, it is to be noted that the values of the considered observables are time-independent for all time $t_i \in (t_{pre}, t_{post})$, and therefore, we expect a unique answer. If the particle is in arm A then it must have polarization $|H\rangle$. Thus, about 50% of the particles in such a pre- and post-selected ensemble would go through port d of the second beam splitter, all of which would get transmitted through the PBS_{post}. So it is tempting to claim that all particle in the intermediate time t_i must have the 'which-path' state as $|A\rangle$. If the 'which-path' state was $|B\rangle$, its polarization state would be $|V\rangle$ and hence, they would not be post selected. But can the intermediate state for the 'which-path' observable be a superposition state of being in both arm A and arm B ? To

[12] Instead of $\frac{1}{\sqrt{2}}$ for both arms we need to have $\sqrt{T_2}$ and $\sqrt{R_2}$ as the coefficients where T_2 and R_2 are the transmissivity and reflectivity for the considered ports of the beam splitter BS_2

answer this question we resort to the use of the ABL rule as discussed next.

The operators of interest for the 'which-path' observable are as follows:

$$\Pi_A = |A\rangle\langle A| \qquad \text{(path A projector)} \tag{3.17}$$

$$\Pi_B = |B\rangle\langle B| \qquad \text{(path B projector)} \tag{3.18}$$

$$\Sigma_z = \Pi_A - \Pi_B \qquad \text{('which-path' unitary hermitian observable)} \tag{3.19}$$

The eigenvalues of Π_A are 1 and 0 corresponding to be the particle being in eigenstate $|A\rangle$ and $|B\rangle$ respectively. Similarly, The eigenvalues of Π_B are 1 and 0 corresponding to be the particle being in eigenstate $|B\rangle$ and $|A\rangle$ respectively. The observable Σ_z has eigenvalues 1 and -1 corresponding to the eigenstates $|A\rangle$ and $|B\rangle$.

If the value of the projector Π_A (or Π_B) is certainly[13] 1 for the intermediate time t_i, then we can claim that the particle subjected to pre-and post selection is in arm A (or B). If the probability $0 < P(\Pi_A = 1|\langle\phi|\ ||\psi\rangle) < 1$ and $0 < P(\Pi_B = 1|\langle\phi|\ ||\psi\rangle) < 1$, then we cannot comment on whether the particle is definitely in on of the paths. In general, it could be in a superposition state. Additionally, in a two path interferometer, the 'which-path' observables are dichotomic. Consequently, knowing the value of one, gives us the value of other due to the following sum rule.

$$\Pi_A + \Pi_B = \mathbb{1} \qquad \text{(sum rule)} \tag{3.20}$$

$$\Rightarrow \Pi_A = \mathbb{1} - \Pi_B \qquad \Pi_B = \mathbb{1} - \Pi_A \tag{3.21}$$

Since the projectors of path A and path B sums up to identity operator, knowing the value of Π_A(Π_B) for a given state[14] also gives us the information about the value Π_B (Π_A). Thus, if the value of the projector Π_A (or Π_B) is certainly determined to be 0 for the intermediate time t_i, then we can claim that the particle subjected to pre-and post selection is in arm B (or A) as well.

Let us now formally evaluate the probability that the intermediate value of $\Pi_A = 1$

[13] is determined with probability 1 i.e., is guaranteed to be only that value

[14] The sum rule describes the operator equation and hence is meaningful only when applied on a test function. However, when we have two-states i.e., pre- and post-selection, the relevant values would be the weak value instead of the eigenvalues. The consequences hold true for the two-state vector as well and will be discussed in the next section

when subjected to pre- and post-selection as per the ABL rule.

$$P(\Pi_A = 1 | \langle\phi| \, ||\psi\rangle) = \frac{|\langle\phi|A\rangle\langle A|\psi\rangle|^2}{|\langle\phi|A\rangle\langle A|\psi\rangle|^2 + |\langle\phi|B\rangle\langle B|\psi\rangle|^2} \tag{3.22}$$

The numerator of the above expression evaluates to $\frac{1}{4}$

$$|\langle\phi|A\rangle\langle A|\psi\rangle|^2 = \left|\frac{1}{\sqrt{2}}\langle H|(\langle A| + \langle B|)|A\rangle\langle A|\frac{1}{\sqrt{2}}(|A\rangle|H\rangle + |B\rangle|V\rangle)\right|^2 \tag{3.23}$$

$$= \frac{1}{4}(\langle H|H\rangle) = \frac{1}{4} \tag{3.24}$$

The second term in the denominator of Eqn. 3.22 evaluates to 0.

$$|\langle\phi|B\rangle\langle B|\psi\rangle|^2 = \left|\frac{1}{\sqrt{2}}\langle H|(\langle A| + \langle B|)|B\rangle\langle B|\frac{1}{\sqrt{2}}(|A\rangle|H\rangle + |B\rangle|V\rangle)\right|^2 \tag{3.25}$$

$$= \frac{1}{4}(\langle H|V\rangle) = 0 \tag{3.26}$$

Therefore, we have, the probability that the intermediate path is $|A\rangle$ is 1.

$$P(\Pi_A = 1 | \langle\phi| \, ||\psi\rangle) = \frac{1/4}{1/4+0} = 1 \tag{3.27}$$

Similarly, we evaluate the probability of the intermediate values of the following observables as per the ABL rule.

$$P(\Pi_A = 0 | \langle\phi| \, ||\psi\rangle) = 0 \tag{3.28}$$

$$P(\Pi_B = 0 | \langle\phi| \, ||\psi\rangle) = 1 \tag{3.29}$$

$$P(\Pi_B = 1 | \langle\phi| \, ||\psi\rangle) = 0 \tag{3.30}$$

$$P(\Sigma_z = 1 | \langle\phi| \, ||\psi\rangle) = 1 \tag{3.31}$$

$$P(\Sigma_z = 0 | \langle\phi| \, ||\psi\rangle) = 0 \tag{3.32}$$

All the above conditional probabilities, which are equivalent to each other, tells us that the particle subjected to pre- and post-selection is certainly in arm A and definitely not in arm B in the intermediate time. Loosely speaking, if we perform any ideal non-demolition non-invasive measurement at the intermediate time, we shall always find the particle in

arm A and never in arm B. Technically, since the pre- and post-selected states are written in $2 \times 2 = 4$ dimensional Hilbert space of 'which-path' and 'polarization' degree of freedom, we also need to represent the 'which-path' observables in 4-dimensions. Since, these observables are independent of the polarization state they can be represented as direct product of the two-dimensional observables with the identity operator in polarization degree of freedom. Thus, the observables described in Eqns. 3.17 - 3.19 needs to be modified as follows:

$$\Pi_A \otimes \mathbb{1} \qquad \Pi_B \otimes \mathbb{1} \qquad \Sigma_z \otimes \mathbb{1} \tag{3.33}$$

All the above three observables remain dichotomic in the higher dimension and hence the conclusions drawn from the use the sum rule[15] remains unchanged. However, we have to account for the fact that we have degenerate eigenstates for the eigenvalues. The eigenvalue of Π_A is 1 for the both the eigenstates $|A\rangle|H\rangle$ and $|A\rangle|B\rangle$. If we consider the intermediate eigenstate to be $|A\rangle|H\rangle$, then the ABL rule gives us the probability 1. But if we choose the intermediate state to be $|A\rangle|V\rangle$, the ABL rule gives us the probability 0. Thus, in the ABL rule, we must specify the eigenstate for which the probability is obtained.

$$P\left(\Pi_A \otimes \mathbb{1} = 1; |A\rangle|H\rangle \,\middle|\, \langle\phi| \,||\psi\rangle\right) = 1 \tag{3.34}$$

Hence, the ABL rule hints that, for the 'which-path' observable or the path projectors, we can associate an element of physical reality with the particle at intermediate time being in arm $|A\rangle$ with the polarization $|H\rangle$. As a corollary, we find the particle subjected to pre- and post-selection is always in arm A and never in arm B whenever we query for the 'which-path' observable. (See Table A.1, A.2 and A.3).

Now if we ask, what is the value of σ_x polarization component[16] in arm B subject to

[15] $\Pi_A \otimes \mathbb{1} + \Pi_B \otimes \mathbb{1} = (\Pi_A + \Pi_B) \otimes \mathbb{1} = \mathbb{1} \otimes \mathbb{1}$

[16] In general, we are looking for some polarization component in arm B. The choice of σ_x is not arbitrary, but will be clear once we compute the weak values for this polarization component. We also could aim for signatures of σ_y polarization component. Later, we shall see that the interpretation of σ_x weak value which turn out to be real becomes less complicated than the weak value of the σ_y polarization component which turns out to be imaginary number. However, at this point, the choice of σ_x could well be regarded as a guess, but with the following thought process. The polarization component in arm B for the pre-selected state ($|V\rangle$) is orthogonal to that in the post-selected state ($|H\rangle$). Thus, the eigenstates σ_z in arm B would not satisfy pre- and post-selection. Thus, we can next attempt any other polarization component. Among the Pauli observables the eigenstates of σ_x are linear and are easier to treat as compared to the eigenstates of σ_y which are circular. Hence, we start with query for the σ_x polarization component.

pre- and post-selection? This might sound absurd as we have just concluded that given the conditions the particle is never located in arm B. But, we can pose the question with rigour if we arrive at the pre- and post-selection as a limiting case. Consider a slightly modified[17] pre-selected state $|\widetilde{\psi}\rangle = |\psi\rangle + \varepsilon\,|\psi\rangle^{\perp}$ or a slightly modified post-selected state $\left|\tilde{\phi}\right\rangle = |\phi\rangle + \varepsilon\,|\phi\rangle^{\perp}$. For the two-state vector $\left\langle\tilde{\phi}\right|\ |\widetilde{\psi}\rangle$, there would be some non-zero probability of finding the particle in arm B. For those particles, we can ask what is the value of σ_x polarization component in arm B. We would eventually consider the limit $\varepsilon \to 0$ and if under this the value of σ_x polarization component converges irrespective of the direction from which we approach the limit, then formally we can assign the converging value to σ_x polarization component in arm B for the pre- and post-selected state $\langle\phi|\ |\psi\rangle$.

Note that posing the above question as we approach a limit, we run the risk of trivial analogies which are not counter-intuitive. For example, consider a stream of horizontally polarized photons, the intensity of which can be controlled by placing neutral density filters (absorbers). As a function of absorption, the polarization state of whatever fraction of photons is transmitted, remains the same horizontal. Under the limit that absorption goes to 100%, the polarization would remain constant. But if we phrase this as “existence of horizontal polarization in a region where no photon is transmitted", it makes us wonder the utility of even considering the limit. Consider another example where we avoid the use of absorption s that the evolution can be treated with unitary operators. If we linearly polarized beam is made incident on a polarizing beam splitter(PBS), a fraction of the beam would get transmitted and the polarization in the transmitting port would be horizontal. Similarly, the polarization of photons which gets reflected would be vertical. If using a half-wave plate(HWP) placed before the PBS, we make the polarization incident on the PBS to be horizontal, then ideally, none of the photons would get reflected. As a function of the HWP angle, the polarization of photons in the reflected arm would remain constant at vertical. Under the limit, the incident polarization is horizontal, the polarization of the beam in the reflecting port would trivially approach to be vertical. Would this be an example of

[17] Here, the modification to the pre- and post-selected states may seem arbitrary with the introduction of the ε parameter. Later, once we discuss the experimental setup, the parameters that modify these states would be physically motivated. Mainly, in the preselected state, the parameter φ describing the relative phase shift can be varied. The polarization in the post-selected state can be modified by placing a HWP before the PBS_{post} and after BS_2. The angle of the HWP would serve as another parameter that can be used to approach towards the desired post-selected state.

having vertical polarization in the reflecting arm of the PBS even when no photon gets reflected ? Of course not. The reason is that, if by any means, we detect the polarization of the photon in the reflecting port we would also find the photon in the reflecting port. If only, we could avoid having any information about 'which port' while asking about its polarization in the reflecting port, the question would be non-trivial. To do so, we may combine the beams in reflected and transmitted port of the PBS using a beam splitter(BS), so that any measurement of polarization after that do not necessarily project the photon in a given path. This is, in essence, what the Mach-Zehnder interferometer serves, as shown in the schematic[18] Fig. 3.1. Note that, the polarization after *BS* is no longer constant as the incident polarization is rotated. In fact, when the incident polarization approaches towards being horizontal, the polarization after the beam combination after BS would also be horizontal. Thus, we would not detect vertical polarization under the limit. However, for the setup given in the schematic Fig. 3.1, any intermediate measurement of polarization must not affect the chances for post-selection to be successful. Therefore, the value of the polarization observable that we may obtain in arm B as the pre- and post selected state approaches $|\psi\rangle$ and $|\phi\rangle$ needs more scrutiny in its interpretation.

First, let us analyze what would be the impact of projective measurement of σ_x polarization component in arm B. We need to detect the photon after successful post-selection. Hence, we may consider an ideal linear polarizer in arm B oriented along the diagonal or anti-diagonal direction as our means of inferring the intermediate polarization. If the polarization in arm B is not diagonal or anti-diagonal, the use of polarizer would affect the post-selection probability, i.e, the ensemble when polarizer is placed would be altered from the ensemble when polarizer was not inserted. Nevertheless, let us assess what we could infer from such an exercise. The operator description of σ_x polarization component in arm B is given by $\Pi_B \otimes \sigma_x$. Its eigenvectors are $\{|A\rangle|+\rangle, |A\rangle|-\rangle, |B\rangle|+\rangle, |B\rangle|-\rangle\}$, with $|A\rangle|+\rangle$ and $|A\rangle|-\rangle$ being degenerate with eigenvalue 0. For the eigenvalue 0, $|A\rangle|H\rangle$ and $|A\rangle|V\rangle$ an also be considered as the eigenvectors. $|B\rangle|+\rangle$ and $|B\rangle|-\rangle$ are distinguishable as they have eigenvalues 1 and -1 respectively. When we have a strong measuring device i.e., for instance, the polarizer oriented diagonally placed in arm B coupled to the

[18]The difference between the setup discussed here and the schematic are as follows: (a) Incident polarization being varied instead of being $|H\rangle$. (b) The PBS_{post} is absent and hence there is no polarization filtering (c) The BS_1 and HWP in arm B are replaced by the PBS in the current discussion.

system operator $\Pi_R\sigma_x$, the pre-selected state can reduce to $|B\rangle|+\rangle$ with 25% probability ($|\langle\psi||B\rangle|+\rangle\rangle|^2 = 1/4$). Once the state is reduced, then it has 25% probability of getting successfully post-selected ($|\langle\phi||B\rangle|+\rangle\rangle|^2 = 1/4$). Thus, there is non-zero (=1/16) probability that one of the σ_x polarization component in arm B can satisfy both pre and post-selection. Similarly, we can find that the component $|B\rangle|-\rangle$ can satisfy both pre- and post-selection with probability 1/16 as well. Does that mean given pre- and post-selection the particle (the cat) is located in arm A but it has, sometimes, a σ_x polarization component (the grin) in arm B ? Do we have the Cheshire cat ?

Unlike the case of 'which-path' observable, the ABL rule tells us that none of the eigenstates of the observable $\Pi_B \otimes \sigma_x$ would occur with certainty as an intermediate state (See Table A.7 in the Appendix B). The conditional probability of obtaining the intermediate state to be $|B\rangle|+\rangle$ or $|B\rangle|-\rangle$ is $1/6$ among all the possible intermediate eigenstates of the observable $\Pi_B \otimes \sigma_x$. The rest $2/3$ of the times the particle would be eigenstate corresponding to arm A i.e, $|A\rangle|H\rangle$. But whenever we obtain the intermediate state to be $|B\rangle|+\rangle$ or $|B\rangle|-\rangle$, can we not claim the evidence for diagonal polarization component in arm B even though earlier we have shown that the particle goes through only A when subjected to pre- and post-selection ? No. For the observable $\Pi_A \otimes \mathbb{1}$, we obtain that the eigenstate $|A\rangle|H\rangle$ occurs as an intermediate state with certainty. But for the observable $\Pi_B \otimes \sigma_x$, the same state $|A\rangle|H\rangle$ occurs only with probability $2/3$ and $|A\rangle|V\rangle$ never occurs[19]. Thus, the probability of the particle being in arm A changes if we attempt to make measurement of a different observable. In quantum mechanics, measurements disturb each other. The paradox of quantum cheshire cat arises if we apply the conclusions of two different measurements simultaneously. If we cannot make simultaneous measurements or upon simultaneous measurement the individually obtained results would never occur together for the same particle, then combining the results in the two scenarios is akin to counterfactual reasoning.

Combining the conclusions from two different measurements is in general not allowed in quantum mechanics. In order to imply the conclusions together, it would require us to perform the 'which-path' measurement and σ_x polarization component measurement in

[19]See Appendix B for detail table of probabilities evaluated using the ABL rule

arm B simultaneously. The 'which-path' operator Π_B[20] commutes with $\Pi_B\sigma_x$ and hence order of operation is not a concern. If we have one measuring device coupled to $\Pi_B\sigma_x$ and another coupled to Π_B, we find that whenever the state would reduce to $|B\rangle$ due to Π_B it also would then reduce to either $|B\rangle|+\rangle$ or $|B\rangle|-\rangle$ (the eigen states of $\Pi_B\sigma_x$). If we individually measure Π_B, we expect the particle to be never found in arm B (See Table A.2 in Appendix B for the probabilities computed with ABL rule). However, when $\Pi_B\sigma_x$ measurement is made simultaneously, sometimes we would obtain the particle in arm B. Thus, even though two observables commute, they still can be incompatible i.e, disturb each other when acting of ensemble subjected to pre- and post-selection [21].

The operator Π_A is incompatible [22] with $\Pi_B\sigma_x$ even though the operators may seem to commute. So if the state reduces to $|A\rangle|H\rangle$, it never reduces to $|B\rangle|+\rangle$ or $|B\rangle|-\rangle$. Thus, subjected to pre- and post-selection, either the particle is located in arm A or it has a σ_x polarization component in arm B. They would never occur simultaneously, and hence the paradoxical nature of the cheshire cat effect seems to be trivially resolved, as its origin is traced to counterfactual reasoning.

Strong projective measurements of two observables, even if commuting, disturb the system subjected to pre- and post-selection. Hence, the quantum cheshire cat effect cannot be realized with strong measurements which lures us to dismiss the paradox as a result of logical fallacy of counterfactual reasoning applied to quantum mechanics. However, with weak measurements, we can leave the pre- and post-selected ensemble undisturbed. Further, the weak values of two different observables, even if non-commuting, can be simultaneously obtained. In the next section, we shall discuss how the concepts of weak measurements can hold us back from letting the quantum cheshire cat effect go to oblivion as a counterfactual illusion.

[20] We can always infer the value of Π_A from the knowledge of the value of Π_B using the sum rule

[21] (see footnote 7 of [3.16])

[22] It may seem that $\Pi_A\Pi_B\sigma_x = \Pi_B\sigma_x\Pi_A = 0$ and hence they commute. But for commutation, we need a test function from the Hilbert space that satisfies the operator relation. Since $\Pi_A\Pi_B$ acting on any test function would give us 0 (the state can never be normalized). This implies the incompatibility of the observables i.e., there would not be an instance where we would have a projection in arm A and arm B. this trivially follows from the orthogonality of $|A\rangle$ and $|B\rangle$

3.3 Weak measurements for the Cheshire cat effect

3.3.1 Need for weak interaction

Intermediate measurements or strong interactions coupling any pointer state to the observable of interest disturbs the system subjected to pre- and post-selection. Only those observables, with which we can associate elements of physical reality as per the ABL rule, can be inferred. For example, the conditional probability for the intermediate state to be $|A\rangle|H\rangle$ is 1 for the observable $\Pi_A \otimes \mathbb{1}$. Thus, if we block the arm B, there would not be any noticeable change in the pre- and post selected ensemble. Similarly, if we place a polarizer oriented along the horizontal in arm A, there would not be any disturbance to the system. We can only infer the intermediate state to be this by the absence of any effect when we try to block or absorb the components orthogonal to the predicted intermediate state. We also can have strong coupling, i.e., say place a thick tilted glass block that causes a transverse shift of the beam in arm A. All the beam which is post-selected would get shifted by the amount we introduce in arm A only, since the beam in arm B would be filtered out by the PBS_{post}. But any other strong interaction that couples to say the σ_x polarization component in arm A would disturb the ensemble, as none of the eigenstates of the observable occurs with certainty as per the ABL rule (See Table A.4 in Appendix B). If we have a polarizing beam displacer in arm A, which displaces the diagonal polarization component in one transverse direction and the anti-diagonal component in the opposite direction, we would notice that the number of particles or the intensity of light that got successfully post-selected would change. Without the beam displacer all the light from arm A which get to port d of the second beam splitter BS_2 would get transmitted through the PBS_{post}. The introduction of the polarization beam displacer now creates two ensembles distinguishable by transverse spatial separation: one with diagonal polarization and the other with anti-diagonal polarization. This does not affect the probability of reflection towards port d as BS_2 is ideally polarization independent [23]. But now, from the diagonally polarized sub-ensemble only 50% gets transmitted through the PBS_{post}. The

[23] For strong interaction, the separation would be larger than the beam width. Hence, at BS_2, the vertically polarized beam coming from arm B would not overlap with the diagonally or anti-diagonally polarized beams coming from arm A. Thus, even though the polarization in arm B is not orthogonal to either of the polarization in arm A, no interference would occur at BS_2 because lack of transverse spatial overlap

same fraction gets transmitted from the anti-diagonally polarized sub-ensemble. Thus, overall, introduction of the polarizing beam displacer which couples the transverse spatial degree of freedom to the σ_x polarization component decreases the intensity which gets post-selected by half. We may get some answer for the value of the observable for different sub-ensembles, but the very ensemble for which we pose the question changes. Similarly, if we place the polarizing beam displacer in the arm B and separate the diagonal and anti-diagonal components, 50% of the particles in arm B which gets transmitted to port d of the second beam splitter BS_2 would now get transmitted through the PBS_{post}.

The strong interaction makes the ensemble into a mixed state of two-distinguishable sub-ensembles of the polarization components, which behaves differently compared to a coherent superposition state of two polarization components. The polarization state $|H\rangle$ or $|V\rangle$ which are equal superposition of diagonal $|+\rangle$ and anti-diagonal $|-\rangle$ gets completely transmitted or completely reflected from the PBS_{post}. But a particle from an equal mixture of $|+\rangle$ and $|-\rangle$, would get transmitted or reflected with probability $1/2$. The mixedness arises due to the in principle distinguishability of the individual component. Based on the transverse shift at the time of detection in the camera, we can conclude the polarization of each photon, whether diagonal or anti-diagonal. If we can reduce the distinguishability, we can retain the coherence between the two components and the thus avoid substantial disturbance to the system.

To minimize the disturbance, we can deploy weak interaction so that the state of the system does not reduce to the eigenstate of the observable. Under the limit that spatial separation between the two orthogonal polarization component is much smaller than their beam width, i.e., the weak interaction limit, we can extract information from the transverse spatial pointer shift resulting from the interference during post-selection. The pointer shift would correspond to the weak values for the weakly coupled observables in the given pre- and post-selected state. Now, let us compute the weak value of the observables of interest.

3.3.2 Weak measurement of path projectors

The weak value of the projector for path A in the interferometer turns out to be 1 for the given pre- and post-selected state.

$$\langle \Pi_A \otimes \mathbb{1} \rangle^{(w)} = \frac{\langle \phi | \Pi_A \mathbb{1} | \psi \rangle}{\langle \phi | \psi \rangle} \tag{3.35}$$

$$= \frac{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) | A \rangle \langle A | \otimes \mathbb{1} \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)}{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)} \tag{3.36}$$

$$= \frac{\langle H|H\rangle}{\langle H|H\rangle} = 1 \tag{3.37}$$

What does this weak value imply? In terms of pointer shift, it implies that if we weakly couple the beam in the arm A and introduced a small transverse shift, we would see the same small transverse shift in the same direction at the camera. This would indicate the presence of photons in arm A conditioned to successful pre- and post-selection. But does this guarantee the absence of photons in arm B ? To comment further on the claim, we have to resort to the fact that the weak value of 1 is the extreme non-degenerate eigenvalue of the dichotomic observable $\Pi_A \otimes \mathbb{1}$. When the weak value equals an extreme eigenvalue of a dichotomic observable, it implies that we would have obtained the corresponding eigenvalue had we performed a strong measurement with that observable (See the discussion in [3.17] and [3.18]). In other words, when the result of a weak measurement i.e., the weak value and the result of an ideal strong measurement, the eigenvalue are identical, we can associate reality with the intermediate eigenstate of the observable. The ABL rule would predict the intermediate eigenstate with certainty. From this, we can claim that the eigenstate orthogonal to what is predicted with certainty by the ABL rule would never occur. And thus, the eigenstate $|B\rangle$ corresponding to the eigenvalue 0 for the observable Π_A can be ruled out.

We can of course directly compute the weak value of $\Pi_B \otimes \mathbb{1}$ and find it to be 0.

$$\langle \Pi_B \otimes \mathbb{1} \rangle^{(w)} = \frac{\langle \phi | \Pi_B \mathbb{1} | \psi \rangle}{\langle \phi | \psi \rangle} \tag{3.38}$$

$$= \frac{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) | B \rangle \langle B | \otimes \mathbb{1} \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)}{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)} \tag{3.39}$$

$$= \frac{\langle H|V\rangle}{\langle H|H\rangle} = 0 \tag{3.40}$$

0 is also an extreme eigenvalue of the observable, Π_B which corresponds to the eigenstate

$|A\rangle$. Thus, the argument regarding the weak value being the extreme non-degenerate eigenvalue of the dichotomic shall sustain. Hence, we can infer the intermediate state of the particle to be in arm A with certainty.

But the interpretation in terms of the pointer shift needs some care. Coupling to Π_B can be achieved by having a tilted glass plate in arm B. The weak value of 0, implies that upon imparting a slight transverse shift in the beam in arm B, the transverse profile of the detected beam remains unaffected. Thus, one may be tempted to claim that, the particle in the pre- and post-selected ensemble are absent in arm B because if they were in arm B, we would have some signature of the shift. The lack of signature of interaction can often lead us to conclude the absence of particles. All we can claim from the lack of pointer shift is that we have no evidence for the pre- and post-selected particles to be present in arm B. The absence of evidence for the presence of particles is not necessarily the evidence for the absence of the particles.

3.3.3 Null weak values and pointer shift

Given the coupling with a given observable, the null weak value of that observable would correspond to no pointer shift [24]. Let us consider various cases which would give us no pointer shift so that we can analyze the reasons based on which the interpretation of the null weak value will be made. In two-dimensional Hilbert space, the weak value would be zero if an operator evolves the preselected state to a state orthogonal to the post-selected state, provided the pre- and post-selected state are not orthogonal to begin with. Given any pre pre-selected state $|\psi\rangle$ and post-selected state $|\phi\rangle$, projection operators of the form $\Pi_{\psi^\perp}$ or $\Pi_{\phi^\perp}$ where, $\langle\phi|\phi^\perp\rangle = 0$ and $\langle\psi^\perp|\psi\rangle = 0$, would give us the null weak value [25]. Although, projection onto state orthogonal to pre-selected state or post-selected state gives a weak value of 0, they are also the extreme eigenvalues of the operator. In two-dimensional Hilbert space, the eigen-spectrum associated with the null weak value of the dichotomic projector observable is non-degenerate and hence we can assign physical meaning to the

[24] We need to be careful here to analyze the exact observable with which the coupling occurs. The statement does not hold true from equivalent observables. For example, physically coupling with Π_A and Π_B are different, although they provide us equivalent information due to the sum rule $\Pi_A + \Pi_B = \mathbb{1}$

[25] One may be tempted to try operators of the form $|\phi\rangle^\perp\langle\phi|^\perp$ as the transform the pre-selected state to a state orthogonal to the post-selected state and vice versa. But such class of operators would be non-hermitian and would not be regarded as observables.

intermediate state corresponding to the 0 eigenvalue. In higher dimensions, the eigenvalue 0 of the projector is, in general, degenerate and thus, we need to additionally verify if and which of the eigenstate occurs at the intermediate time with probability 1 as computed using the ABL rule.

Let's consider two observables such that $\langle \mathcal{O}_1 \rangle^{(w)} = 1$ and $\langle \mathcal{O}_2 \rangle^{(w)} = -1$. Since weak values are additive, we would have an observable $\mathcal{O}_1 + \mathcal{O}_2$ with null weak value.

$$\langle \mathcal{O}_1 + \mathcal{O}_2 \rangle^{(w)} = \langle \mathcal{O}_1 \rangle^{(w)} + \langle \mathcal{O}_2 \rangle^{(w)} = 1 - 1 = 0 \tag{3.41}$$

The weak value of 1 can be obtained if the observable is projector onto either the preselected state or the post-selected state. The operator[26] of the form $\mathcal{O} = \Pi_\phi - \Pi_\psi = |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|$ would give us null weak value.

$$\langle \mathcal{O} \rangle^{(w)} = \frac{\langle\phi|(\Pi_\phi - \Pi_\psi)|\psi\rangle}{\langle\phi|\psi\rangle} \tag{3.42}$$

$$= \frac{\langle\phi|(|\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|)|\psi\rangle}{\langle\phi|\psi\rangle} \tag{3.43}$$

$$= \frac{\langle\phi|\psi\rangle - \langle\phi|\psi\rangle}{\langle\phi|\psi\rangle} \tag{3.44}$$

$$= 0 \tag{3.45}$$

As an example, consider the preselected state $|-\rangle$ and the post-selected state $|H\rangle$ which are not orthogonal to each other. The weak value for the observable σ_x and σ_z turn out to be -1 and 1 respectively. It implies that, we would have $\langle \sigma_x + \sigma_z \rangle^{(w)} = 0$. Now $\sigma_x = \Pi_+ - \Pi_- = \mathbb{1} - 2\Pi_-$ and $\sigma_z = \Pi_H - \Pi_V = 2\Pi_H - \mathbb{1}$. Thus, we have $\sigma_x + \sigma_z = 2(\Pi_H - \Pi_-)$ i.e, the observable of the operator form like $\mathcal{O}$.

For such observables of the form $\mathcal{O}$, 0 is not one of the eigenvalue. Then what does the null weak value imply? Does it imply that the polarization component along $\sigma_{\pi/4} = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_y)$ is absent? If we do an intermediate strong measurement with $\sigma_{\pi/4}$ allowing for pre- and post-selection, sometimes we would find the eigenvalue to be 1 and other times -1. Thus, in such a scenario, where the null weak value is not an eigenvalue, we may not associate any physical reality and claim the absence of the corresponding

[26] The operator considered here is hermitian because its difference of two projectors which are hermitian and hence can be considered as an observable

polarization component. Consequently, as per the ABL rule, the eigenvectors of $\mathcal{O}$ would occur with conditional probability strictly less than 1.

When $\langle\phi|\psi\rangle \to 0$, the weak value for the observable $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ is indeterminate. For example, if the pre-selected state is $|H\rangle$ and post-selected state is $|V\rangle$, the weak value of σ_z under the limiting case can either approach 1 or -1 depending on the direction of approach.

$$\lim_{|\phi'\rangle\to|V\rangle,|\psi'\rangle\to|H\rangle} \frac{\langle\phi'|(\sigma_z|\psi'\rangle)}{\langle\phi'|\psi'\rangle} = 1 \lim_{|\phi'\rangle\to|V\rangle,|\psi'\rangle\to|H\rangle} \frac{\langle\phi'|\psi'\rangle}{\langle\phi'|\psi'\rangle} \tag{3.46}$$

$$= 1 \quad (\sigma_z \text{ acting on } |\psi'\rangle) \tag{3.47}$$

$$\lim_{|\phi'\rangle\to|V\rangle,|\psi'\rangle\to|H\rangle} \frac{(\langle\phi'|\sigma_z)|\psi'\rangle}{\langle\phi'|\psi'\rangle} = -1 \lim_{|\phi'\rangle\to|V\rangle,|\psi'\rangle\to|H\rangle} \frac{\langle\phi'|\psi'\rangle}{\langle\phi'|\psi'\rangle} \tag{3.48}$$

$$= -1 \quad (\sigma_z \text{ acting on } \langle\phi'|) \tag{3.49}$$

In such scenarios, the pointer shift would average out to zero. But that does not imply that a counterfactual strong measurement would result in the eigenvalue 0, since 0 is not an eigenvalue of σ_z. Just that half of the times, the measurement of σ_z would give $+1$ and in the other runs it would give -1. In such a case, the pointer shift would give us the statistical average. And hence, we cannot apply the result of no pointer shift to imply that each particle in the ensemble did not have the associated property.

Therefore, when we observe no pointer shift, it could imply that either the weak value turns out to be zero or it is indeterminate. Even when we have a null weak value, the observable may not have 0 as an eigenvalue. In that case, we may not claim any intermediate state with certainty. Only when the null weak value corresponds to one of the extreme eigenvalues with non-degenerate spectrum, we can associate physical reality the corresponding eigenstate. When we have degeneracy, we have to check with the ABL rule to verify if the conditional probability of intermediate occurrence of any of the eigenstate with null weak value turns out to be 1.

For the observable $\Pi_B \otimes \mathbb{1}$, we conclude that the null weak value implies absence of photon in arm B because the ABL rule predicts the occurrence of the eigenstate $|A\rangle|H\rangle$ with certainty. This conclusion is not simply based on absence of pointer shift.

3.3.4 The sum rule for 'which-path' observable

Now let us consider what would consist a measurement of the 'which-path' observable in a Mach-Zehnder interferometer $\Sigma_z = \Pi_A - \Pi_B$? We shall ignore the polarization degree of freedom in this paragraph for the sake of simplicity. Ideally, one would use a tilted glass plate in the arm A so that the beam gets a horizontal transverse shift along $\hat{x}$ say by amount a . The beam in the arm B needs to be similarly coupled with glass plate tilted in opposite direction so that at the detector/camera plane the transverse shift is along $-\hat{x}$ and preferably by the same amount. If the phase shift between the two arms of the interferometer is maintained at 0, then if at the detector plane we observe the resultant centre of the transverse profile of the beam to be at a, we can claim that all the light is coming from arm A. If some component of light were coming from arm B, then the centre c of the beam would be $-a < c < a$. If all the light were coming from arm B, then the centroid would be at $-a$. If the relative phase shift between the two arms changes, then, the pointer shift could go beyond the range $[-a, a]$ due to weak value amplification.

But what really helps us distinguish if the beam is coming from arm A or arm B is the relative separation. If only the glass plate is present in arm A (arm B) and not in B (arm A), then no pointer shift would imply the beam coming from arm B (arm A) and if the shift is a $(-a)$, then we can conclude that the beam is from arm A(arm B). Thus glass plate in any one of the arm suffices, corresponding with a coupling with the observable Π_A or Π_B. Also, since all the three observables Π_A, Π_B and Σ_z have the same eigenstates with different eigenvalues, practical implementation of any one of them suffices to determine the weak value of all the three observables. This is because of the sum rule and additivity of weak values.

$$\Pi_A + \Pi_B = \mathbb{1} \Rightarrow \langle \Pi_A \rangle^{(w)} + \langle \Pi_B \rangle^{(w)} = \langle \mathbb{1} \rangle^{(w)} \tag{3.50}$$

$$\Rightarrow \langle \Pi_A \rangle^{(w)} = 1 - \langle \Pi_B \rangle^{(w)} \tag{3.51}$$

$$\rightarrow \langle \Pi_B \rangle^{(w)} = 1 - \langle \Pi_A \rangle^{(w)} \tag{3.52}$$

$$\langle \Sigma_z \rangle^{(w)} = \langle \Pi_A - \Pi_B \rangle^{(w)} = \langle \Pi_A \rangle^{(w)} - \langle \Pi_B \rangle^{(w)} \tag{3.53}$$

$$= \langle \Pi_A \rangle^{(w)} - \langle \mathbb{1} - \Pi_A \rangle^{(w)} = 2 \langle \Pi_A \rangle^{(w)} - 1 \tag{3.54}$$

$$\equiv \langle \mathbb{1} - \Pi_B \rangle^{(w)} - \langle \Pi_B \rangle^{(w)} = 1 - 2 \langle \Pi_B \rangle^{(w)} \tag{3.55}$$

The sum rule and hence all the above results extends to higher dimension, i.e., when we have the direct product of the above observables with identity operator in the polarization degree of freedom. Hence, we do not need to place a tilted glass plate in arm A and arm B separately. By placing the tilted glass plate in any arm, we would get the weak value of all the relevant observables. The interpretation of the weak value for the 'which-path' observables (and path projectors) are in accordance to the ABL rule as the weak value for the given pre- and post-selection happens to be eigenvalues of the corresponding observables.

3.3.5 Polarization component in a given arm B

Unlike the 'which-path' observables, the polarization component in arm B (or A) are not dichotomic. The observable $\Pi_B \otimes \sigma_x$ corresponds to the σ_x polarization component in arm B and has three eigenvalues $+1, 0, -1$. The eigenvalue $+1$ corresponds to the eigenstate $|B\rangle |+\rangle$ i.e., diagonal polarization in arm B. The eigenvalue -1 corresponds to the anti-diagonal polarization in arm B i.e., the state $|B\rangle |-\rangle$. The eigenvalues 0 has degenerate eigenstates which corresponds to path A and any basis vectors for polarization degree of freedom. Thus, the eigenstates $\{|A\rangle |H\rangle, |A\rangle |H\rangle\}$ are one possible set while $\{|A\rangle |+\rangle, |A\rangle |-\rangle\}$ is another. The weak value for the pre- and post-selected state becomes 1.

$$\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = \frac{\langle \phi | \Pi_B \sigma_x | \psi \rangle}{\langle \phi | \psi \rangle} \tag{3.56}$$

$$= \frac{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) | B \rangle \langle B | \otimes \sigma_x \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)}{\frac{1}{\sqrt{2}} \langle H | (\langle A | + \langle B |) \frac{1}{\sqrt{2}} (|A\rangle |H\rangle + |B\rangle |V\rangle)} \tag{3.57}$$

$$= \frac{\langle H | (\langle A | + \langle B |) | B \rangle \langle B | (|A\rangle |V\rangle + |B\rangle |H\rangle)}{\langle H | (\langle A | + \langle B |) (|A\rangle |H\rangle + |B\rangle |V\rangle)} \tag{3.58}$$

$$= \frac{\langle H | H \rangle}{\langle H | H \rangle} = 1 \tag{3.59}$$

This weak value is the same as the eigenvalue corresponding to the intermediate state $|B\rangle |+\rangle$. But, given the pre- and post-selection, the ABL rule tells us that the probability of the $|B\rangle |+\rangle$ being the intermediate state when we make ideal measurement is $1/6$ (See Table A.7 in Appendix B). Thus, we may not associate the element of physical reality to the diagonal polarization in arm B in the spirit in which we claimed the particle to be in arm A at any intermediate time.

The interpretation of weak values in such a scenario is debatable. If we believe that the weak value truly represents the intermediate value between pre- and post-selection for each particle in the ensemble, then the weak value $\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = 1$ screams that the polarization in arm B is diagonal $|+\rangle$, the eigenstate of σ_x. However, in general, we may not associate an intermediate state when the weak value is not equal to one of the eigenvalues. Thus, even if we accept the weak value to represent physical reality, it is in terms of associating a value to the observable.

If we treat the weak value as a conditional average, we allow the possibility of a statistical distribution of intermediate states. Since here, the weak value obtained is one of the extreme eigenvalues, we are tempted to claim that the only possible intermediate state is the diagonal polarization in arm B. But note that weak values can go beyond the range of eigenvalues and in general can be complex. Thus, associating an intermediate state, although convenient for the concerned observable, cannot be a generic principle.

Lastly, we resort to the fact that the non-zero real weak value will cause a physical shift of the pointer coupled to the observable. The pointer shift would indicate the presence of the property in which coupling is made. Thus, the pointer shift would indicate the presence of intermediate σ_x polarization property in the arm B.

The weak values of σ_y polarization component in arm B is $\langle \Pi_B \otimes \sigma_y \rangle^{(w)} = -i$. Although the weak value is non-zero, it is purely imaginary. The real part of weak value being zero, implies that the pointer would not shift. Instead, the momentum conjugate to the pointer variable will get shifted. [27]. This exemplifies that we cannot treat the absence of pointer shift as null weak value. On the other hand, we have $\langle \Pi_B \otimes \sigma_z \rangle^{(w)} = 0$. The weak value for σ_z polarization component in arm B turns out to be one of the eigenvalue with degenerates. The ABL rule predicts that the eigenstate $|A\rangle |H\rangle$ occurs with certainty, and all eigenstates never occur as the intermediate state (See Table A.9). Thus, in this case, we can claim that the σ_z polarization component is absent in arm B when the particle is subjected to pre- and post-selection. Consequently, the coupling would not shift the pointer.

3.3.6 Polarization component in a given arm A

Given pre- and post-selection, the σ_z polarization component is absent in arm B. In arm A, the weak value of the σ_z observable evaluates to $\langle \Pi_A \otimes \sigma_z \rangle^{(w)} = 1$. This reaffirms the intermediate state to be $|A\rangle |H\rangle$ (See Table A.6). The weak value of σ_y polarization component in arm A is zero. Although 0 is one of the eigenvalues of the observable $\Pi_A \otimes \sigma_y$, the ABL rule suggests that the eigenstates corresponding to that eigenvalue never occurs as an intermediate state (See Table A.5). In fact, the states corresponding to the eigenvalue 1 and -1 each occurs with 50% probability. Thus, in this case, the weak value indicates the conditional average of intermediate values of the observable rather than the true value of each of the particle in the pre- and post-selected ensemble.

We also have $\langle \Pi_A \otimes \sigma_x \rangle^{(w)} = 0$. Just like the case with σ_z polarization component in arm A, the σ_x component will also not show any pointer shift. But the eigenstates corresponding to the null weak value do not occur as intermediate states as per the ABL rule (See Table A.4). Again, the diagonal and anti-diagonal polarization states in arm A corresponds to the eigenvalue of $+1$ and -1 respectively, which occurs as intermediate states with 50% probability. Hence, the null weak value $\langle \Pi_A \otimes \sigma_x \rangle^{(w)} = 0$ and the consequent absence of pointer shift do not necessarily imply the absence of the σ_x

[27] Some interpret this as the back action of the pointer to the observable which affects the post-selection probability. See https://doi.org/10.48550/arXiv.1909.04806

polarization component in arm A.

3.3.7 Polarization in the interferometer

From the weak value $\langle \Pi_A \rangle^{(w)} = 1$ and use of ABL rule, we could conclude the absence of photon in arm, B which manifested as null weak value. Once Π_A is measured, we do not need to separately measure Π_B. Similarly, can we infer the weak value of $\Pi_A \otimes \sigma_x$ by measuring the weak value of $\Pi_B \otimes \sigma_x$? We have the following relation between the observables.

$$\Pi_A \sigma_x = (\mathbb{1} - \Pi_B)\sigma_x \tag{3.60}$$

$$= \mathbb{1} \otimes \sigma_x - \Pi_B \otimes \sigma_x \tag{3.61}$$

$$\Rightarrow \langle \Pi_A \sigma_x \rangle^{(w)} = \langle \sigma_x \rangle^{(w)} - \langle \Pi_B \sigma_x \rangle^{(w)} \tag{3.62}$$

The above equation tells us that knowing the weak value of a particular polarization component in one arm does not give us the weak value of the polarization component in the other arm. To make the inference, we need to also know the weak value of polarization component irrespective of the 'which-path' information. The weak measurement for the $\mathbb{1} \otimes \sigma_x$ observable would require polarizing beam displacers to be placed in each arm of the interferometer so that the resultant pointer shift do not carry any information about in 'which-arm' arm the coupling to the σ_x polarization component occurred.

The weak value of σ_x in the interferometer would be the sum of weak values of polarization component in arm A and arm B respectively. Hence, $\langle \sigma_x \rangle^{(w)} = 1$. If we make simultaneous ideal strong measurements of σ_x observable in both the arms, we would obtain the diagonal or anti-diagonal polarization in arm A or arm B, all with 25% probability (See Table A.10). Note that, in this case, the weak value cannot be regarded as a statistical conditional average of the possible intermediate values. The same can be concluded about σ_y (See Table A.12).

For the σ_z polarization component, we have the weak value of $\langle \mathbb{1} \otimes \sigma_z \rangle^{(w)} = 1$. The eigenstates $|A\rangle |H\rangle$ corresponding to the eigenvalue 1 emerges as the intermediate state with certainty (see Table A.12).

3.3.8 Joint weak measurements

For the various observables considered so far, the only eigenstate to emerge as the intermediate state with certainty is $|A\rangle|H\rangle$. Measuring the weak value of the observable $\Pi_A \otimes \mathbb{1}$ to be 1 confirms the same by indicating presence of the particle subjected to pre- and post-selection in arm A, which manifests in the pointer shift. For a stream of identically pre- and post-selected particles, this would imply that the particle is absent in arm B. But weak measurement of the σ_x polarization component in arm B would result in a pointer shift corresponding to the weak value of 1. Considering the two observations simultaneously, one would conclude that we have a signature of σ_x polarization component in arm B which is not taken by the particle given pre- and post-selection.

The argument in favour of simultaneous consideration of the consequences of two separate weak measurements is that the results can be applied together because weak measurements do not disturb the system appreciably, they do not disturb other weak measurements as well. However, unless we perform the weak measurements simultaneously, the obtained pointer shift would correspond to two separate runs of the experiment with coupling introduced between two different observables in the separate runs of the experiment. This may imply two different experimental contexts, which may affect the intermediate values. Without simultaneous weak measurement of both the observables, the criticism of counterfactual reasoning cannot be ruled out.

Simultaneous weak measurements of two different observables requires coupling with two independent pointers. We can weakly couple Π_A to one pointer state variable (say displacement of beam along x-axis) and $\Pi_B\sigma_x$ with another (displacement of beam along y-axis). Since the weak values of the two observables considered are 1, we expect to observe diagonal pointer shift as a signature of joint weak measurement.

Ideally, to tell the story of quantum cheshire cat in the language of weak values, we would ideally desire the simultaneous weak measurement of the following observables:

- *[necessary]* Ideally both or at least one of Π_A or Π_B (preferably Π_A) : to verify that the particle is in arm A and not in arm B.
- *[necessary]* $\Pi_B \otimes \sigma_x$: to obtain the signature of the σ_x polarization component in arm B via the pointer shift

- *[additional]* $\Pi_A \otimes \sigma_x$: to signify absence of signature of σ_x polarization component in arm A
- *[additional]* $\Pi_A \otimes \sigma_z$: to assert that signature of some polarization component can still be found in arm A and thus the quantum chehsire cat is partial.

If we simultaneously measure the necessary observables, we would require two independent pointers. They can be the centre of the beam along the two independent directions (i.e., along $\hat{x}$ and $\hat{y}$) in the transverse spatial profile. From the pointer shifts, we shall only be able to infer the real of the weak value. Thus, establishing a null weak value from absence of pointer shift itself requires some input from theory. But the presence of pointer shift can always be regarded as presence of the coupling to the property. Therefore, given two independent pointers, it is best to jointly measure Π_A to verify presence of photon in arm A, which for a single photon source would imply the absence of photon in arm B, and $\Pi_B \otimes \sigma_x$ to obtain a signature of presence of σ_x polarization component in arm B. This is done in the next chapter where we discuss the experiment with single photons.

Later in this chapter, we discuss the experimental implementation of joint weak measurements with laser source. There, we perform joint measurement of the following pairs of observables in different runs: $\{\Pi_A, \Pi_B\sigma_x\}$, $\{\Pi_B, \Pi_A\sigma_x\}$,$\{\Pi_A, \Pi_A\sigma_x\}$ and $\{\Pi_B, \Pi_B\sigma_x\}$. We aim to experimentally obtain weak values from the pointer shifts by implementing weak coupling and then verifying that simultaneous weak measurement of two observables do not disturb each other. Hence, the objective is to experimentally realize the partial quantum chehsire cat without resorting to counterfactual reasoning.

Often, the quantum cheshire cat is presented by interpreting the weak value of 1 as the presence of the property and the weak value of zero as the absence of the property. The weak values $\langle \Pi_A \rangle^{(w)} = 1$ and $\langle \Pi_B \rangle^{(w)} = 0$ are considered to signify the presence of the particle in arm A only and the absence of the particle in arm B. The conclusion is correct and is consistent with the ABL rule. However, $\langle \Pi_A \otimes \sigma_x \rangle^{(w)} = 0$ and $\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = 1$ are also considered to imply the absence of the σ_x polarization component in the arm A and its presence in the arm B. As discussed in this section, the interpretation of such weak values is not consistent with the ABL rule, i.e., the ABL rule does not predict the presence or absence of the σ_x polarization component in either arm with certainty. Also, the interpretation of the weak value do not extend to other observables with complex weak

values. Interestingly, the paradox of the quantum chehsire cat effect survives but with a subtle reasoning that when we aim to perform joint weak measurements as we would obtain signature of σ_x polarization component in arm B while the which-path observable would be consistent with the particle being in arm A.

3.4 Has quantum Cheshire cat already been experimentally demonstrated ?

3.4.1 Literature on demonstration of cheshire cat effect

The Quantum Cheshire Cat effect was first described with a photon and its polarization subjected to pre-selection and post-selection in an interferometer [3.19, 3.16], where the photon is located in one arm of the interferometer and one of the polarization component is located in the other arm. This effect is termed as *partial* Cheshire Cat [3.20] because some polarization component still resides in the arm where the photon is located and thus the polarization degree of freedom is not completely separated from the photon. Quantum Cheshire Cat effect has led to the theoretical development of spin currents without mass that may explain non-local interactions with underlying local dynamics [3.21]. The fact that the QCC effect can be robust to certain kind of decoherence [3.22] makes it an ideal candidate for quantum paradoxes that survives in the classical limit.

Just like the double slit interference build up disappears if we infer the trajectory of each particle [3.23, 3.24], many phenomena based on quantum interference are not realizable when measurement strongly disturbs the system. However, weak measurements can be implemented to infer properties of system subjected to pre- and post-selection [3.25, 3.26]. The weak value, the property of the system as inferred from the shift of a pointer coupled to it [3.27, 3.28], can in general be amplified beyond the range of eigenvalues and thus has found numerous applications like precision metrology [3.29, 3.30] and enhancing non-linearity [3.31]. Weak measurements have been implemented in a wide variety of systems ranging from atom-electron to NMR [3.32, 3.33, 3.34, 3.35] and are a tool to probe foundational issues in quantum mechanics [3.36, 3.37, 3.38, 3.39, 3.40, 3.41, 3.42]. Despite their utility, the meaning of weak values and weak measurement have been debated

[3.43, 3.44, 3.45]. The controversy spills over to demonstrating Quantum Cheshire Cat using weak measurements. In reference [3.15], much of the controversy is addressed.

The first experiment [3.18] aimed at demonstrating the Quantum Partial Cheshire Cat effect involved neutrons claimed to be separated from its magnetic moment. The conclusion of this experiment was contested with concerns of quadratic magnetic field interactions destroying the implications that would have been obtained from weak values [3.46] and also was interpreted as "simple quantum interference" [3.47] that do not necessarily imply the separation of property from the photon. The quantumness of Cheshire Cat effect have been debated based on simulations using Discreet Event Models [3.48] as well as optics based experiments claiming Classical Cheshire Cats [3.49]. The above experiments have limitations in implying the Quantum Cheshire Cat effect as the weak values were not obtained from shift of pointer state weakly coupled to the observable [3.15]. Experiment on Quantum Cheshire Cat with single photon source [3.50] still had used absorbers and lacked implementation of weak measurements. Proper implementation of weak measurement was employed in [3.51], but joint weak measurements could not be performed in a single run of the experiment (For a detailed analysis, see Appendix C). Hence, a loophole-free demonstration of the Quantum Cheshire Cat effect, where joint weak measurement of 'which-path' and polarization in a given path observables are simultaneously measured, is still pending in the literature.

3.4.2 Use of absorbers

The experiments reported in [3.18, 3.49, 3.50] have used absorbers or neutral density filters to probe the presence of the particle in a given path. The central idea is that if under pre- and post-selection, the particle always takes arm A, then if we place an absorber[28] in arm B it would not affect the number of particles getting post-selected. When we place a perfect absorber in arm A, we would see none of the particles getting post-selected. This is consistent with the prediction of ABL rule that the intermediate location of the particle is in the arm A. One can also vary the transmissivity of the absorber in arm A and observe that the number of post-selected particles increases proportionally to the transmissivity.

[28]The absorbers or neutral density filters absorb a fraction of the particles incident on them and transmit the rest

However, changing the transmissivity of absorber in arm B has no effect.

The main concern with the use of absorbers is that it changes the pre- and post-selected ensembles for which the property of the system is being queried. Fundamentally, if a particle is absorbed in the path, it cannot satisfy the post-selection. Thus, the amount of particles absorbed do not tell us anything about the property of the particles satisfying pre- and post-selection. An absorber acts as a strong projector which demolishes the particle itself. The use of absorber makes the entire ensemble a mixture of two sub-ensembles: those satisfying pre- and post-selection and those which get absorbed in between. Although the statistical expectation that any absorber would decrease the post-selection probability is satisfied, for the subset of particles that get post-selected we do not have the 'which-path' information. If the ensemble itself is disturbed, one cannot non-invasively extract the property of the pre- and post-selected ensemble. But if the absorption is low i.e., transmission approaches 1, the change in detected intensity is also very small. This is argued as a 'weak' interaction. However, here we argue that it is not a weak measurement for a pre- and post selected ensemble, as the absorbed particle do not get subjected to post-selection. Further, the use of absorber[29] is not a unitary evolution.

Consider the case in which we have a perfect absorber in arm A. In that case can be measure the σ_x polarization component in arm B ? If the earlier result consistently applies, then none of the particles would get post-selected. Hence, asking the σ_x polarization component in arm B of the pre- and post-selected ensembles becomes counter-factual and meaningless if none of the particles are post-selected. When the absorber has a certain transmission and the particle is successfully post-selected, can we claim that the particle definitely went through the arm A? No. Thus, with the use of absorber, we cannot comment on which path the particle went through in the ensemble that satisfies the successful post-selection takes. Thus, any measurement of $\Pi_B\sigma_x$ for a particle belonging to post-selected ensemble would not have the 'which-path' information. Therefore, we are back to the use of counterfactual reasoning, if we wish to apply the conclusions to all the particles in the

[29]If one considers the state of the absorber as well, say ground or excited state of an atom that absorbs the photon, then it is possible to describe the unitary operation in a higher dimensional Hilbert space that corresponds to absorption. But the premise of Chehsire cat involves existence of one particle considered as 'system' in the interferometer. If we include the pointer states, in the state description, the overall evolution of joint system-pointer state is unitary. However, if the system itself is absorbed, the paradox itself becomes ill-posed

ensemble.

3.4.3 Use of polarization/spin-rotators

If we insert a half-wave plate in arm B, depending on the orientation of fast-axis angle θ_h from the horizontal, the polarization will change from $|V\rangle$ to $\sin(2\theta)|H\rangle - \cos(2\theta_h)|V\rangle$. The horizontal component of the beam in arm B would interfere with the beam in arm A, which has polarization $|H\rangle$ at the second beam splitter BS_2. Thus, the intensity of light in port d can get affected by the phase shift φ between the two arms of the interferometer. The horizontal component of the beam coming from arm B will transmit through the PBS_{post} and therefore, the intensity detected after post-selection would get affected. The post-selected intensity would change as a function of the phase shift φ due to interference. This is presented as evidence for the presence of polarization component in arm B. However, it is not clear why the HWP would be regarded as interacting with only σ_x the polarization component. Depending on θ_h, there would exist polarization components, along the eigenstates of the Jones matrix representing the HWP evolution operator, that would not get affected. For example, if $\theta_h = 0$, the eigenstates of σ_z would not get effected. But the eigenstates of σ_x and σ_y, if present at the intermediate, time, would change in principle and thus as a function of phase shift φ any of those components would lead to a change in post-selected intensity. Similarly, the HWP would affect σ_z and σ_y components when $\theta_h = \pi/8$ and the σ_x polarization component would remain unaffected.

If we place the HWP oriented at θ_h in the arm A, the polarization in arm A would change from $|H\rangle$ to $\cos(2\theta_h)|H\rangle + \sin(2\theta_h)|V\rangle$. The vertical component of the beam in arm A would now interfere with the beam coming from arm B as a function of φ. But since the vertical component would be blocked by the PBS_{post}, the post-selected intensity would not vary as a function of φ. This is presented as signature of absence of polarization component in arm A. But let's note that for any fixed angle of θ_h of the HWP in arm A, the detected intensity would not change as a function of the phase shift φ. If we place a quarter-wave plate in arm A along with the HWP, we can convert the polarization in arm A from $|H\rangle$ to any arbitrary elliptic polarization. Even then, as a function of φ, the post-selected intensity would not change. Could this be then regarded as the signature for absence of any polarization component in arm A ? This is in sharp contrast with

not only the weak value of $\langle \Pi_A \otimes \sigma_z \rangle = 1$ but also the ABL rule predicting $|A\rangle |H\rangle$ as an intermediate state with certainty (See Table A.6).

Irrespective of the phase difference, if we rotate the angle θ_h of the HWP in arm A, the detected intensity would change as function of θ_h. How this can be treated as absence of polarization component in arm A? For a given φ, if we rotate the HWP placed in arm B, we would also observe a similar change in intensity as a function of θ_h.

But in the mentioned references for the experiments, often, the HWP angle at 0° is implicitly referred to as the “weak interaction". In weak interaction limit, the intensity after post-selection is not supposed to change any faster than linear dependence on the interaction strength . Since the polarization component in arm A and arm B in the pre-selected state happen to be eigenstates of σ_z, the Jones matrix when $\theta_h = 0$, the changes in the post-selected intensity are small if the θ_h is close to 0. However, let alone the ‘weak’ portion, we argue that the use of HWP is not a ‘measurement’ of polarization at all.

3.4.4 Distinction between evolution operators and measurement operators that couples the pointer states to the observable

The HWP is a unitary operator that evolves one polarization into another. It happens to be a Hermitian operator as well. For $\theta_z = 0$, the Jones matrix for the HWP turns out to be σ_z and for $\theta_x = 0$, it turns out to be σ_x. But placing a HWP with Jones matrix with σ_z or σ_x configuration do not measure the σ_z or σ_x polarization components by themselves. If we have the HWP with $\theta_h = 0$, and place a detector after it, just by making the detection, we cannot know what is the polarization of the incident photon. In this sense, it is not a measurement at all.

If we make $\theta_h = \pi/4$ and have the operator corresponding to σ_x, it would be regarded as a ‘strong interaction’ as it would change the post-selection intensity by the maximum possible amount, when placed in arm A or in arm B (with $\varphi = 0$). But when placed in arm B (with $\varphi = 0$), does this give us any information about whether the polarization in arm B is diagonal or anti-diagonal? A strong measurement of $\Pi_B \sigma_x$ should project the intermediate state either into diagonal or anti-diagonal polarization in arm B or to any polarization in arm A. But we would not get any such answer by placing a HWP at $\pi/4$

in arm B. Hence, it cannot be considered as a strong measurement of any polarization component. A strong measurement interaction would ideally entangle the eigenstates of the observable to distinguishable states of a pointer variable. In the weak interaction limit, the distinguishability can be reduced by lowering the interaction strength. But that presumes the coupling to the pointer states which is not provided by the evolution operator, such as the HWP, alone.

Further, when absorber is placed in one arm A and HWP in the arm B, just by measuring the mere post-selected intensity change, we cannot comment on individual properties in each arm. The change in post-selected intensity as a function of φ is presented as a joint effect of the absorber and the wave-plate, which gives a consistent result when the components are individually placed in separate runs of the experiment. But this cannot be considered as joint 'weak-measurement' as the two observables are not coupled to independent pointers. Both the operators in this case affect the post-selected intensity. Thus, if we place two HWPs in both arms, one would not be able to conclude which HWP affects the intensity without relying on the cases where only one of them is present. Further, we can place two HWPs in each are at $\pi/4$ angle. The polarization in arm A to become $|V\rangle$ and the polarization in arm B would become $|H\rangle$. The post-selected intensity would remain the same as the case when no operators are introduced. If the operators are considered as measurements, then one would erroneously conclude that joint strong measurements do not affect the system. Hence, the experimental procedure performed in [3.49] and [3.50] cannot be regarded as joint weak-measurements since after all, they do not even qualify as von Neumann type measurements of the 'which-path' and 'polarization' in a given path observables.

3.4.5 Inference of weak value from intensity

In the references [3.18, 3.50], the weak value is not observed by its manifestation in the pointer shift as coupling to pointers were not employed. Instead, the weak values are inferred from the change in post-selected intensity.

When we have an absorber with transmission amplitude $0 \leq \alpha^2 = T \leq 1$ in arm A and the HWP rotated at angle θ_h such that it evolves the polarization by the operator $S(\theta_h)$, in arm B, the pre-selected state evolves to the following state $|\psi_e\rangle$.

$$|\psi_e\rangle = (\Pi_A \alpha + \Pi_B S(\theta_h))\,|\psi\rangle \tag{3.63}$$
$$= (\Pi_A \alpha + \Pi_B S(\theta_h))\frac{1}{\sqrt{2}}\left(|A\rangle\,|H\rangle + e^{i\varphi}\,|B\rangle\,|H\rangle\right) \tag{3.64}$$
$$= \frac{1}{\sqrt{2}}\left(\alpha\,|A\rangle\,|H\rangle + e^{i\varphi}\,|B\rangle\,(\sin(2\theta_h)\,|H\rangle - \cos(2\theta_h)\,|V\rangle)\right) \tag{3.65}$$

The post-selection intensity is then obtained as the following which gives the oscillating behavior as a function of phase shift.

$$|\langle\psi_e|\psi\rangle|^2 = \tag{3.66}$$
$$= \left|\frac{1}{\sqrt{2}}(\langle A| + \langle B|)\,\langle H|\,\frac{1}{\sqrt{2}}\left(\alpha\,|A\rangle\,|H\rangle + e^{i\varphi}\,|B\rangle\,(\sin(2\theta_h)\,|H\rangle - \cos(2\theta_h)\,|V\rangle)\right)\right|^2 \tag{3.67}$$
$$= \left|\frac{1}{2}(\langle A| + \langle B|)(\alpha\,|A\rangle + e^{i\varphi}\,|B\rangle\,(\sin(2\theta_h)\,|H\rangle)\right|^2 \tag{3.68}$$
$$= \frac{1}{4}|(\alpha + e^{i\varphi}\sin 2\theta)|^2 \tag{3.69}$$
$$= \frac{1}{4}(\alpha^2 + \sin^2(2\theta_h) + 2\alpha\sin(2\theta_h)\cos(\varphi)) \tag{3.70}$$

When proper weak measurement for the path projector Π_A is employed, the Hamiltonian interaction couples the observable to the conjugate momentum of the pointer variable x_A with the quantum state $|\chi_A\rangle\,(x_A)$. When jointly measured, the observable $\Pi_B\sigma_x$ must be coupled with another independent pointer state $|\chi_B\rangle\,(x_B)$. The interaction Hamiltonian then is given by $H = g_A(t)\gamma_A\Pi_A \otimes \hat{p_A} + g_B(t)\gamma_2\Pi_A\sigma_x \otimes \hat{p_B}$. Here, $\hat{p}_A$ and $\hat{p}_B$ are the momentum operator which generates the translation of the pointer state over time. $g_A(t)$ and $g_B(t)$ are compact functions the time integral of which evaluates to 1. γ_A and γ_B are the peak interaction strength. The interaction The joint system and pointer state i.e., $|\psi\rangle\,|\chi_A\rangle\,|\chi_B\rangle$, is evolved by the unitary $U = \exp(-iHt/\hbar)$.

$$U = \exp\left(-\int_t iHt/\hbar dt\right) \tag{3.71}$$
$$= \exp\left(-\int_t \frac{i}{\hbar}\left(g_A(t)\gamma_A\Pi_A\frac{\hbar}{i}\partial_{x_A}\right) + \left(g_B(t)\gamma_A\Pi_B\sigma_x\frac{\hbar}{i}\partial_{x_B}\right)dt\right) \tag{3.72}$$
$$= \exp\left(-\gamma_A\Pi_A\partial_{x_A} - \gamma_B\Pi_B\sigma_x\partial_{x_B}\right) \tag{3.73}$$
$$\approx \mathbb{1} - \gamma_A\Pi_A\partial_x - \gamma_B\Pi_B\sigma_x\partial_{x_B} \tag{3.74}$$

In the last step, we have used the approximation for weak interaction $\gamma_1, \gamma_2 << 1$. The joint system and pointer states evolves to Ψ_e given by the following.

$$|\Psi_e\rangle = \mathbb{1} - \gamma_A\Pi_A\partial_x - \gamma_B\Pi_B\sigma_x\partial_{x_B}|\psi\rangle \otimes |\chi_A\rangle(x_A)|\chi_B\rangle(x_B) \tag{3.75}$$
$$= |\psi\rangle|\chi_A\rangle|\chi_B\rangle - \Pi_A|\psi\rangle\gamma_A\partial_{x_B}|\chi_A\rangle(x_A)|\chi_B\rangle(x_B) \tag{3.76}$$
$$- \Pi_B\sigma_x|\psi\rangle|\chi_A\rangle(x_A)\gamma_B\partial_{x_B}|\chi_B\rangle(x_B) \tag{3.77}$$

Upon post-selection, we have the shifted pointers states proportional to the weak values.

$$\langle\phi|\Psi_e\rangle = \langle\phi|\psi\rangle|\chi_A\rangle|\chi_B\rangle \tag{3.78}$$
$$- \langle\phi|\Pi_A|\psi\rangle\gamma\partial_{x_B}|\chi_A\rangle(x_A)|\chi_B\rangle(x_B) \tag{3.79}$$
$$- \langle\phi|\Pi_B\sigma_x|\psi\rangle|\chi_A\rangle(x_A)\gamma\partial_{x_B}|\chi_B\rangle(x_B) \tag{3.80}$$
$$= \langle\phi|\psi\rangle\left(|\chi_A\rangle|\chi_B\rangle - \langle\Pi_A\rangle^{(w)}\gamma_A\partial_{x_A}|\chi_A\rangle|\chi_B\rangle - \langle\Pi_B\sigma_x\rangle^{(w)}|\chi_A\rangle\gamma_B\partial_{x_B}|\chi_B\rangle\right) \tag{3.81}$$

For small enough $\gamma_A << 1$ and γ_B, we may approximate that the overlap between the pointer states is negligible. Then, we arrive at the approximate, post-selection intensity, i.e.,

$$I_{pps} = |\langle\phi|\psi\rangle|^2(1 - 2\langle\Pi_A\rangle^{(w)} - 2\langle\Pi_B\sigma_x\rangle^{(w)}) \tag{3.82}$$

Here, we have assumed the weak values to be real and expanded up to the terms proportional to γ before absorbing into the pointer state, which approximately cancels out between the LHS and RHS. When only the absorber is present, by measuring the intensity and

comparing with the above expression, one can infer the weak value. However, when we place the absorber and the HWP in the respective arms simultaneously, we cannot individually determine the weak values by simply measuring the post-selection probability. One has to rely on individual experiments. Of course, if one plots the post-selected intensity as a function of φ, one can fit the data to infer the two weak values separately. It is because while $\langle \Pi_A \rangle^{(w)}$ is independent of φ, $\langle \Pi_B \sigma_x \rangle^{(w)}$ has functional dependence on the relative phase shift φ between the two arms.

$$\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = \frac{\langle \phi | \Pi_B \sigma_x | \psi \rangle}{\langle \phi | \psi \rangle} \tag{3.83}$$

$$= \frac{\frac{1}{\sqrt{2}} \langle H | \left(\langle A | + \langle B | \right) | B \rangle \langle B | \otimes \sigma_x \frac{1}{\sqrt{2}} \left(| A \rangle | H \rangle + e^{i\varphi} | B \rangle | V \rangle \right)}{\frac{1}{\sqrt{2}} \langle H | \left(\langle A | + \langle B | \right) \frac{1}{\sqrt{2}} \left(| A \rangle | H \rangle + e^{i\varphi} | B \rangle | V \rangle \right)} \tag{3.84}$$

$$= \frac{\langle H | \left(\langle A | + \langle B | \right) | B \rangle \langle B | \left(| A \rangle | V \rangle + e^{i\varphi} | B \rangle | H \rangle \right)}{\langle H | \left(\langle A | + \langle B | \right) \left(| A \rangle | H \rangle + e^{i\varphi} | B \rangle | V \rangle \right)} \tag{3.85}$$

$$= e^{i\varphi} \frac{\langle H | H \rangle}{\langle H | H \rangle} = e^{i\varphi} \tag{3.86}$$

The paper [3.49] concludes that the effect can be considered as a simple interference phenomenon and [3.50] echoes the same about measurement statistics. Additionally, [3.50] suggests that because the weak values in their experiment are measured from variation in the post-selected intensities, they represent the average property of an ensemble and do not apply to individual particles. While the conclusion that their experiment cannot comment on the property of individual particle is correct for their experiment, the reasoning incorrectly relies on the interpretation of weak values. The fact that proper weak measurements were not employed limits us to comment on weak values and property of individual particles. Obtaining weak values without weak measurement is possible and can be done even with proper post-selection as shall be discussed in Ch 5. However, merely inferring the weak value without any coupling to the pointer and thus without implementing genuine weak measurements, do not allow us to conclude the above experiments as realization of quantum cheshire cat. Another experiment [3.51, 3.52] implemented genuine weak measurements using two photon interference (See Appendix A) but joint weak measurements of the two observables in the cheshire cat could not be performed in a single

run of the experiment (See Appendix C).

In this chapter, we aim to present a loophole free demonstration of partial Cheshire cat effect where weak values are inferred from pointer shifts after genuine weak measurements are employed. The source for the experiment described in this chapter is coherent (laser) and thus the conclusions applies to individual photons only after consideration of the equivalence of weak values obtained with coherent source vs. a stream of single photons. The experiment performed with single photon source shall be described in the next chapter (Ch 4). The experiment in this chapter focuses on the implementation of joint weak measurements to infer the weak value from the pointer shifts.

3.5 Experimental Implementation with proper weak measurements

The quantum Cheshire cat effect strictly applies to a single particle subjected to pre-selection and post-selection. If there are many particles within the interferometer, some of them may give the signature of being present in arm A and others may give the signature of σ_x polarization component in arm B. Therefore, it seems essential to ensure that there is only one photon within the interferometer at any given point in time. The experimental set up typically would have physical dimensions of the order of 1 m (arm length of the Mach-Zehnder interferometer) and hence the typical time that a photon would spend between pre- and post selection would be of the order of 3 ns. With a laser source of 10 mW power, there would be about 10^7 photons between pre- and post-selection within the time scale of 3 ns. Thus, it seems that attempting to perform Quantum Cheshire cat demonstration with laser light source is a pointless exercise. But, as far as joint weak measurement is concerned, it is possible to employ them with the laser source and obtain the weak values from the pointer shift. The average statistical properties of light are equivalent whether we use an ensemble of single photons or a coherent source [3.53]. Since the pointer shift measured would be the average over ensemble of particles, it is expected to remain the same even if we use a coherent source.

3.5.1 Experimental Set up

We use Helium Neon laser(Melles Griot LHX1-25-LHP991) source ($\lambda = 632.8nm$) with coherence length typically around 20 cm. It has a beam width of about $2220\mu m$ at the beam profiler and has a beam divergence of about $1mrad$. To obtain high purity in polarization, we place a Glan Thompson polarizer GT_{pre}(Thorlabs GTH5-B) before allowing light to enter the Mach-Zehnder interferometer. The Glan Thompson is oriented in such a way that it transmits horizontally polarized light. The Mach Zehnder consists of two beam splitters BS_{pre} and BS_{post} (both Thorlabs BS013). To be able to control path difference between the two arms of the interferometer, we use a Corner Cube Retroreflector CCR (Thorlabs PS975-A) mounted on a piezo in the reflected arm of the first beam splitter BS_{pre}. The CCR is used to minimize any angular deviation as we move the piezo for alignment of the Mach-Zehnder interferometer. After the CCR and the mirror M_R(Thorlabs BB1-E02), the polarization does not remain horizontal, but in general becomes elliptic. To restore linear polarization, we use half wave plate HWP_c (Thorlabs WPH05M-633)and quarter wave plate QWP_c(Thorlabs WPQ05M-633) combination to compensate for the ellipticity. They are oriented in such a way that the polarization after the QWP_c is made Vertical. The usage of CCR causes a macroscopic path delay which is compensated for by creating additional path delay in the transmitted arm of the BS_{pre} using the mirrors M_{T_1}, M_{T_2} and M_{T_3}.

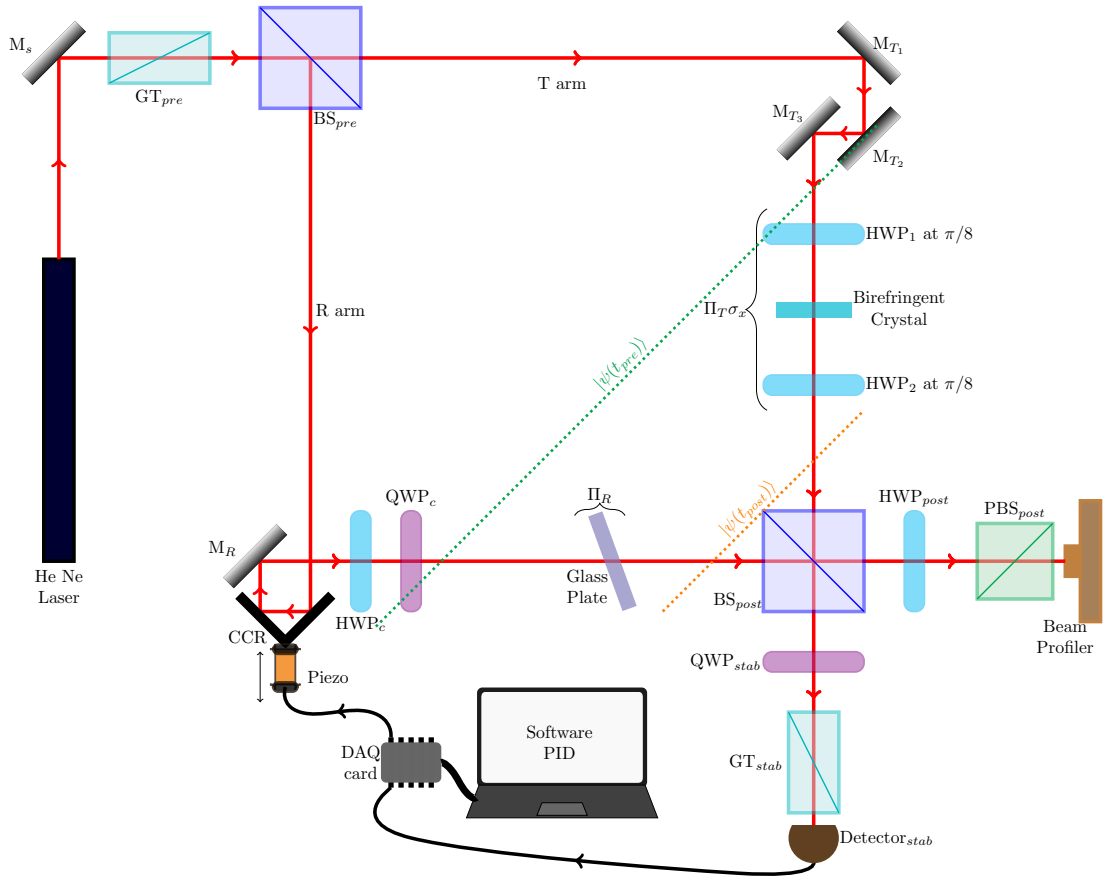


Figure 3.2: Experimental Set up for Partial Cheshire Cat with LASER source. The glass plate serves as 'which-path' projector and is placed in T arm (also called arm A) and R arm (also called arm B) in different runs of the experiment. The birefringent crystal (polarizing beam displacer) serves along with the half wave plates couples σ_x polarization component in the arm in which it is placed. Out of 8 combinations in which one glass plate and/or beam displacer HWP combination can placed in the interferometer, only one is shown for representative purposes although all the 8 experiments were performed

On one output port of the second beam splitter BS_2, we have the half wave plate HWP_{post} mounted on a motorized rotation mount. The beam profiler (Dataray UCD 15.5) is placed in the transmitted arm of the polarizing beam splitter PBS_{post}. The half wave plate HWP_{post} in combination with the PBS_{post} determines the polarization component of the post-selected state. We used transmitted arm of the PBS as it has better extinction ratio. Ideally, a Glan Thompson would have provided purer polarization, but they tend to long and hence slight misalignment tends to cause polarization dependent beam deviations. Care was taken to place the HWP_{post} close to the beam profiler so that the angular deviation due to HWP plate's rotation is minimized.

On the other output port of the BS_{post}, we have a quarter wave plate QWP_{post} with fast axis oriented at 45° from Horizontal to convert the Horizontal and vertical polarization components coming from arm T and arm R respectively into left and right circular polarization. Then we place a Glan Thompson polarizer (Thorlabs GTH10-B) to project the states into Horizontal polarization so that they can interfere.The intensity of light after interference is detected with a fast photo diode $Detector_{post}$. This detector signal collected using the DAQ card (NI USB 6003) is used to infer the phase difference φ between the two arms off the interferometer . A software(LabView) based PID loop is used to stabilize the phase difference between the two arms by giving appropriate Voltage signal generated from the DAQ card to the piezo.

3.5.2 Methods

The pre and post-selected states are considered along the time indicated by the dotted lines in Figure 3.2. Optical components performing weak interactions are placed within these two lines.

'Which-path' Measurement To infer which arm of the interferometer the particle is located at t_i, we place a glass plate (Thorlabs WG41010) in one of the arm of the interferometer and tilt it such that the emergent ray gets displaced horizontally (along x) by $52 \pm 3\mu m$. This displacement is much smaller compared to the standard deviation[30] of the Gaussian beam profile which is $525 \pm 25\mu m$ and thus the interaction can be said to be weak.

When the glass plate is placed in arm T, it is described by the operator $\Pi_T \otimes \mathbb{1}$. Here, Π_T signifies the 'which-path' projection on arm T and $\mathbb{1}$ signifies that the polarization is unaffected by introduction of a glass plate. When the glass plate is placed in arm R, we denote it using the operator $\Pi_R \otimes \mathbb{1}$. The direction in which the glass plate is tilted is inconsequential as long as they re consistent with the operator description and the pointer states from which the eigenstates of the operator are inferred.

[30] at the time of detection the beam width is about $2250\mu m$ which is higher than the beam width at the time of interaction given the beam divergence of $1 \pm 0.05 mrad$

'Polarization in a given path' measurement To perform weak interaction in polarization degree of freedom, we use a birefringent crystal (Newlight Photonics PDC12005) that displaces[31] the vertical polarization by $56 \pm 3\mu m$. The beam displacer was azimuthally rotated such that the displacement occurs along the vertical (y) direction of the pointer. Since the eigenstates of the birefringent crystal are $|H\rangle$ and $|V\rangle$, it acts a σ_z operator in polarization degree of freedom. To convert it into a σ_x operator and still retain the pointer displacement along vertical axis, we place two half-wave plates HWP_1 and HWP_2 before and after the birefringent crystal with fast axis oriented at $\pi/8$ from Horizontal.

However, to ensure that we indeed have σ_x weak interaction, the horizontally and vertically polarized beams emerging out of the birefringent crystal must have the same phase. This is ensured as following: If the birefringent crystal along with the two half wave plates are placed in arm T, we first block arm R. Then since the incident polarization in arm T is $|H\rangle$, we rotate the HWP_{post} such that we detect $|V\rangle$ at the beam profiler. Then the beam displacer is tilted vertically[32] till we observe the destructive interference pattern typical of weak interaction with operator that mediate transition between orthogonal pre and post selected states. Then we unblock arm R and the birefringent crystal acts as the operator $\Pi_T \otimes \sigma_x$.

Similar procedure is followed when making the birefringent crystal act as $\Pi_R \otimes \sigma_x$ operator where arm T is blocked and the HWP_{post} is rotated such that $|H\rangle$ gets detected. The only difference is that the emerging beams from the birefringent crystal do not have same relative phase shift for the same tilt if the incident polarization is changed. However, if we tilt too much the beam separation is also affected and the horizontal component also gets appreciably deviated. To avoid this, we rotate the birefringent crystal azimuthally by π angle when placing it in arm R.

Data Acquisition We perform the following experiments to measure various observables required to demonstrate the Partial Cheshire Cat effect:

[31] The exact beam displacement was not measured with 633 nm source. It was $57 \pm 2\mu m$ with 810 nm. It was theoretically estimated for $633nm$ wavelength

[32] The tilt along the axis about which the pointer gets displaced affects the phase

Experiment	Observable
Glass Plate in arm T	$\Pi_T \otimes \mathbb{1}$
Glass Plate in arm R	$\Pi_R \otimes \mathbb{1}$
Birefringent Crystal in arm T	$\Pi_T \otimes \sigma_x$
Birefringent Crystal in arm R	$\Pi_R \otimes \sigma_x$

In addition to that, to explicitly show that weak measurements do not get affected by simultaneous measurements with independent pointers, we perform an additional experiments by simultaneously placing the Glass plate in arm T and the Birefringent crystal in arm R and vice versa.

The measured quantity is the centroid of beam profile. Since this quantity can change if any optical component is not perfectly aligned, we need to establish that the pointer shifts are indeed because of weak interaction and not any bad alignment of optical components. For this we choose to infer weak value from pointer shifts by varying post-selected state.

For each experiment, the phase difference between the two arms was stabilized at zero. The post selected state was varied by changing the angle of HWP_{post} angle and is given by

$$\left|\psi(t_{post})\right\rangle = \frac{1}{\sqrt{2}}\left(|T\rangle + |R\rangle\right)\left(\cos(2\theta)\,|H\rangle + \sin(2\theta)\,|V\rangle\right) \tag{3.87}$$

The HWP_{post} angle was varied from 0 to 90° in steps of 1° and for each angle beam profiler acquires 10 images. Centroid for each image is computed and the mean and standard deviation of centroid over 30 images are recorded for each angle. The entire process is repeated 5 times. The maximum and minimum of one standard deviation range over 5 sets is presented as the final error bar. The mean of the means is presented as the centroid.

3.5.3 Error Analysis

The weak value[33] of any observable a is computed by rescaling the centroid as follows

$$A_w = \frac{p - p_0}{\gamma} \tag{3.88}$$

[33]Technically, this is only the real part of the weak value. But for partial Cheshire cat experiment, all the concerned weak values are real

Here, p_0 is zero error of the pointer and γ is the beam displacement of the two pointer states that corresponds to the two eigenstates of the concerned operator A. Thus any error in γ for small values causes large error in the weak value.

Other systematic errors comes from beam displacement due to improper alignment or non-ideal surfaces of optical components. Also, since the resultant pointer shift is due to interference, any phase uncertainty in the interferometer will also affect the centroid and hence the weak value.

The theoretical account for error estimates is provided in the Appendix E, where the centroid is computed taking into account most of the experimental parameters like alignment (angular tilt, displacement etc) and uncertainties (in polarization purity , phase stability etc.)

3.6 Experimental Results & Analysis

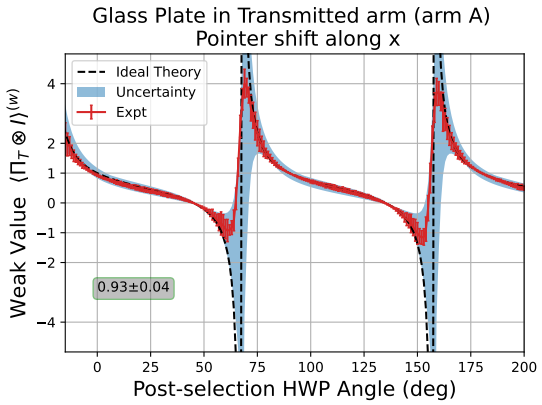
Here, we present the centroid shift for 6 different experimental settings, four of which includes interaction in only one degree of freedom i.e. either glass plate/ beam displacer in one of the arms. Joint measurement of 'which-path' and polarization degree of freedom was also performed by placing glass plate in one arm and beam displacer in the other arm (and vice versa).

The centroid of the beam (after zero correction) is plotted as a function of the orientation of the HWP_{post} half waveplate. The Cheshire Cat effect occurs when the HWP_{post} half waveplate is oriented at angle zero.

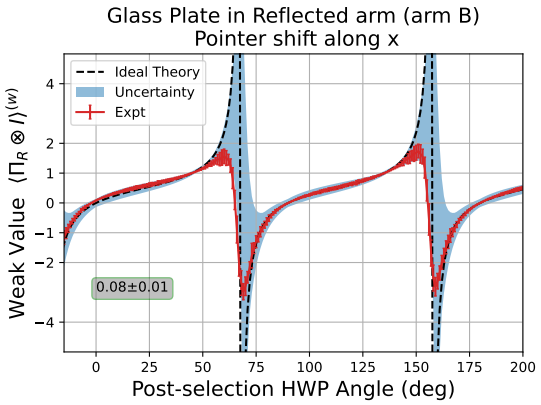
In each of the graphs below, the shaded light blue region denotes the theoretical prediction for centroid taking into account most of the experimental parameters[34]. The blue dots and bars represent the experimental mean and std dev over 5 repetitions of the experiment in which in each run ten images were acquired and the mean centroid and std dev of the centroid was computed.

Glass Plate Tilted glass plate was placed in T-arm of the interferometer.

[34] See Supplementary



(a) Tilted Glass Plate in Transmitted Arm



(b) Tilted Glass Plate in Reflected Arm

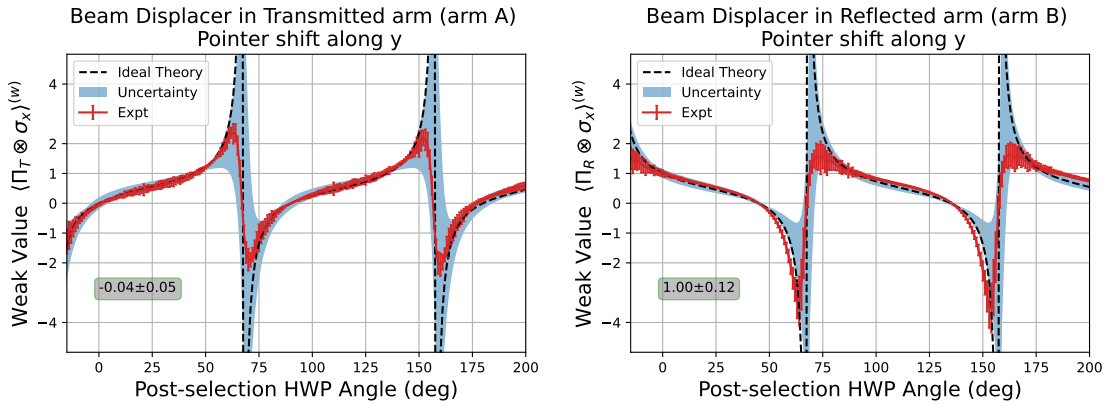
At post selection angle 0, the net displacent of the horizontal component of the centroid amounts to the weak value of 0.93 ± 0.04 and is within the predicted band which arises from the propagation of uncertainties in setting up the experiment . This implies that under subjected to pre and post-selection all the particles go through T-arm where the glass plate is inserted.

As a corollary, we must also have no shift in centroid, when the glass plate is placed in R arm. We experimentally find the weak value to be 0.08 ± 0.01 which although does not overlap with zero, but is still within the theoretically predicted band. From the above two graphs, it is implied that at HWP_{post} angle zero, all the particles satisfying post-selection came through T arm.

Beam Displacer The beam displacer together with the two Half wave plates displaces only the anti-diagonal polarization component only and leaves the diagonal polarization component undeviating. Although, we have treated the interaction as σ_x in polarization degree of freedom, in practice we have to remap the eigenspectrum. A shift of about $28\mu m$ shall lead to the inference that σ_x polarization component is absent and a shift of $-28\mu m$ shall be treated as signature that the particle's σ_x polarization component has manifested in the shift. This is because of the fact that the coupling to the birefringent crystal along with HWP's should be treated as Π_- as only the anti-diagonal component gets displaced. However, since they are related by

$$\Pi_- = \frac{\mathbb{1} - \sigma_x}{2} \tag{3.89}$$

, we can make the above said conclusion.



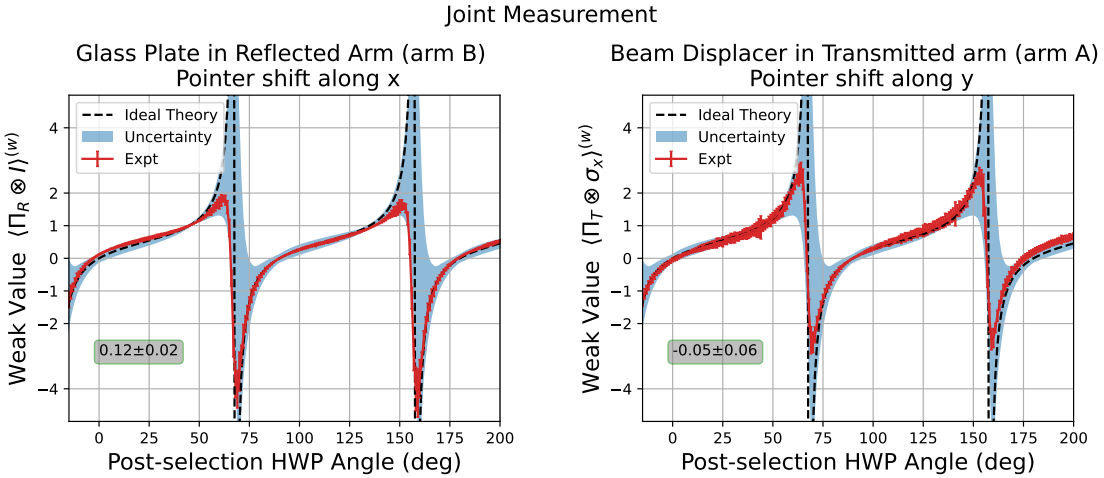
(a) Beam Displacer along with HWPs in Transmitted Arm

(b) Beam Displacer along with HWPs in Reflected Arm

When beam displacer is placed in Transmitted arm, we obtain pointer shift corresponding to zero weak value. And when it is placed in the reflected arm, we have the weak value of 1 ± 0.12.

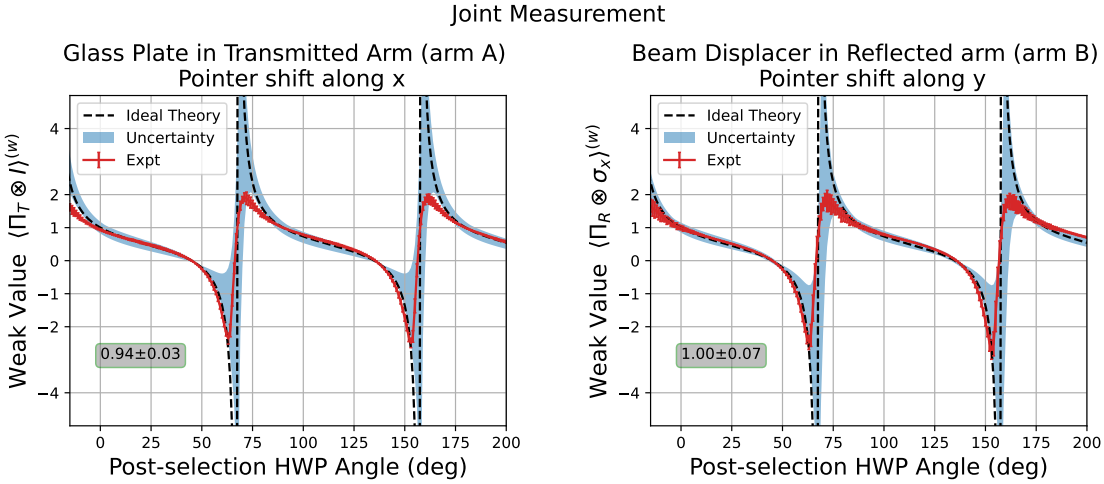
Even though in both cases the beam shifted, the shifts represent the real part of the weak value of the operator Π_- which has eigenvalues 1 and 0. Mapping them to eigenvalues of σ_x, we infer that when the beam displacer is placed in T arm, up to experimental uncertainty, the displacement suggests no interaction with the σ_x polarization component (weak value of σ_x is 0). When the beam displacer is placed in R arm, the displacement is opposite and can be mapped to weak value 1 of the σ_x operator.

Next, we consider joint weak measurements where glass plate is placed in arm B coupled to horizontal beam displacement (along x) and the beam displacer is placed in arm A coupled to vertical beam displacement (along y).

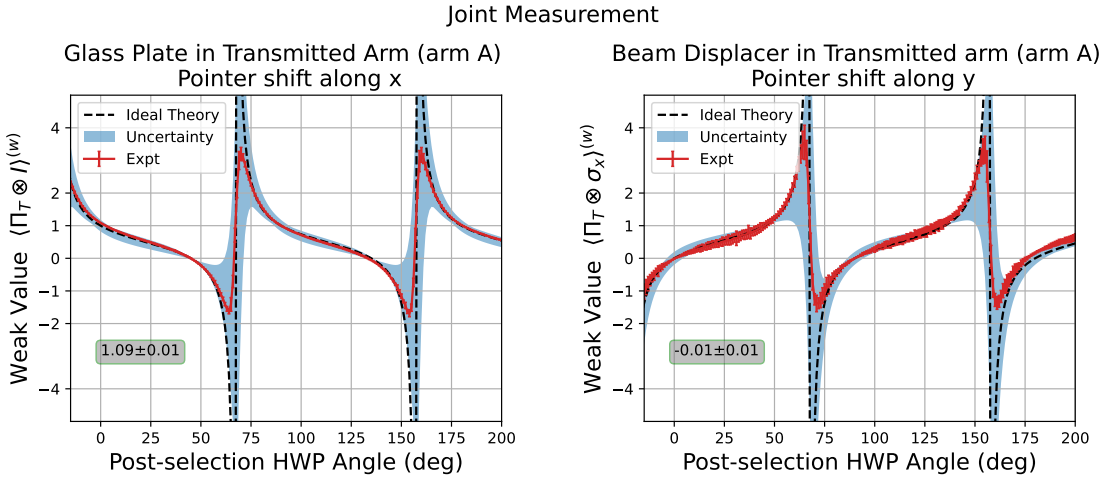


.

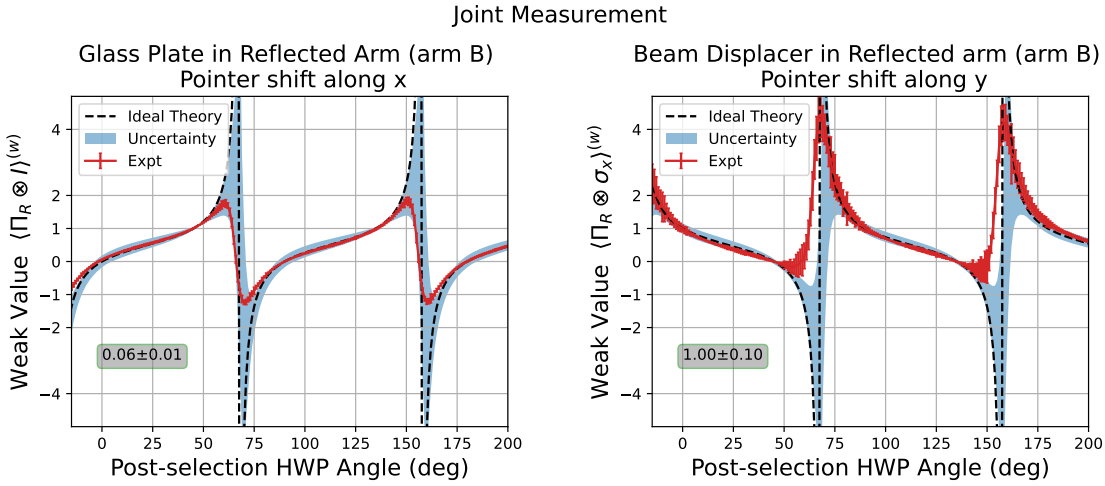
The joint measurement are consistent with the individual measurement case and the experimentally obtained weak value are close to zero.



When the beam displacer is in arm B and the glass plate is in arm A, the respective pointer shifts are consistent weak value of 1 for the respective observables.



When both the glass plate and the beam displacer are placed in the same arm, only the glass plate cause a pointer shift in arm A and the beam displacer leads to a null weak value When both in arm B, the beam displacer causes the pointer shift, while the glass plate does not.



Thus, the joint weak measurements are verified to be consistent with the individual weak measurement in the Mach Zehnder interferometer subjected to pre- and post-selection.

There exists systematic deviation of the experimental data from the theory. In the amplification region, the major factor is the fluctuation in phase which introduces mixedness. The alignment also plays a crucial role in determination of the symmetry about the amplification region. The details of estimating error is presented in the appendix E.

The weak values obtained suggests the experimental realization of the quantum cheshire cat if one applies it to single photons. This can be done as the optical equivalence theorem

would be applicable here i.e, the average results obtained with statistics of single photon source would correspond to that of an experiment performed with coherent laser source [3.54]. This argument has been used in other experimental realization of paradoxes [3.7] in pre- and post selected system. However, for completeness and we also demonstrate the effect with single photon source, the details of which are presented in the next chapter.

In this experiments, we perform weak measurement on two observables jointly and without disturbing the set up within the pre and post -selection. Ideally, one would like to obtain all four weak values to demonstrate the partial cheshire cat effect. However, we are limited by number of independent pointers for a single photon. By using the centre of transverse spatial profile as the pointer we have two independent pointers and thus can obtain the weak value of two observables at a time. But by design of the setup, i.e, by virtue of pre- and post-selection, ABL rule would imply that the photon takes the path A (arm T) and for the single photon version of the experiment, it would imply that the photon does not take the arm B (arm R). This can of course be verified by measuring the weak value of either of the 'which-path' projectors. However, in the same run of the experiment we have also obtained the signature for non-zero σ_x polarization component in arm B. This joint observation, thus suggests signature of presence of property in one arm of the interferometer simultaneously with the signature of the photon being present in the other arm. Thus, a cheshire cat is demonstrated. The paradox is applicable to the case where there is only one photon within the interferometer that is pre- and post-selected. This is experimentally realized with a single photon source as we shall discuss in the next chapter.

References

[3.1] Alberto Peruzzo et al. "A Quantum Delayed-Choice Experiment". In: *Science* 338.6107 (2012), pp. 634–637. DOI: 10.1126/science.1226719.

[3.2] Vincent Jacques et al. "Delayed-Choice Test of Quantum Complementarity with Interfering Single Photons". In: *Phys. Rev. Lett.* 100 (22 2008), p. 220402. DOI: 10.1103/PhysRevLett.100.220402.

[3.3] Vincent Jacques et al. “Experimental Realization of Wheeler’s Delayed-Choice Gedanken Experiment”. In: *Science* 315.5814 (2007), pp. 966–968. DOI: 10.1126/science.1136303.

[3.4] Yoon-Ho Kim et al. “Delayed “Choice” Quantum Eraser”. In: *Phys. Rev. Lett.* 84 (1 2000), pp. 1–5. DOI: 10.1103/PhysRevLett.84.1.

[3.5] A. Danan et al. “Asking Photons Where They Have Been”. In: *Phys. Rev. Lett.* 111 (24 2013), p. 240402. DOI: 10.1103/PhysRevLett.111.240402.

[3.6] Zong-Quan Zhou et al. “Experimental observation of anomalous trajectories of single photons”. In: *Phys. Rev. A* 95 (4 2017), p. 042121. DOI: 10.1103/PhysRevA.95.042121.

[3.7] K.J. Resch, J.S. Lundeen, and A.M. Steinberg. “Experimental realization of the quantum box problem”. In: *Physics Letters A* 324.2 (2004), pp. 125–131. ISSN: 0375-9601. DOI: https://doi.org/10.1016/j.physleta.2004.02.042.

[3.8] Ming-Cheng Chen et al. “Experimental demonstration of quantum pigeonhole paradox”. In: *Proceedings of the National Academy of Sciences* 116.5 (2019), pp. 1549–1552. DOI: 10.1073/pnas.1815462116.

[3.9] Lewis Carroll and Hugh Haughton. *Alice in Wonderland ; and, Through the looking glass and what Alice found there*. Penguin Classics, 2013.

[3.10] M S Leifer and R W Spekkens. “Logical Pre- and Post-Selection Paradoxes, Measurement-Disturbance and Contextuality”. In: *International Journal of Theoretical Physics* 44.11 (2005), pp. 1977–1987. DOI: 10.1007/s10773-005-8975-1.

[3.11] Matthew F. Pusey. “Anomalous Weak Values Are Proofs of Contextuality”. In: *Phys. Rev. Lett.* 113 (20 2014), p. 200401. DOI: 10.1103/PhysRevLett.113.200401.

[3.12] Jeff Tollaksen. “Pre- and post-selection, weak values and contextuality”. In: *Journal of Physics A: Mathematical and Theoretical* 40.30 (2007), pp. 9033–9066. DOI: 10.1088/1751-8113/40/30/025.

[3.13] L. Vaidman. “Weak value controversy”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2106 (2017), p. 20160395. DOI: 10.1098/rsta.2016.0395.

[3.14] Lev Vaidman et al. “Weak value beyond conditional expectation value of the pointer readings”. In: *Phys. Rev. A* 96 (3 2017), p. 032114. DOI: 10.1103/PhysRevA.96.032114.

[3.15] Q. Duprey et al. “The Quantum Cheshire Cat effect: Theoretical basis and observational implications”. In: *Annals of Physics* 391 (2018), pp. 1–15. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2018.01.011.

[3.16] Yakir Aharonov et al. “Quantum Cheshire Cats”. In: *New Journal of Physics* 15.11 (2013), p. 113015. DOI: https://doi.org/10.1088/1367-2630/15/11/113015.

[3.17] Y Aharonov and L Vaidman. “Complete description of a quantum system at a given time”. In: *Journal of Physics A: Mathematical and General* 24.10 (1991), pp. 2315–2328. DOI: 10.1088/0305-4470/24/10/018.

[3.18] Tobias Denkmayr et al. “Observation of a quantum Cheshire Cat in a matter-wave interferometer experiment”. In: *Nature Communications* 5.4492 (2014).

[3.19] *Quantum Paradoxes, Ch-Weak Values and Entanglement*. John Wiley & Sons, Ltd, 2005. Chap. 17, pp. 249–263. ISBN: 9783527619115. DOI: https://doi.org/10.1002/9783527619115.ch17.

[3.20] Yelena Guryanova, Nicolas Brunner, and Sandu Popescu. “The Complete Quantum Cheshire Cat”. In: *ArXiV* (2012), arXiv:1203.421.

[3.21] Sandu Popescu Yakir Aharonov Eliahu Cohen. “A Current of the Cheshire Cat’s Smile: Dynamical Analysis of Weak Values”. In: *ArXiV* (2015), arXiv:1510.03087.

[3.22] Monika Richter, Bartosz Dziewit, and Jerzy Dajka. “The Quantum Cheshire Cat Effect in the Presence of Decoherence”. In: *Advances in Mathematical Physics* 2018.7060586 (2018).

[3.23] William K. Wootters and Wojciech H. Zurek. “Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr’s principle”. In: *Phys. Rev. D* 19 (2 1979), pp. 473–484. DOI: 10.1103/PhysRevD.19.473.

[3.24] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman lectures on physics. Vol. 3: Quantum mechanics*. 1965.

[3.25] Yakir Aharonov, David Z. Albert, and Lev Vaidman. “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100”. In: *Phys. Rev. Lett.* 60 (14 1988), pp. 1351–1354. DOI: 10.1103/PhysRevLett.60.1351.

[3.26] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. “The sense in which a "weak measurement" of a spin-½ particle’s spin component yields a value 100”. In: *Phys. Rev. D* 40 (6 1989), pp. 2112–2117. DOI: 10.1103/PhysRevD.40.2112.

[3.27] Lev Vaidman et al. “Weak value beyond conditional expectation value of the pointer readings”. In: *Phys. Rev. A* 96 (3 2017), p. 032114. DOI: 10.1103/PhysRevA.96.032114.

[3.28] A. Matzkin. “Weak values and quantum properties”. In: *ArXiV* (2018), arXiv:1808.09737.

[3.29] P. Ben Dixon et al. “Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification”. In: *Phys. Rev. Lett.* 102 (17 2009), p. 173601. DOI: 10.1103/PhysRevLett.102.173601.

[3.30] Onur Hosten and Paul Kwiat. “Observation of the Spin Hall Effect of Light via Weak Measurements”. In: *Science* 319.5864 (2008), pp. 787–790. ISSN: 0036-8075. DOI: 10.1126/science.1152697.

[3.31] Amir Feizpour, Xingxing Xing, and Aephraim M. Steinberg. “Amplifying Single-Photon Nonlinearity Using Weak Measurements”. In: *Phys. Rev. Lett.* 107 (13 2011), p. 133603. DOI: 10.1103/PhysRevLett.107.133603.

[3.32] Stephan Sponar et al. “Weak values obtained in matter-wave interferometry”. In: *Phys. Rev. A* 92 (6 2015), p. 062121. DOI: 10.1103/PhysRevA.92.062121.

[3.33] J. T. Muhonen et al. “Coherent control via weak measurements in ^{31}P single-atom electron and nuclear spin qubits”. In: *Phys. Rev. B* 98 (15 2018), p. 155201. DOI: 10.1103/PhysRevB.98.155201.

[3.34] G. J. Pryde et al. “Measurement of Quantum Weak Values of Photon Polarization”. In: *Phys. Rev. Lett.* 94 (22 2005), p. 220405. DOI: 10.1103/PhysRevLett.94.220405.

[3.35] Dawei Lu et al. “Experimental realization of post-selected weak measurements on an NMR quantum processor”. In: *New Journal of Physics* 16.5 (2014), p. 053015.

[3.36] Jeff S. Lundeen et al. “Direct measurement of the quantum wavefunction”. In: *Nature* 474 (2011), pp. 188–191.

[3.37] J. S. Lundeen and A. M. Steinberg. “Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy’s Paradox”. In: *Phys. Rev. Lett.* 102 (2 2009), p. 020404. DOI: 10.1103/PhysRevLett.102.020404.

[3.38] Sacha Kocsis et al. “Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer”. In: *Science* 332.6034 (2011), pp. 1170–1173. ISSN: 0036-8075. DOI: 10.1126/science.1202218.

[3.39] Q. Duprey and A. Matzkin. “Null weak values and the past of a quantum particle”. In: *Phys. Rev. A* 95 (3 2017), p. 032110. DOI: 10.1103/PhysRevA.95.032110.

[3.40] A. Matzkin. “Observing Trajectories with Weak Measurements in Quantum Systems in the Semiclassical Regime”. In: *Phys. Rev. Lett.* 109 (15 2012), p. 150407. DOI: 10.1103/PhysRevLett.109.150407.

[3.41] Arun Kumar Pati, Uttam Singh, and Urbasi Sinha. “Measuring non-Hermitian operators via weak values”. In: *Phys. Rev. A* 92 (5 2015), p. 052120. DOI: 10.1103/PhysRevA.92.052120.

[3.42] A. Danan et al. “Asking Photons Where They Have Been”. In: *Phys. Rev. Lett.* 111 (24 2013), p. 240402. DOI: 10.1103/PhysRevLett.111.240402.

[3.43] Christopher Ferrie and Joshua Combes. “How the Result of a Single Coin Toss Can Turn Out to be 100 Heads”. In: *Phys. Rev. Lett.* 113 (12 2014), p. 120404. DOI: 10.1103/PhysRevLett.113.120404.

[3.44] L. Vaidman. “Weak value controversy”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2106 (2017), p. 20160395. DOI: 10.1098/rsta.2016.0395.

[3.45] D. Sokolovski. “Weak measurements measure probability amplitudes (and very little else)”. In: *Physics Letters A* 380.18 (2016), pp. 1593–1599. ISSN: 0375-9601. DOI: https://doi.org/10.1016/j.physleta.2016.02.051.

[3.46] W.M. Stuckey, Michael Silberstein, and Timothy McDevitt. “Concerning Quadratic Interaction in the Quantum Cheshire Cat Experiment”. In: *International Journal of Quantum Foundations* 2.1 (2016), pp. 17–31.

[3.47] Raul Corrêa et al. “‘Quantum Cheshire Cat’ as simple quantum interference”. In: *New Journal of Physics* 17.5 (2015), p. 053042.

[3.48] H. De Raedt K. Michielsen Th. Lippert. *Mysterious quantum Cheshire cat: an illusion*. 2015. DOI: 10.1117/12.2185770.

[3.49] David P. Atherton et al. “Observation of a classical Cheshire cat in an optical interferometer”. In: *Opt. Lett.* 40.6 (2015), pp. 879–881. DOI: 10.1364/OL.40.000879.

[3.50] James M. Ashby, Peter D. Schwarz, and Maximilian Schlosshauer. “Observation of the quantum paradox of separation of a single photon from one of its properties”. In: *Phys. Rev. A* 94 (1 2016), p. 012102. DOI: 10.1103/PhysRevA.94.012102.

[3.51] Yosep Kim et al. “Observing the quantum Cheshire cat effect with noninvasive weak measurement”. In: *npj Quantum Information* 7.1 (2021), p. 13. ISSN: 2056-6387. DOI: 10.1038/s41534-020-00350-6.

[3.52] Yosep Kim et al. “Observing the “quantum Cheshire cat” effect with noninvasive weak measurement”. In: (2020). DOI: 10.1038/s41534-020-00350-6.

[3.53] E. C. G. Sudarshan. “Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams”. In: *Phys. Rev. Lett.* 10 (7 1963), pp. 277–279. DOI: 10.1103/PhysRevLett.10.277.

[3.54] “Optical Coherence and Photon Statistics”. In: *Quantum Theory of Optical Coherence*. John Wiley & Sons, Ltd, 2006. Chap. 2, pp. 23–182. ISBN: 9783527610075. DOI: https://doi.org/10.1002/9783527610075.ch2.

Chapter 4

Experiment with Single Photons

Contents

4.1 Single Photon source 207
4.1.1 Source Characterization 207
4.1.2 Quantifying Multi-photon events 213
4.2 Experimental Setup 215
4.2.1 Mach Zehnder Interferometer 215
4.2.2 Phase Stabilization 218
4.2.3 Coupling with obervables 219
4.2.4 Alignment of Setup 220
4.3 Data Analysis 222
4.3.1 Data acquisition 222
4.3.2 with pulsed laser 222
4.3.3 with single photon source 223
4.3.4 Results 226

In the previous chapter, we demonstrated the experimental realization of quantum Cheshire cat with coherent laser light source. Although, the result of weak value obtained from the ensemble of single photons would be same as the one obtained with coherent source, it is important to perform the quantum cheshire cat experiment with single photons due to the conceptual formulation of the paradox presuming particle nature to which we associate localized properties. In this chapter, we first characterize the nature of the

spontaneous parametric down conversion (SPDC) source and establish that the heralded photons counts (coincidences) do exhibit single photon characteristics. Then we discuss the experimental implementation mostly concentrating on the changes that needed to be done in comparison with the single photon experiments. Certain details on data analysis and characterization of the setup are in the Appendix. For this chapter, the schematic of the setup is similar to the laser experiment in the previous chapter and is as shown below :

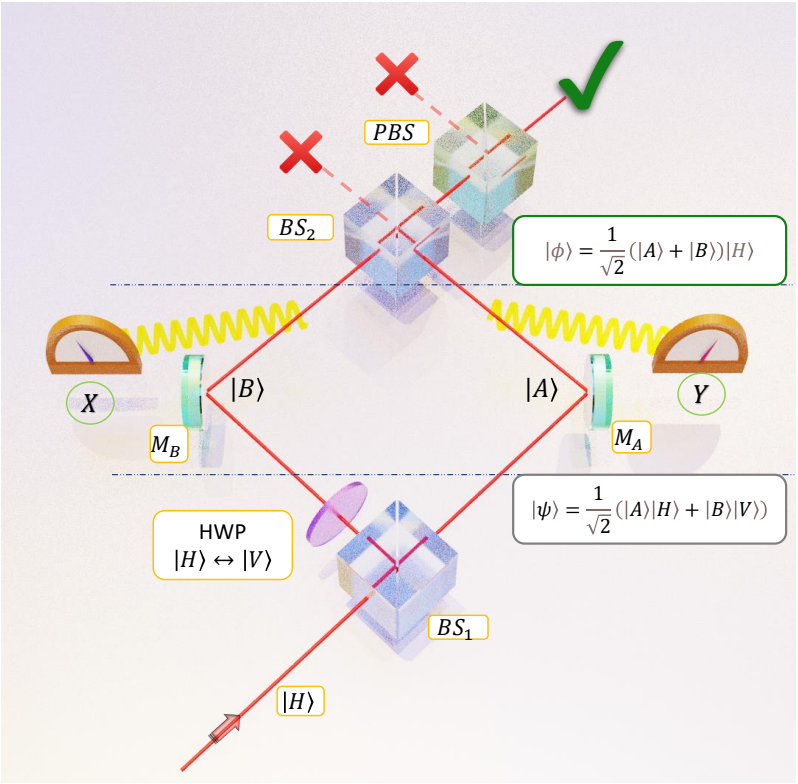


Figure 4.1: Schematic of the partial cheshire cat setup with pre and post selection denoted. The coupling of the observables occur within the interferometer between pre and post-selection.

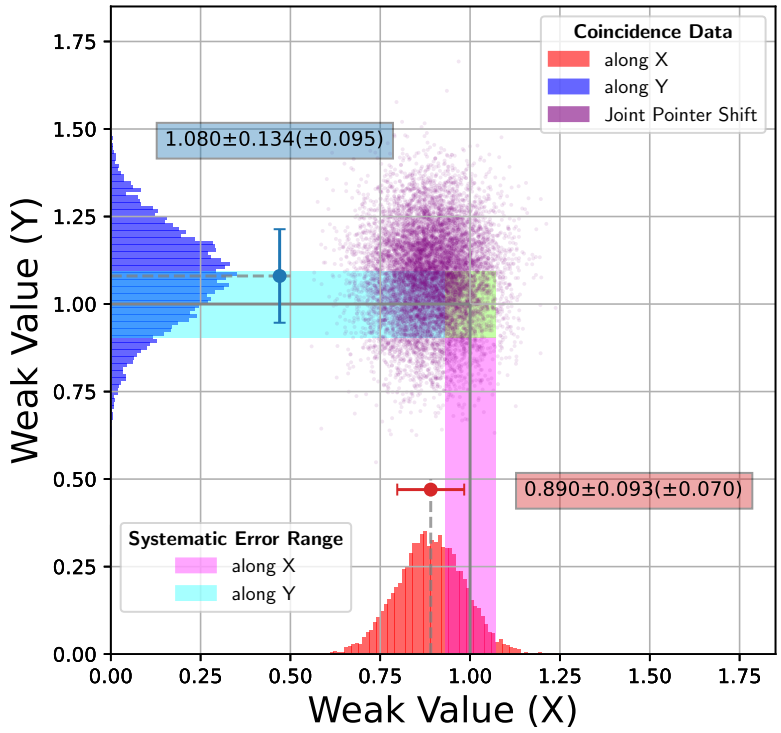


Figure 4.2: The partial Cheshire Cat: The distribution of the pointer shifts corresponding to the weak values. The glass plate in arm A causes the shift along y and the apparatus in arm B causes the shift along x for the pre-and post-selected heralded single photons. This indicates signature of presence of property in the arm B even we verify the expectation from the ABL rule by measuring the weak value of the projector in arm A

With the 'which-path' interaction in arm A and interaction with 'σ_x' polarization component in arm B, we have the distribution of weak values along x and y (the horizontal and vertical transverse profile). The joint detection results in a diagonal shift of the transverse centroid. Quantitatively, this indicates the presence of photon in arm A and simultaneous signature of diagonal polarization component in arm B. Using ABL rule we can argue that the photon subjected to pre and post selection takes the path A. Using the sum rule of $\Pi_A + \Pi_B = \mathbb{1}$, we can infer the weak value of the projector in arm B by knowing the weak value of the projector in arm A.

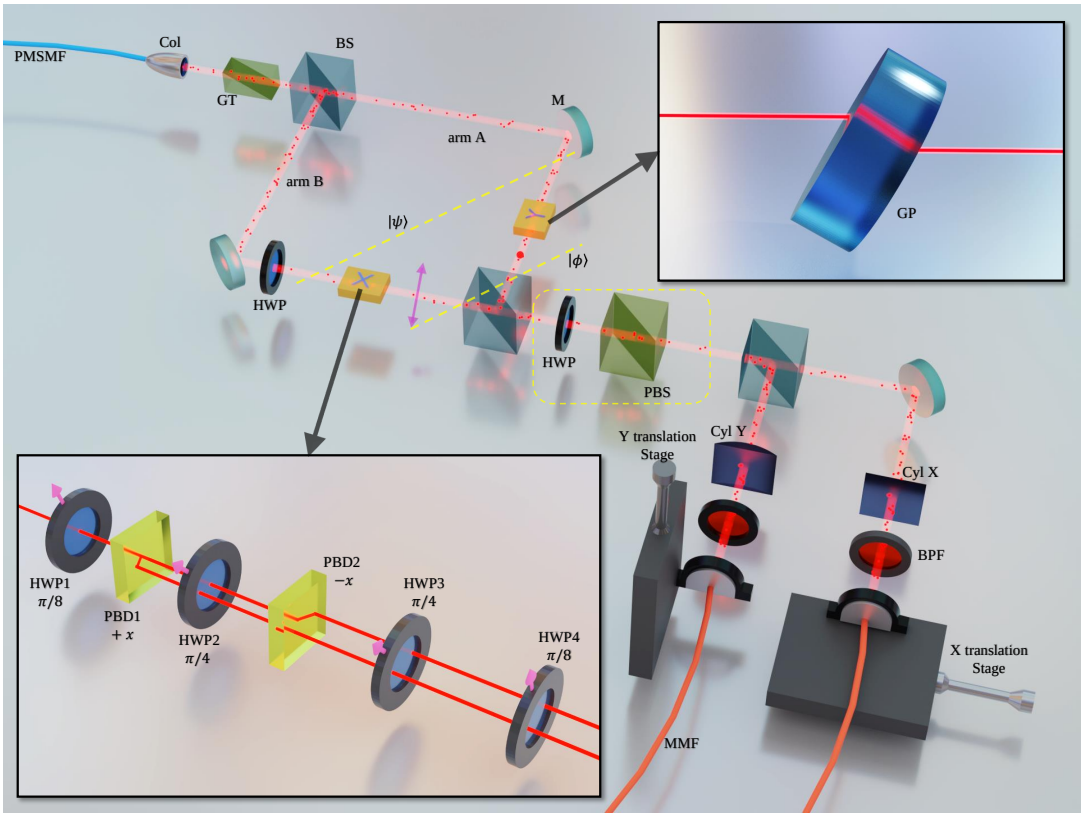


Figure 4.3: In the single photon version of the partial cheshire cat experiment, we use the glass plate in arm A to cause a vertical shift (in contrast with the laser vesrion of the experiment discussed in the previous chapter.) This is because, the experimental realization of the σ_x weak interaction requires more component and it is easier to adjust the tilt of the components horizontally and thus σ_x couples to the shift of the centre of the beam along the horizontal of the table.

The simultaneous measurement in a single experimental setup, i.e, where no components change between the pre and post-selection, indicates that we can apply the results for interaction with a single particle and hence realize the partial quantum cheshire cat effect. For this, with the given SPDC source, we measure the g_2 value which indicates the quantum properties of light.

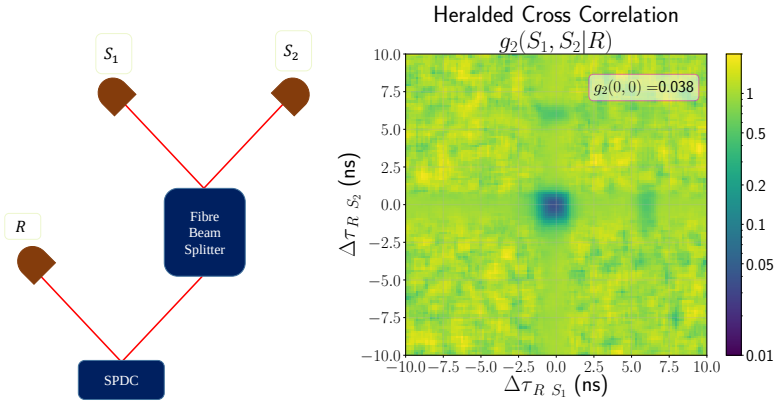


Figure 4.4: Experimental Setup and result to measure the g_2 value for the heralded SPDC source (triple coincidence). The low value of g_2 indicates the particle nature of light.

In the next section, we shall discuss the characterization of the source. We shall determine its coherence properties as well as weak value which is inherently a result of interference. Any incoherence would not give us expected results as discussed in appendix I. Later, we discuss the experimental methods in which we ensure coherent overlap of all the components of the beam, besides methods for the experimental implementation. Finally, we show the pointer shifts as a function of post-selection to further substantiate the claim.

4.1 Single Photon source

4.1.1 Source Characterization

The paradox of partial Chehsire cat (PCC) is posed for a single particle. The weak value computed from a statistics of ideal single photon source is same as the weak value obtained from a coherent source. Thus, we could have claimed the demonstration of partial cheshire cat from the laser version of the experiment. However, the result may explained by interference of electromagnetic waves and do not necessarily pose a paradox in the domain of quantum mechanics. Therefore, it is important to ascertain that the source used is indeed as single photon source. At the same time, since the PCC experiment is based on quantum interference, we also need to establish that the experiment is designed such that the single photons can interfere.

Coherence Length

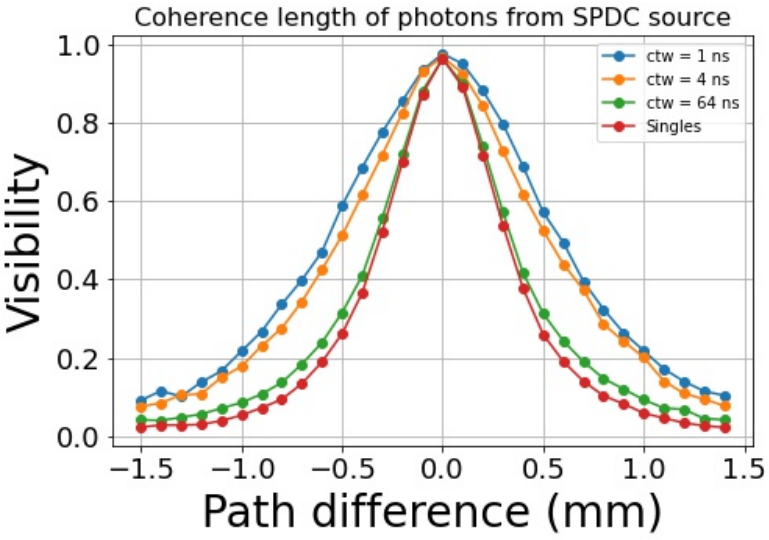


Figure 4.5: Visibility as a function of path delay. The coherence length of the source is inferred from this graph.

Anti-bunching

From the sagnac photon source, we have two streams of photons using various combinations of fibers. About 3-4 fibers are used per stream covering a total length of about 30 m. Due to insertion losses, the counts at the detector would be slightly lower. In one of the streams of the photons, we used a 50:50 fiber beam splitter FBS [Fused 22-810-5/125 50:50 PM, yellow jacket, S+K)]. The outputs of the FBS are connected to detectors D_1 and D_2 and the signals are labeled as S_1 and S_2 (signal 1 and 2). The other stream which goes to D_3 is the heralding arm and is called reference R. The count rates at the three detectors were 1.043 MHz, 538.6 kHz, and 569.3 kHz for R, S_1, and S_2 respectively when the pump power was set to 20 mW from the laser. Hardware delays were added to the channels corresponding to S1 and S2 so that they have coincidences with R at the same delay which we can translate and use as our zero delays. The raw time tags acquired using UQD for 100s were saved to disk.

In any given time duration for data acquisition, the second-order coherence function can be computed for a given coincidence time window ctw as follows:

$$g_2(\tau_1, \tau_2; ctw) = N_R \frac{C_{S_1,S_2|R}(\tau_1, \tau_2)}{C_{S_1|R}(\tau_1) C_{S_2|R}(\tau_2)} \tag{4.1}$$

Here, Here, N_R is the number of detections in R during the acquisition. $C_{(S_1,S_2|R)}(\tau_1, \tau_2)$ is the number of triple coincidences for the given time window ctw when S_1 is delayed by τ_1 and S_2 is delayed by τ_2. Similarly, $C_{(S_1|R)}(\tau_1)$ and $C_{(S_2|R)}(\tau_2)$ are coincidences computed with the same time window ctw between S_1 and S_2 with delays τ_1 and τ_2 from R respectively.

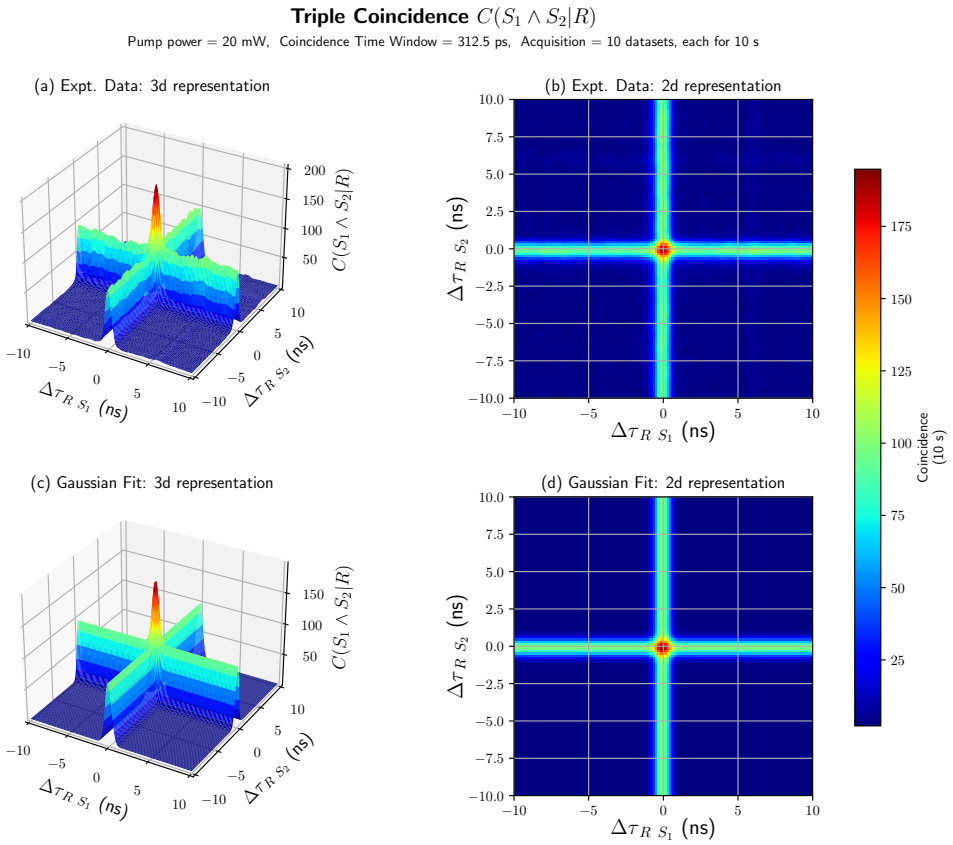


Figure 4.6: Triple coincidences

When delays τ_1 and τ_2 are large, then g_2 is expected to be 1. However, this would require the background of the triple coincidence to be reasonably high. However, since the probability of triple coincidence is low, for the background to be uniform and non-zero, one may need to acquire it for a longer time, increase the flux or increase the coincidence time window ctw. Here, we stick to 20 mW of pump power and 312.5 ps (2 uqd units) of the coincidence time window. We divide the 100s data into 10 sets of 10s and present the

mean calculated over all datasets. The acquisition time is important as one would need some non-zero background triple coincidence in that time scale.

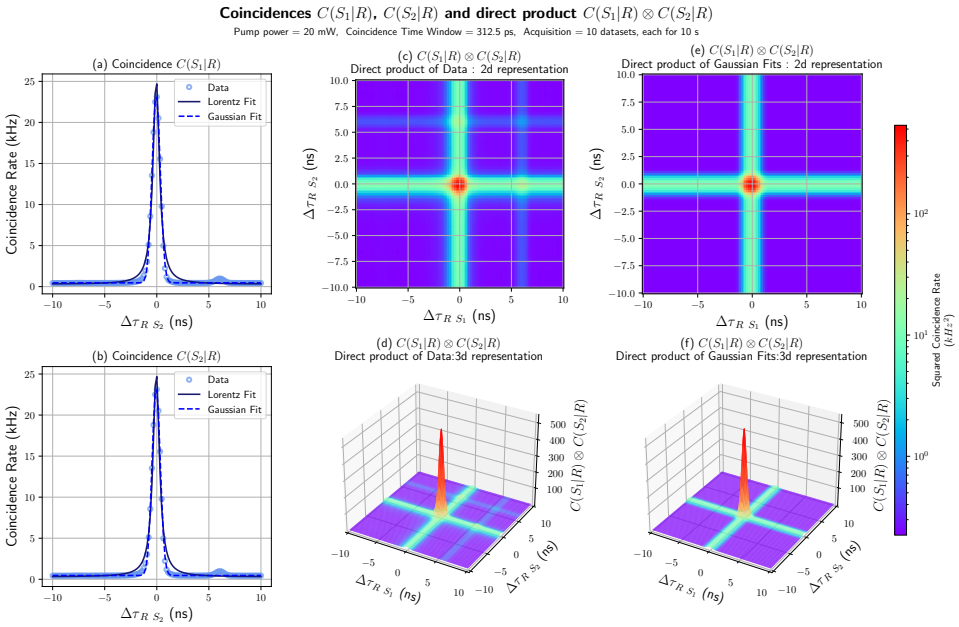


Figure 4.7: Direct product of individual coincidences as function of delay

The triple coincidence is presented as a function of delays τ_1 and τ_2. In (a) and (b), we present the mean (over 10 sets) of triple coincidence accumulated in 10s. In (c) and (d), we fit the data with the sum of two Gaussians $g_1(\tau_1) + g_2(\tau_2)$. The triple coincidence implies at zero delays, the unconditional probability of detecting photons is higher. At about 6 ns along τ_1 and τ_2, we see a faint above average background in (b). The effect of this will be pronounced in the g_2 plots.

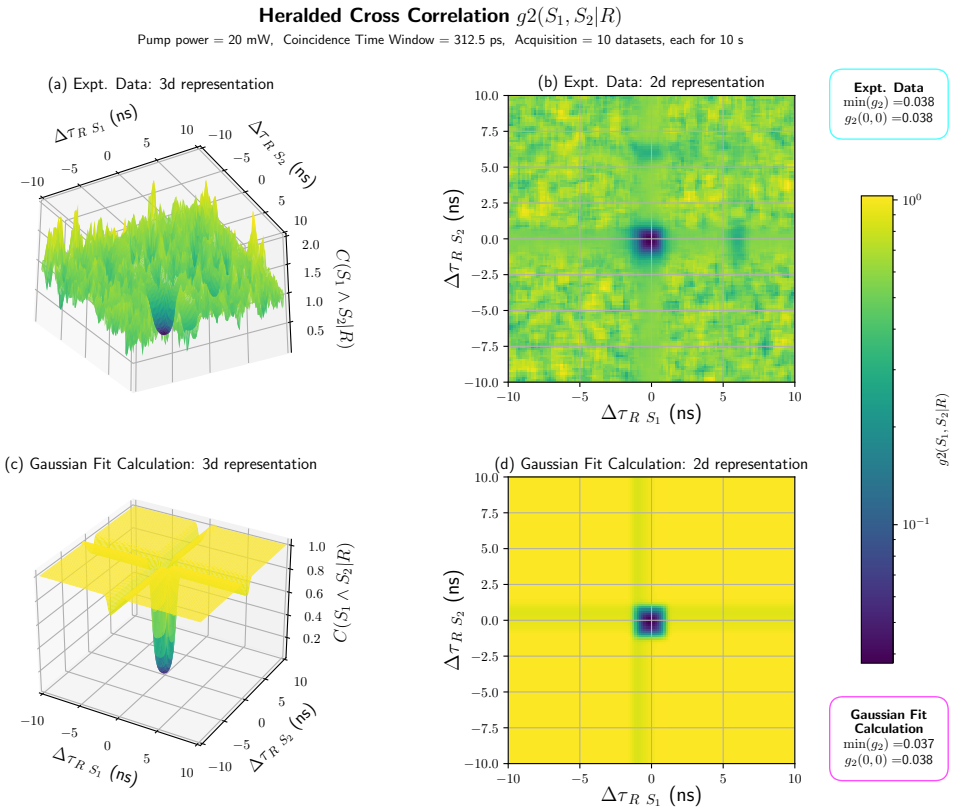


Figure 4.8: g_2 as a function of delay between two signals

Denominator of the coherence function is the direct product of coincidence of S_1 with R and coincidence of S_2 with R. In the individual plots (a) and (b), we can see that the coincidence with the chosen ctw is about 20-25 KHz. Also, a Gaussian fit matches the data better than the Lorentzian fit. The σ turns out to be about 155 ps. This is of the order of 350 ps timing resolution/jitter quoted by Count detectors (which is not clear whether it is RMS, FWHM, also its dependent upon wavelength). Thus, the width entirely arises from electronic jitters in the process of photodetection.) The maximum allowed bandwidth is governed by 3 nm filters (even without PCC setup, filters are in source setup as well). From $\Delta E\Delta t \geq \hbar/2 \Rightarrow \hbar\Delta\nu\Delta t \geq \hbar/2 \Rightarrow \Delta t \geq \frac{1}{2\Delta\nu} = \frac{\lambda^2}{2cd\lambda} \approx 0.36ps$. Given the resolution of UQD is 156.25 ps, arrival time uncertainty due to bandwidth is unobservable.

At about 6 ns delay in plots (a) and (b), we see there is a small peak. The location is quite close to the specified dead time of the UQD time tagger (5 ns). However, in principle, it could be after pulsing from the detector but 6 ns is much less than the dead time of the detector. It could also be the sideband if spectrally the beam hits the edge of the spectral window of the bandpass filter. It also could be because of multiple reflections from any

of the fiber connectors. The effect of this tiny peak can be seen, when we take the direct product in (c) and (d). Although the peak at 6 ns seems tint compared to the peak at 0 delays, the peak is significant compared to the background (about twice). This would mean that when we compute g2, at the region where the tiny peak is, we would obtain a lower g2.

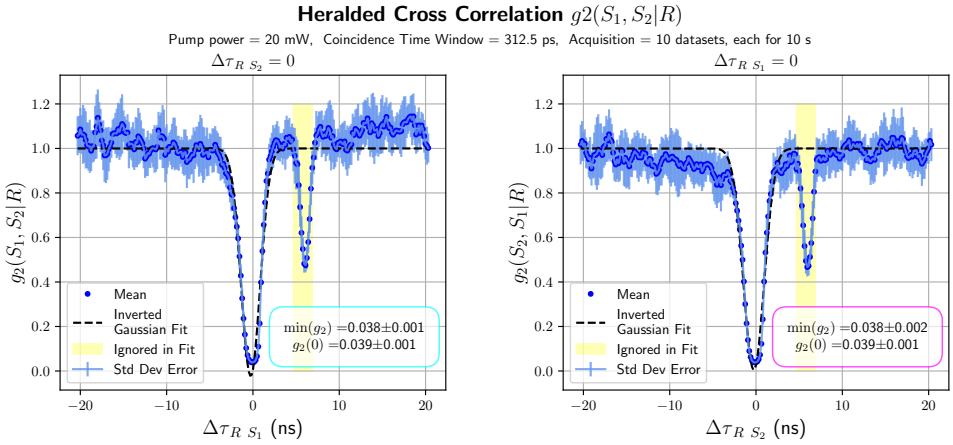


Figure 4.9: g_2 as a function of delay one signal keeping the other fixed

The minima of $g_2(\tau_1, \tau_2)$ is 0.038 which is the same as $g_2(0,0)$. From the g_2 calculated from Gaussian fit, we also get very similar values. Heralded g_2 is computed from the raw data (avg. over 10 sets) and is shown in (a). The background is noisy and not uniformly 1 due to fluctuations (Need 10 or 100 times longer acquisition and computation to make it smooth). In (b) we use the log scale of color function so that the effects of the tiny peak at 6 ns are not visually dominant. In (c) and (d), we compute g2 from the Gaussian fits obtained from the earlier two plots. 10) However, when we plot the 1d slice of g_2, i.e $g_2(\tau_1, \tau_2 = 0)$ or $g_2(\tau_(1 = 0), \tau_2)$ in linear scale, we see that at 6 ns the value of g_2 becomes less than 0.5 as well. The Gaussian fit ignores the data in the yellow shaded region. The Gaussian fit here is on the g_2 data itself, unlike the earlier figure. The error bars on data are the standard deviation of g_2 computed of 10 datasets. The minima of g_2 occurs very close to the value of $g_2(\tau_1 = 0, \tau_2 = 0)$ and is about 0.038.) However, $g_2(0,0)$ is proportional to the mean photon number (count rate) and the chosen coincidence time window. We can arbitrarily go to lower and lower $g_2(0)$ by choosing smaller and smaller values of ctw and by choosing a lower flux (and consequently acquire data for ages). Since, technology in timing resolution limits us to above 3 orders of magnitude, the value of $g_2(0)$ does not

reflect how good the single-photon source is but how precise the timing electronics are. If we can resolve femtoseconds then only $g_2(0)$ would be limited by the photon source characteristics.

4.1.2 Quantifying Multi-photon events

For the measurement of the pointer variable, we need to compute the centre of the transverse profile that emerges from the ensemble of single photons detected over time and as the multi-mode fiber is scanned. The paradox of 'Quantum Cheshire Cat' is posed when there is single particle in the interferometer. If there are multiple particles inside the interferometer one could argue that some of the particles interacted with the glass plate i.e., the 'which-path' projector in one arm where as others interacted with the σ_x polarization component in the other arm. Thus, the phenomenon of Cheshire Cat could not be applicable of each and every particle and could potentially be dismissed as a statistical manifestation rather than an attribute to each particle. Therefore, it is important to quantify the contribution of the multi-photon events to the construction of the transverse beam profile.

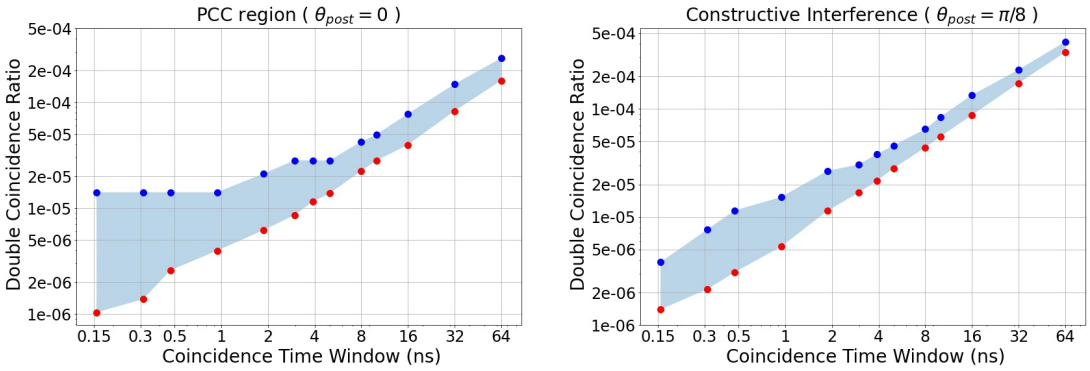
The length of each arm of the interferometer is about 1 m. The time taken by light to travel 1 m is about $3ns$. In fact, the interaction of the photon with σ_x components takes at most 1 ns because all the components necessary for that interaction are placed with 12 inches or 30 cm in the reflected arm after the corner cube retro-reflector and the mirror. The glass plate in the transmitted arm of the interferometer is only 1 mm thick and hence the interaction time is much smaller i.e. 3 ps. Therefore, ideally the single photon source should not have two emissions or more within 3 ns for us to claim that no two particles were there inside the interferometer. However, even if there are two photons but at least separeted by 1 ns in time, we can claim that only one particle was interacting with components that gives rise to the weak value of the observables.

However, the one stream of particles from the SPDC source do not behave as a true single photon source. There is contamination of multi-photons. Typically, one only considers the subset of particles that gets heralded to obtain the statistics of what would be close to the ideal single photons source. However, if the count rate is small enough the contribution of multi-photon events remains negligibly small in a small time period. Since we collect the photons using multi-mode fiber with 50 micron core diameter the maximum

rate of photon detection is of the order of kHz with which we can claim that the probability of multi-photon events to be really small if we consider the time interval of the order of ns. There is a beam-splitter which divides the beam so that two multi-mode fibers can collect it. One of the fibers is placed after cylindrical lens that compresses the vertical profile and the fiber is scanned horizontally to reconstruct the profile. The other is placed after cylindrical lens that compresses the horizontal profile and the fiber is scanned vertically to reconstruct the vertical profile. The core of the fiber being smaller than the beam do not collect the entire beam. But only the portion of the beam that gets collected eventually gives us the pointer variable and hence to quantify the multi-photon events we would use only the subset of photons that are collected.

To maximize the count rates, the two fibers were placed at the peak of horizontal and vertical gaussian profile respectively. The time stamps of arrival of photon with the two detectors where recorded with UQD Time Tagger with a timing resolution of 156.25 ps. The typical count rate of the Horizontal MMF was about 1.74 kHz and that of the Vertical MMF was 1.14 kHz when the post-selection HWP angle was set to zero i.e., at the post-selection where we expect the partial Cheshire Cat phenomenon to occur. When the post-selection angle was made to be $\pi/8$, we had the maximum in terms of the count rate which were 3.23 kHz and 2.12 kHz for the Horizontal and the Vertical MMF respectively.

The coincidence between the photo-detections corresponding to the two MMFs was computed as a function of the coincidence time window which was varied from 0.15 ns to 64 ns. The coincidence was computed for a delay in one of the fibers which was varied on either side with a range of 20 times the coincidence time window used. As a function of delay, coincidence was almost a flat line. We divide the maximum of the coincidence as a function of delay and the average with the $\sqrt{(N_{s_H} \times N_{s_V})}$, where $N_{s_{H,V}}$ denotes the singles count rate in the respective Horizontal/Vertical Channel. The acquisition time over which the data was collected and analyzed is 100 s. However, note that, since both the numerator and denominator have the dimensions of the time-rate, the acquisition time do not affect the analysis.



(a) Ratio of Double Coincidence to Singles as function of coincidence time window with post-selection HWP angle set to 0. Here, we expect the partial cheshire cat phenomenon. The shaded region shows the region between average and maximum of double coincidence ratio.

(b) Ratio of Double Coincidence to Singles as function of coincidence time window with post-selection HWP angle set to $\pi/8$. Here, we expect maximum singles rate. Therefore, this would give us the upper limit of mult-photon contribution.

From the graphs above, we see that up to the interval of 16 ns, the contribution of two-photons to the collected statistics is less than 10^{-4}. This is low enough to affect the centroid of the beam.

4.2 Experimental Setup

4.2.1 Mach Zehnder Interferometer

The stream of single photons of wavelength 810 nm (bandwidth about 2 nm) is made incident on the Mach Zehnder Interferometer (MZI) from a polarization maintaining single mode fiber $PMSMF_s$ [PM780-HP, Thorlabs] to minimize pointing fluctuation of the beam. A suitable collimating lens COL [F240FC-780, Thorlabs] is used to get a beam size of about 1.5 mm and minimize divergence. The beam is passed through a Glan-Thompson polarizer GT [GTH5-B, Thorlabs] to ensure the polarization is horizontal with high degree of purity before entering the 50:50 beam splitter BS_1 [BS014, Thorlabs]. To ensure that the path difference between two arms of the MZI is negligible compared to the coherence length of the stream of detected photons, we use a corner cube retroreflector CCR [PS976M-B, Thorlabs] mounted on an actuator [ZST225-B, Thorlabs] attached to a 3D translation stage in one of the arms (here in arm B). The actuator is adjusted so that the path difference is ensured to be within about 2 microns. The CCR is used instead of a

mirror assembly to avoid angular beam deviation upon translation. The CCR is attached to a piezo (osi-stack) which is used to stabilize the phase difference between the two arms of the MZI described later. However, since the CCR itself creates an additional path difference, this is macroscopically compensated using the three mirrors M_{T_1},M_{T_2} and M_{T_3} [NIR 5102, Newport] in the A arm of the MZI. The CCR, however, introduces ellipticity in the polarization of the beam in arm B. This is corrected using a half-wave plate HWP_c [WPA03-H-810, Newlight Photonics] followed by quarter-wave plate QWP_c [WPA03-Q-810, Newlight Photonics] and the polarization is made vertical. Additional polarizing beam splitter PBS_T is introduced after the three mirrors in arm A to further purify the polarization.

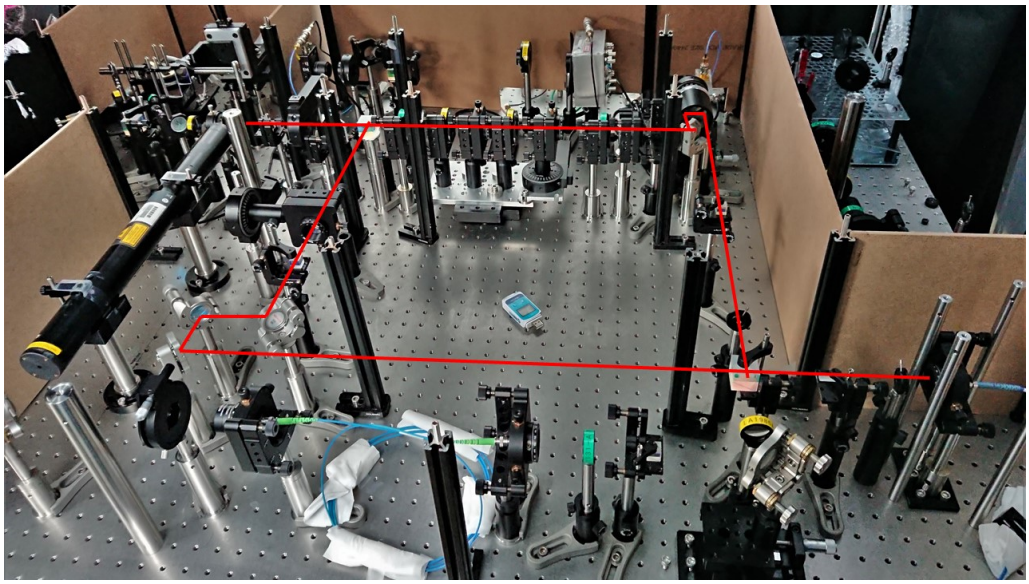


Figure 4.11: Picture of Experimental set up

The second beam splitter BS_2 [BS014, Thorlabs] is fixed at the intersection point of the two beams and oriented to roughly ensure collinearity of MZI. For fine alignment, the tip tilt degrees of freedom of the mirror mounts are used to ensure collinearity and the overlap of the beams is ensured by translation of the CCR. The overlap of beams is ensured down to 5 microns (although the precision of measuring the centre is much higher at 0.1 microns the spatial noise in the beam limits the accuracy) and angle between the two beams emerging out of the second beam splitter (BS_2) is ensured to be less than 10^{-5} radians (actual value may be around 10^{-7} radians as estimated from the fringe stability of the interferometer). A half wave plate HWP_{post} [WPZ0-200-L/2-810, Castech] mounted

in a motorized rotation mount [PRM1/MZ8, Thorlabs] is placed before a polarising beam splitter (PBS_{post}) at one of the exit ports of the second beam splitter (BS_2). The post selected state changes with the change in the angle of the HWP_{post}. In addition to the angle of interest two more post selection angles are chosen to measure the reference pointer positions associated with the eigen values 0 and 1 respectively. This enables in evaluating the weak value for a particular post selection from the measured pointer shift.

The pre-selected state ($|\psi\rangle$) is prepared after the compensating waveplate QWP_c on arm B and after PBS_T in arm A .The post-selected state is obtained by back evolving the transmitted component of the post-selection PBS (PBS_{post}) to a time before the second beam splitter BS_2 of the MZI. Such a state, which is thus guaranteed to be transmitted in the post-selection PBS (PBS_{post}), is given by $\phi(\theta) = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle) \otimes S(\theta)|H\rangle$. Here $S(\theta) \equiv HWP(\theta)$ is the Jones matrix for HWP whose fast-axis is orientated at angle θ from the Horizontal in the post selection.

A band pass filter BP10 [FB810-10, Thorlabs] followed by another BP3 [LL01-810-25, Semrock] are used after the post-selection PBS (PBS_{post}) to prevent light of other wavelengths from being detected and also to narrow the linewidth down to at most 3 nm. The bandwidth of the filter determines the minimum coherence length. Visibility of the interference upon a suitable (diagonal, at $\theta = 22.5^\circ$) polarization post-selection is measured as function of the path length difference by moving the actuator on which the CCR is mounted. Finally, the actuator is left at the position where maximum visibility is obtained.

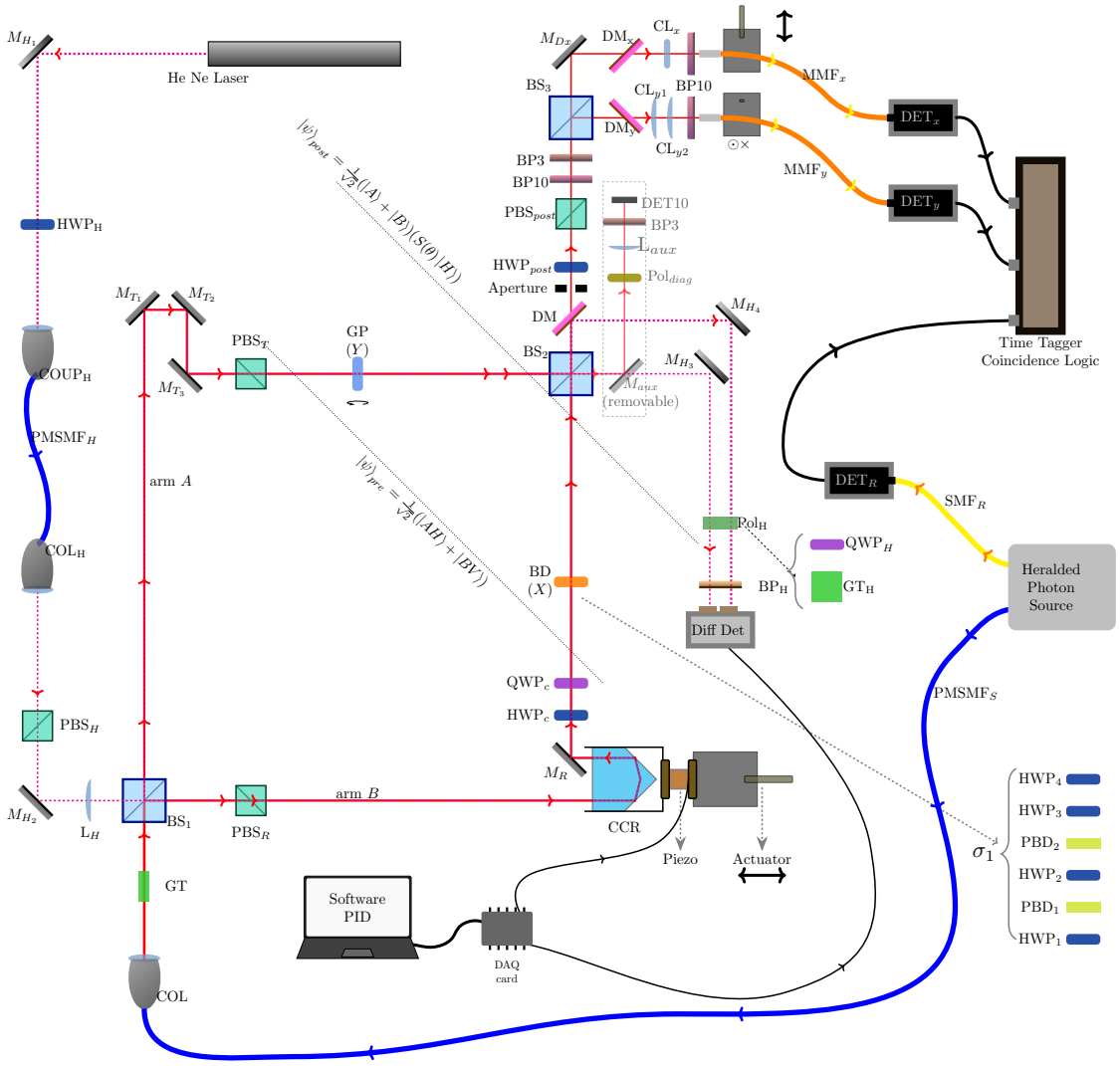


Figure 4.12: Schematic of the experimental setup

4.2.2 Phase Stabilization

The relative phase between the two arms of the interferometer can drift due to mechanical and acoustic vibrations and therefore the CCR needs to be moved accordingly in order to maintain a constant phase relationship between the two arms. This is achieved by using the piezo (attached to the CCR) which contracts or expands depending on the voltage provided to it, thus causing the CCR to move. For this a PID algorithm is implemented on a computer along with a DAQ card [USB-6003] which is used to generate and receive voltages. The phase shift in the interferometer is monitored by measuring the difference in power in the two output ports of the second beam splitter BS_2 for a Helium Neon laser (He

Ne) beam of wavelength 633 nm, inserted into the MZI from the other input port of the first beam splitter (BS_1). This beam is mostly blocked by the Dichroic Mirror DM [DMLP735, Thorlabs] placed after BS_2 in addition to the band pass filters (BP10 and BP3). Before the differential detector (Diff Det), the polarizations of the He Ne beams coming from two output ports of BS_2 are transformed into circular basis using quarter-wave plate QWP_H and then projected to horizontal polarization using GT_H to achieve maximum visibility. The differential intensity signal obtained from 633 nm beam is calibrated with intensity of 810 nm laser measured at the other output port (at which the post selection is not performed) of BS_2 with diagonal polarization projection using photodetector [DET10/M, Thorlabs] as a function of voltage given to the piezoelectric transducer. When the path difference is within one wavelength (See Appendix D) these two signals become almost linear and this is how the differential intensity signal is used to stabilize the phase difference within low uncertainty.

4.2.3 Coupling with obervables

'which-path' observabe A parallel window GP [WG41010-B, Thorlabs] is placed in arm A and is tilted vertically (along y) to cause a vertical shift of about 50 microns in the beam. Tuning the angle of tilt of the glass plate, the shift in the beam can be controlled. If the shift were more than the beam width of 1.8 mm, observing this shift after the post-selection would have indicated whether the photon came from arm A or arm B. The shift of about 50 microns being much less than the beam width ensures that the observable ($Y_A = \Pi_A$) is weakly coupled.

Polarization component in a given arm A polarising beam displacer (PBD) allows the ordinary component of polarization of the beam incident on it to pass through without any deviation and causes a lateral shift (depending upon its thickness) in the path of the extraordinary component of polarization of the incident beam. Thus the operation of a PBD can be considered as σ_z measurement operation. σ_x measurement operator can be constructed from the σ_z measurement operation. Here the σ_z measurement operator is created using two beam displacers PBD_1 and PBD_2 [PDC 12005, Newlight Photonics] oriented in such a way that one of them causes about 50 microns shift in the extraordinary

component along one direction (say +x) and the other shifts the extraordinary component by about 50 microns in the opposite direction (-x) horizontally. A half wave plate HWP_2 (fast axis at $\pi/4$) is inserted between the two beam displacers so that the extra-ordinary beam for the first PBD becomes the ordinary component for the second PBD and vice versa. This is followed by another half wave plate HWP_3 (with fast axis at $\pi/4$ so that the ordinary and the extra-ordinary polarized beams have the same phase in the description of the σ_z operator. The whole σ_z measurement operator is placed between two half wave plates HWP_1 and HWP_4 [WPA03-H-810] with their fast axes oriented at angle $\pi/8$ to realise the coupling σ_x operator in the polarization degree of freedom.

4.2.4 Alignment of Setup

Ideally, the zero reference of the beam, on a 2D plane after the post-selection would be at the position where both the beams from arm A and arm B merge. They are expected to be on top of each other at the detection plane and collinear. However, when we insert components for weak interaction with the σ_x polarization component in arm B, we need to tilt the beam displacers in order to adjust the phase shift between the emergent e-ray and o-ray so that in the limit of the beam displacement between them going to zero the evolution operator due to the interaction remains identity. Because of the tilt as well as slight angular deviation of beam from components, the beam in the arm B gets refracted after the components are inserted. This causes a slight non-collinearity and translation of the beam. The collinearity change is very fine and can be adjusted by maximizing the visibility when post selection is either diagonal or anti-diagonal. This is achieved by the tip/tilt of the mirror. The displacement between diagonal and anti-diagonal component is fixed by the alignment in components for the σ_x interaction. However, they should ideally be on either side equidistant from the centre of the beam coming from arm A. The above cannot be simply verified by measuring the centre of the beam coming from arm A with arm B blocked and vice versa. This is because, when arm A is blocked, the pre-selected polarization in arm B is orthogonal to the post-selected polarization. Due to the displacement between diagonal and anti-diagonal components, we have the destructive interference pattern after post-selection. The centre of the destructive interference pattern is very error prone. Therefore, the fact that the centre of the beam in arm A and arm B

coincides we have to use the resultant interference pattern formed when the post-selection is diagonal and anti-diagonal. From theory, these two displacements needs to be symmetric about the centre of the beam of arm A. This is achieved by translating the CCR and measuring the resultant centroid at the three post-selection configurations. Once the above alignment is ensured, the glass plate in arm A which was back aligned, is now tilted so that it causes the desired shift along the vertical direction when arm B is blocked. All the above steps are performed with pulsed laser and data was acquired with beam profiler. The 2D centroid of the beam, after being scaled to be the weak value, as the post-selected state is changed is plotted below. The scaling can be achieved by subtracting the ensured zero position as mentioned earlier and scaling it with the pointer shift corresponding to the eigenvalues with individual arms blocked. This can be done with beam profiler data only as blocking and insertion of components are all performed with beam profiler as the monitoring tool for the pulsed laser source. However, once we change the measuring device, we lose the information about absolution position in different configuration. Hence, for consistency, the plot below is obtained after considering the pointer shift at post-selection HWP angle as 0 and the same at 90 deg as 1. Note that 90 deg is just a repetition of the 0 deg for a HWP and hence it appears that we are calibrating the very point that we are trying to measure. However, in the next section, we show that the data agrees with the weak value calculated theoretically not only the PCC point (HWP angle 0) but also for all post-selected states. Further, we achieve weak value amplification, in which the pointer shift is beyond the eigenvalue range. This unambiguously shows that the resultant pointer shift is due to coherent weak interaction and not by any other incoherent means that may occur if the measurement interactions were disturbing.

Then we take the beam profiler out and use the multimode fibers mounted on a translation stage. The weak value as a function of post-selection angle is measured to ensure that the alignment of multi-mode fibers is consistent with the beam profiler data. Once the above is ensured, we move to single photon source to measure singles and coincidences. Here, we scale the pointer shift to weak values by subtracting the pointer shift at post-selection angle 45 deg which from we know should correspond to the zero pointer shift. Similarly, displacement by weak value 1 is when post-selection HWP angle is 90 deg. We need to take these values as reference, as with the single photon source,

changing any thing in between the pre and post-selection may change the alignment.

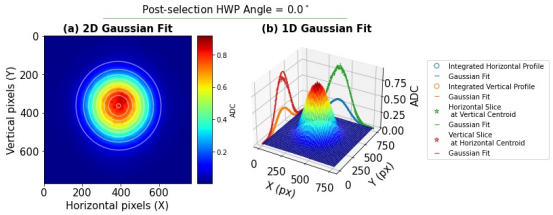
For further details on how the pulsed laser was used to align the interferometer, how the extra path length due to the CCR was calculated and compensated, kindly refer to the Appendix F. One of the major challenges to measure centroid of the beam is the mechanical stability of various components which results in pointing fluctuations, the characterization of which is presented in Appendix H.

4.3 Data Analysis

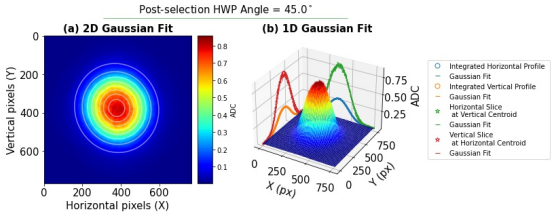
4.3.1 Data acquisition

4.3.2 with pulsed laser

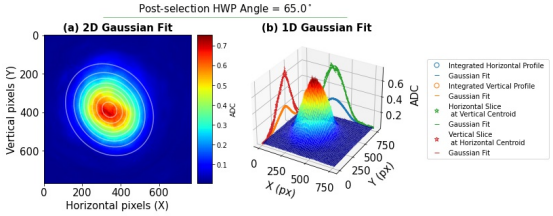
In order to subsequently use a single photon source, a 100 fs pulsed laser [Mira 900 D Ti:Sapphire, rep rate: 76 MHz] is used with average power of about 2 mW to ensure good alignment. Although the coherence length of the pulsed laser is lower (than CW), it is slightly enhanced by the use of the bandpass filter so that the pulsed laser has almost the same central wavelength and the bandwidth as the single photon source to be used later. A beam profiler [Dataray UCD -15] is placed after the bandpass filter. The post-selection HWP angle is rotated to the angle 67.5° that creates the destructive interference profile along the diagonal. The phase difference is adjusted to make this destructive interference pattern as symmetric as possible along the diagonal. The slight spatial noise in the beam profile i.e., deviation from perfect Gaussian profile typically would dominate the destructive interference pattern. The overall power incident on the beam profiler also reaches the minimum which is another indicator for the correct phase difference to which the Mach Zehnder interferometer is locked to. The post selection HWP angle is rotated and 5 images are captured for each HWP angle. The first order centroid of the 2D images is computed. The pointer shifts are scaled to the weak value using the pointer shifts at 45° and 90° as reference.



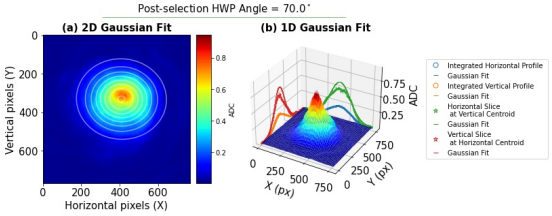
(a) Post Selection angle 0° (PCC region)



(b) Post Selection angle 45°



(c) Post Selection angle 65°



(d) Post Selection angle 70°

Figure 4.13: The destructive interference between the two beams coming from two arms occurs when the post-selection angle is 67.5*degrees*. We observe the weak value amplified diagonal shift if we compare the beam profiles for post-selection angles on either side of it. The same diagonal shift does exists if we compare the beam profile in the PCC region with the beam profile when the post-selection angle was 45°. However, this is not easily noticeable given the shift is much less compared to the beam width.

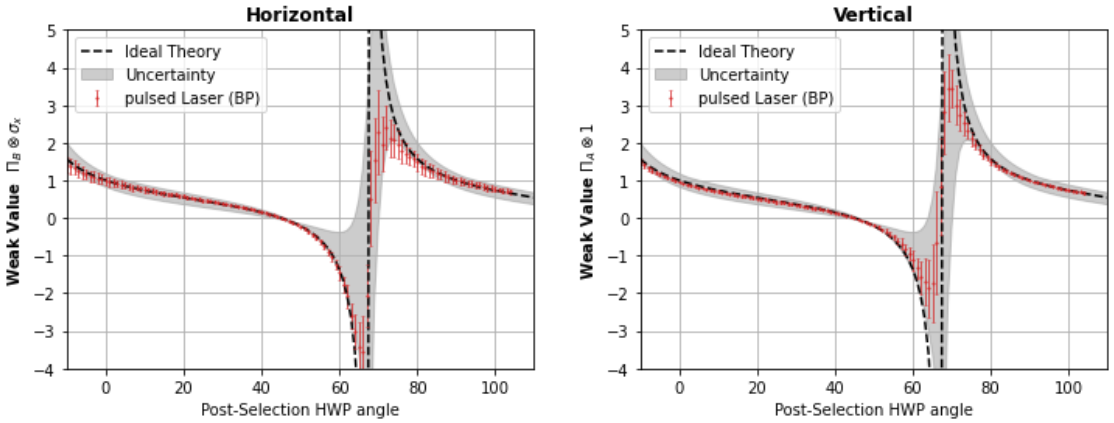
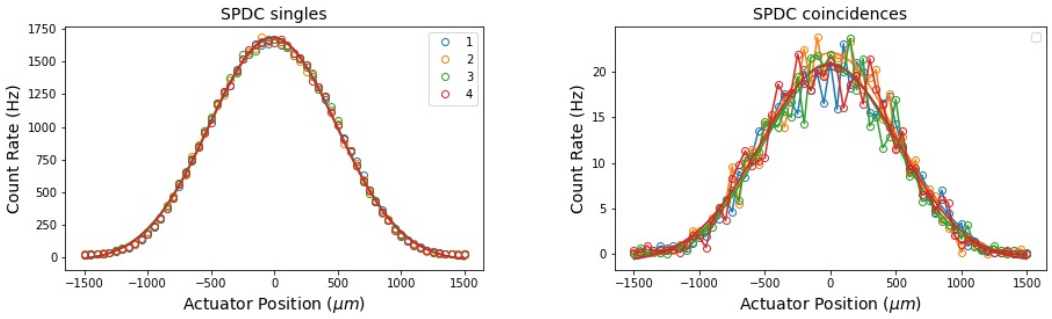


Figure 4.14: Weak Value obtained with the pulsed laser source as a function of post-selection HWP angle

4.3.3 with single photon source

For the single photon source, the timing information along with the spatial profile is needed to be measured. The timing information ensures that the contribution of multi-photon events to the pointer shift is negligible. For this SPAD detectors [Tau SPAD-20, Pico quant] are used. The photons are collected using a bare multimode fiber [M42L02, Thorlabs] with

core diameter of 50 microns. These fibers (MMF_x and MMF_y) are moved in steps of 50 microns to sample the Gaussian profile. A width of about 3 mm is covered with 61 points. Although the use of such fibres averages the intensity over the 50 microns, the precision with the centroid is much better than this and depends on the total number of photons collected. To enhance the collection, a (combination of) cylindrical lens (CL_x, CL_{y_1} and CL_{y_2}) is used to compress one spatial dimension. The multi-mode fiber tip is placed at the focus of the cylindrical lens and is translated to obtain the spatial profile. Since the objective is to jointly observe the two pointers, a 50:50 beam splitter BS_3 [BS014] is used to divide the beam. One beam is used to reconstruct the horizontal profile (X) and the other beam is used to reconstruct the vertical profile (Y). Since the beam splitter sends the photon towards the horizontal or the vertical detector with inherent randomness and there is no change of experimental settings within the interferometer (i.e., between pre and post-selection), the horizontal and vertical centres of the beam are jointly obtained for the ensemble of photons for a given post-selected state.



(a) Profile of Singles

(b) Profile of coincidences

Figure 4.15: Singles and coincidences over four datasets and their gaussian fit

We compute the center of the beam by fitting the profile with a Gaussian.

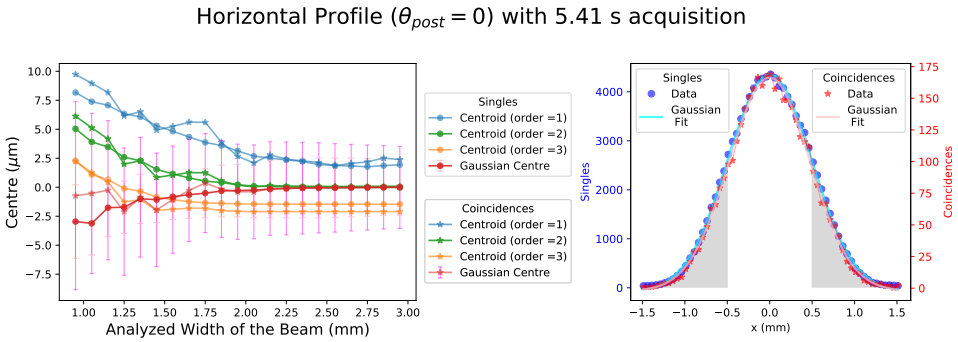


Figure 4.16: The second order centroid

With the single photon source, we had two datasets to generate the weak value curve as a function of post-selection angle. Since they were taken on different days, we can only combine them after converting them to weak values. For the method being used, let us consider the data for the PCC region i.e. for the post-selection angle zero.

The data consists of the following:

- 5 scans with 1 repetition per position with acquisition time 5 sec per reading (actual time varied from 5.4-5.6 sec)
- 1 scan with 5 repetition per position with acquisition time 5 sec per reading (5.4 sec)
- 1 scan with 3 repetition per position with acquisition time 3 sec per reading (3.4 sec)
- 1 scan with 3 repetition per position with acquisition time 5 sec per reading (5.4 sec)

Here, the scan refers to the movement of actuator to different positions so as to generate the beam profile. The repetition refers to data acquisition at a given detector position.

We have overall 16 readings per position of the multi-mode fiber. However, these 16 readings were taken over different acquisition times. The readings of singles and coincidences after being divided by the acquisition time gives us the singles and coincidence rates as a function of 61 positions of multi-mode fibers. The count rates may drift over time as well because the single photon data was taken over10 hrs or so. Thus, the photon rates are multiplied with R_i/R where R was median count rate in the reference (heralding) detector over the entire dataset and R_i is the reference counts for a given reading. Thus, overall, there are 8 scans of the profile for the PCC point. 3 of those scans have 3 or 5 readings per position. If we treat each position of multi-mode fiber as columns, then effectively we have 16 rows as we have 16 readings. The difference between each rows could be only inadvertent systematic errors which would be accounted for later. Therefore,

it is not needed that we have to consider the beam profiles (Singles/Coincidences vs position) only as the rows.

For each position of MMF, we have about 16 readings. We randomly choose any one of the 16 readings per position and do this for all the positions to create a profile. There are 16^{61} such possible profiles. We consider 10^4 such Gaussian profiles. Note that, we are not resampling the centers of the Gaussian. We are resampling the Singles/Coincidence rates so as to have a Gaussian profile. We fit each such profile with a Gaussian function to determine the centre. This gives us the distribution position X for a given post-selection.

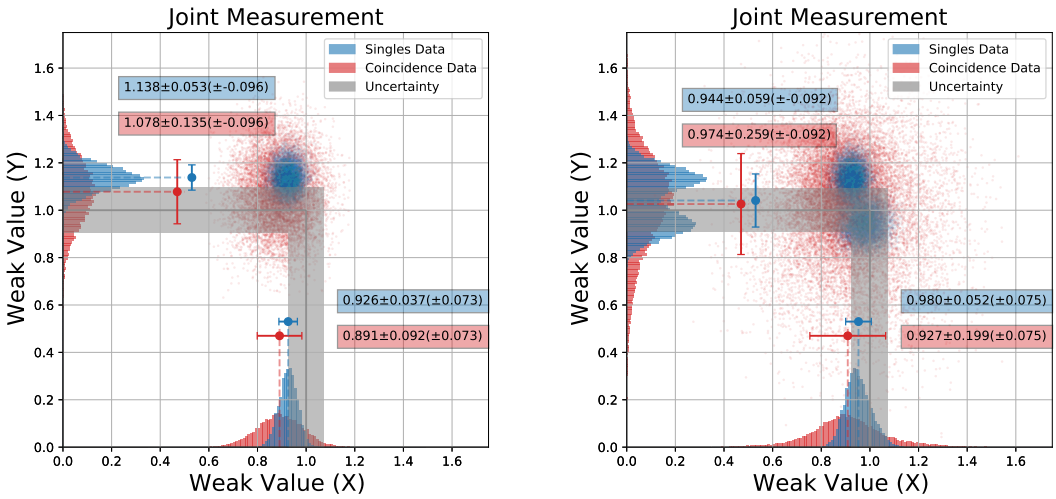
The distribution centres of the beam is obtained for the case when the post-selection angle is 45 degree and let's call that distribution X_0. The 10^4 profiles for the post-selected angle 45 degree were considered from one scan with 3 readings (thus we have 3^{61} possible profiles). Similary, we have one scan with 3 readings at the post-selection angle 90. The distribution of centres of the beam is say X_1 which again is obtained from Gaussian fit of 10^4 profiles. Since the profiles were drawn at random, we can create the distribution of weak values by taking $\langle X \rangle^w = \frac{(X - X_0)}{\langle X_1 - X_0 \rangle}$ The statistical mean is considered in the denominator to avoid affecting the distribution of $X^{(w)}$ when few of the values of from X_1 could be close to few of the values of X_0. The random errors in the experiment are entirely captured in the numerator. The above distribution is computed with singles and coincidences and for both horizontal and vertical shift.

Further details on statistical analysis of the centroid data obtained are presented in Appendix J

4.3.4 Results

Note that the distribution of singles is narrower because the count rate is higher and hence the centre of the Gaussian is inferred more accurately compared to the coincidences. The centre as inferred from singles can differ from the centre as inferred from coincidences because of coincidence do have a slight better coherence compared to the singles.

Systematic error range The gray region around the grid line denoting 1 along x and y represents the range over which the experiment can have systematic error arising from drift of the beam centre. Basically, if we measure few profiles at a particular post-selection angle



(a) One dataset

(b) Weak values of two datasets combined

Figure 4.17: The singles and coincidence distribution shown on X axis in blue and red colors respectively. The mean and 1σ errors from the distribution are shown above graphically and the numbers are mentioned alongside. Similarly, the distribution along Y is represented with histogram with vertical orientation (horizontal bars). Since both the pointer shift occurs simultaneously and are measured simultaneously in the experiment (i.e., without changing anything between pre and post-selection), the overall shift of the pointer would be along the diagonal as inferred from the shift on the projections along horizontal X and vertical Y. The distribution of these diagonal shifts, in the scale of weak value, are represented by the scatter plots.

and then change the post-selection angle to expect a certain shift, it could happen that the observed shift is more or less because the beam's position has drifted over time. This drift typically arises from the effect of temperature, pressure and humidity on opto-mechanics and fibers. The exact beam drift cannot be directly monitored during the experiment (even if we tap off a portion of beam it does does pass through same optics) and hence can only be estimated. We estimate after repeatedly measuring the beam in arm A after post-selection for a long period. Since arm B is blocked, the centroid would represent the drift of the beam's centre over time. Each scan below represents the mean and std deviation of Gaussian centre obtained from 5 readings. We take 5 scans for a given acquisition time and then vary the acquisition time.

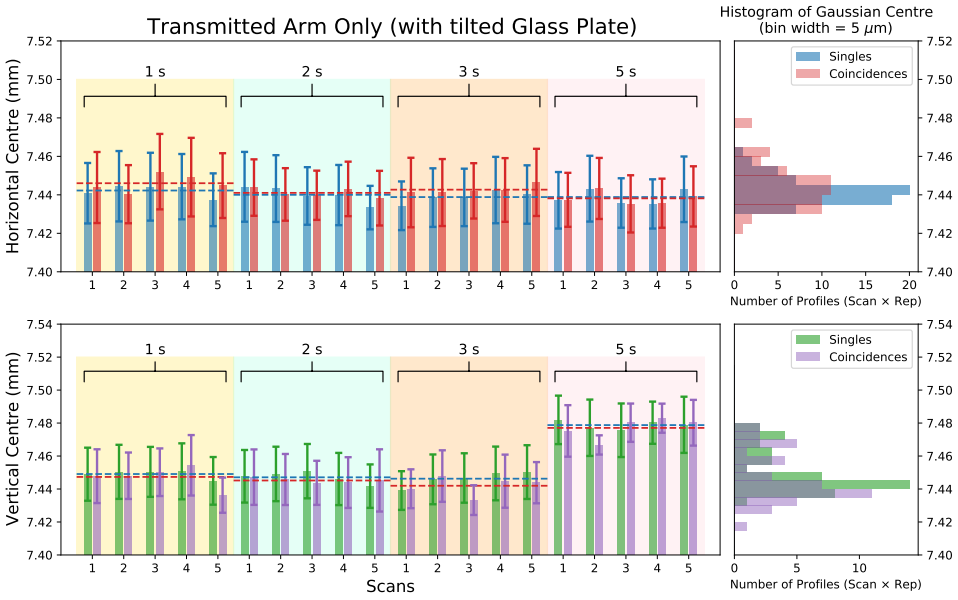


Figure 4.18: **Transmitted Arm only Data** Stability of Singles and coincidences

Since the vertical of the 5s acquisition data is an outlier with significant shift, we consider rest of the data to compute what is the standard deviation of the inferred centers. This turns out to be 8.4 and 10.4 microns for coincidences along X and Y respectively (about 2.2 and 3.1 microns for singles respectively). This uncertainty for coincidences , after being scaled in terms of weak value, is shown as the systematic error range in the plot.

The experimental data are in agreement with the weak value $\langle \Pi_A \rangle^{(w)} = 1$ along with $\langle \Pi_B \sigma_x \rangle^{(w)} = 1$, owing to the joint measurement without changing experimental setup

between pre and post-selection. This implies that we do have signatures of the diagonal polarization component of the photon in the arm B, whereas the photon seems to be present in arm A.

To validate that the above result are indeed weak values resulting from weak measurement, we show the weak value as a function of post-selection angle. We observe weak values for the two observables to be beyond the eigenvalue range.

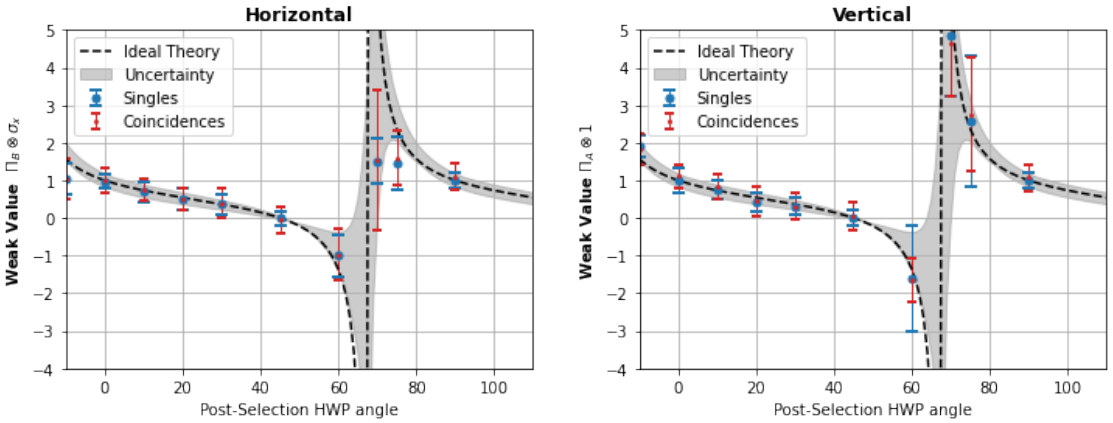


Figure 4.19: Combined data sets after the pointer shift is calculated in terms of the weak value. certain weak values are beyond the eigen value range. Here, the mean and error bars are obtained by propagation of standard deviations from the statistics of individual datasets.

Thus, this experiment demonstrates the partial cheshire cat phenomenon with single photon source by joint measurement of weak values inferred from pointer shifts resulting from weak interaction of the observables with transverse spatial pointers. Therefore, this extends the conclusion of experiment performed in Ch 3 and makes it applicable to a single photon level. This presents us with the experimental realization of the paradoxical signature of polarization property of a photon in an arm in which the photon do not seem to take subjected to boundary conditions in time. The paradox arises if one presumes particle nature of light. In other interpretations of quantum mechanics such as the pilot wave theory where the particle and the wave are separate entity, it could have interesting implications.

Chapter 5

Inferring Expectation value of non-Hermitian operator

Contents

5.1 Polar decomposition of Operators 234
5.1.1 Mathematical Relation with Weak Values 234
5.1.2 Example: Spin Ladder Operators 235
5.2 Inferring Expectation values using Mach Zehnder Interferometers 236
5.2.1 Collinear interferometer 236
5.2.2 Non-Collinear interferometer 239
5.3 The apparent 'non-unitary' evolution 242
5.4 Experimental Implementation 245
5.5 Results 250
5.5.1 Using Mach Zehnder Interferometer 250
5.5.2 Comparison with weak measurement 251
5.6 Outlook 255

In quantum mechanics, we define an observable to be a real dynamical variable whose eigenstates form a complete set [5.1]. The above definition by Dirac is not without caveats that he himself identifies. In a theory, we need to specify what are the observables so that we can relate the possible set of outcomes to what the result an experimentalist gets after making a measurement. For example, we can design an experiment to measure the charge

of a particle, say electron. Of course, we would get a real number as an answer. Whether to consider this dynamical variable, charge of a particle, as an observable or a mere physical constant is ambiguous. The two descriptions are equivalent however if we represent it as an observable with only one possible eigenvalue[1]. Treating it as an observable comes with the burden of describing the mathematical operator whose eigenvalue we have measured. But claiming it to be not an observable would contradict the fact that we have experimentally measured the number. Surely, Dirac argues, it is reasonable to argue that such physical parameters are observables, even though a proof is missing in the mathematical framework.

But can we only observe real values? We can argue that we can always design an experiment where we can measure the real part and imaginary part of a quantity separately. We can combine the two results to claim observation of a complex number. This is routinely done when we report measurements of the complex refractive index of materials [5.2]. The real part has a signature in the phase shift and the imaginary part has the signature in attenuation. So why do we demand that the result of measurement must be a real number in quantum mechanics? It is because, in quantum mechanics, we cannot combine the results of two different experiments and claim a value. In general, let's say the eigenvalue of an operator A is given by the complex number z. It is often possible to decompose A into its real and imaginary parts $A = A_r + iA_i$ such that we also can write $z = z_r + iz_i$ where z_r turns out to be the eigenvalue of A_r and z_i turns out to be the eigenvalue of A_i. We cannot claim observation of A by making measurements of A_r and A_i. In general, they may not commute and hence cannot share simultaneous eigenstates. They cannot be measured one after the other as the first measurement, say of A_r would disturb the state for which we would like the measure the value of A_i. As an example, consider the spin ladder operator for a spin-1/2 particle. We can write $\sigma_\pm$, the spin ladder operator in terms of Pauli observables as $\sigma_\pm = \sigma_x \pm i\sigma_y$. Surely, we can design experiments to measure the value of σ_x and σ_y separately. But since the commutator of the two Pauli observables $[\sigma_x, \sigma_y] = i\sigma_z \neq 0$ is not zero for a spin half particle, we may not measure them simultaneously. If we measure σ_x first then the state of the system would reduce to one of the two eigenstates of σ_x. The following measurement of σ_y would indeed be on the eigenstate of σ_x instead of the

[1] If from theory, we know that something has one possible eigenvalue, we do not get any information upon measurement.

original prepared state.

From the previous example, we may be tempted to conclude that complex numbers, in general, cannot be measured as an outcome of measurement corresponding to a single observable. But on a closer look, we find that the eigenvalue(s) of the spin ladder operator[2] $\sigma_{\pm}$ is 0. Indeed, it is a real quantity! Then why did quantum mechanics conspired against it being an observable ? The reason is truly embedded in the second part of the definition of an observable as described by Dirac. The only non-trivial[3] eigenstate of σ_{+} is $|\uparrow\rangle$ (and σ_{-} is $|\downarrow\rangle$). So the eigenstate does not span the Hilbert space to form a complete set. Hence, even though the eigenvalue is real, the spin ladder operators cannot be considered as observable.

The requirement that the eigenstates of an observable must form a complete set comes from the fact that we would like to represent any state as a linear combination of the eigenstates of the observable. Hermitian operators fit the bill by offering a set of orthonormal eigenstates for non-degenerate eigenvalues, in addition to the fact that the eigenvalues themselves are real. Thus, often, observables are represented by Hermitian operators [5.3]. Hermitian operators, however, form a subset of a more generic class of operators called the Normal operators which satisfies a real spectrum as well as have a complete set of eigenstates. An operator N is normal if it commutes with its hermitian conjugate, i.e, $[N,N^{\dagger}] = 0 \Rightarrow NN^{\dagger} = N^{\dagger}N$. We can see that a hermitian operator H with $H^{\dagger} = H$ trivially satisfies to be a normal operator as well. It has been argued that observables should indeed be formulated as normal operators [5.4]. But Hermitian operators do provide additional perks. In von Neumann description of quantum measurement, the observable of the system couples to the pointer variable and the evolution is unitary. If we let go Hermiticity in general, we may have to deal with non-unitary time evolution which would describe non-conservative systems. If we need a real spectrum, the self-adjoint condition can be replaced with a much weaker condition known as $\mathcal{PT}$-invariance [5.5]. Let's consider $A = a\sigma_x + ib\sigma_z$ as an example of a non-hermitian operator. Here, we can interpret $\mathcal{T}$-symmetry as the invariance under complex conjugation. But the concept of parity for non-spatial states is not trivial. We can consider σ_x to be $\mathcal{P}$ operator. Thus, by performing

[2]The spin ladder matrices are not diagonalizable. The eigenvalues here are the solutions to the characteristic polynomial equation, i.e. obtained from solving λ for $\det(A - \lambda\mathbb{1}) = 0$

[3]The trivial eigenstate is (0,0) of course

$\sigma_x A \sigma_x$ and then taking the complex conjugate, we can show that A is $\mathcal{PT}$ symmetric. The two eigenvalues of A are $\pm\sqrt{a^2-b^2}$ corresponding to the eigenstates $(-b\pm\sqrt{a^2-b^2},a)$. The eigenvalues are real for $a>b$. When $a=b$ however, the two eigenstates, which in general are non-orthogonal, become equal. This condition breaks the symmetry in the sense that the eigenstates now do not span the Hilbert space. The transition about these *exceptional points* has a variety of applications in non-hermitian optics [5.6].

Another criteria for the spectrum to be real is pseudo-Hermiticity [5.7]. A operator A is pseudo Hermitian if there exists a transformation η such that $A^\dagger = \eta A \eta^{-1}$. Note that Hermiticity is its special case when $\eta = \mathbb{1}$. With this background, lets return to our example of a spin ladder operator? Is it $\mathcal{PT}$ invariant ? No, because $\sigma_x\sigma_-\sigma_x = \sigma_+$. Is it pseudo Hermitian ? Yes, with $\eta = \sigma_x$. Also, note that $\sigma_\pm$ are not normal as well because $[\sigma_\pm^\dagger, \sigma_\pm] = \mp\sigma_z$.

This begs the question of how to measure some properties on pseudo hermitian operators or in general non-hermitian operators. Since the spin ladder operators, even if pseudo hermitian are still not a candidate for observables owing to the lack of complete eigenspectrum, we cannot experimentally realize them. Thus, properties of the operator in a given state, such as the expectation value of the operator cannot be computed from the statistical distribution of the outcomes of a measurement. Why one would be interested in the expectation value of an operator which is not an observable ? Well, it depends on the exact operator. The expectation value of the class of column non-Hermitian operators can be useful for quantum state reconstruction [5.8, 5.9]. The spin ladder operators for a qubit form a specific case of the column operators.

Further, the non-Hermitian annihilation operators have found application in the description of multi-particle weak values [5.10]. Weak values are in general complex and hence seem to be the suitable candidate to express the expectation value of non-hermitian operators, which in general are complex [5.11]. In this chapter, we shall discuss the theory behind representing the average of non-Hermitian operators with weak values. We shall demonstrate the principle with weak measurement. In addition, we show an interferometric technique in which we can directly evaluate the expectation value of a non-hermitian operator which also serves as a method to experimentally compute weak values without weak measurement or traditional post-selection.

5.1 Polar decomposition of Operators

5.1.1 Mathematical Relation with Weak Values

Consider the operator A, which in general can be non-Hermitian. The expectation value of A in the state $|\psi\rangle$ could, in general, be a complex number $z = \langle\psi|A|\psi\rangle$. In this section, we shall discuss how polar decomposition can be used to recast the complex expectation value of any non-Hermitian operator in terms of a complex weak value of a Hermitian operator. The complex weak value of a Hermitian operator can be experimentally obtained by measuring its real part from the shift of the pointer variable and its imaginary part inferred from the shift of the momentum conjugate to the pointer variable.

Given any operator A, we can always polar decompose as $A = UR$, where U is unitary operator and R is the Hermitian semi-definite operator obtained as $R = \sqrt{A^\dagger A}$.

The expectation value of A can now be expressed in terms of the weak value of R as follows

$$
\begin{aligned}
z &= \langle\psi|A|\psi\rangle && (5.1)\\
&= \langle\psi|UR|\psi\rangle && (5.2)\\
&= \frac{\langle\psi|UR|\psi\rangle}{\langle\psi|U|\psi\rangle}\langle\psi|U|\psi\rangle && (5.3)\\
&= \frac{\langle\phi|R|\psi\rangle}{\langle\phi|\psi\rangle}\langle\phi|\psi\rangle && (5.4)\\
&= {}_{\phi}\langle R\rangle_{\psi}^{(w)}\langle\phi|\psi\rangle\,. && (5.5)
\end{aligned}
$$

Here, $|\phi\rangle = U^\dagger|\psi\rangle$. Given the polar decomposition, we can experimentally measure the weak value of R in the pre-selected state $|\psi\rangle$ and the post-selected state $|\phi\rangle$. The post-selected state is uniquely determined by the unitary U which in turn is unique for a given R. Thus, the expectation value of a non-Hermitian operator can be restated as the transition amplitude mediated by the Hermitian polar component R from the state $|\psi\rangle$ in which the expectation value is to be measured to a state $|\phi\rangle$ into which $|\psi\rangle$ can evolve into with the unitary polar component U. Note that $\langle\phi|\psi\rangle$, in general, can be complex but since $|\psi\rangle$ and $|\phi\rangle$ will be the pre and post-selected state in the weak measurement of R and hence can be always computed. Knowing the weak value of R and the expectation value of U in $|\psi\rangle$,

i.e., $\langle\psi|U|\psi\rangle$, we can infer the complex expectation value $z = \langle A\rangle$.

Conversely, if by some other method, we obtain the expectation value of A and also the expectation value of U, we can infer the weak value of R. Then, we can experimentally obtain a weak value without performing a weak measurement. The protocol for determining weak values does not necessarily require the introduction of the non-Hermitian operator. Our proposal can measure the weak value of any Hermitian operator without involving the weak measurement and the post-selection. As a further application, we also demonstrate how the expectation value of the non-Hermitian operator can be determined from the experiment. Although the method uses optics as an example, but is applicable, in general, to all other quantum systems.

5.1.2 Example: Spin Ladder Operators

In polarization degree of freedom, one interesting class of non-Hermitian operators are the spin ladder operators $\sigma_\pm = \frac{1}{2}(\sigma_x \mp i\sigma_y)$. Let us consider the following operator,

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{5.6}$$

as somewhat analogous to the ladder operators. It converts $|H\rangle \to |V\rangle$ and annihilates $|V\rangle$. This can be polar decomposed with

$$R = \sqrt{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}} = \sqrt{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{5.7}$$

Solving for U, we have $U = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

It turns out that expectation value of this non-hermitian operator A will always be real in all states and hence is not necessarily a generic case where we can demonstrate complex expectation value. To get the expectation value of A, we need the weak value of R, which is Hermitian but not unitary.

5.2 Inferring Expectation values using Mach Zehnder Interferometers

5.2.1 Collinear interferometer

In the collinear Mach-Zehnder, let us assume the polarization state out of the laser to be $|\psi\rangle$. After the first beam splitter(assumed to be a symmetric beam splitter that does not affect polarization) BS_1 the spatial and polarization state can be written as

$$\frac{1}{\sqrt{2}}\left(|t\rangle + i\,|r\rangle\right)|\psi\rangle \tag{5.8}$$

Here $|t\rangle$ & $|r\rangle$ represents the spatial states transmitted and reflected from BS_1 respectively. After passing through the optical elements U and V , which are, as of now generic, in the respective paths can write the state as

$$\frac{1}{\sqrt{2}}\left(V\,|t\rangle + iU\,|r\rangle\right)|\psi\rangle \tag{5.9}$$

After the second beam splitter BS_2, we have

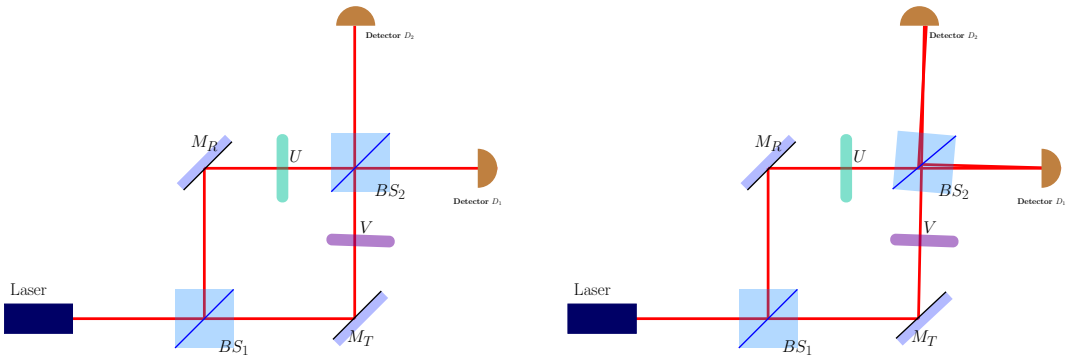
$$|r\rangle \rightarrow \frac{1}{\sqrt{2}}(|1\rangle + i\,|2\rangle) \tag{5.10}$$

$$|t\rangle \rightarrow \frac{1}{\sqrt{2}}(|2\rangle + i\,|1\rangle) \tag{5.11}$$

Here, $|1\rangle$ & $|2\rangle$ represents the state before the respective detectors.

Now since the path difference between the transmitted and reflected arm of the beam splitter may not be in general equal[4], we will absorb the effect by considering a relative phase shift ε between the two arms of the interferometer.

[4]By geometry, both the paths should be equal for ideal stable components. However, since the path difference considered are a fraction of the wavelength, even errors due to the surface flatness of optical components will matter.



(a) Colinear Mach Zehnder (b) Non Colinear Mach Zehnder

Figure 5.1: (a) Collinear and (b) Non collinear configuration of the Mach Zehnder. Note that to get to the non-collinear, after configuring the Mach Zehnder Interferometer in collinear configuration, we rotate the mirror M_T by a very small angle. But only this makes the beam diverge after the BS_2. So to combine the beams we have to rotate the BS_2 slightly as well. The rotations are adjusted to get the desired fringes.

Thus the resultant state can be written as

$$\frac{1}{\sqrt{2}}\left(V\frac{1}{\sqrt{2}}(|2\rangle+i|1\rangle)+iUe^{i\varepsilon}\frac{1}{\sqrt{2}}(|1\rangle+i|2\rangle)\right)|\psi\rangle \tag{5.12}$$

$$=\frac{1}{2}\left(V(|2\rangle+i|1\rangle)+iUe^{i\varepsilon}(|1\rangle+i|2\rangle)\right)|\psi\rangle \tag{5.13}$$

$$=\frac{1}{2}\left(i(V+Ue^{i\varepsilon})|1\rangle+(V-Ue^{i\varepsilon})|2\rangle)\right)|\psi\rangle \tag{5.14}$$

The intensity that can be measured at the respective detectors are given as

$$4I_1=\left\langle(V^\dagger+U^\dagger e^{-i\varepsilon})(V+Ue^{i\varepsilon})\middle|\psi\right\rangle \tag{5.15}$$

$$=\langle\psi|(V^\dagger V+V^\dagger Ue^{i\phi}+U^\dagger Ve^{-i\varepsilon}+U^\dagger U)|\psi\rangle \tag{5.16}$$

$$4I_2=\langle\psi|(V^\dagger-U^\dagger e^{-i\varepsilon})(V-Ue^{i\phi})|\psi\rangle \tag{5.17}$$

$$=\langle\psi|(V^\dagger V-V^\dagger Ue^{i\varepsilon}-U^\dagger Ve^{-i\varepsilon}+U^\dagger U)|\psi\rangle \tag{5.18}$$

Now, the difference between these two intensities gives us

$$2(I_1-I_2)=\langle\psi|V^\dagger Ue^{i\varepsilon}+U^\dagger Ve^{-i\phi}|\psi\rangle \tag{5.19}$$

$$=\langle\psi|V^\dagger U|\psi\rangle e^{i\varepsilon}+\langle\varepsilon|U^\dagger V|\varepsilon\rangle e^{-i\phi} \tag{5.20}$$

Now if $\langle\psi|V^\dagger U|\psi\rangle = z$, where z is a complex number, we have $\langle\psi|U^\dagger V|\psi\rangle = \langle\psi|V^\dagger U|\psi\rangle^\dagger = z^*$

Hence,

$$2(I_1 - I_2) = ze^{i\varepsilon} + z^* e^{-i\varepsilon} \tag{5.21}$$
$$= ze^{i\varepsilon} + (ze^{i\varepsilon})^* \tag{5.22}$$
$$= 2\,\mathrm{Re}(ze^{i\varepsilon}) \tag{5.23}$$
$$= 2\,\mathrm{Re}(|z|e^{i\varphi}e^{i\varepsilon}) \tag{5.24}$$
$$\Rightarrow (I_1 - I_2) = |z|\cos(\varphi + \varepsilon) \tag{5.25}$$

Note that here $|z|$ is the visibility of the interference pattern obtained by variation of ε. Here φ is the phase change caused due to the operations U and V together.

- In the simplest case where U and V are *Hermitian* operators which *commute*, we have

$$\langle\psi|V^\dagger U|\psi\rangle^\dagger \tag{5.26}$$
$$= \langle\psi|VU|\psi\rangle^\dagger \tag{5.27}$$
$$= \langle\psi|UV|\psi\rangle \tag{5.28}$$
$$= \langle\psi|VU|\psi\rangle \tag{5.29}$$
$$= \langle\psi|V^\dagger U^\dagger|\psi\rangle \tag{5.30}$$

  Thus, we would have $z = z^*$ and thus $\varphi = 0$.
- The next case to be considered when they are *Hermitian* operators which however *do not commute*. In this case, we have

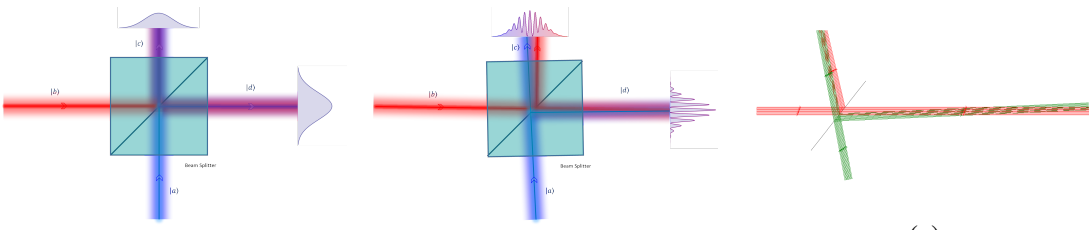
$$(V^\dagger U)^\dagger = U^\dagger V = UV \neq V^\dagger U \tag{5.31}$$

  Thus UV will be a non-hermitian operator.

5.2.2 Non-Collinear interferometer

In the collinear configuration is that we need to measure the visibility and phase change χ by varying ϕ. Since ϕ is difficult to vary controllably without proper stabilization schemes, we make use of the non-collinear configuration.

To go from the collinear configuration to the non-collinear configuration, experimentally we rotate one of the mirrors say M_T by a very small angle. Note that in the experiment the pivot about which the mirror rotates is not the same as the point of incidence on the mirror and specification of how much to rotate the mirror is practically done by later looking at the fringes (since it is hard to measure the exact distance between the point of incidence and the pivot). However, this only will make the two beams diverge after the second beam splitter BS_2. To have some overlap, we need to rotate the beam splitter BS_2 as well. This is shown, in an exaggerated manner in the following figure.



(a) (b) (c)

Figure 5.2: (a) Collinear geometry: The two beams overlap to give us a gaussian beam shape. As a function of phase difference the power in each port changes. (b) Non-Collinear geometry: The interferences give rise to double-slit like fringes. As the phase changes, the power in each port of the beam splitter remains the same, but the fringe shifts. (c) Non-Collinear Rays: The green and red rays depict the rays emerging from the mirrors M_R and M_T respectively. Note that the green rays are tilted (by rotating M_2) and to get the two rays to overlap we have tilted the BS_2 (depicted by the oblique line) as well. software.

Notice that (on right) the red and green rays overlap only in a certain region and then they diverge again. We ideally wish to place the detector when the overlap between the green and red is maximum. However, in practice, we place the camera at a certain position and adjust the angles to get a maximum overlap.

In a non-collinear configuration the interference is spatial (and hence a camera in one arm is enough). For calculation, let us consider the two beams after the BS_2 emerging towards the detectors.

Let the gaussian beam's amplitude of part which is transmitted be given as

$$g_t(x,y,z) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \exp\left(-\frac{(x_t-\mu)^2}{4\sigma^2}\right) \exp\left(-\frac{y^2}{4\sigma^2}\right) \exp\left(ikz_t\right) \tag{5.32}$$

Here, the coordinates z_t and x_t used above are the coordinates obtained by rotation in $z-x$ plane about an angle ξ_t that denotes the angle the beam makes with the normal($\hat{z}$) to the detector plane (x-axis).

$$\begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} \cos(\xi_t) & -\sin(\xi_t) \\ \sin(\xi_t) & \cos(\xi_t) \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \tag{5.33}$$

Similarly, we have, for the part which is reflected

$$g_r(x,y,z) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \exp\left(-\frac{(x_r-\mu)^2}{4\sigma^2}\right) \exp\left(-\frac{y^2}{4\sigma^2}\right) \exp\left(i(kz_r+\phi)\right) \tag{5.34}$$

Here also, we have

$$\begin{pmatrix} x_r \\ z_r \end{pmatrix} = \begin{pmatrix} \cos(\xi_r) & -\sin(\xi_r) \\ \sin(\xi_r) & \cos(\xi_r) \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \tag{5.35}$$

Note that g_r also contains ϕ that takes care of the path difference between the two beams as they reach the detector.

We compute the intensity as

$$I = |g_t + g_r|^2 \tag{5.36}$$

Choosing the angles $\xi_t \& \xi_r$ and fixing wavelength λ (and hence $k = 2\pi/\lambda$), ϕ, the centre of the transmitted and reflected beam μ_t, μ_r and the beam width σ, we get the interference pattern as shown in 5.3.

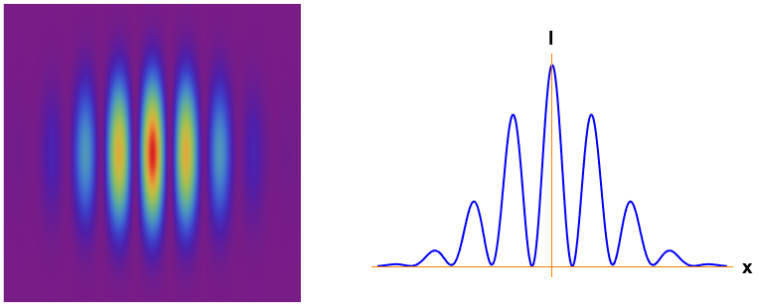


Figure 5.3: On left, we have the (theoretically constructed) image of the interference pattern as seen on the camera (xy plane). On right, we have the interference pattern, i.e. intensity vs x-axis taken at the central section (y=0).

Note that unlike the collinear case here we see the interference pattern because of the wave propagation vectors for the transmitted and reflected beam has some angle, just like the two beams emerging out of the double slit.

If, we fix all other parameters and change ϕ, we see that the fringes shift within the gaussian envelope.

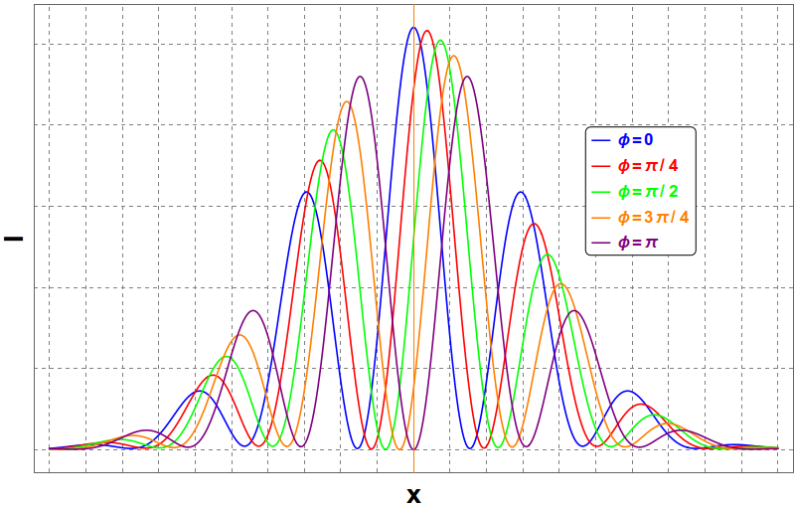


Figure 5.4: On left, we have the (theoretically constructed) image of the interference pattern as seen on the camera (xy plane). On right, we have the interference pattern, i.e. intensity vs x-axis taken at the central section (y=0).

Now, if we have U in one arm and let U be a half-wave plate rotated at an angle α. Let the input beam be horizontally polarized. Then if U is placed in the reflected arm we have

$$I = |g_t\,|H\rangle + g_r \cos(2\alpha)\,|H\rangle\,|^2 + |g_r \sin(2\alpha)\,|V\rangle\,|^2 \tag{5.37}$$

The rotation of the half-wave plate, as shown in the diagram below, changes the visibility of the fringes.

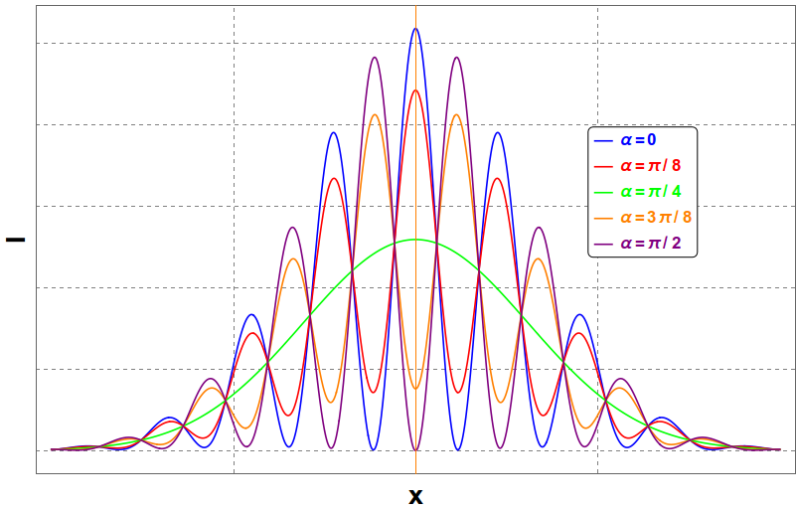


Figure 5.5: The blue and purple lines are for the fast and slow axis respectively which globally have a $\pi/2$ phase difference. The green line denotes when the HWP plate turns $|H\rangle \rightarrow |V\rangle$. At other angles we have visibility in between 0 to 1

.

5.3 The apparent 'non-unitary' evolution

Introduction to the Confusion

In standard quantum mechanics, we deal with two kinds of operators that concerns with evolution. The evolution is usually described by *Unitary* operators until and unless a *measurement* is made, which usually is described with a *Projector*.

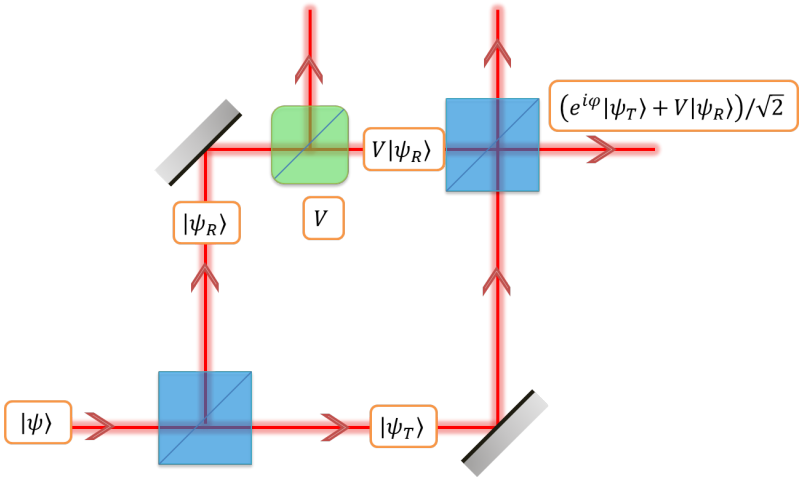


Figure 5.6: Schematic of the Experimental Setup

In the experiment where we compute visibility and phase shift, we used the operator V, acting on the polarization degree of freedom $|\psi\rangle$, in one arm of the Mach-Zehnder Interferometer which is non-Unitary.

Non-unitary operators should violate the conservation of probability and thus should be impractical. The corresponding evolution Hamiltonian should be non-Hermitian as well. So how do we create such an operator in the experiment?

As an example, we choose the operator V to be

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{5.38}$$

, which basically is a Polarizer with the transmission axis aligned to Horizontal.

Solution to the Confusion

The overall operator needs to be unitary. Thus, the non-unitarity must arise from losses only, i.e. there must be a part of the beam that does not reach the detector giving us the apparent non-unitarity.

This can be clearly seen if we replace the Polarizer with a Polarizing Beam Splitter which are almost equivalent in their usage as long as all we are concerned with is the transmitted beam. The difference is that with a Polarizer the rest is absorbed where as in case of a PBS it is reflected.

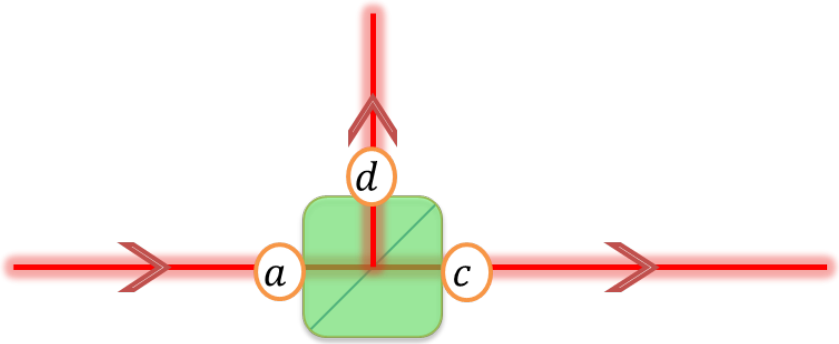


Figure 5.7: Schematic of the action of the polarizing beam splitter

If we carefully look at the operator, we realize that there is an additional degree of freedom that shall concern us. That is the port of the polarizing beam splitter.

When we write the evolution as

$$|\psi_R\rangle \rightarrow V|\psi_R\rangle \tag{5.39}$$

,

what is implied is that we are considering only the transmitted port.

Practically what the operator does is to transmit the Horizontally polarized component and reflect the vertically polarized component.

The incident beam can be written as

$$\begin{pmatrix} \psi_H \\ \psi_V \end{pmatrix} |a\rangle \tag{5.40}$$

The output state can be written as

$$\begin{pmatrix} \psi_H \\ 0 \end{pmatrix} |c\rangle + \begin{pmatrix} 0 \\ \psi_V \end{pmatrix} |d\rangle \tag{5.41}$$

Here, ψ_H and ψ_V are the Horizontal and the vertical component of ψ_R. Note that the overall evolution till now is unitary (probability conserved).

But since we only couple the transmitted portion after V to go the 2nd beam splitter (and subsequently the detector), additional projection $\Pi_c = |c\rangle\langle c|$ is implied.

Thus, with that consideration, we have

$$|c\rangle\langle c|\left(\begin{pmatrix}\psi_H\\0\end{pmatrix}|c\rangle+\begin{pmatrix}0\\\psi_V\end{pmatrix}|d\rangle\right)=\begin{pmatrix}\psi_H\\0\end{pmatrix}|c\rangle \tag{5.42}$$

This explains the apparent non-unitarity although the overall evolution is unitary.

5.4 Experimental Implementation

We use Ti-sapphire laser continuous-wave single-mode fiber output at $810(\pm 2)$ nm and pass it through a polarizing beam splitter (PBS) to make the polarization linear. Then we use a half-wave plate(HWP) to make the polarization state $|\psi\rangle$ input to the Mach Zehnder as diagonally polarized state $|+\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$. The HWP was rotated to make the transmitted and reflected beam of the PBS placed in one arm of the interferometer have equal power. Thus, by virtue of the alignment procedure, we verify that

$$|\psi\rangle = |+\rangle \Rightarrow \langle\psi|R^2|\psi\rangle = \frac{1}{2} \tag{5.43}$$

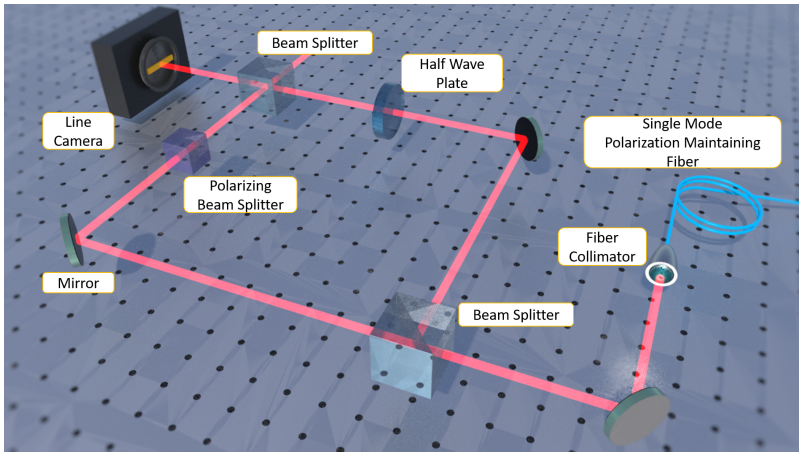


Figure 5.8: Experimental Mach Zehnder Setup: We have polarizing beam splitter(PBS)/polarizer oriented along Horizontal in one arm. To compensate for elliptic polarization components after reflection from the first beam splitter, we place a quarter wave-plate ((QWP) before the Half wave-plate (HWP).

One can controllably vary path difference to obtain I_d as a function of ε so that $|z|$ and phase shift φ can be obtained. But, for our choice of the non-Hermitian operator and the $|\psi\rangle$, the expectation value turns out to be real i.e.,

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}=\frac{1}{2} \tag{5.44}$$

At this point, it is worth mentioning that indeed our method is applicable in general to obtain a complex weak value, but our aim here is to prove the efficacy of this novel method, so we did not do the otherwise necessary phase stabilization.

Since we already know $\langle R^2\rangle$, if we can directly know visibility V, we can obtain the expectation value of the non-Hermitian operator $|z|$. To experimentally obtain the visibility, we use an easier technique, i.e., to align the Mach Zehnder interferometer in a non-collinear configuration, with a slight angle between the two beams incident at the final beam splitter, so that we see spatial fringes. This is attained by first aligning the interferometer in an almost collinear configuration and then iteratively tweaking the mirrors to create a small displacement between the two beams incident at the final beam splitter. Then the final beam splitter is tilted to make the two beams overlap at the detector so that we have spatial fringes similar to a double slit interference pattern. Along the transverse axis of the fringes, we obtain I_d vs. φ as the path taken by two beams coming from two arms to reach the same detector point varies within the Gaussian envelope.

To measure the average value of the non-Hermitian operator A, we place a PBS (R) in one arm and a HWP(U) in the other. To have $U=\sigma_x$, we need to have the HWP making an angle 45° with the Horizontal. But for completeness, the HWP is rotated from 0° to 360° in steps of 2°, so that we can comment on the weak value of R as a function of the post-selected state $|\phi(\theta)\rangle$. θ is the angle that the fast axis of the HWP makes with the horizontal. Although, we are terming the state $|\phi\rangle$ as the post-selected state, note that there is no actual post-selection performed. Conventionally, post-selection is made after weak interaction at a later time. Here, the interaction of the beam with PBS and HWP do not have a specific time ordering and the results are independent of at which distance PBS and HWP are placed from the first beam splitter. We call $|\phi\rangle$ the post-selected state in the sense that, had we performed a conventional weak measurement of the operator R, we would have chosen $|\phi\rangle$ as our post-selected state.

Also, varying the post-selected state as a function HWP angle enables us to find the

expectation value of a class of operators given by

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & 0 \\ \sin(2\theta) & 0 \end{pmatrix}. \tag{5.45}$$

We record a horizontal slice of the interference pattern using a Line Camera (Thorlabs LC100M). The Line camera is placed on the exit port of the second beam splitter of the Mach Zehnder interferometer where the beam converges (see Figure **??**) at a distance where the beams maximally overlap.

We then obtain visibility as described in Figure 5.10. But, we could also obtain phase shift by fitting the beam profile with the following function

$$A_0 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)(1+V\cos(kx+\alpha)). \tag{5.46}$$

Here, A_0 is the amplitude of the Gaussian envelope with the standard deviation of σ centred at μ. k here represents the inverse of fringe width and x is in term of pixels. α gives the phase shift of the cosine from the centre of the Gaussian envelope. If we stabilize the interferometer, we can have the zero error for ε fixed and then we can also determine $\varphi = \alpha + \varepsilon$. Since the interferometer was not stabilized α can change randomly. If the beams have slight transverse separation at the line camera, which can occur due to slight angular displacements caused when we insert or rotate optical components, then as the amplitudes of two beams vary, the beam profile can have an asymmetric envelope. In such cases, the model described in expression (5.46) does not perfectly capture the effect of two Gaussian envelopes. Nevertheless, the visibility obtained from the model with a single Gaussian, is fairly accurate and computationally inexpensive to fit, when we have slight asymmetry.

Since visibility is obtained from spatial fringes, pixel size would be lowering it due to spatial averaging. If the amplitude at any position x is given by $\mathcal{E}(x)$, due to pixel centred at x having size $\mathcal{A}$ will record the intensity

$$I(x) = \int_{x-\mathcal{A}}^{x+\mathcal{A}} \mathcal{E}^*(\chi)\mathcal{E}(\chi)d\chi. \tag{5.47}$$

Hence, we tweak the final beam splitter to create fringes of various width to verify that detector averaging do not affect our experiment significantly. The Figure 5.9 below contains fringes of different visibility and fringe width.

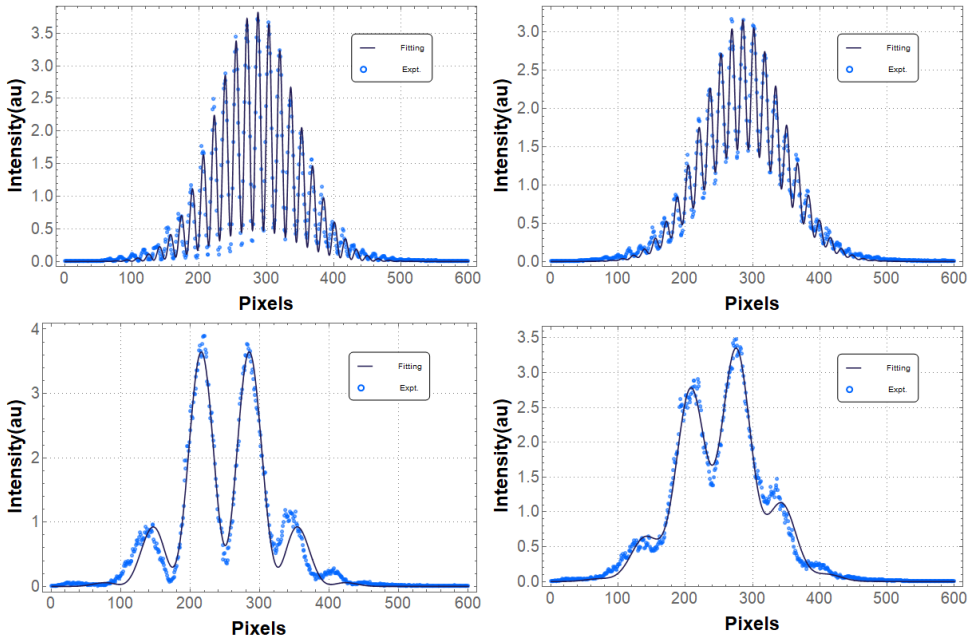


Figure 5.9: The top row has profiles with lower fringe width compared to the bottom row. The left column has higher visibility compared to the right column.

Since the interferometer was not phase stabilized, ε may have random zero errors. Given an interference profile, we can obtain the visibility by knowing the individual Gaussian profiles enveloping the peaks and dips.

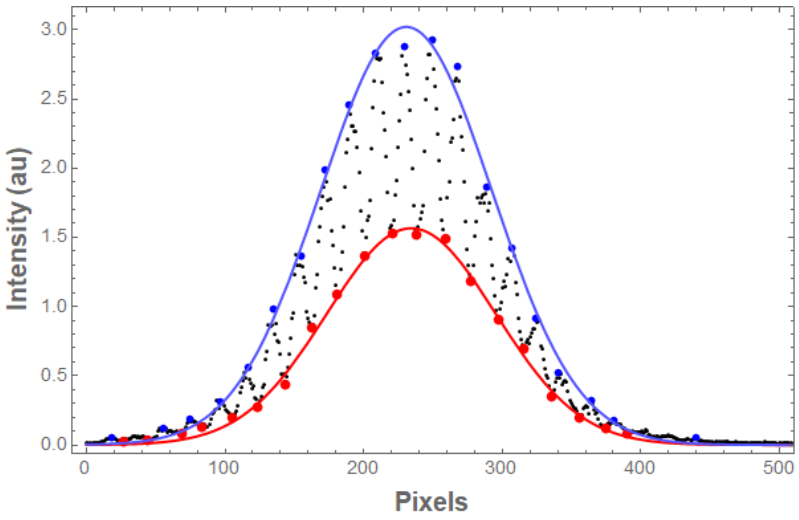


Figure 5.10: Obtaining visibility from spatial fringes: We first find the peaks (blue dots) and dips (red dots) of the interference profile. Then we fit the peaks and dips individually with Gaussian profile $A\exp\left(-a(x-x_0)^2\right)$ to obtain the blue solid curve and red solid curve respectively. If A_p and A_d are the amplitudes of the Gaussian obtained from fitting the peaks and dips respectively, then the visibility is obtained by $V = \frac{A_p - A_d}{A_p + A_d}$

.

However, since fitting the entire curve suffices to obtain visibility, this method of finding visibility was only used as a verification exercise.

We consider various datasets with different fringe widths to verify that our results do not suffer from detector averaging. In some data sets, we also replace the Line camera with Beam profiler (WinCamD-UCD15) to vary the pixel size. In each dataset, we take 100 profiles for each Half wave plate angle to eliminate random errors like fluctuation in the beam profile. We fit the beam profiles and obtain the mean visibility as a function of half-wave plate angle. PBS (Thorlabs PBS 122) is also replaced with a polarizer (Thorlabs LPVIS100MP) to ensure that the results are unaffected. The weak value of R can be obtained by dividing $\langle\psi|UR|\psi\rangle$ with $\langle\phi|\psi\rangle$. To obtain $\langle\phi|\psi\rangle$, we remove the PBS so that only HWP is placed in one arm.

5.5 Results

5.5.1 Using Mach Zehnder Interferometer

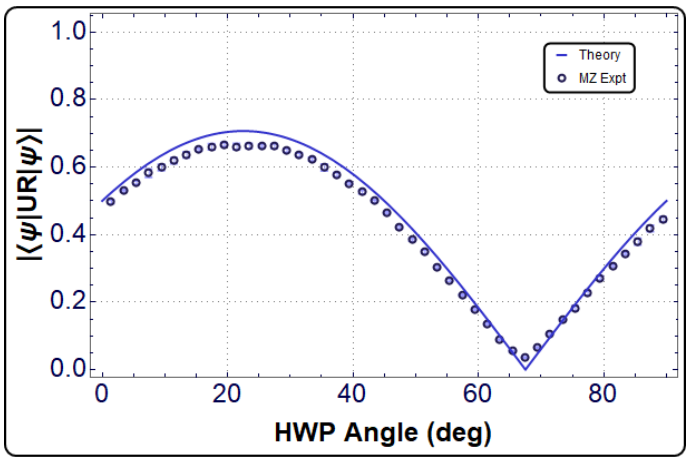


Figure 5.11: The expectation value of a class of non-Hermitian operators described in Eq. (5.45) as a function of HWP angle placed in one arm of the MZ interferometer when the other arm has a polarizing beam splitter in place

For HWP angle of 45°, we infer $\langle A \rangle$ from above graph to be 0.480 ± 0.004.

The error bars in Figure 5.11 are the standard deviation of visibility over 100 profiles. These errors include the effect of any disturbance in the beam profile originating due to mechanical vibrations in the setup. Also, imperfect beam shape or dust may also give rise to slightly fitting results, which contributes to these error bars. The expectation value obtained in experiment do not overlap the same obtained in theory because of elliptic polarization introduced at each reflections moderates the visibility from going too high or too low.

The weak value of R can be obtained by dividing $\langle \psi | UR | \psi \rangle$ with $\langle \phi | \psi \rangle$. To obtain $\langle \phi | \psi \rangle$, we remove the PBS so that only HWP is placed in one arm.

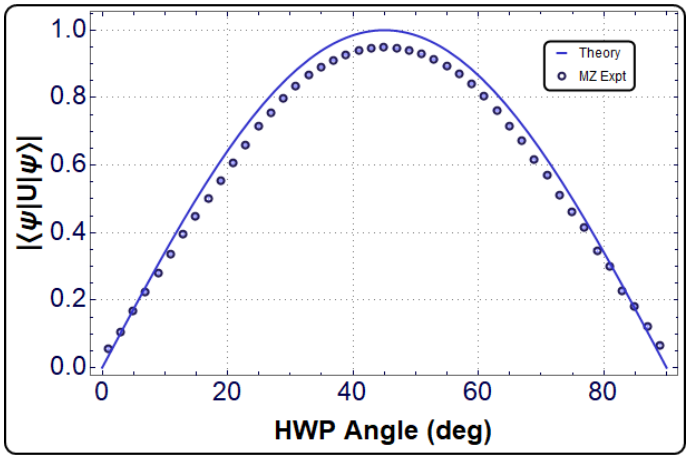


Figure 5.12: Visibility as a function of HWP angle placed in one arm of the MZ interferometer

We obtain the weak value of R from the ratio of $\langle\psi|UR|\psi\rangle$ obtained in Figure 5.11 to visibility obtained in Figure 5.12 as given in Eq. (5.1).

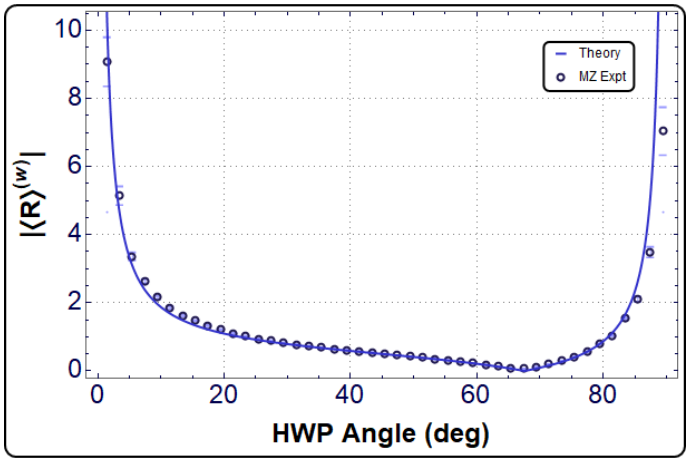


Figure 5.13: Weak Value of R inferred from MZ set up

The weak value of R with pre-selected state $|\psi\rangle$ and post-selected state $|\phi\rangle = \sigma_x|\psi\rangle$ occurs at 45 ° HWP angle. From the above graph, we infer the weak value at that angle as 0.501 ± 0.005, very close to the theoretical value of $1/2$.

5.5.2 Comparison with weak measurement

Using the MZI, we obtain the weak value without having any weak interaction, i.e., there was no element that coupled the system degree of freedom (polarization) with any pointer

states as per von-Neumann measurement scheme. In this section, we obtain the weak value of R by conventional weak measurement.

The weak value of any operator can be reconstructed if we know the weak values of all Pauli operators for a given state. Consider any operator M given by

$$M = \alpha\mathbb{1} + \beta\sigma_x + \gamma\sigma_y + \delta\sigma_z. \tag{5.48}$$

Then, the average value of M in the state $|\psi\rangle$ can be expressed as

$$\langle\psi|M|\psi\rangle = \alpha + \beta\langle\psi|\sigma_x|\psi\rangle + \gamma\langle\psi|\sigma_y|\psi\rangle + \delta\langle\psi|\sigma_z|\psi\rangle. \tag{5.49}$$

Thus, knowing the expectation value of all Pauli operators helps us reconstruct the expectation value of any operator for qubit systems. For the case of spin ladder operator, we would need to measure the expectation value of σ_x and σ_y and for the state $|+\rangle$, we have

$$\langle+|\sigma_-|+\rangle = \frac{1}{2}(\langle+|\sigma_x|+\rangle + i\langle+|\sigma_y|+\rangle). \tag{5.50}$$

The weak value of R can be given by

$$\langle R\rangle^{(w)} = \frac{1+\langle\sigma_z\rangle^{(w)}}{2}. \tag{5.51}$$

Experimentally, the preselected state is achieved by placing a half wave plate at an angle $\pi/8$ (measured from Horizontal) after a polarizing beam splitter to convert the horizontally polarized light into diagonally polarized. We vary the post-selection using a half-wave plate placed before another half-wave plate fixed at $\pi/8$ followed by a PBS. Here, we refer to the state before both the wave plates used for post-selection in Figure 5.14, which guarantees transmission through PBS (and hence guarantees detection with the beam profiler), as the post-selected state.

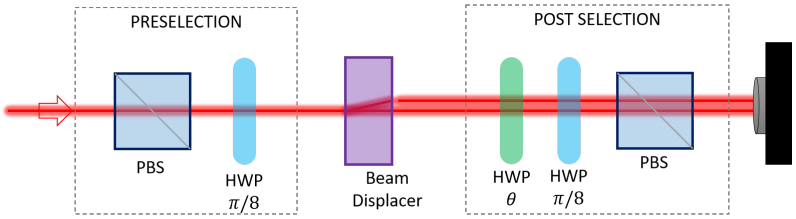


Figure 5.14: Experimental setup to measure the weak value of R with pre-selection $\langle +|$ and post-selected state given by $\langle +|U(\theta)$, where $U(\theta)$ is the HWP angle that is used to vary post-selection

The weak interaction occurs when the calcite beam displacer (Newlight Photonics PDC12005) lets the horizontally polarized light go undeviated and displaces the vertical component. This displacement is small compared to the standard deviation of the Gaussian beam (from He-Ne laser) which is $500\mu m$ at the time of detection (beam divergence $< 1mrad$). Although, its the vertically polarized beam that gets displaced, the operator description can be $R = \Pi_H$ as it is related to Π_V by $\Pi_H = \mathbb{1} - \Pi_V$. The weak value is experimentally obtained from the centroid shift along the direction of beam displacement normalized with the beam displacement.

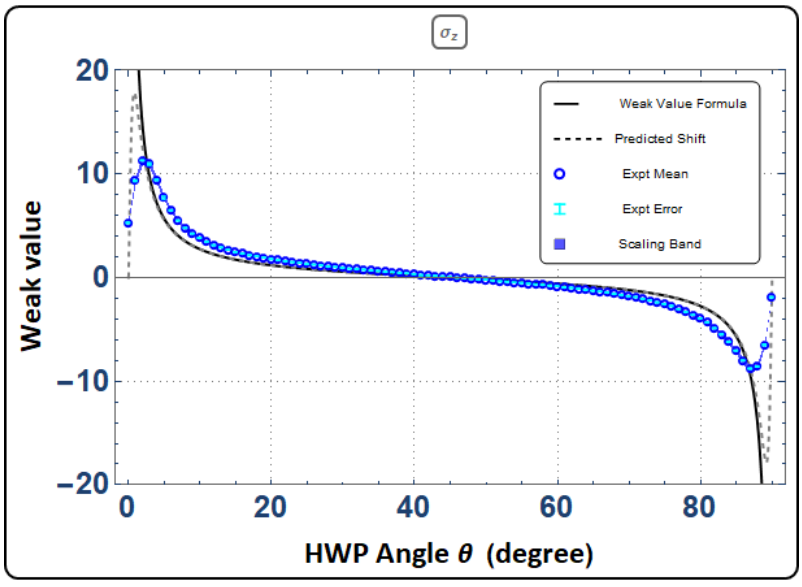


Figure 5.15: Weak value of σ_z in with pre-selection $|+\rangle$ and post-selection $\langle\phi|U(\theta)$.

We compute $R^{(w)}$ from the $\sigma_z^{(w)}$, which is plotted in the graph below. The error bars are propagated from the error bars obtained in Figure 5.15.

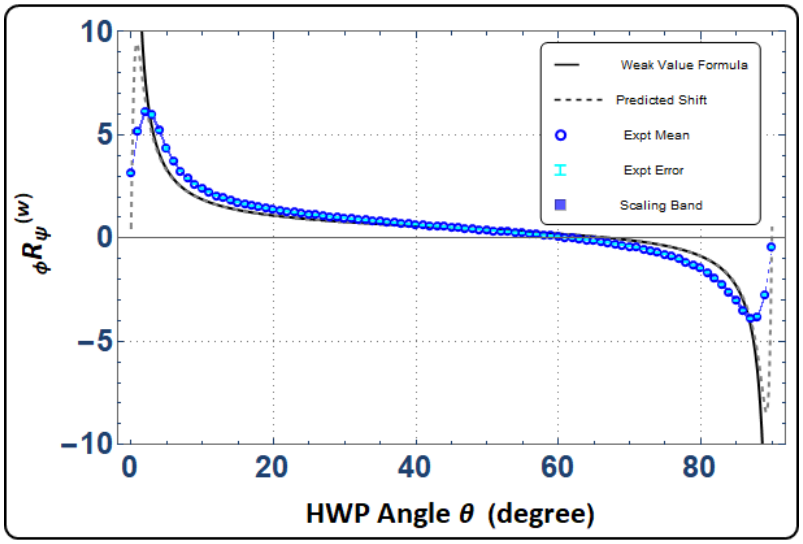


Figure 5.16: Absolute of weak value of R in with pre-selection $|+\rangle$ and post-selection $|\phi\rangle = U(\theta)|\psi\rangle$.

The solid black line indicates the weak value of R computed from the definition of the weak value 1.59 in the preselected state $|\psi\rangle$ and post-selected state $|\phi(\theta)\rangle$. The dashed line represents the weak value if we consider the more realistic a/σ ratio. In the limit $a/\sigma \to 0$, this curve approaches the solid black line. In the experiment, at each half-wave plate angle, 10 images were obtained and the standard deviation of the centroid is represented as error bars. The experimentally obtained centroid needs to be mapped into the eigen-range and the uncertainty involved in the mapping is represented as the light blue band.

By multiplying experimentally obtained $R^{(w)}$ with theoretically known $\langle\phi|\psi\rangle$, we obtain the expectation value of A in Figure 5.17. The expectation value of A obtained in this method is not robust due to slight changes in weak value. Systematic errors in centroid propagate and make the value obtained from experiment deviate from the theoretical value. We should have $\langle A\rangle = 0.5$ at $\theta = 0$.

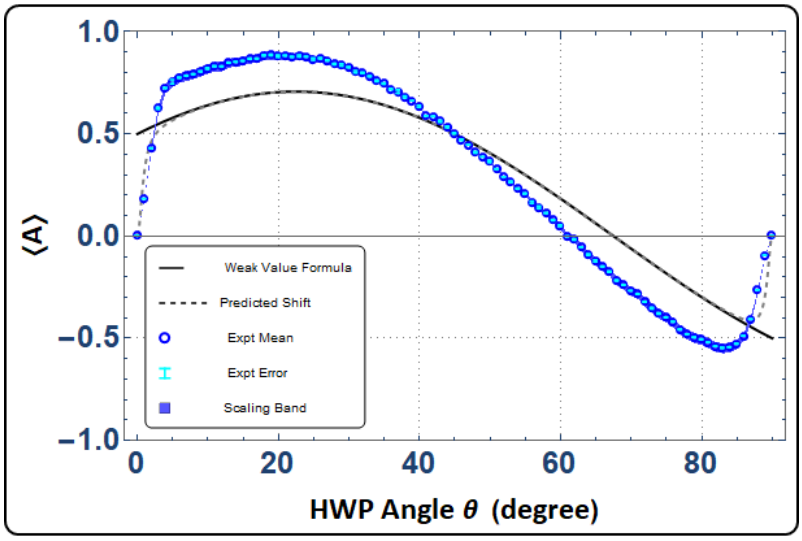


Figure 5.17: Absolute value of Expectation value of A in the state $|+\rangle$.

5.6 Outlook

The chosen operator A had a real expectation value in the given state and consequently, the weak value of R turns out to be real. But in general, both the methods are well applicable to obtain complex expectation values of any non-Hermitian operator.

Theoretically, the weak value approaches infinity when pre and post-selection are orthogonal. Experimentally, this occurs only when the ratio of beam displacement to beam width tends to zero. But finite beam displacement, along with the finite extinction ratio of various polarization components make the experimentally obtained weak value finite. This when multiplied with the overlap $\langle\phi|\psi\rangle$, we, therefore, get zero instead of the desired expectation value of A. The expectation value of non-Hermitian operator for $\theta = 0$ is 0.5, but from weak measurements (Figure 5.17) we obtain 0. Thus, weak measurement is not a good method to infer $\langle A\rangle$ in the region of amplification of weak value of R. For all other non-Hermitian operators expressed in Eq. 5.45 parametrized by θ, weak measurement provides reasonably accurate expectation value.

The MZI method circumvents the above problem as visibility directly gives a finite expectation value. However, since reflection from beam splitters/ mirrors usually introduces elliptic polarization component, additional QWP is needed to compensate for ellipticity to obtain Figure 5.11. Also, computing the weak value from visibility gives us accuracy in

the amplification region (Figure 5.13).

The Mach Zehnder Interferometric method can be used to infer weak value without performing any weak measurement. The weak value obtained using weak measurement gives us information about the property of a particle at any time in between pre-selection and post-selection. Although both the experiments discussed here were performed with laser sources, the results would be the same with average statistics of single photon experiments. In the Mach Zehnder set up, however, there is no weak interaction. If the experiment is performed with a single photon source, the interaction with R would not necessarily occur before the interaction with U. In fact, the temporal ordering of interactions is irrelevant to the visibility as long as the delay is within the coherence time. Thus, the MZI simulates the weak value obtained from weak measurement.

References

[5.1] P.A.M. Dirac. *The Principles of Quantum Mechanics*. Clarendon Press, 1981. ISBN: 9780198520115.

[5.2] M H Chiu, Ji Yong Lee, and Der-chin Su. "Complex refractive-index measurement based on Fresnel's equations and the uses of heterodyne interferometry." In: *Applied optics* 38 19 (1999), pp. 4047–52.

[5.3] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. 3rd ed. Cambridge University Press, 2018. DOI: 10.1017/9781316995433.

[5.4] Meng-Jun Hu, Xiao-Min Hu, and Yong-Sheng Zhang. "Are observables necessarily Hermitian?" In: *Quantum Studies: Mathematics and Foundations* 4.3 (2017), pp. 243–249. DOI: 10.1007/s40509-016-0098-2.

[5.5] Carl M. Bender and Stefan Boettcher. "Real Spectra in Non-Hermitian Hamiltonians Having $\mathcal{PT}$ Symmetry". In: *Phys. Rev. Lett.* 80 (24 1998), pp. 5243–5246. DOI: 10.1103/PhysRevLett.80.5243.

[5.6] Ramy El-Ganainy et al. "The dawn of non-Hermitian optics". In: *Communications Physics* 2.1 (2019). DOI: 10.1038/s42005-019-0130-z.

[5.7] Ali Mostafazadeh. “Pseudo-Hermiticity versus PT symmetry: The necessary condition for the reality of the spectrum of a non-Hermitian Hamiltonian”. In: *Journal of Mathematical Physics* 43.1 (2002), pp. 205–214. DOI: 10.1063/1.1418246.

[5.8] Eliot Bolduc, Genevieve Gariepy, and Jonathan Leach. “Direct measurement of large-scale quantum states via expectation values of non-Hermitian matrices”. In: *Nature Communications* 7.1 (2016). DOI: 10.1038/ncomms10439.

[5.9] Surya Narayan Sahoo et al. “Quantum State Interferography”. In: *Phys. Rev. Lett.* 125 (12 2020), p. 123601. DOI: 10.1103/PhysRevLett.125.123601.

[5.10] J.S. Lundeen and K.J. Resch. “Practical measurement of joint weak values and their connection to the annihilation operator”. In: *Physics Letters A* 334.5 (2005), pp. 337–344. ISSN: 0375-9601. DOI: https://doi.org/10.1016/j.physleta.2004.11.037.

[5.11] Arun Kumar Pati, Uttam Singh, and Urbasi Sinha. “Measuring non-Hermitian operators via weak values”. In: *Phys. Rev. A* 92 (5 2015), p. 052120. DOI: 10.1103/PhysRevA.92.052120.

Chapter 6

Measuring deviation from naive application of superposition principle

Contents

6.1 Superposition principle & Born rule . . . 265

6.2 Multi-order interference . . . 267

6.3 Superposition principle & boundary conditions . . . 270

6.4 κ for the Triple Slot Case . . . 272

6.5 Experiment . . . 274

6.6 Results . . . 278

6.7 Error Types and their effect in the experiment . . . 281

6.8 Sources of error . . . 287

6.8.1 Detector non linearity analysis . . . 287

6.8.2 Effect of detector size on κ . . . 290

6.8.3 Quantifying the effect of Horn Source and Reciever distance . . . 291

6.8.4 Quantifying the effect of Reflection from the Slots . . . 294

6.9 Comparison of Path integral result with Finite Difference Time Domain simulations . . . 298

6.10 Mutual Inductance in Electromagnetism & κ . . . 303

6.10.1 Antenna Configuration . . . 304

6.10.2 Comparison of κ obtained in simulation with Path Integral calculation . . . 305

The dynamics of a quantum particle is governed by the Schrödinger's wave equation, which describes the unitary time evolution of the wavefunction $|\psi\rangle$ under a given Hamiltonian. However, quantum particles when detected, appear to manifest their particle-like localized behaviour. So which property of the particle the quantum state $|\psi\rangle$ was describing ? For the particular case of the electron, Schrodinger emphasized on regarding them as continuous electric charge density [6.1, 6.2], given by $e|\psi|^2$, rather than a classically idealized point particle. The fact that the "intensity" of the electron wave is given by the square modulus of the complex wavefunction is a straight forward extension of the scalar wave theory or electromagnetic wave theory for optics [6.3]. What was debatable was what kind of intensity does it represent? By then it was experimentally possible to detect particles one by one [6.4]. Thus, foregoing the particle nature was not an easy swallow. However, the quantum wave theory had remarkable success in calculating atomic spectra [6.5]. For bound states, the pilot wave theory pioneered by de Broglie, where there exists a particle that is guided by a physical wave was also a compelling interpretation [6.6]. Signatures of wave particle duality was then beginning to emerge with the diffraction pattern of electrons [6.7]. Max Born analyzed the scattering problem[6.8, 6.9] of a swarm of electrons incident on a heavy atom and showed that the relative probability of obtaining the electron is given by the square modulus of the wavefunction ψ. This was, as per Born, following the lead from Einstein who comprehended the wave particle duality of light by interpreting the square of the amplitude of the optical wave as the probability density function with which the manifestation of photons occurs. Once the statistical interpretation of quantum mechanics [6.10, 6.11] became dominant, the method to compute the probability of finding a particle for a given wavefunction came to be referred to as the Born rule [6.12].

The Born rule has two aspects to it. The first deals with computation of the probability distribution $P(x,t)$ of obtaining the particle with a given complex wavefunction $\psi(x;t)$. The rule [1] states that $P(x,t) = \psi^*(x,t)\psi(x;t) = |\psi(x,t)|^2$. Here, one may question the validity of the Born rule, i.e., one may ask how do we know that the probability is given by the square modulus of the state rather than something else, like quadratic power. This often motivates the attempt to derive the Born rule from certain other fundamental assumptions or

[1] Of course, it can be generalized to discreet quantum states. But here we would presume wavefunctions unless otherwise specified

physical principles. The derivation of Born rule has been attempted using the assumption of non-contextuality applied to theory of measurements in quantum mechanics [6.13, 6.14].It has also been suggested that Born rule follows from non-probabilistic classical decision theory along with other axioms of quantum theory [6.15]. Envariance and environment induced superselection were used as principles with which one can arrive at the rule [6.16, 6.17]. Recently, some derivations have based their premise on symmetry requirements demanding independence of transition probabilities from dimension of the Hilbert space[6.18], while others attempt to arrive at the same consequence from a different interpretation of quantum mechanics [6.19]. Nevertheless, Born rule is not a natural consequence of other axioms of quantum theory and would require additional assumptions for its derivation [6.20]. Thus, the square law in the Born is more or less agreed upon as postulate of quantum mechanics [6.21].

The second aspect of Born rule is the statistical interpretation. We talk about probability of obtaining something without referring to any physical interpretation of the wavefunction, thus leaving the description of quantum state as amplitude for a mathematical probability wave. For a given density matrix ρ, the probability of obtaining the outcome a_i is given by $P(a_i) = Tr(\rho A_i)$, where A_i satisfies $\sum A_i = 1$ and represents the positive operator valued measure corresponding to the outcome a_i. For pure state ψ, of course, the probability simplifies to $\langle\psi|A_i|\psi\rangle$. The statistical interpretation, owing to its minimalism, is generally not applied in the context of what happens in the context of detection of a single particles.

So given that the mathematical description of Born rule requires it to be postulated in some form or to be derived from principles assumed to be fundamental, should we question the validity of Born rule? Apart from curiosity driving the case, there are repercussions of violation of Born rule on other aspects in physics. For example, if Born rule gets modified to some $P(x) = |\psi|^{p\neq 2}$, such that the power some real number other than 2, then computational complexity of many problems can get affected, besides the need for an updated quantum paradigm [6.22]. It has been claimed that violation of Born rule would lead to apparent deviation in energy conservation and break the correpondence between quantum electrodynamics and classical electrodynamics [6.23]. Violation of Born rule as we know it could allow higher order interference [6.24]. Violation of higher order sum rules allows contextual probability measures and such theories could allow superluminal

signalling [6.25]. Therefore, when we aim to construct a theory for uncharted territories such as quantum gravity, it is essential to be confident of the principles to which the new theory can reduce to with appropriate approximations. Thus, test of Born rule, specifically the square power, is important for quantum foundations and its extensions.

But can we test the rule without resorting to any particular interpretation of quantum mechanics? For this, we would need to rely on empirically determined quantities only i.e. quantities which are accessible in an experiment. Therefore, we cannot compare experimentally obtained probability density function $P(x)$ against square of the theoretically reconstructed wavefunction $\psi(x)$, which presumably cannot be obtained directly[2] in the experiment. The formulation of test of Born rule becomes easier if we follow the hierarchy of sum rules[6.24] for quantum measure[3] (say μ_A) associated with set of histories (A). This is because probability measures can be empirically determined in an experiment without any further assumptions on the underlying state. When there are two possible histories, the sum rule would imply that $\mu_{A\cap B} = \mu_A + \mu_B$. However, we know that in quantum mechanics[4] this sum rule does not hold true. We have the interference term defined by $I_2(A,B) = \mu_{A\cap B} - (\mu_A + \mu_B)$. In quantum mechanics, $I_2(A,B) \neq 0$ in general. But we can extend the concept to create a hierarchy of interference terms as shown below:

$$I_1(A) = \mu_A \tag{6.1}$$

$$I_2(A,B) = \mu_{A\cap B} - (\mu_A + \mu_B) \tag{6.2}$$

$$I_3(A,B,C) = \mu_{A\cap B\cap C} - (\mu_{A\cap B} + \mu_{B\cap C} + \mu_{C\cap A}) + (\mu_A + \mu_B + \mu_C) \tag{6.3}$$

$$I_n(\{A_i\}_{n\ terms}) = \mu_{\cap_{i=1}^{n} A_i} - \sum_{perm\ A_i} \mu_{\cap_{i=1}^{n-1} A_i} + \sum_{perm\ A_i} \mu_{\cap_{i=1}^{n-2} A_i} - \cdots \sum_{i=1}^{n} \mu_{A_i} \tag{6.4}$$

Here, I_1 refers to the interference when say only slit [5][6.24] A is open. When two slits are open, we know the interference term $I_2(A,B)$, which is the difference between the measure when both slits are open and the sum of measures when individual slits are open, is non-zero. The interference term in the third order sum rule is given by the measures for

[2] Here directly means, for a single particle, not from an ensemble of identically prepared particles

[3] non negative real number associated with the history

[4] In any context where we can have interference this sum rule does not hold true, for example : classical wave theory

[5] slit, here represents a set of histories

the sum of combination of three slits and the three measures for individual slits and its difference with the sum of measures for all possible two-slit configuration. Generalizing this, the interference term I_n in nth order sum rule if given by the sum and difference of various possible combinations of i slits from given n slits.

The interference term I_{n+1} the $n+1$-th order sum rule is related to the interference terms I_n in n-th order by the following:

$$I_{n+1}(\{A_i\}) = I_n(A_p \cap A_q, \{A_i\}; i \neq p,q) - (I_n(A_i; i \neq p) + I_n(A_i; i \neq q)) \quad (6.5)$$

The interference term in the second order sum rule in 6.2 can be seen as the motivation for the generalized n-th order sum rule given in expression 6.5. As a corollary, one can see that if $I_n = 0$ then $I_{n+1} = 0$, which inductively would ensure that all interference terms in sum rules of order higher than n are also 0. We trivially know that $I_1 \neq 0$, i.e, clearly there is a difference in observed intensity between opening and closing the only slit even for classical particles. We expect for classical particles that the measure when two slits would be open to be the sum of the meausures when individual slits are open. But for classical waves and quantum particles we can have interference resulting in a non-zero I_2. The natural question to ask now is whether I_3 also is non-zero ? It can be shown that I_3 can be predicted to be zero if one of the assumptions, along with few others, is the validity of the Born rule [6]. If this is true, then all the interference terms in higher order sum rules are also zero. However, if I_3 turns out to be non-zero in an experiment, then one of the assumptions- the application of Born rule may come under scrutiny.

The Born rule was reported to be valid [6.26, 6.27] and thus higher order interference terms was ruled out up to the precision of 10^{-2} in a triple-slit experiment. This was done by measuring the quantity κ - called the *Sorkin parameter*, which is a normalized equivalent of I_3 after background correction, which was not found to be non-zero given the precision and systematic errors in the experiment. This experiment was performed with heralded single photon source, attenuated laser and laser to show the validity of Born rule in quantum regime as well as its corresponding consequence in semi-classical physics. Three path interference experiments was also performed in NMR systems to test the consistency of the Born rule[6.28] and the experimental bound of the deviation was lowered to 10^{-3}.

[6] See next section for details of the assumptions

However, soon it was realized that multiorder interferences could be generally non-zero[6.29]. This is because another assumption that went in claiming $I_3 = 0$, besides the validity of Born rule, is the principle of superposition of individual solutions. Numerical solutions to Maxwell's equations using Finite Difference Time Domain (FDTD) methods show that the electromagnetic fields when both slits are open is not always the superposition of the two fields when individual slits are open. This would result in κ or I_3 to be non-zero in general for slit based experiments. In fact, for the geometry and parameters used in the experiment for test of Born rule[6.27], the value of κ obtained from numerical solutions to Maxwell's equations [6.29] was of the order of 10^{-5} which was three orders of magnitude lower than the bound obtained in the experiment.

Around the same time, it was realized that one could get a non-zero value of κ in an experiment if systematic errors, like detector's non-linearity creeps in the experiment[6.30]. In this experimental approach the three beams gets created using a tritter consisting of a diffraction grating, as opposed to slits. Using a lens the various diffracted beams are made parallel and shutters are used to block and unblock the beams as opposed to the use of a mask in triple-slit experiment [6.27]. The diffraction grating based experiment was later extended to realize a 5-path interferomter[6.31] which tightened the bound of deviation from Born rule to 10^{-5}, 10^{-4} and 10^{-3} in classical, semi-classical and quantum regimes respectively. Few other experiments, one using orbital angular momentum of light[6.32] and another using third-order quantum interference on a single spin in solids [6.33] also attempted to put the Born rule to test and concluded its validity. The test for multi-path interference has been performed using large molecules as well [6.34].

But, as demonstrated via numerical solutions to Maxwell's equation [6.29], the superposition principle do not apply to different slit configurations because they represent different boundary conditions to the partial differential equation. Using Feynman's path integral method [6.35, 6.36], the value of the non-linear interference terms[6.37] in the Young's double slit experiment was already calculated in 1986 for a particle with mass. However, for light, one needs to use a propagator[6.38] that satisfies Huygen-Fresnel principle as well the Helmholtz equation. Such propagators were used to find the contribution of non-classical[7] feynman paths to the quantity κ for the triple-slit geometry[6.39]. It

[7]Here, non-classical paths are described to the paths that cross or touches the slit plane more than once.

was found the contributions of the leading order non-classical feynman paths makes the order of κ to be 10^{-5}, the same as demonstrated by the FDTD simulations [6.29]. The path integrals were analytically solved which agreed with the FDTD numerical solutions for small angles [6.40]. The maximum value of κ was expressed analytically in terms of the slit parameters and wavelength as $\max(|\kappa|) \approx 0.03\frac{\lambda^{3/2}}{d^{1/2}w}$. This gives the insight regarding the design of experiment and choice of parameters so that κ can be observed in an experiment. The details of how the *Sorkin parameter* κ is computed from path integrals and how it compares to the FDTD simulations is discussed in a later section .

One of the challenges in an experiment to detect a non-zero quantity is to ensure that the value is much about the noise levels. The noise or experimental errors could be systematic or random. If κ is of the order of 10^{-5} or less one needs to reduce the errors to much lower. Most of the devices in optical domain do not offer linearity or accuracy to that level. This was one of the reason why the κ was bounded within 10^{-2} in the triple slit experiment [6.27]. To overcome this limitiation, the experiment was designed to be in microwave domain (wavelength $\lambda = 5$ cm) with blocks of width 10 cm with edge to edge gap of 3 cm between them. This is a triple slot experiment as opposed to triple slit, but using Babinet principle one can also express the Sorkin parameter in terms of intensities one would observe in the detection plane. The details of how this experiment was carried out[6.41], the error analysis and comparision with various theoretical and numerical modelling forms the basis of this chapter. We, in this experiment [6.41] reported a non-zero *Sorkin paramter* in microwave domain. However, around the same time, non-zero κ was reported [6.42] to claim demonstration of exotic looped trajectories of photons in a triple slit experiments. We compare and contrast methods in the appendix K.

Now, what does obtaining a non-zero κ and hence non-zero I_3 imply? Since to predict that I_3 should be zero, we require the validity of Born rule as well as superposition principle, a non-zero κ would indicate that one or both of them are not applicable. Given that we have a theoretical understanding of why the superposition principle is not applicable when the boundary condition changes and estimate κ from the corrected terms in the superposition principle, we may not comment much on the test of validity of Born rule from these experiments. For now, we have no evidence to doubt its validity as the deviation are well attributed to the naive application of the superposition principle.

The higher order interference can be induced by optical non-linearities[6.43] or non-linearities in the interaction of multiple quantum particles[6.44]. The test for higher order interference have also been formulated in different systems like tripod energy level configurations in alkaline atoms [6.45], neutrino oscillations [6.46], cavity and atoms [6.47] , quantum computers[6.48]. Also the concept has been extended to multi-particle states and signature of higher orders path have been found [6.49, 6.50].

6.1 Superposition principle & Born rule

The measurement problem in quantum mechanics and the associated wave-particle duality are exemplified in the double-slit experiment. The quantum particle is described be in a superposition state of being in two slits which at the time of detection collapses to a definite position. The probability of the position state to reduce to a particular location is given by the Born rule:

$$P_{AB}(x) = |\psi_{AB}(x)|^2 \tag{6.6}$$

Here, $\psi_{AB}(x)$ is the transverse wave-function describing the state of the particle along the detector plane when both the slits, slit A and slit B are open. Usually, $\psi_{AB}(x)$ is described as the superposition of the individual wave functions when each slit is open. That is, we often apply the superposition principle to arrive at the following:

$$\psi_{AB}(x) = \psi_A(x) + \psi_B(x) \tag{6.7}$$

Here, ψ_A is the wave-function when only slit A is open and slit B is closed. ψ_B is the wave-function when only slit B is open and slit A is open. Using Born rule, the probability of finding the particle when both slits are open is given by

$$\begin{aligned} P_{AB}(x) &= |\psi_{AB}(x)|^2 && (6.8)\\ &= |\psi_A(x) + \psi_B(x)|^2 && (6.9)\\ &= \langle \psi_A(x) + \psi_B(x) | \psi_A(x) + \psi_B(x) \rangle && (6.10)\\ &= |\psi_A(x)|^2 + |\psi_B(x)|^2 + 2\langle \psi_A | \psi_B \rangle && (6.11) \end{aligned}$$

While we add the wavefunctions, the probabilities do not add up i.e., $P_{AB} \neq P_A + P_B$. This is because of the presence of the interference term $2\langle\psi_A|\psi_B\rangle$. This is a consequence of the two premises:

- The superposition principle is applicable to the wavefunctions i.e., the probability amplitudes and not to the probabilities themselves.
- The probabilities are obtained from the wavefunction using the Born rule i.e, the probabilities are obtained from the wavefunctions using the square law.

The above two premises are almost universally used whereever the description of quantum mechanics is applicable. Almost every experiment seems to be consistent with the above two premises. There have been attempt to directly test these premises over the last decade and most of them have not found any conclusive deviations .

But why question the Born rule at the first place as it has been established over the last century? The assertion that the probability is given by the square modulus of the wave-function is not straight forward. The conceptual problem begins with the what meaning we assign to the wave-function itself. If we claim that the wave-function has mere epistemic significance as a probability amplitude and no ontological existence, we are simply avoiding the question itself. Further, it has to be separately asserted that these probability amplitudes superpose and not the probability itself. However, if the wave-function has ontological existence, then the superposition of states is as natural as the physical superposition of waves. But, the relation of probability and wave-function needs to be established and cannot be presumed. It becomes important to consider whether the Born rule is a fundamental aspect of nature or an emergent phenomenon in the statistical limit. For example, in the de Broglie Bohm pilot wave theory, the wave and particle are two separate entities with physical existence. Therefore, in principle, there can be the wave in a region without the particle being there. Then, the probability of finding the particle has no correlation with the wavefunction itself. The probability distribution given by the Born rule may be considered as an emergent phenomenon. But it needs additional constraints, known as the quantum equillibrium hypothesis. It requires the distribution of the hidden variable, the position of the particles, to be given by the Born rule as an initial condition. Nevertheless, this initial condition need not be satisfied and hence there exist possibility of Born rule not being applicable. Exploration of such physical systems, is

not only important for the fundamental tests of quantum mechanics, but also for all other theories which extends the applicability of quantum mechanics to physical systems beyond what is experimentally established. In the absence of experimental tools, the theories of quantum gravity and the dream of unified theory of fundamental forces relies on being consistent with quantum mechanics. Therefore, it is imperative that the assumptions and valididty of premises used in quantum mechanics undergoes stringent fundamental tests.

Testing the Born rule directly in an experiment requires the knowledge of the wavefunction and the probability distribution. Only the latter is directly accessible in an experiment. The wavefunction is often reconstructed from probability distribution various observables. There are experiments titled "Direct measurement of wave function" , however, here the term "direct" is a coined adjective to the experimental method. It is another reconstruction method using weak-values. The quantum description of weak values itself relies on the applicability of Born rule when the pointer state is projected. Thus, we do not yet have the direct experimental access to what the wavefunction is and without it we would not be able to verify the Born rule without invoking further assumptions.

6.2 Multi-order interference

The consequence of Born rule manifests itself as the interference term in a double-slit experiment. If we restrict ourselves to deal only with probabilities and not wavefunctions in the fundamental test for Born rule, the interference terms appears as the difference between the sum of individual probabilities with either slit open and the probability with the two-slits open simultaneously.

$$\Delta_2 = P_{AB} - (P_A + P_B) \tag{6.12}$$

Note that the difference Δ_2 is zero if the probablity was proportional to ψ and not to its square. From any two-path quantum interference experiment, we know that $\Delta_2 \neq 0$. But what exactly its value is depended on what law we use to relate probability and the wavefunction.

For instance, if instead of $P = |\psi|^2 =$, we had $P = |\psi|^r$, where r is any other integer, we would get a different Δ_2. Since all the terms in the RHS of Eqn. 6.12 can be experimentally

determined, the computed value of Δ_2 can be compared with the experimental observations. But the value of Δ_2 would depend on experimental parameters as well. For example, in a double slit experiment, Δ_2 would be a function of slit width and inter-slit separation. This would need accurate and precise measurement of experimental parameters. The error propagation of these experimental parameters would provide the uncertainty with which we can comment on Δ_2. Since, we have been using the Born rule for almost a century, the deviation in theoretically computed Δ_2 and experimentally obtained $P_{AB} - (P_A + P_B)$ should be very small, if at all non-zero. Without any theoretical basis for predicting the deviation, any experiment to test the Born rule must establish that deviation if any are not due to idealizations in the experiment. For example, in an experiment, the slits may not be perfectly rectangular and the discrepancy between the calculated Δ_2 and experimentally obtained $P_{AB} - (P_A + P_B)$ could just very well be because of that. Yes, one can then try to precisely model the shape of the slits, but then there are other parameters to consider like the reflectivity of edges, the thickness of slits, the complex refractive index of the material etc. All these can contribute to the discrepancy observed if any. Further, the exact spectral and transverse profile of the incident beam needs to be known. Due to the above considerations, any attempt to test the Born rule with a double-slit interferometer, would need precise modeling of the experimental scenario without which it would always have loopholes.

Luckily, if we consider a triple slit experiment however, it is possible to construct a quantity that is independent of any experimental parameters if Born rule is satisfied. We have, the expression for the probability when three slits A, B and C are open as follows:

$$P_{ABC} = |\psi_A + \psi_B + \psi_C|^2 \tag{6.13}$$

$$= |\psi_A|^2 + |\psi_B|^2 + |\psi_C|^2 + 2\langle\psi_A|\psi_B\rangle + 2\langle\psi_B|\psi_C\rangle + 2\langle\psi_C|\psi_A\rangle \tag{6.14}$$

$$= P_A + P_B + P_C + 2\langle\psi_A|\psi_B\rangle + 2\langle\psi_B|\psi_C\rangle + 2\langle\psi_C|\psi_A\rangle \tag{6.15}$$

We don't have the expression for probability when three slits are open in terms of the probability of individual slits. There are still the interference terms which are parameter dependent. But note that, the interference terms are always as pair of two slits. This

is a consequence of the square law in the Born rule, i.e., all interference terms can be written as a sum of interference terms of a pair of slits. The aim would be to consider the corresponding double slit interference and consider the possibility of eliminating all the terms which cannot be directly expressed as a probability.

If we consider the slit C to be closed and consider the interference due to the remaining two slits we have the following:

$$P_{AB} = |\psi_A + \psi_B|^2 \tag{6.16}$$
$$= |\psi_A|^2 + |\psi_B|^2 + 2\langle\psi_A|\psi_B\rangle \tag{6.17}$$
$$= P_A + P_B + 2\langle\psi_A|\psi_B\rangle \tag{6.18}$$

Similarly, if we close only slit B and slit C in separate runs, we would obtain the following two expressions.

$$P_{BC} = |\psi_B + \psi_C|^2 \tag{6.19}$$
$$= |\psi_B|^2 + |\psi_C|^2 + 2\langle\psi_B|\psi_C\rangle \tag{6.20}$$
$$= P_B + P_C + 2\langle\psi_B|\psi_C\rangle \tag{6.21}$$

$$P_{CA} = |\psi_C + \psi_A|^2 \tag{6.22}$$
$$= |\psi_C|^2 + |\psi_A|^2 + 2\langle\psi_C|\psi_A\rangle \tag{6.23}$$
$$= P_C + P_A + 2\langle\psi_C|\psi_A\rangle \tag{6.24}$$

Now, if we subtract the sum of all the probabilities corresponding to the three possible double slit intereferences from the probability with triple-slit interference, we can eliminate all the interference terms.

$$P_{ABC} - (P_{AB} + P_{BC} + P_{CA}) = (P_A + P_B + P_C + 2\langle\psi_A|\psi_B\rangle + 2\langle\psi_B|\psi_C\rangle + 2\langle\psi_C|\psi_A\rangle) - (P_A + P_B + 2\langle\psi_A|\psi_B\rangle + P_B + P_C + 2\langle\psi_B|\psi_C\rangle + P_C + P_A + 2\langle\psi_C|\psi_A\rangle) \tag{6.25}$$
$$= -(P_A + P_B + P_C) \tag{6.26}$$

Therefore, we can use the quantity Δ_3 defined as follows to test Born rule.

$$\Delta_3 = P_{ABC} - (P_{AB} + P_{BC} + P_{CA}) + (P_A + P_B + P_C) \tag{6.27}$$

From the Eqn. 6.26, if Born rule indeed is valid and the probability is given by the square modulus of the wavefunction, then Δ_3 must be zero and hence independent of any experimental parameter. Thus no precise modeling of the experiment is needed to establish a deviation. All the terms in the RHS of the Eqn. 6.27 are probabilities which can be directly obtained in an experiment. Thus, with a triple slit, this fundamental test of Born rule can be carried out without any uncertainties that we had to deal with with only a double-slit interferometer. The only requirement is that all the other parameters like the intensity of the incident beam should remain the same for all the individual experiments where the probabilities of the above 7 combinations of slits opened and closed is determined one by one. However, the lack of any theoretical prediction means that, with what precision the experiment needs to be carried out is unknown.

6.3 Superposition principle & boundary conditions

Apart from the Born rule, there is one more implicit assumption while arriving at 6.26. We have applied the superposition principle. We have considered that the wavefunction when all the three slits are open is the sum of the wavefunctions when each slits are individually opened with the other two closed. The superposition principle has been routinely applied in textbooks of quantum mechanics dealing with the double-slit experiment. However, there is mathematical inconsistency while using the superposition principle in scenarios where the governing equations are given by linear partial differential equations. If ψ_1 and ψ_2 are two solutions to a partial differential equation, then any linear combination $\alpha\psi_1 + \beta\psi_2$ is also a solution when the domain and the boundary conditions are maintained to be the same. As an example, consider solving the Schrodinger equation with infinite square well potential with width L. If ψ_1 is a solution to the equation when the well is shifted slightly right by an amount δL as the width of the well spans from $[-L/2+\delta L, L/2+\delta L]$ and ψ_2 is a solution when the width of the well spans $[-L/2-\delta L, L/2-\delta L]$, can we use $\psi_1+\psi_2$ as the solution when the width of the well spans from $[-L/2, L/2]$?. We can see that, the

solution $\psi_1 + \psi_2$ would be incompatible with the boundary conditions. However, if δL is small enough, the $\psi_1 + \psi_2$ could be a very good approximation to the actual solution. Similarly, in a double slit interferometer, opening and closing of slits changes the boundary conditions associated with the partial differential equation under consideration. Therefore, the solution ψ_A when slit A is open and slit B is closed cannot be superposed on the solution ψ_B when slit A is closed and slit A is open. But so far, many experiments with double slit interferometers have been carried out and none have reported any significant deviation from what is expected. Therefore, the change in boundary condition must be small enough so that the solution obtained by applying the superposition principle naively would be close to the actual solution.

It is important to quantify these deviations in an experiment. If the actual solution when two-slits are open is different from the superposition state that we have assumed while arriving at Eqn. 6.27, then Δ_3 could be non-zero in an experiment. In that case, Δ_3 being non-zero cannot tell us anything about validity of Born rule directly. It would only imply that either the use of Born rule or the application of the superposition principle in this context is not correct.

The maximum value of $\kappa = \Delta_3 / \max(P_{ABC})$ has been quantified to be of the order of $0.03\frac{\lambda^{3/2}}{d^{1/2}w}$ using path integral formulation , where w and d are the slit width and inter slit distance respectively and λ is the wavelength of the source. If the slit widths and inter-slit distances are significantly larger than λ, i.e. say about 10λ, the maximum of κ is of the order of $10^{-}4$. This explains why in most experiments, the superposition principle approximately works. To be able to measure such small quantities either one can aim to increase the precision with which the experiment is carried out or chooses λ, w and d such that the maximum of κ becomes larger. The former is difficult to achieve given the accuracy with which power can be measured and the stability with which the source can be maintained at constant power. Therefore, a microwave version (with $\lambda = 5cm$) of the experiment was designed, where we can have $w \approx \lambda$. But, the slit version of the experiment would require very large block otherwise the microwaves can diffract from the edges of the block. Therefore, the slit version of the experiment is replaced with the slot version. Although, we the entire experiment can be treated using Classical electrodynamics using electric and magnetic field E and H respectively, it is important to note that the

consequences applies to any experiment where superposition principle is used with change in boundary conditions. Thus, the qualitative results would also the applicable too the quantum version of the experiment.

6.4 κ for the Triple Slot Case

The formula to compute kappa in a triple slit experiment was as follows:

$$\kappa = \frac{p_A + p_B + p_C - p_{AB} - p_{BC} - p_{CA} + p_{ABC}}{\max(p_{ABC})} \tag{6.28}$$

This normalization is with the idea that $\max(p_{ABC}$ occurs at the centre of the detector plane. This is true for a triple slit experiment in the far field. However in near field, the centre position can be a minima instead. Further, if we use slots instead of slits, the centre position, even though a local maxima , is not a global maxima.

Hence for a triple slot experiment at near field it is imperative that the quantity of interest is normalized with something that is robust while comparing different regimes of the experiment. The maxima of the triple slot interference pattern is not such a quantity. On the other hand the global maxima of the radiation pattern lies at the centre (for a horn source / dipole source/ gaussian source).

Regarding the numerator, we need a quantity which will be zero if we incoorectly apply the superposition principle. In path integral language, the contribution of classical paths only should make the numerator zero.

In a triple slot experiment, we can divide the slot plane in 5 regions.

- r1 : the region left to the left edge of slot A
- r2 : the region where slot A is placed
- r3 : the region between right edge of slot A and left edge of slot B
- r4 : the region where slot B is placed
- r5 : the region between right edge of slot B and left edge of slot C
- r6 : the region where slot C is placed
- r7 : the region right to the right edge of slot A

The probabilities at the detector plane corresponding various slot combinations when

the superposition principle is incorrectly applied is as follows

- $p_{BG} \propto (\vec{E}_1+\vec{E}_2+\vec{E}_3+\vec{E}_4+\vec{E}_5+\vec{E}_6+\vec{E}_7)\times(\vec{H}_1+\vec{H}_2+\vec{H}_3+\vec{H}_4+\vec{H}_5+\vec{H}_6+\vec{H}_7)$
- $p_A \propto (\vec{E}_1+\vec{E}_3+\vec{E}_4+\vec{E}_5+\vec{E}_6+\vec{E}_7)\times(\vec{H}_1+\vec{H}_3+\vec{H}_4+\vec{H}_5+\vec{H}_6+\vec{H}_7)$
- $p_B \propto (\vec{E}_1+\vec{E}_2+\vec{E}_3+\vec{E}_5+\vec{E}_6+\vec{E}_7)\times(\vec{H}_1+\vec{H}_2+\vec{H}_3+\vec{H}_5+\vec{H}_6+\vec{H}_7)$
- $p_C \propto (\vec{E}_1+\vec{E}_2+\vec{E}_3+\vec{E}_4+\vec{E}_5+\vec{E}_7)\times(\vec{H}_1+\vec{H}_2+\vec{H}_3+\vec{H}_4+\vec{H}_5+\vec{H}_7)$
- $p_{AB} \propto (\vec{E}_1+\vec{E}_3+\vec{E}_5+\vec{E}_6+\vec{E}_7)\times(\vec{H}_1+\vec{H}_3+\vec{H}_5+\vec{H}_6+\vec{H}_7)$
- $p_{BC} \propto (\vec{E}_1+\vec{E}_2+\vec{E}_3+\vec{E}_5+\vec{E}_7)\times(\vec{H}_1+\vec{H}_2+\vec{H}_3+\vec{H}_5+\vec{H}_7)$
- $p_{CA} \propto (\vec{E}_1+\vec{E}_3+\vec{E}_4+\vec{E}_5+\vec{E}_7)\times(\vec{H}_1+\vec{H}_3+\vec{H}_4+\vec{H}_5+\vec{H}_7)$
- $p_{ABC} \propto (\vec{E}_1+\vec{E}_3+\vec{E}_5+\vec{E}_7)\times(\vec{H}_1+\vec{H}_3+\vec{H}_5+\vec{H}_7)$

We see that if we compute the numerator of kappa defined for slits

$$num(\kappa) = \frac{p_A+p_B+p_C-p_{AB}-p_{BC}-p_{CA}+p_{ABC}}{} = pBG \tag{6.29}$$

(Please find the calculation in the mathematica notebook in this folder)

Thus, we can define the new numerator to be

$$p_{BG} - num(\kappa_{slits}) \tag{6.30}$$

which is zero when we apply the superposition principle incorrectly.

Thus, we define kappa for slots as

$$\kappa_{SLOT} = \frac{p_{BG}-(p_A+p_B+p_C-p_{AB}-p_{BC}-p_{CA}+p_{ABC})}{\max(p_{BG})} \tag{6.31}$$

Now onwards κ_{SLOT} will be denoted by κ only

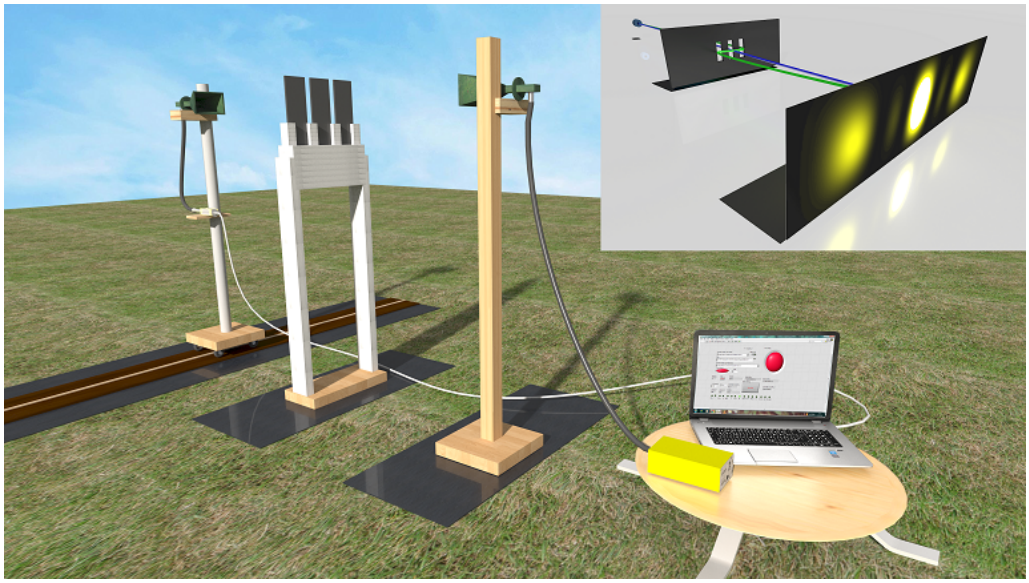


Figure 6.1: A schematic of the experimental set-up. The green antennas on either side are pyramidal horn antennas which act as the source and detector of electromagnetic waves respectively. The detector antenna is placed on a moving rail to enable measuring of diffraction patterns. The three slots are placed between the source and the detector. The inset shows a triple slit schematic where the blue line is a representative classical path and the green line a representative non classical path in path integral formalism.

6.5 Experiment

Figure 6.1 shows the details of the experimental set-up. We have a pyramidal horn antenna acting as a source of electromagnetic waves at 5 cm wavelength. These waves are incident on slots which are 10 cm wide having an inter-slot distance of 13 cm (centre to centre). The slots are composite structures which consist of two layers of specially manufactured materials which act as near perfect absorbers of microwave frequency (Eccosorb SF6.0) sandwiched with an aluminium layer in between to enable even more perfect absorption especially of back reflected beams which may make their way back to the source antenna from the detector. We have done some rigorous analysis of such back reflection and concluded that they do not affect our experiments .

The detector antenna is also a horn antenna similar to the source but housed on a moving rail to enable collection of data as a function of detector position. We use a high frequency power probe to record power values for the different combinations of slots required for measuring κ as a function of detector position. The measurement involves eight separate experiments which measure the individual contributions to equation 6.31. A

ninth measurement involves measuring with source off and detector on to confirm that the antenna does not pick up any comparable signal from unknown emitters in the environment. This was several orders of magnitude lower than the measured values with the source on leading to a very low stray signal level. In the Methods section, we have included technical details on how we aimed at achieving perfect alignment which plays a crucial role in a precision experiment like this one.

Alignment procedure

One of the most crucial steps which enables us to measure a convincing non-zero for κ involves precise alignment of the various components in the experiment. While there is alignment at a basic level, there is also finer alignment using dedicated tools.

- The first condition to be ensured is perfect levelling of the ground. A spirit level is used at various points on the ground to check the ground level and all unevenness is filled with sand. Once preliminary levelling is achieved, we place a marble-like stone on the ground to ensure further smoothness. These stones need to be settled into the ground using water so that once set, they do not sag any further.
- The experimental set-up consists of a rail, a motor for horn antenna detector movement, two horn antennas, slot stand and slots. The source horn antenna is connected to an Anapico signal generator which generates microwaves at 6 GHz. The receiver horn antenna is connected to a high frequency power probe from Agilent. The data acquisition is done by using Labview.
- The next alignment involves alignment of transmitter and receiver horn antennas.
- The transmitter horn antenna is fixed while the receiver antenna moves on a rail. For initial alignment, the receiver antenna is placed at the centre of the rail directly in front of the source antenna. It is ensured that the two antenna centres coincide with each other. This is done using a plumb line.
- Once the perpendicular alignment is done, one needs to ensure that the distance between the edge of the rail on either side is the same from the source. This is done using a Laser Distance metre (LDM).
- Next, the slot stand (made of high density thermocol which is separately checked to be a perfect transmitter for the 6 GHz microwaves upto the desired accuracy)is

placed between the source and receiver. The distance between the source antenna probe wire and the receiver antenna probe wire is 2.50 metres and the slot stand is exactly 1.25 metres from each. The distances are again ensured using LDM.

- Water level is used to ensure that the height of both the source and receiver antennas are 1.75 metres from the ground.
- Figure below shows the final set up which is housed in an appropriate tent for protection against wind and rain.



Figure 6.2: Final experimental set up which is housed in a tent

- One of the considerations that plays a role in experiments which are performed in open field conditions is the possibility of reflection from the ground affecting the source radiation. Other than the back reflections from metal and other structures on the field (which we have discussed in detail in the supplementary material), we also need to confirm that reflections from the earth's surface do not cause any difference in measured power. One can reduce the effective reflection from the ground by raising antenna height to an extent that these reflections do not play any role. Figure 9 shows a plot of measured power vs height of source and receiver antenna at different source powers (0 dBm, 10 dBm, 15 dBm) when the source and receiver antenna are at a distance of 8 metres from each other. As one can see, beyond 145 cm height of the antennas, there is no appreciable change in power measured. This implies that beyond this height, there is no relevant change caused due to specular reflection. Our experiment was conducted at source and receiver antenna heights of 175 cm and at a smallest source-receiver distance of 2.5 metres where the specular reflection

component will be even more negligible.

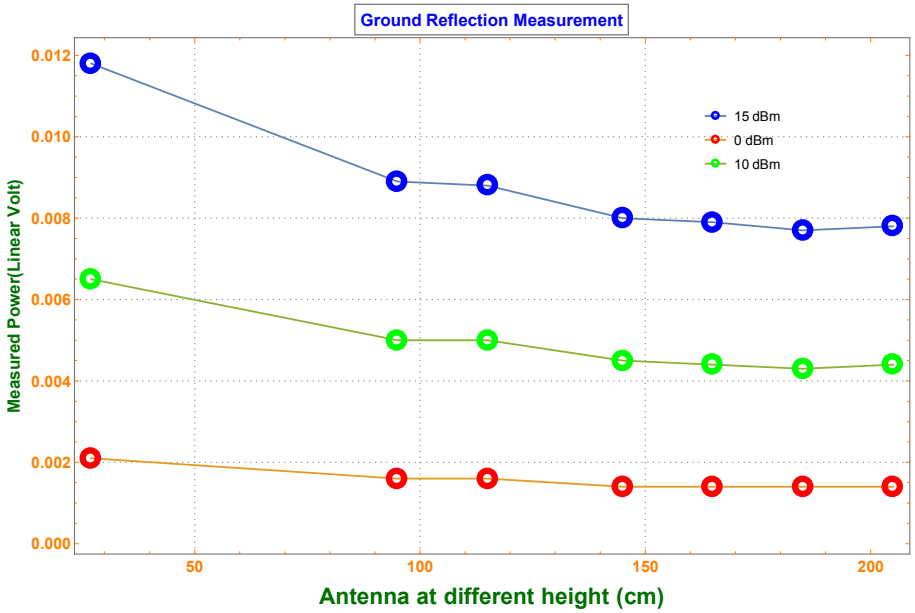


Figure 6.3: Measured power values as a function of antenna height for source-receiver distance of 8 metres. The values saturate to almost the same value beyond 145 cm.

In order to substantiate this point further, we include below a geometry based argument which proves that ground reflections do not play a role in our experiment. Consider the two horn antennas 2.5 m apart at a height of 1.75 m from ground as shown in Figure 10. From laws of reflection, both source and detector antennas will see the waves reflected from the ground at a distance of 1.25 m from it. Since both the antennas are at 1.75 m above the ground level, the angle subtended by the ray that gets detected after ground reflection from the line joining the two detectors is

$$\arctan\left(\frac{1.75\ m}{1.25\ m}\right) = 54.46^{\circ} \tag{6.32}$$

If we consider that the direct beam has a unit gain from the antenna and factor in the gain at 54.46° appropriately, under realistic ground reflection percentages (typically chosen to be 10%), the change in power due to ground reflection (compared to the direct beam) will be about 0.0001. This is less than experimental errors due to source fluctuations (typically 0.003) and hence is not a major source of error in the experiment.

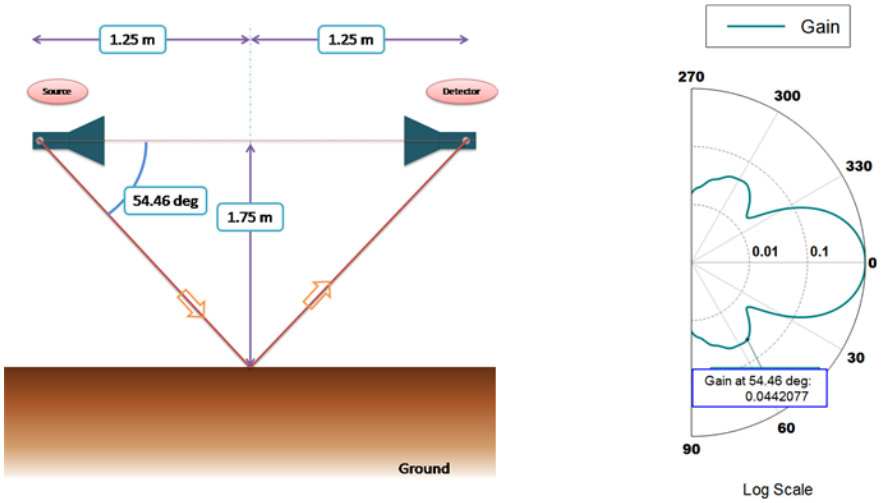


Figure 6.4: Figure on left shows two horn antennas facing one another at a distance of 2.5 metres and at a height of 1.75 metres from the ground. The source and detector antennas both will see waves reflected from the ground at an angle of 54.46°. Figure on right shows the antenna gain pattern.

6.6 Results

The measured graphs of κ as a function of detector position have good agreement with theoretically simulated plots using the Method of Moments (MOM is a 3D simulation technique in which exact horn detector and slot material parameters can also be simulated). These results demonstrate the importance of taking proper boundary conditions into consideration while applying the superposition principle in slit/slot based interference experiments.This is essentially a boundary value problem and need not need a quantum mechanical explanation per se. They also describe experimental situations in which the non-zeroness of the Sorkin parameter need not be a good measure for Born rule violation. Our experiment is actually testing the length scale dependent boundary condition effect on the application of the Superposition principle in interference experiments. Thus, it also serves as a guide to future experiments intending to test Born rule violations which need to be properly designed to minimize these effects. They also exemplify a situation in which not just the classically dominant paths in the Path Integral formalism has been used to explain experimental observations. They thus bring forth the experimental reality of the winding paths in Path Integrals.

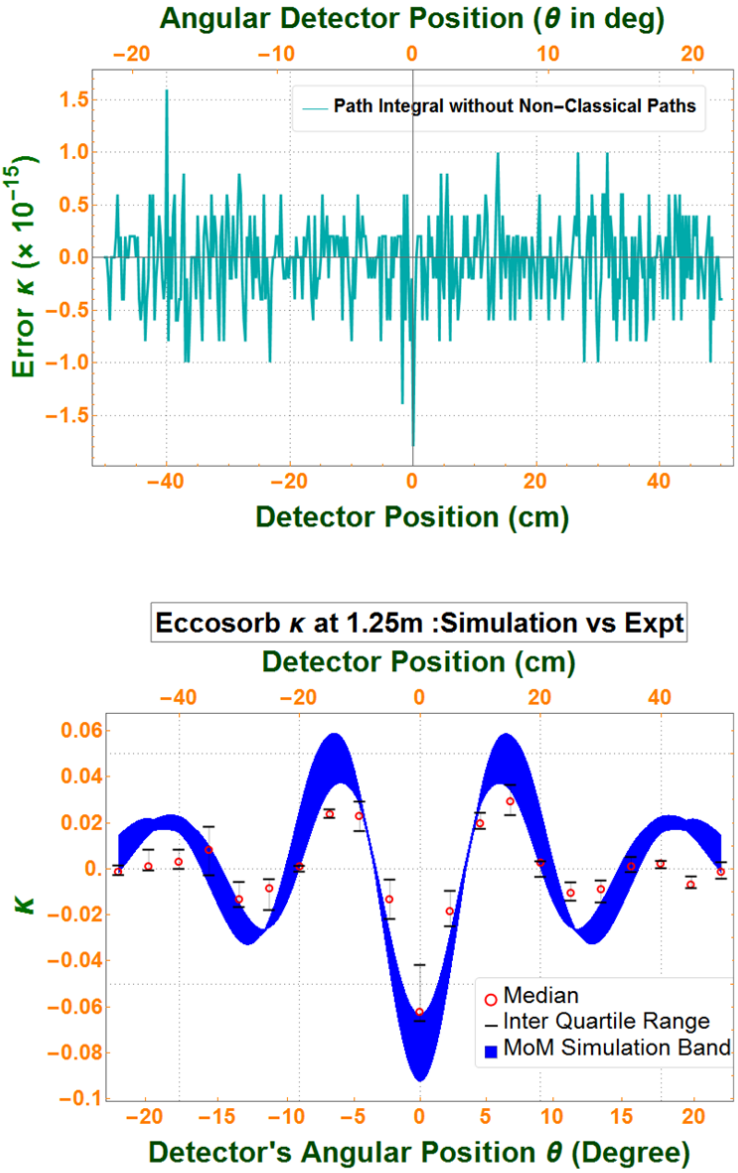


Figure 6.5: On top is a simulation result which is a plot of κ as a function of detector's angular position (which is the angle subtended at the detector plane by the centre of the slot plane) when the superposition principle is incorrectly applied. We have obtained this by taking into account only classical paths in the path integral formalism whereby κ is manifestly zero. The plot below is a representative experimentally measured κ as a function of detector's angular position. The red markers represent the median κ at each position. The black lines denote the interquartile range with outliers removed. A few detector positions had one or two outliers. The blue band represents the theory band obtained from MOM based simulation. To create the theory band, we simulated κ for the experimental parameters taking into account uncertainties in the parameters. Major contributions to the band came from antenna probe wire height, distance from backplate, interslot distance as well as the material refractive index based uncertainties.

Figure above shows a representative plot of κ as a function of detector position at a source detector separation of 2.5 metres (1.25 metres between source and slots plane and same distance between slot plane and detector plane). The detector movement is controlled using a 10 RPM DC motor. At each detector position, all eight contributions to equation 6.28 are measured by placing and removing slots as required before the detector is moved to the next position. This way, we ensure that the κ value at a certain detector position is measured on a short time scale and fluctuations due to the environment do not affect the individual contributions in such a way that it can affect the value of κ measured. We have ensured that our measurement time scale is sufficient by actually measuring an antenna radiation pattern as a background measurement before measuring the pattern due to a particular slot combination. These background patterns overlap during the κ measurement time indicating that our noise fluctuation time scale is much longer than the measurement time. The background corresponding to a certain slot combination is taken to be the average of the background value measured before the combination and after. The formula for κ measured experimentally is modified as follows.

$$\kappa = \gamma(P_{ABC} - P_{AB} - P_{BC} - P_{CA} + P_A + P_B + P_C) \tag{6.33}$$

where $P_\alpha = \frac{p_{BG\alpha} - p_\alpha}{p_{BG\alpha}}$, $\alpha = A, B, C, ...ABC$ slots being present and $p_{BG\alpha}$ refers to the background corresponding to each combination. $\gamma = \frac{p_{BG(at\ x=x_D)}}{Max(p_{BG})}$, x_D being a certain detector position. As can be seen, at the position corresponding to maximum of the background, which is usually the center of the radiation pattern, $\gamma = 1$ and the equations 6.28 and 6.33 become equivalent. By defining individual background contributions to the different terms in equation 6.28, we can take care of varying source power, if any, between combinations. The background value is measured using a reference detector and also by averaging between background values measured before and after each combination which turn out to be equivalent in our case.

For each combination, 3000 data points are collected, the median of which contributes the measured power value. Thus, each κ value has 3000 measured values for each combination. The number 3000 was arrived at after sampling for both lower and higher number of data points (it was found that the κ values converge to the same value for 2000 data points per combination itself so we decided to measure 3000 data points for each combination

to overcompensate). 10 Measurements of κ were done at each detector position and the median value has been chosen as the representative value. The errors for each value have been represented by box plots [**box**]. Further, details on choice of measurement statistics as well as error analysis are given in the Methods section. We have also randomized the order in which slot combinations are measured and ensured that κ remains constant.

Figure 3 shows measured κ as a function of detector position for three different source-detector distances. This was done to ensure that the match between theory and experiment persists even on changing some changeable parameters in the experiment.

6.7 Error Types and their effect in the experiment

In any experiment we can classify the errors as

- **Random errors** These tend to increase the uncertainty in the variable of interest but do not change the mean value provided we have enough statistics.
- **Systematic Error** These do not increase the uncertainty in the experiment but shifts the mean of the variable of interest.

For example, suppose the source power fluctuates a lot about a mean power , then the power in each combination will have more uncertainty and hence the error propagated to κ will have more uncertainty. However, the mean value wont be affected much.

But in case the source has a drift only (assume less fluctuation), the uncertainty of power in each combination is less. So the error propagated to κ will be less. But since the source drifts the last combination will see a very different source compared to the initial combinations. This will lead to a change in the mean kappa compared to the case when the source had no drift.

Whether a contribution to error is random or sytematic depends on time scale of data aqcuisition and number(and order) of repeatation of the experiment. For example consider the case when the source has a uniform drift. Suppose we take one kappa data set with a particular sequence of combinations, we will have one mean kappa value with a (small) uncertainty. Now if we take one more kappa data set with the same sequence of slot combinations , we will end up with the same mean kappa for the second set (because the drift is uniform). Thus, the uniform drift causes an error than goes unnoticed (and hence

a systematic error). However, if the drift was random in the time scale between the two data sets we will notice a change in the mean kappa. Even otherwise if the drift is uniform we can randomize the sequence of combinations and this will result in a change of mean kappa from set 1 to set 2. A design of experiment is better if most systematic errors can be converted into random errors.

However, there could be some parameters which we dont(or cannot) directly measure in the experiment. For instance, it may so happen that the slot stands tilt a bit due to the weight of the slot. So when slot A is placed it tilts a bit towards left and when slot C is placed it tilts towards right. If we dont keep a track of it, this error will be consistently repeated and the change in sequence of slot combinations taken would not help. Another kind of error might occur if the slot does not fit tightly into the slot stand. This error has nothing to do with time, unlike source drift, (and hence wont be accounted by randomizing sequence of slot combinations) However this error is expected to be randomized when we repeat more data sets, by placing the slots more number of times. In the kind of experiment like triple slot experiment, the challenges are even higher. The background radiation is not constantly monitored nor is the weather like humidity and temperature. Change in humidity change the refractive index of air. Change in Temperature will change resistance of some components and hence might affect the readings. Unless monitored one can only hope that if there are errors due to these causes, they do show up as random errors.

Till now, we have just talked about errors occuring in an experiment, that increases the uncertainity or shifts the value obtained in the experiment from the true value. But in case we want to compare data with theory, we must ensure that all the parameters used in theory agree with the experiment. For instance, if we have use , in theory, a different refractive index of the slot than what is used in the experiment, this will be an error in comparing the theory and experiment. However, this is neither a random error or a systematic error (these terms are reserved to be used in the context of experimenatal error.)

Computing Error Propagation to Kappa

The formula used in the experiment is as follows

$$\kappa = P_A + P_B + P_C - P_{AB} - P_{BC} - P_{CA} + P_{ABC} \tag{6.34}$$

where

$$P_\alpha = \frac{P_{BG_\alpha} - p_\alpha}{P_{BG_\alpha}} \tag{6.35}$$

Now, we calculate P_{BG_α} by measuring background before and after the slot combination α. Each measurement of background consists of taking 3000 readings (about 45 seconds). To compute the mean $\mu(P_{BG_\alpha})$ we average over the 6000 readings (combining the readings before and after). To compute the $\sigma(P_{BG_\alpha})$ we take the std. deviation of 6000 readings. We do the same thing for p_α, the slot combination. From here, we use 6.35 to compute mean $\mu(P_\alpha)$ where the RHS of 6.35 are the mean quantities.

We compute the uncertainty in P_α as follows:

$$\delta P_\alpha = \sqrt{\left(\frac{\partial(\frac{P_{BG_\alpha}-p_\alpha}{P_{BG_\alpha}})}{\partial P_{BG_\alpha}}\delta P_{BG_\alpha}\right)^2 + \left(\frac{\partial(\frac{P_{BG_\alpha}-p_\alpha}{P_{BG_\alpha}})}{\partial P_\alpha}\delta P_\alpha\right)^2} \tag{6.36}$$

In the above equation we choose to ignore the higher order terms and further assume that background and slot combination powers are independent (in terms of its variation). In the above equation the partial derivatives are calulated by first taking the partial derivative and then substituting the means. For the variations, the std deviation is used.

This gives us error in each effective slot combination. Then the mean kappa for a data set is computed using 6.34 where in RHS all the quantities are mean for each effective slot combination.

The error in κ is simply computed as

$$\sigma_\kappa = \sqrt{\sum(\sigma_{P_\alpha})^2} \tag{6.37}$$

How much statistics to take ?

Till now we have talked about computing mean and error of kappa for single data set. If we take 10 kappa data sets at position, we compute the mean kappa at that detector position as the mean of all 10 mean kappa(over 3000 readings). But how to know if 10 is a good number.

Lets see what taking more kappa data sets leads to. First of all, for a basic reproducability we should take at least three(or five, there is no cross checking of statistics with just one data set!!!) readings. Now if the std deviation of the mean of three readings is very less compared to the average error bars of each data set, then we can say the experiment is reproducible. And we can say that the true value is well contained within the error bars (or all the systematic errors has not been randomized by a careful design of experiment).

However, the other extreme is that the errorbars of two datasets dont overlap. In this case, the two values are inconsistent with each other i.e. the probability that the results were driven by just random error is very very low. If this happens we can conclude that there should be a systematic error that is randomized over different data sets. In such a case, if we take number of data sets such that the $\sigma(\mu(\kappa)) < \mu(\sigma(\kappa))$ for every reasonable subset of the statistic (i.e take 3 to 10 values randomly of of the ten and for all such random selections we should satisfy the condition consistently) we can claim that we have taken enough statistics such that the systematic errors in each data set are averaged out being random over many data sets and the mean of the kappa data sets is close to the true value. However, if we cannot satisfy the above condition, that would means we have more uncertainty in κ than what we estimate using error propagation formula for each data set. In that case, we should report the higher error (but the requirement to report this error could be 30 datasets)

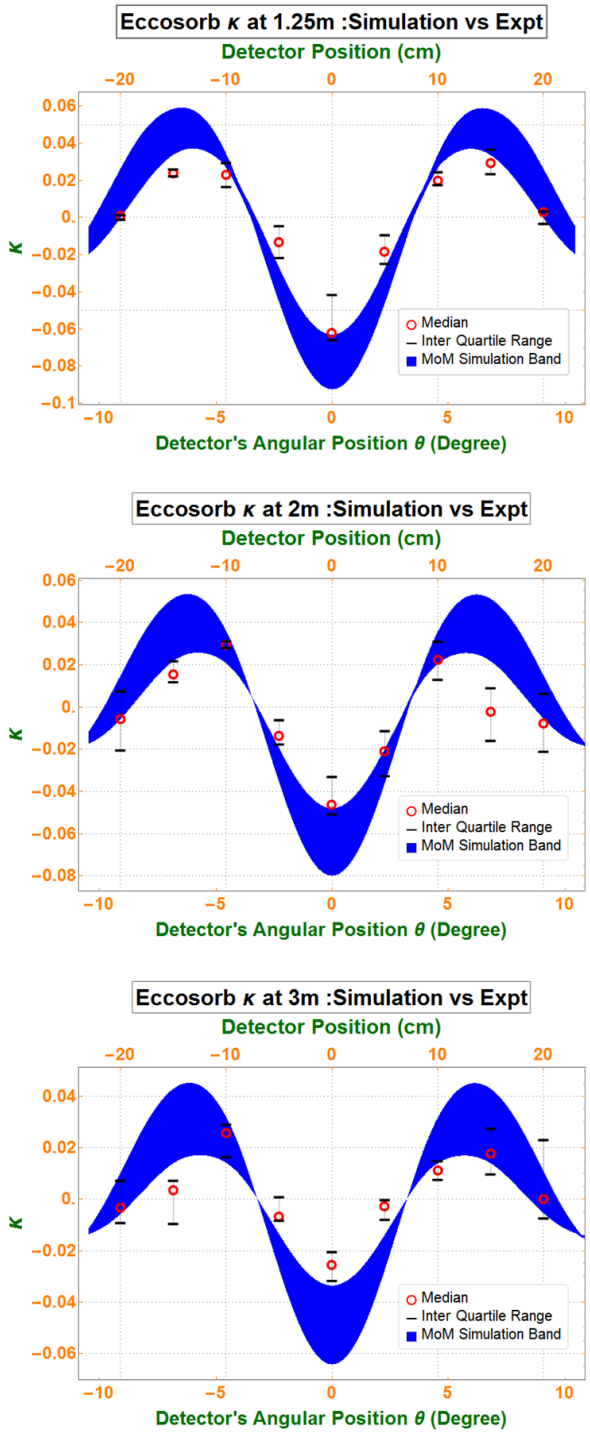


Figure 6.6: The plot on top is for source-slot plane distance of 1.25 m. The middle plot is for source-slot plane distance of 2 m. The bottom plot is for source-slot plane distance of 3 m. Slot plane - detector plane distance is kept 1.25 m in all three cases. As can be seen, experiment and theory match very well in all cases. Theory predicts a drop in the κ values with increasing source-slot plane distance which is corroborated by experiment. As distance increases, the noise remains similar but signal drops making signal to noise go down which results in generally bigger error bars at larger distance.

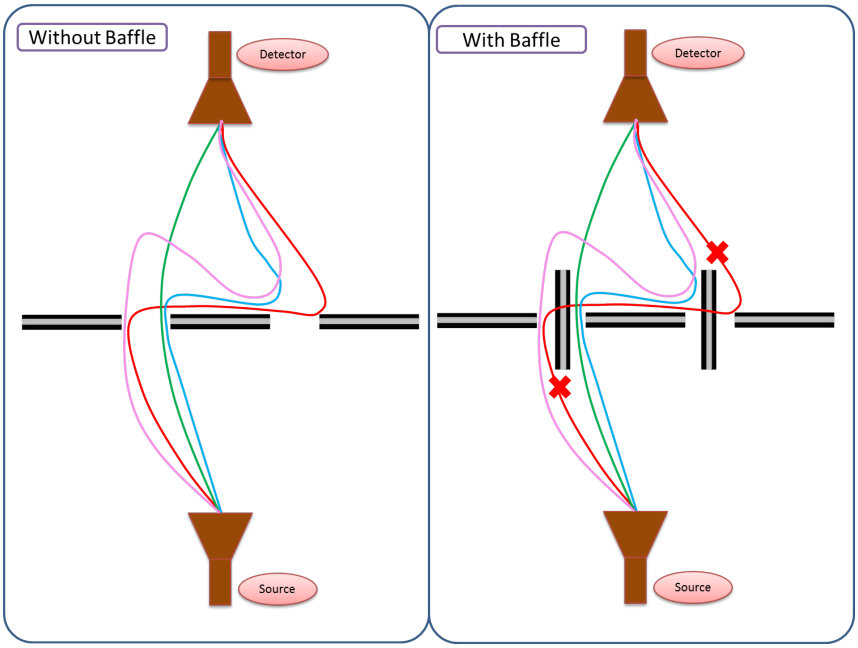


Figure 6.7: Red lines indicate class of non classical paths that get blocked by baffle, Blue lines indicate those that are still allowed and may contribute to κ, Pink lines indicate those that get suppressed with increase in baffle size, Green lines indicate classical paths.

Killing κ with a baffle

In order to have an independent check on the fact that κ as measured is not due to some unaccounted for errors, we have added absorbers perpendicular to the slots (in between the slots) to in principle “slowly kill" the effects due to non classical paths (see Figure 4). It had postulated that the paths which cross the slit plane twice [6.39] i.e. the hugging paths would have maximum contribution to κ. By placing such perpendicular baffles and measuring κ as a function of increasing baffle size, we find that κ decreases with increase in baffle size and show that it is indeed the hugging paths that contribute maximally to non-zero κ. Figure 5 shows $|\kappa|$ at central detector position as a function of increasing baffle size. This is definitive alternative proof that what we are observing is a real physical effect and not a result of some unaccounted for error. This makes ours a tunable experiment where the baffle size gives the tunability parameter. In the absence of baffle, both classical and non-classical paths contribute to κ. As the baffle size increases, the contribution due to non-classical paths diminishes and finally becomes zero (which is observed by κ becoming zero). Thus, the baffle experiment demonstrates the true effect due to the exotic non-classical paths and brings forth the reality of Feynman paths in a classical domain.

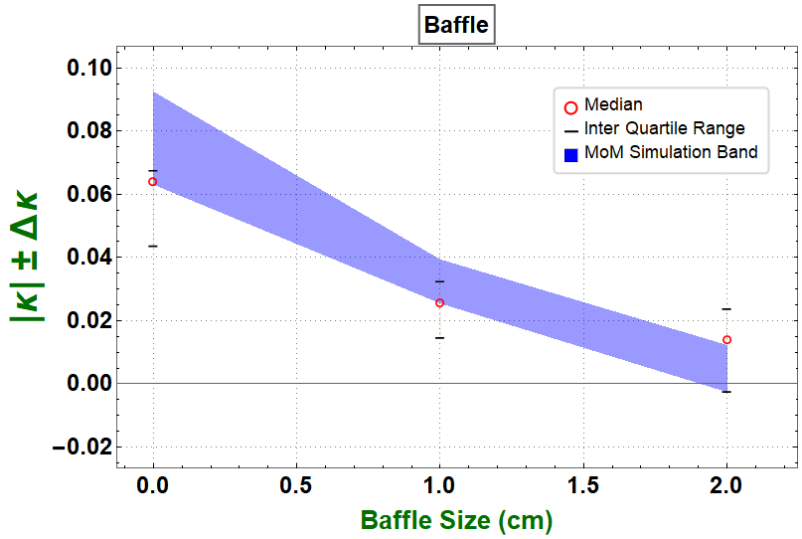


Figure 6.8: $|\kappa|$ at central detector position as a function of baffle size. Red markers represent medians of 10 measured $|\kappa|$ values for different baffle sizes, black lines denote the interquartile range, shaded blue region is the theory band from MOM simulation. We have used p_{BG} at the central detector position as the normalization factor. In order to plot $|\kappa|$, we take the absolute value of the median but keep the relative distance of interquartile range.

6.8 Sources of error

6.8.1 Detector non linearity analysis

One of the main errors which can lead to observation of a non-zero κ will be the non-linearity if any of the detector. If a detector behaves non-linearly with increase in incident power, then the quantity in equation 6.31 will automatically be a non-zero just from such errors. We have done detailed analysis of detector non-linearity and found that our measured κ cannot be explained by any such non-linearity, effects due to which happens to be much below the measured κ. In our analysis, we have derived a non-linear function for the detector using both spline interpolation method as well as polynomial fit and derived the resultant κ from this function. We have found the κ value so derived to be much lower than the measured κ thus indicating that the non linearity effects do not play a major role in our experiment.

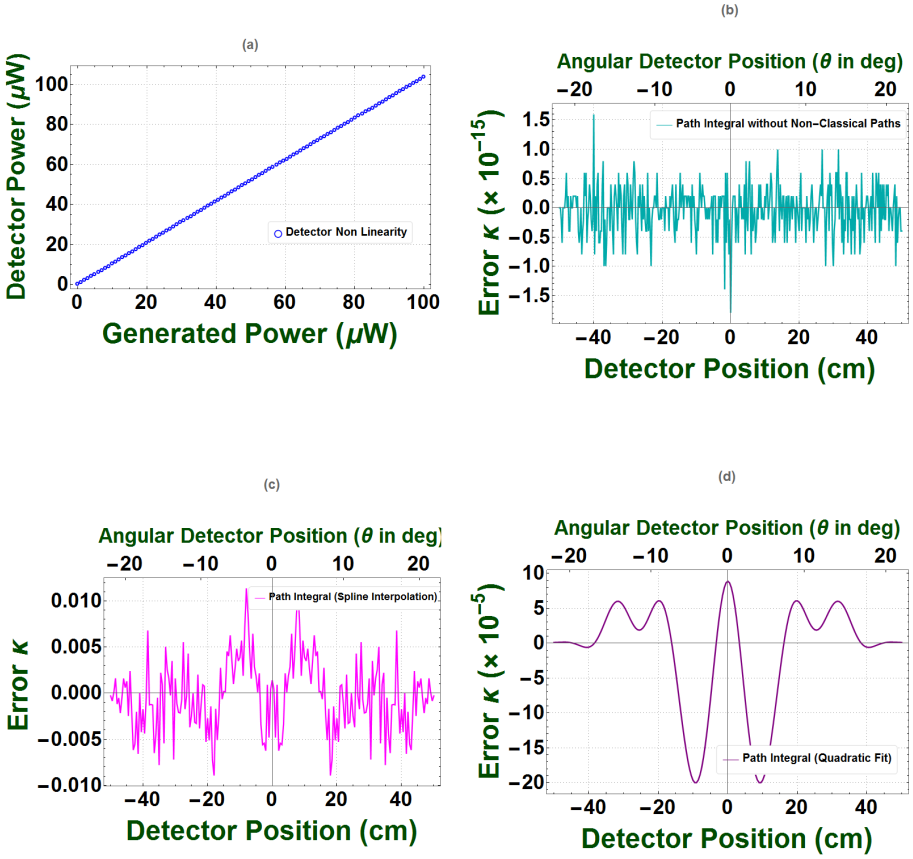


Figure 6.9: a) measured power values versus input power b) κ generated from classical paths only c) κ generated from spline interpolation method d) κ generated from polynomial fit.

Figure 6.9 shows four plots. Plot (6.9a) shows measured power values in our Agilent power probe with an Agilent signal generator acting as a source. If the detector is perfectly linear, then the measured value will be exactly the same as the input value. However, no detector can be linear up to arbitrary accuracy. We have used this plot to generate non linear functions from both spline interpolation method as well as polynomial fit. Plot (6.9b) shows the κ that is generated using only classical paths in the path integral formalism. As is expected, κ is identically zero (10^{-16} which is the accuracy of our solver, i.e. float precision for Mathematica.). We have used the power values that lead to this zero value for κ and feeding them to the non linear function generated above, derived the power values that would have been measured. The measured value will vary from the input value due to various non linearity effects.

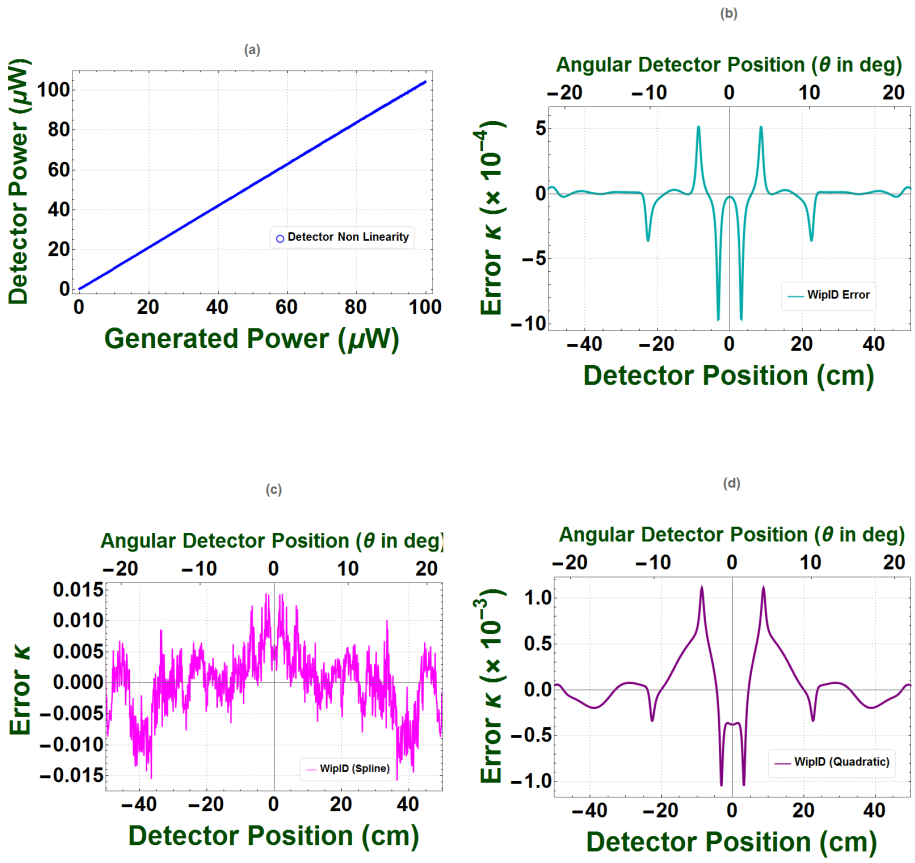


Figure 6.10: a) measured power values versus input power b) κ generated from the naive application of the superposition principle using MOM c) κ generated from spline interpolation method d) κ generated from polynomial fit.

Plot (6.9c) and Plot (6.9d) show the resultant κ as a function of detector position. These values represent what we define to be *error* κ. What we are verifying in this experiment is the deviation from the naive application of the superposition principle. While the naive application gives us an expected zero value for κ, the correct application brings forth the non-zeroness. Following references [6] and [8], we define κ only in the presence of "classical" paths or in other words when the Superposition principle is naively applied. This is expected to be zero in ideal theory. However, in case of real experimental/simulation scenarios which involve non-ideal conditions, this quantity can be a non-zero. One has to appreciate that this non-zero is simply due to different sources of error as the case may be and has nothing to do with the correction to the application of the Superposition principle which is a"real" non-zero as opposed to an error bound. We call such an error bound"error κ." This quantity derived simply from the terms involving classical paths comes in very

handy as it tells us whether some source of error has a competing effect with the actual non-zero value. Thus if error κ is lower order in magnitude than actual κ, we need not worry about a particular error playing a dominant role in explaining the non-zeroness of κ. In case of Plot (6c) and (6d), as they have been generated from taking into account only the contribution from the classical paths in path integral formalism, they should have been zero. However, non linearity effects make them non-zero.

Figure 6.10 shows a similar analysis done using WIPL-D (our MOM solver) in which the accuracy of the solver is 10^{-4}.

Two very important points should be noted here.

- The non linearity effects captured here reflect the maximum non linearity that can affect our experiment which is of course not representative. Even in this worst case scenario based simulation, the values of κ are many times smaller for the interpolation method and two to three orders of magnitude lower for the polynomial fit. Thus, they do not in any way explain the results obtained in the experiment.
- Plot (a) captures the non linearity not only due to the power probe but also the source signal generator itself. There is no trusted device that one can assume is perfectly linear and use as a source such that only the measurement device non linearity can be captured. In our experiment, the source is used at a constant power and thus non linearity due to the source does not affect us. The effective non linearity seen in our experiment is thus lower than what we have been able to estimate. The issue of a trusted device also existed in previous work [**gregorone**] as there the attenuator was assumed to be a trusted device.

6.8.2 Effect of detector size on κ

We have simulated the effect on κ from changing detector size inspired by observations in [15]. The detector size is varied as a fraction of the mean experimental aperture size value. It is varied uniformly for both E and H plane. Zero aperture size represents a screen detector. The uncertainty band has been formed by taking into account uncertainties in experimental parameters and the plot in Figure 13(a) indicates that our measured κ value does not depend significantly on detector aperture size.

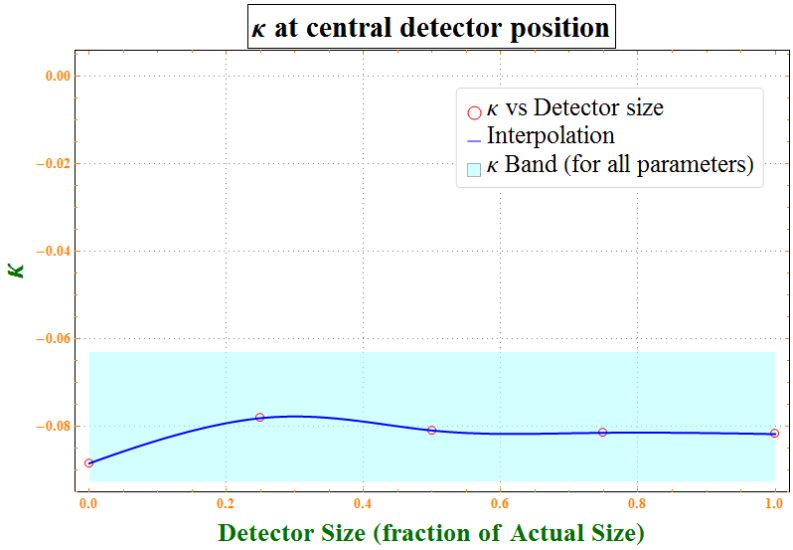
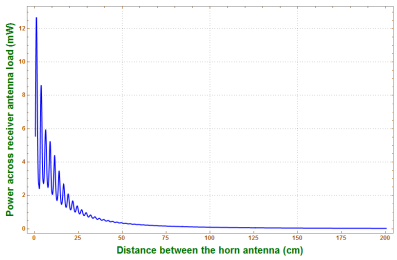


Figure 6.11: κ at the central detector position as a function of varying detector aperture size

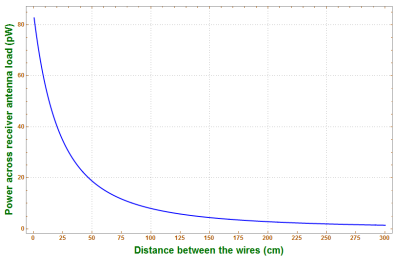
6.8.3 Quantifying the effect of Horn Source and Reciever distance

The reciever antenna in any experiment does not perform ideally i.e. it partly changes the source by its mere presence. The effects due to reflection from the horn structure may go back to the source. Therefore its important to maintain the distance between horn source and reciever such that these effects do not dominate our experimental results. In this section our approach would be to estimate what is the maximum contribution to kappa from various effects. By showing that the maximum contribution from these effects is less than the maximum kappa obtained in our experiment, we can claim that the kappa obtained in our experiment is not due to any of these effects alone.

Reciever Power as a function of Horn source and receiver distance The horn source antenna was fixed and one end of the probe was fed with a voltage source of peak to peak voltage of $3.556V$, which is equivalent to giving $15dBm$ power with 50Ω impedance. The horn structure as a receiver has 50Ω load across its probe wire as well. We record the power across this load and varied the distance between the horn antenna. Here all the distance, unless otherwise mentioned, are measured from face of the flare aperture.



(a) When horn antenna are closer to each other, we clearly see the effect due to interference of the reflected beam. At higher distances the oscillation drops (but so does the power)



(b) We see that, with the wires only, the power drops off quite smoothly without any oscillations

At very close distance, we see the interference pattern typical to a cavity. Here, the two back plates of the waveguide part of the horn antenna act as the reflectors that give rise to such effect. To ensure that this effect is not simply due to the mutual coupling between the probe wires, we removed the horn structure and kept just the probe wires and repeated the above procedure.

From the above figure 6.12b two things are evident

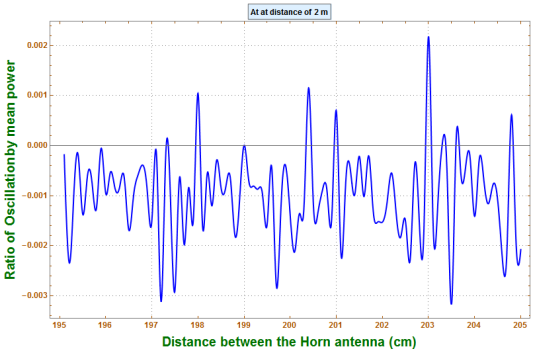
- The oscillations are caused by reflections from the horn structure and not due to mutual coupling between the probes.
- The horn antenna is useful to increase the gain by orders of magnitude. Both the source and receiver horn antenna together increase the power by a factor of about 10^9. Hence, with probe only, the signals[8] would be below the noise floor of our detector(Agilent Power Probe)

For the purpose of maximum contribution to kappa due to the above effect, its not enough that the oscillations in power are very low. What is important is the ratio of these oscillation to the mean power. For these purpose, we computed the following

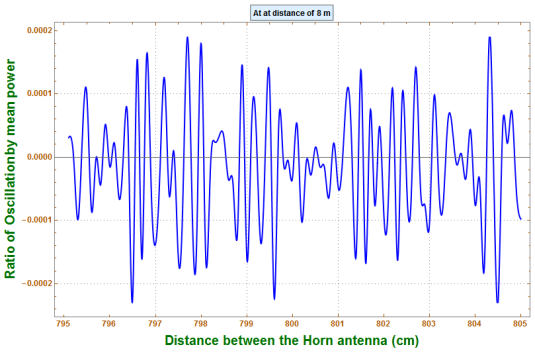
$$\frac{\text{Difference between power values between neighbouring points}}{\text{Average of the power values between neighbouring points}} \tag{6.38}$$

and we got the following plot

[8] In simulation , we have not taken care of any cable loss or other non idealness. Further, the gain of horn is very sensitive to wire height and hence we may not directly compare the absolute power from the simulation with the expt.



(a) Change in Power at a position normalised to the average power at that position



(b) The errors now decrease by another order

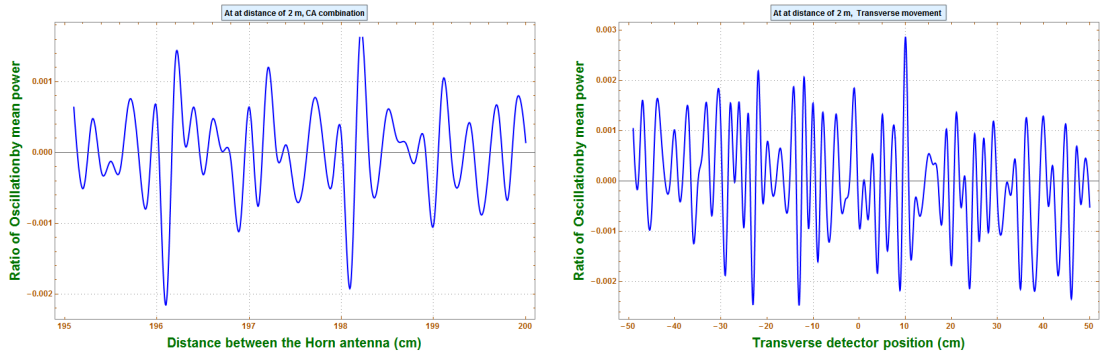
Few things noteworthy from the above graph

- The difference of power between two horn positions depends on what is the step size with which we measure the power. For instance if we measure power by variying the distance at 10 cm interval the difference would be different than the case if we vary with 1 mm interval. In this particular case, we choose the step size to be 1 cm because, it is higher than the maximum error than we can have in distance measurements(between source and detector plane) from our experiment and it is less than half of the wavelength and hence within each oscillations we are guaranteed at least two samples. Thus this yields the maximum error than we can have.
- The maximum error that we have is about 0.003. Now, the worst case scenario is the conspiracy in which errors will add up to give us a kappa of about $8 \times 0.003 = 0.024$. Still this is lower than the maximum value of kappa. Hence even the worst case scenario from this effect does not explain kappa and thus this is not the reason. On an average, this oscillations are about 0001 and if the errors add up randomly, this would yield an error[9] in kappa which is less than 0.003.We repeat the above but at about $8m$ source to detector distance
- Note that although the computed ratio has decreased by an order, so has the actual power. Since, our source fluctuation itself is about 0.001 and above. the source fluctuations forms a bottleneck and ruins the advantage of having the source and detector at a higher distance. Further, the lowest power, the device can read is about $1nW$. Thus, to achieve greater than 3 orders of precision, the detected power should be μW and above, which is hard to achieve with the detector at 8 m. By choosing

[9] $2\sqrt{2} \approx 2.8$

the distance to be 2 m, instead of 8, the detected power roughly increases by a factor of 16 (and also kappa is predicted to be larger).

To see that for most other combinations, this ratio is well below 0.003, below we also plot the same for slot combinations CA where the detector is moved at about 2 m.



(a) For CA combination the above stated ratio is less than 0.002

(b) The horn detector is moved in detector plane at 2m distance

Horn detector moving in transverse direction Since there would be also interference effects when horn detector moves in the transverse direction, it is also important to check that the above mentioned ratio is not higher. Below, we have plotted the power recorded by the horn detector for the background combination.

We again see that, the maximum value is about 0.003 and hence its maximal propagation to kappa is still less than 0.06.

All the above results suggest that, kappa greater than magnitude of 0.025 cannot be explained by the above effects due to reflection from the horn structure.

6.8.4 Quantifying the effect of Reflection from the Slots

When we place the slots directly in front of the horn source, there would be reflections back to the source. Now the horn source wire may pick up these reflections and the emitted source profile may change. Here, using simulations, we quantify the upper bound that the reflections would have on κ.

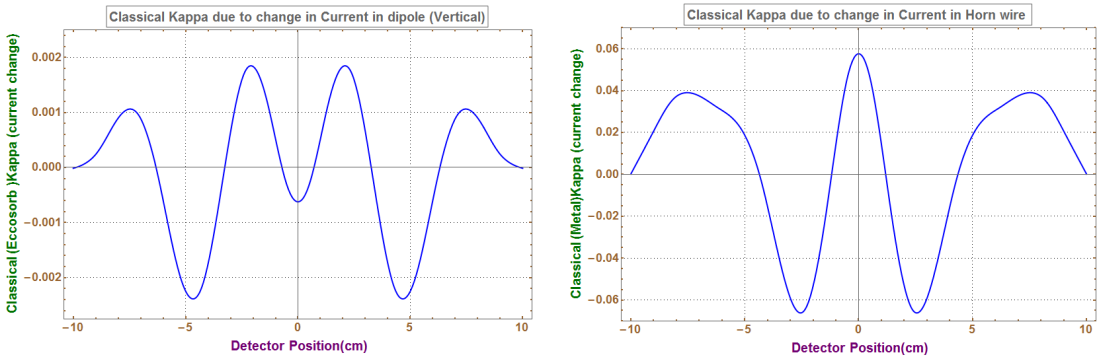
Tests based on currents in the source wire Hence we would like to compute maximum contribution to kappa because of change in source, if any. First, we would assume that reflection leads to change in current in the source wire, which in return leads to change in

radiated field. Here, we would calculate the following In path integral, we multiply the source field with the complex current in the source wire and then obtain the power patterns at the detector plane. From there, we compute kappa.This (multiplied by 3) gives us a bound.

The algorithm is as follows

- Using MoM method, compute the complex current in the wire for various slot combinations.
- In path integral, when non classical paths are turned off, scale the source wavefunction by the complex current for a given combination. Since current in the wire for different slot combinations is different different combination sees a different source power and hence we have a kappa, which we plot below.

Eccosorb simulations were performed with the NAL parameters with $\varepsilon_0 = 10.01 - 0.078078i$ and $\mu = 1.5 - 0.5i$. The source was a vertically oriented wire.



(a) Eccosorb Kappa due to current change in dipole

(b) Metal: Kappa due to change in current due to horn source wire

We see that kappa, if at all is due to change in source, is 0.0025 or less (almost 50 times less than the actual kappa, see next part for kappa for this configuration). Note that while computing the above classical kappa due to change in current, we have used an effective slot width of 7.5 cm and a multiplication factor of 3.

For metals, we take the horn source and carry current in the horn wire probe. We found the kappa computed from this method is five times less than simulated kappa(about 0.3) and less than the experiment (about 0.17) by a factor of two or more.

Note that, kappa from path integral for metals, i.e. slot width of 10 cm, still has a peak at centre and has a value of 0.1(after multiplication factor of 3).

Test with Equivalent Sources without Structure The previous section revolves around the argument that there could be some systematic error if there is reflection back from the slots onto the source and the source profile changes either due to mutual coupling or due to reflection from the horn structure (of the reflected waves from the slot) to the slot/detector. In this section we would argue that, if we dont have any horn structure, there is no question of reflection from the horn structure. Since there are no wires as well, there is no question of coupling either. The algorithm is as follows

- For the background combination, the complex electric and magnetic fields at the horn source aperture were calculated . (The aperture was sampled in a grid of 101 and 51 points along H-plane and E-plane respectively)
- The above fields are impressed in the simulation (i.e. they are assumed to be there without any source/structure). Since the same source is impressed for each combination and the source is guaranteed not to change, any difference in kappa between the case where the equivalent source is impressed and the actual structure acts as a source would mean effects due to reflection/coupling etc.

We follow the above procedure for metal slots, screen detector, and found that except at the centre where kappa differs a little, elsewhere kappa due to impressed source/due to back reflection do not change.

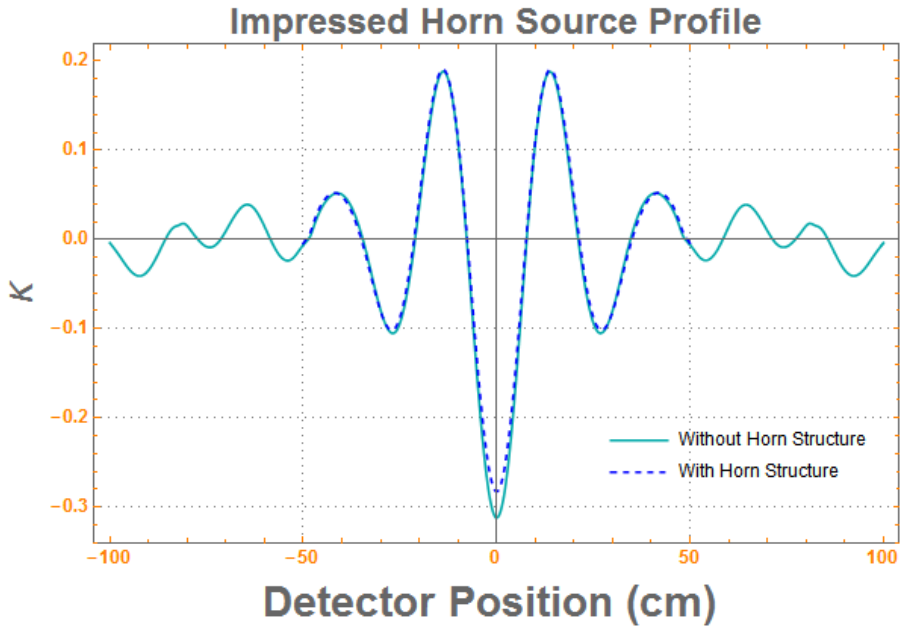
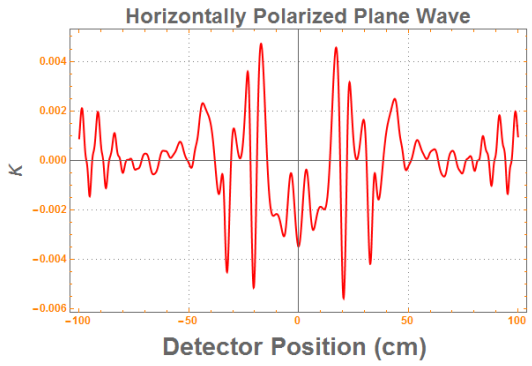


Figure 6.16: Metal Slots, Screen detector at 100 cm from slot plane. Kappa with the impressed source and the actual horn structure are almost the same except at the centre even where the inclusion of horn structure shifts the dip towards zero.

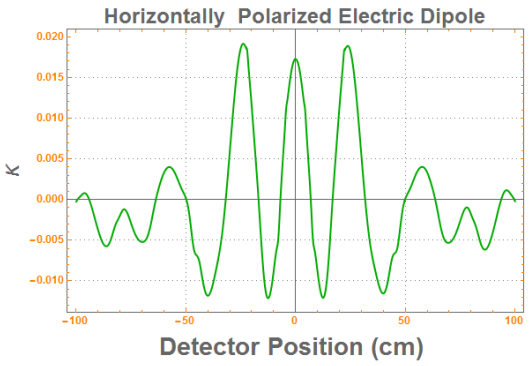
This seals the case that reflections are not playing a dominant part in kappa.

Polarization and Source Profile

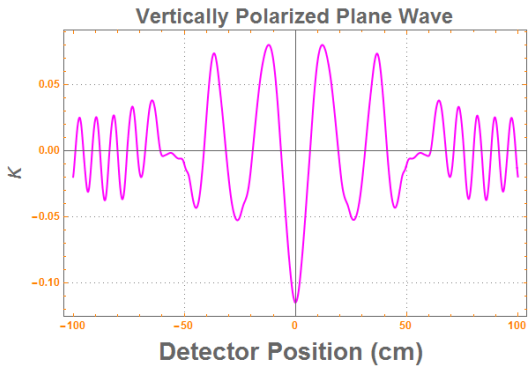
For metal slots and screen detector, we compute kappa for ideal plane wave source , an ideal dipole source (situated 100 cm from the horn) , and for both cases we made the polarization horizontal and vertical.



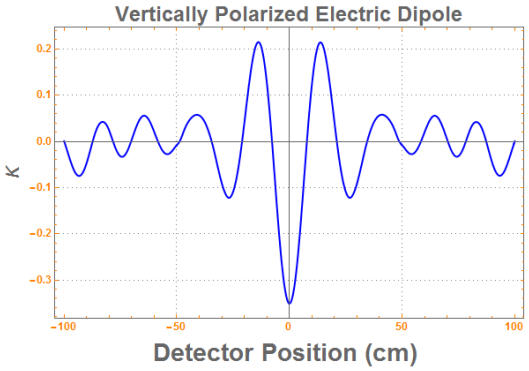
(a)



(b) HorizontallyPolarizedElectricDipole



(c)



(d) HorizontallyPolarizedElectricDipole

- The horizontally polarized source consistently produce kappa of lesser magnitude than the vertically polarized counterparts. This could be due to the fact the larger length scale of the slots is along vertical direction or may be because the slits combinations are changed horizontally.
- Changing the source from Plane wave to dipole does not change the modulation of kappa close to centre.

6.9 Comparison of Path integral result with Finite Difference Time Domain simulations

The parameters used for the FDTD simulations were vertically polarized point dipole source, source to slot plane distance of 1m, slot plane to detector plane distance of 1m, simulation box = 2m along slot plane * 2.5m along the beam propagation direction, composite slots of 3.0mm aluminium sandwitched between 2.2.mm Eccosorb material, slot width of 10cm, inter-slot distance of 13cm. Eccossorb material model was for a

paramagnetic material with permittivity of 11.107 and permeability of 1.912. For the path integral simulations, we used slot width of 7cm, inter-slot distance of 13cm, wavelength of 5cm, point source with same source plane-slot plane and slot-plane-detector plane distances as FDTD.

We select a point source at $\vec{r}_S = (x_S =, y_S = 0.0, z_S = 10m)$ and point detector at $\vec{r}_D = (x_D, y_D = 0.0, z_D = 14inch)$. The intermediate plane is the triple slit placed at $\vec{r}_s = (x_s, y_s = 0.0, z_s = 0.0)$

Propagators in Triple Slit Expt In path integral we use the Huygen Fresnel propagator

$$K_0(r_f, r_i) = -\frac{i}{\lambda}\frac{e^{ikr}}{r} \tag{6.39}$$

Here, $r = |r_f - r_f|$ the displacement vector from initial to final position.

After normalizing all length scales with λ, we perform the subsequent computation in 2D.

The first order "classical" propagators which cross the slit plane once while propagating from source r_S to detector r_D is given by

$$K1[r_S, r_D] = \int dr_s K_0(r_S, r_s) K_0(r_s, r_D) \tag{6.40}$$

Here, the coordinate over the slit plane r_s is integrated over once.

The contribution from non classical hugging paths can be computed from the second order propagator with

$$K2[r_S, r_D] = \int d\ r_{s_1} K_0(r_S, r_{s_1}) K_0(r_{s_1}, r_{s_2}) K_0(r_{s_2}, r_D) \tag{6.41}$$

Here, s_1 and s_2 are the two distinct skits integrated over.

For the single combinations A, B or C, no non classical paths are considered. For two slit combinations, AB, BC, and CA, we consider two non classical paths: one path that goes from source to first slit (say A) to second slit (say B) to detector and the other path is from source to second slit (say B) to first slit (say A) to detector.

The three slit combination has 6 non classical paths: S to A to B to D,S to A to C to D,S to B to A to D ,S to B to C to D, S to C to A to D and S to C to A to B;

Adding thickness to slit Here we change the propagators to the following

$$K1 = K1[r_S, r_D] = \int dr_{s_1} \int dr_{s_2} K_0(r_S, r_{s_1}) K0(r_{s_1}, r_{s_2}) K_0(r_{s_2}, r_D) \tag{6.42}$$

Here s_1 and s_2 is the same slit at two different planes. That is in 2D vector form $s_1 = (x_A, zc = st/2)$ and $s_2 = (x_A, zc + st/2)$ Here, zc is centre of the slot plane and st is the thickness. In the above example, s_1 represents the back plane of slit A and s_2 represents the front face of slit A.

Similarly, we change K2 to involve additional integration.

$$K2f[r_S, r_D] = \int d\ r_{s_1} \int d\ r_{s_2} \int d\ r_{s_3} K_0(r_S, r_{s_1}) K_0(r_{s_1}, r_{s_2}) K_0(r_{s_2}, r_{s_3}) K_0(r_{s_3}, r_D) \tag{6.43}$$

Here, $K2f$ stands for the path hugs the front of slit plane. So for AB combination, this path would be $s_1 = (x_A, zc - st)$, $s_2 = (x_A, zc_s t)$ and $s_3 = (x_B, zc + st)$. Note that the slit changes in the second plane. Similarly we define K2b where the path hugs the back surface. So the slit planes are as $s_1 = (x_A, zc - st)$, $s_2 = (x_B, zc - st)$ and $s_3 = (z_B, zc + st)$.

Overalll $K_2 = K2f + K2b$

Treating as Effective four Slits We can treat the three slot problem as equivalent 4 slits. Indeed, we can treat the region to have 7 gaps:

- LAg = gap from left edge of A onwards
- Ag = gap where slot A is present
- ABg = gap between A and B slot
- Bg = B slot
- BCg = gap between B and C
- Cg = C slot
- RCg = gap from right of A onwards

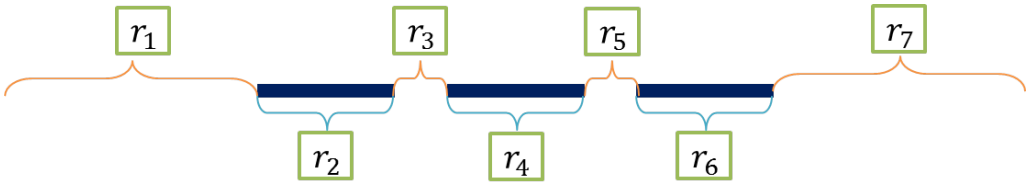


Figure 6.18: Treating triple slots as 4 slits

For background, the integration over K_1 is done over union of all the above regions. For slot A combination, we have K2 contributions for paths that goes from LAg to $ABg \cup Bg \cup BCg \cup Cg \cup RCg$ vice versa. Similarly, we view each of single slot as an effective two slit problem.

For two slots we see the problem as effective 3 slits. For instance, for combination AB, we have the following non classical paths :

- LAg to ABg
- LAg to $BCg \cup Cg \cup RCg$
- ABg to LAg
- ABg to $BCg \cup Cg \cup RCg$
- $BCg \cup Cg \cup RCg$ to LAg
- $BCg \cup Cg \cup RCg$ to ABg

So just like 3 slits, for 2 slots we have to consider 6 hugging paths.

When all the three slots are present, we have 12 such paths to consider

- LAg to ABg
- LAg to BCg
- LAg to RCg
- ABg to LAg
- ABg to BCg
- ABg to RCg
- BCg to LAg
- BCg to ABg
- BCg to RCg
- RCg to LAg
- Rcg to ABg
- RCg to BCg

All the boundaries of the above 7 seven regions are decided by the slot width and inter slot distance except LAg an RAg. For this we arbitrarily choose a high number (say 100 lambda) for the left and right edge from the centre respectively. The number is chosen such that there is no change in central kappa (say up to 4 decimal places) when we change the limits by a factor of 2

Using Babinet principle Using the equivalent 4 slit case has the computational issue that for all slot combinations , we needed to an integration over infinite domain, which we

do truncate. Further, we cannot extend it to a thick slot as it would involve two integrals over very large domains and the convergence will be slow.

Using Babinet principle we evaluate the power for a slot combination as follows.

- First we calculate the back ground field, BGf) by evaluating K1 integration over $(-1000\lambda, 1000\lambda)$
- Compute the power for slit case. For instance, when slot A and B are placed, treat A and B as slits and compute the field. F_{slit}
- Compute the field for slot as $F_{slot} = BGf - F_{slit}$
- absolute square of the above gives the power and then apply the slot kappa formula

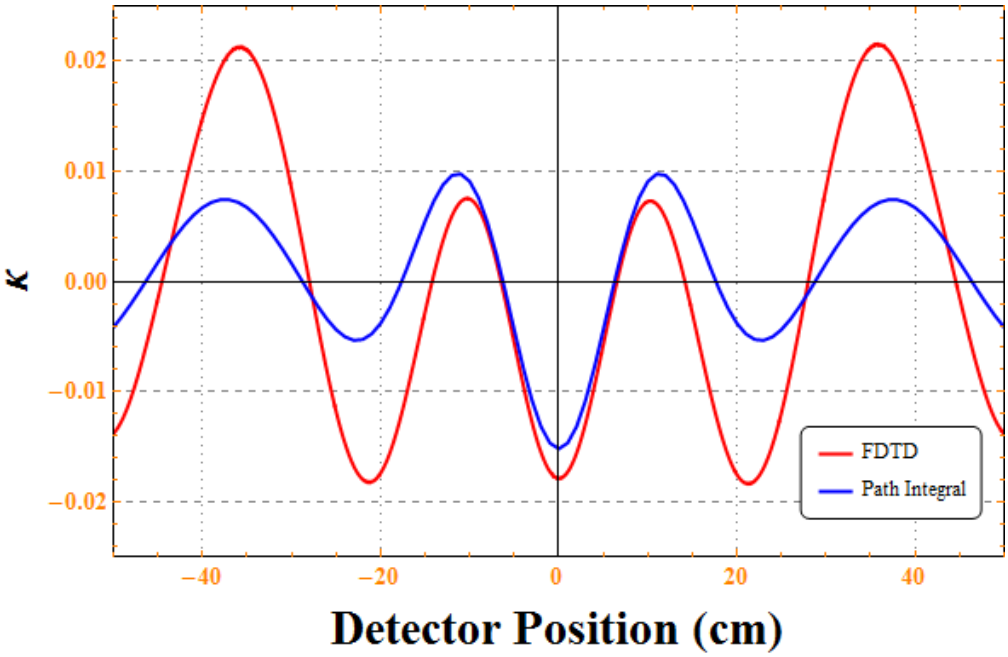


Figure 6.19: The red line shows κ as a function of detector position obtained using FDTD. The blue line is generated using path integral formalism. The effective slot width in path integral has been taken to be 7 cm as opposed to the actual width of 10 cm. As path integral does not capture material properties, it fails to capture the effect due to waves penetrating the material leading to an effective slot width which is smaller than actual one. This has been discussed in detail in . FDTD on the other hand has material parameters as input so does not require the concept of effective slot width.

While FDTD and Path Integral show reasonable match in magnitude and modulation, the plot does not match very well in magnitude with experiments as well as simulations based on Method of Moments (MOM). Both FDTD and Path integral have several shortcomings as compared to the full wave MOM based simulation. While these are 2D

simulations, MOM is 3D. In MOM, one can define the horn source and horn detector while in both these methods, we use point source and point detector. Moreover, being 2D simulations, both FDTD and path integral assume an infinite slot height which is not the case in reality (slot height is 30 cm). Moreover, both path integral and FDTD will have errors as one moves away from the centre. Errors in path integral are due to finite integration domain whereas in FDTD, one has errors due to PML reflections. Over and above these, path integral suffers from some additional limitations. The path integral formalism used in this paper as well in [6] and [8] is based on scalar field theory whereas FDTD can take into account source polarization. In path integral, material properties need to be accounted for by the concept of “effective width". Also, we use the thin slot approximation whereas the slot actually has finite thickness.

Inspite of the above limitations, path integral formalism gives the same order of magnitude for κ as FDTD as well as similar modulation as experiment. It also serves as a useful aid to distinguish between actual non-zeroness of κ and error contributions as the contributions from the non classical paths can be turned off and on at will. It is thus a handy theoretical tool which is perhaps slightly too ideal to expect perfect match with experiments.

6.10 Mutual Inductance in Electromagnetism & κ

The three slits in a triple slit experiment can be considered as three virtual secondary sources. The non-classical paths used in path integral formulation can then be thought of mutual coupling between the slits, i.e., the emission through slit A is affected by whether slit B or C are open. This can be better illustrated without getting into complications of optics and simply considering electromagnetic radiators.

We consider a wire of radius $\lambda/100$ and length $\lambda/2$ with a centre fed port to be an array element.There are segments [10] along the length of wire due to meshing.

When we want to make the wire act as a source, a voltage source is attached to the port which give $1kV$ voltage across the port. Also, the load across the port is kept to be 50Ω which active. When the source is inactive the load across the port is made to be $1M\Omega$

When we want to model the wire as a receiving antenna we don’t attach any voltage

[10]The material is segmented into a number of volumes/surfaces which gets solved using MoM. However, since we use PEC for the material of wire, segmentation does not cause any accuracy issue

source across the port. But it does have a load of 50Ω when considered active and a load of $1M\Omega$ when considered inactive.

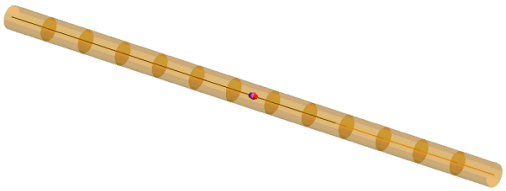


Figure 6.20: A single wire with center fed port

6.10.1 Antenna Configuration

We define baseline to be the vector $\vec{b}$ that joins the centr of the array elements. If the length of the wires are along the baseline we call the configuration to be “inline". If the length of the wires are perpendicular to the baseline, we call the configuration to be parallel.

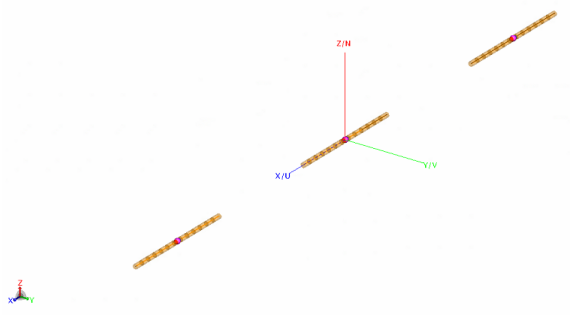


Figure 6.21: Inline configuration

In the coordinate system described in 6.21, we place another wire of the same width and thickness at $z = 10^3\lambda$ oriented such that its length is along x. When the three wires are treated as source, this wire is treated as a reciever and vice versa. This wire is moved and when it acts as a source, we call it the case that the source is moving and the three wires in combination act as a detector. We take the current at load for each wire and coherently[11] add it to get the total current in the detector. We take the absolute square of it to find the power[12]. According to reciprocity, this power has the same value when the three wires act

[11] complex addition i.e. considering both magnitude and phase
[12] $P = I^2R$ We can have any test resistance after we have added the currents and the reciprocity will hold

as source and we measure the power from currents across the load in the detector that is at $z = 10^3\lambda$.

6.10.2 Comparison of κ obtained in simulation with Path Integral calculation

For comparison with analytical path integral formula, we need to move the detector in equal intervals of theta along a circle with a large radius. However, since the power picked up by a metal wire is polarization depended, we need to orient the wire along x for maximum coupling. Thus, there would be an additional $\cos\theta$ factor coming in. To avoid the above complication, we compare the kappa obtained from power obtained from Poynting vector (all component) rather than that of a single polarization. For inline configuration, kappa computed from analytical formula and numerical MoM method do have similar modulation and magnitude.

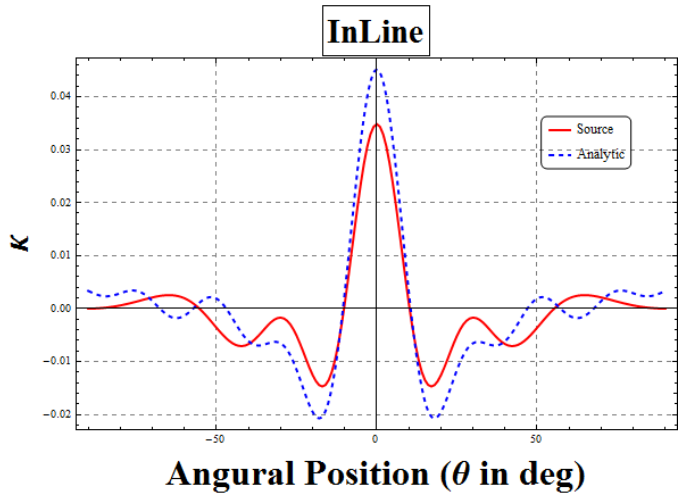


Figure 6.22: Comparision of κ computed using path integral and MoM simulation

References

[6.1] E. Schrödinger. “Quantisierung als Eigenwertproblem”. In: *Annalen der Physik* 385.13 (1926), pp. 437–490. DOI: https://doi.org/10.1002/andp.19263851302.

[6.2] E. Schrödinger. *Collected Papers on Wave Mechanics*. AMS Chelsea Publishing Series. AMS Chelsea Pub., 2003. ISBN: 9780821835241.

[6.3] M. Born and E Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999. ISBN: 9780521642224.

[6.4] Ernest Rutherford and H Geiger. "An electrical method of counting the number of α-particles from radio-active substances". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 81.546 (1908), pp. 141–161. DOI: 10.1098/rspa.1908.0065.

[6.5] E. Schrödinger. "An Undulatory Theory of the Mechanics of Atoms and Molecules". In: *Phys. Rev.* 28 (6 1926), pp. 1049–1070. DOI: 10.1103/PhysRev.28.1049.

[6.6] Guido Bacciagaluppi and Antony Valentini. *Quantum Theory at the Crossroads*. 2013.

[6.7] C. Davisson and L. H. Germer. "Diffraction of Electrons by a Crystal of Nickel". In: *Phys. Rev.* 30 (6 1927), pp. 705–740. DOI: 10.1103/PhysRev.30.705.

[6.8] Max Born. "Zur Quantenmechanik der Stoßvorgänge". In: *Zeitschrift für Physik* 37 (), pp. 863–867.

[6.9] Max Born. "Statistical Interpretation of Quantum Mechanics". In: *Science* 122.3172 (1955), pp. 675–679. DOI: 10.1126/science.122.3172.675.

[6.10] L. E. Ballentine. "The Statistical Interpretation of Quantum Mechanics". In: *Rev. Mod. Phys.* 42 (4 1970), pp. 358–381. DOI: 10.1103/RevModPhys.42.358.

[6.11] A. Pais. "Max Born's Statistical Interpretation of Quantum Mechanics". In: *Science* 218.4578 (1982), pp. 1193–1198. DOI: 10.1126/science.218.4578.1193.

[6.12] Nicolaas P. Landsman. "Born Rule and its Interpretation". In: *Compendium of Quantum Physics*. Ed. by Daniel Greenberger, Klaus Hentschel, and Friedel Weinert. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 64–70. ISBN: 978-3-540-70626-7. DOI: 10.1007/978-3-540-70626-7_20.

[6.13] Andrew Gleason. “Measures on the Closed Subspaces of a Hilbert Space”. In: *Indiana Univ. Math. J.* 6 (4 1957), pp. 885–893. ISSN: 0022-2518.

[6.14] Alexia Auffèves and Philippe Grangier. “Deriving Born’s Rule from an Inference to the Best Explanation”. In: *Foundations of Physics* 50.12 (2020), pp. 1781–1793. DOI: 10.1007/s10701-020-00326-8.

[6.15] David Deutsch. “Quantum theory of probability and decisions”. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 455.1988 (1999), pp. 3129–3137. DOI: 10.1098/rspa.1999.0443.

[6.16] Wojciech Hubert Zurek. “Environment-Assisted Invariance, Entanglement, and Probabilities in Quantum Physics”. In: *Phys. Rev. Lett.* 90 (12 2003), p. 120404. DOI: 10.1103/PhysRevLett.90.120404.

[6.17] Wojciech Hubert Zurek. “Probabilities from entanglement, Born’s rule $p_k = |\psi_k|^2$ from envariance”. In: *Phys. Rev. A* 71 (5 2005), p. 052105. DOI: 10.1103/PhysRevA.71.052105.

[6.18] Sabine Hossenfelder. “A derivation of Born’s rule from symmetry”. In: *Annals of Physics* 425 (2021), p. 168394. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2020.168394.

[6.19] Charles T. Sebens and Sean M. Carroll. “Self-locating Uncertainty and the Origin of Probability in Everettian Quantum Mechanics”. In: *The British Journal for the Philosophy of Science* 69.1 (2018), pp. 25–74. DOI: 10.1093/bjps/axw004.

[6.20] Lev Vaidman. “Derivations of the Born Rule”. In: *Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky*. Ed. by Meir Hemmo and Orly Shenker. Cham: Springer International Publishing, 2020, pp. 567–584. ISBN: 978-3-030-34316-3. DOI: 10.1007/978-3-030-34316-3_26.

[6.21] R. Shankar. *Principles of Quantum Mechanics*. Springer US, 2012. ISBN: 9781475705768.

[6.22] Scott Aaronson. “Quantum computing, postselection, and probabilistic polynomial-time”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 461.2063 (2005), pp. 3473–3482. DOI: 10.1098/rspa.2005.1546.

[6.23] Ruth Kastner. *Violation of the Born Rule: Implications for Macroscopic Fields*. 2016.

[6.24] Rafael D. Sorkin. "QUANTUM MECHANICS AS QUANTUM MEASURE THEORY". In: *Modern Physics Letters A* 09.33 (1994), pp. 3119–3127. DOI: 10.1142/S021773239400294X.

[6.25] Karthik S. Joshi, R. Srikanth, and Urbasi Sinha. "Violation of no-signaling in higher-order quantum measure theories". In: *International Journal of Quantum Information* 14.5, 1650024 (2016), p. 1650024. DOI: 10.1142/S0219749916500246.

[6.26] Urbasi Sinha et al. "Testing Born's Rule in Quantum Mechanics with a Triple Slit Experiment". In: *AIP Conference Proceedings* 1101.1 (2009), pp. 200–207. DOI: 10.1063/1.3109942.

[6.27] Urbasi Sinha et al. "Ruling Out Multi-Order Interference in Quantum Mechanics". In: *Science* 329.5990 (2010), pp. 418–421. DOI: 10.1126/science.1190545.

[6.28] D K Park, O Moussa, and R Laflamme. "Three path interference using nuclear magnetic resonance: a test of the consistency of Born's rule". In: *New Journal of Physics* 14.11 (2012), p. 113025. DOI: 10.1088/1367-2630/14/11/113025.

[6.29] Hans De Raedt, Kristel Michielsen, and Karl Hess. "Analysis of multipath interference in three-slit experiments". In: *Phys. Rev. A* 85 (1 2012), p. 012101. DOI: 10.1103/PhysRevA.85.012101.

[6.30] Immo Söllner et al. "Testing Born's Rule in Quantum Mechanics for Three Mutually Exclusive Events". In: *Foundations of Physics* 42.6 (2011), pp. 742–751. DOI: 10.1007/s10701-011-9597-5.

[6.31] Thomas Kauten et al. "Obtaining tight bounds on higher-order interferences with a 5-path interferometer". In: *New Journal of Physics* 19.3 (2017), p. 033017. DOI: 10.1088/1367-2630/aa5d98.

[6.32] J. M. Hickmann, E. J. S. Fonseca, and A. J. Jesus-Silva. "Born's rule and the interference of photons with orbital angular momentum by a triangular slit". In: *EPL (Europhysics Letters)* 96.6 (2011), p. 64006. DOI: 10.1209/0295-5075/96/64006.

[6.33] Fangzhou Jin et al. “Experimental test of Born’s rule by inspecting third-order quantum interference on a single spin in solids”. In: *Phys. Rev. A* 95 (1 2017), p. 012107. DOI: 10.1103/PhysRevA.95.012107.

[6.34] Joseph P. Cotter et al. “In search of multipath interference using large molecules”. In: *Science Advances* 3.8 (2017), e1602478. DOI: 10.1126/sciadv.1602478.

[6.35] Richard Phillips Feynman, Albert R. Hibbs, and Daniel F. Styer. *Quantum Mechanics and path integrals*. Dover Publications, 2017.

[6.36] R. P. Feynman. “Space-Time Approach to Non-Relativistic Quantum Mechanics”. In: *Rev. Mod. Phys.* 20 (2 1948), pp. 367–387. DOI: 10.1103/RevModPhys.20.367.

[6.37] H. Yabuki. “Feynman path integrals in the young double-slit experiment”. In: *International Journal of Theoretical Physics* 25.2 (1986), pp. 159–174. DOI: 10.1007/bf00677704.

[6.38] L. D. Landau and E. M. Lifschitz. “59”. In: *The classical theory of Fields*. Addison-Wesley Pub. Co., 1959, pp. 156–170.

[6.39] Rahul Sawant et al. “Nonclassical Paths in Quantum Interference Experiments”. In: *Phys. Rev. Lett.* 113 (12 2014), p. 120406. DOI: 10.1103/PhysRevLett.113.120406.

[6.40] Aninda Sinha, Aravind H. Vijay, and Urbasi Sinha. “On the superposition principle in interference experiments”. In: *Scientific Reports* 5.1 (2015). DOI: 10.1038/srep10304.

[6.41] G Rengaraj et al. “Measuring the deviation from the superposition principle in interference experiments”. In: *New Journal of Physics* 20.6 (2018), p. 063049. DOI: 10.1088/1367-2630/aac92c.

[6.42] Omar S Magaña-Loaiza et al. “Exotic looped trajectories of photons in three-slit interference”. In: *Nature Communications* 7.1 (2016). DOI: 10.1038/ncomms13987.

[6.43] Lee A. Rozema et al. “Higher-Order Interference in Quantum Mechanics Induced by Optical Nonlinearities”. In: *Conference on Lasers and Electro-Optics*. OSA, 2020. DOI: 10.1364/cleo_at.2020.jtu2a.22.

[6.44] Lee A. Rozema et al. “Higher-order interference between multiple quantum particles interacting nonlinearly”. In: *Phys. Rev. A* 103 (5 2021), p. 052204. DOI: 10.1103/PhysRevA.103.052204.

[6.45] Kai Sheng Lee et al. “Atomic test of higher-order interference”. In: *Phys. Rev. A* 101 (5 2020), p. 052111. DOI: 10.1103/PhysRevA.101.052111.

[6.46] Patrick Huber et al. “Neutrino Oscillations at JUNO, the Born Rule, and Sorkin’s Triple Path Interference”. In: (2021).

[6.47] J. O. de Almeida, M. Lewenstein, and J. Q. Quach. “Significant nonclassical paths with atoms and cavities in the double-slit experiment”. In: *Phys. Rev. A* 102 (4 2020), p. 042225. DOI: 10.1103/PhysRevA.102.042225.

[6.48] Simanraj Sadana, Lorenzo Maccone, and Urbasi Sinha. “Quantum computers to test fundamental physics or viceversa”. In: (2021).

[6.49] Marc-Oliver Pleinert, Joachim von Zanthier, and Eric Lutz. “Many-particle interference to test Born’s rule”. In: *Phys. Rev. Research* 2 (1 2020), p. 012051. DOI: 10.1103/PhysRevResearch.2.012051.

[6.50] Marc-Oliver Pleinert et al. “Testing Higher-Order Quantum Interference with Many-Particle States”. In: *Phys. Rev. Lett.* 126 (19 2021), p. 190401. DOI: 10.1103/PhysRevLett.126.190401.

Summary & Outlook

This thesis comprises mainly of three questions fundamental to quantum mechanics

- Can the property of the particle be found in a spatial region where the particle itself is absent? This sheds light on aspects of the wave-particle duality in ontological foundations of quantum mechanics.
- Can we measure the expectation value of a non-hermitian operator ? In the quest for why non-hermitian operators are usually not considered eligible as an observable, we found a method to experimentally compute their expectation values.
- Can we test for deviation from higher order interference terms? Initially, one may think that finding no deviation would validate Born rule up to the experimental uncertainties. However, in slit based experiment, owing to the boundary conditions being different, we could measure the deviation from the naive application of the superposition principle due to geometric effects.

Quantum Cheshire Cat

In this thesis, we attempt to obtain signatures of the photon being present in one arm of a Mach-Zehnder interferometer while simultaneously obtaining signatures of presence of its diagonal polarization component in the other arm. The weak value or pointer shift forms the signature which are the result of joint weak measurement on an ensemble of identically prepared and post-selected single photons. Earlier attempts by other groups have either resorted to strong demolition measurements using absorbers or requiring change of experimental parameter between pre-and post selection. Here, without changing any experimental condition between pre and post-selection, we can perform joint measurements using the centers of transverse profiles along two directions as two independent pointers.

The weak value $\langle \Pi_A \rangle^{(w)} = 1$ suggests that $\langle \Pi_B \rangle^{(w)} = 0$. Since the weak values are equal to the eigenvalues of the dichotomic observables we may associate element of physical reality with them. Thus, we can claim that, in the limiting case of pre and post selection being satisfied no photons went through arm B. However, we obtained pointer shift resulting from the weak interaction of $\Pi_B \otimes \sigma_x$, i.e. the diagonal spin component in the arm B. Quantitatively, we obtain $\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = 1$ up to experimental uncertainties and errors. This forms the demonstration of partial cheshire cat, where some polarization component is found in a region where the particle is absent, albeit for a short duration in between pre and post-selection.

The quantum cheshire cat appears as a paradox when we presume that the quantum entity is a particle like. However, without diving deep into the ontology, the pointer shifts on average can be understood with classical electrodynamics or physical optics. Even with single photon source, one can invoke the optical equivalence theorem to understand the results as consequences of quantum mechanics. The paradox emerges as we pick one ontological premise over the other. For example, if we back the deBroglie-Bohm interpretation where the particle and the waves are separate entities, it would not be surprising at all to find polarization - oscillation of waves in a region of space separated from where the particle is located. Thus unambiguous realization of cheshire cat in the context of various interpretations of quantum mechanics may provide us directions in addressing question on foundations of quantum mechanics.

Expectation Value of Non-Hermitian Operators

The expectation value of the non-Hermitian operators cannot be measured from the statistics of outcomes of projective measurements. However, experimentally measuring the expectation value of a certain class of non-Hermitian operators for an unknown state can enable the reconstruction of the quantum state. We relate the average of the non-Hermitian operator to the weak value of a Hermitian operator into which it can be polar decomposed into. Using weak measurements we experimentally obtain the weak value and compare it with the calculated expectation value of the non-Hermitian operator. Using an interferometric technique, we could directly obtain the expectation value from the interference

pattern, which later evolved into a technique now known as quantum state interferography - a method to determine unknown quantum state. This technique could also be used to determine the weak value without performing weak measurements or perform conventional post-selection.

On the application of superposition principle

We have been applying the superposition principle ever since we were taught the Young's double slit experiment. Using a triple slit experiment and assuming the superposition principle to be applicable to experiments where wavefront gets divided, we could design an experiment to test the validity of the Born rule and as a corollary search for the existence of higher order interference terms. However, since different slits being open and closed refers to different boundary conditions for the partial differential equation, the superposition principle strictly does not hold true. But the deviation is usually small in optical domain and thus has not been highlighted much in the literature or textbooks. The cause for deviation can be attributed to the higher of Feynman paths and estimating the same quantitatively tells us that to magnify the deviation we need to perform the experiment in a regime where the slit dimensions are not much larger than the wavelength. This led to the experimental demonstration of the same in the microwave regime by measuring non-zero κ.

But to explain the deviation as such we do not need any new physics. Numerical simulation of Maxwell's equation for the geometry of the experiment explains the deviation within experimental uncertainties. Also, the path integral formalism with the inclusion of higher order paths provides a close estimation of the deviation. This experiment serves as a reminder to be careful while applying superposition principle in the foundational aspects of quantum mechanics. It also paves the way for what kind of experiments need to designed so that the validity of Born rule and higher order sum rules can be tested. Experiment's to test Born rule should have amplitude division only and avoid or suppress wavefront division so that the superposition principle would be applicable and any deviation from the sum rule can lead us to deeper theories.

Appendix

A Determining weak value using two-photon interference

So far we have considered the pointer state to be another degree of freedom (mostly the transverse spatial degree of freedom) of the particle which itself bears the system degree of freedom. The interaction Hamitonian coupled the two degrees of freedom of the same particle and there is often a classical analogue of the interaction. The weak value is inferred from the pointer shift of large number of identically prepared photons, for most observables, is the same as one would obtain from a coherent light source which is often deemed as classical. Thus, to claim the weak value is quantum in nature, in experiments one need to perform additional experiments to certify the quantumness of the state of light.

However, there exists a method to obtain weak value of polarization component using two-photon interference. Since the interaction itself is quantum in nature and do not have direct classical analog, it shows that the obtained weak values can be regarded as a quantum property. In this thesis, we do not perform any experiments that uses this method. However, other experiments have used this method to demonstrate Quantum Cheshire Cat . Therefore, the discussion here aims to provide an overview of the method followed by comparison with the method we use with spatial degree of freedom as the pointer.

A.1 The Measurement Interaction

For the measurement of an observable we need to entangle the system and the pointer so that the pointer shift gives information about the value of the observable. For weak interaction, we need tune some parameter or the strength of interaction so that the system

and pointer are almost separable in the limit of the strength of interaction going to zero. The problem is that two photons do not interact with each other in a linear media. Thus, in general, non-linearity in atoms can be used for effective interaction of two photons. However, two-photons do interfere in a linear media as evident from the HOM effect at the beam-splitter. . This two-photon interference can be used to effectively generate measurement induced non-linearity that entangles two-photons. Since, it is not a genuine interaction mediated by forces, we do not start with a well defined interaction Hamiltonian. Rather, we describe the unitary evolution operator directly which is responsible for the entanglement.

The system is the polarization degree of freedom of the photon whose polarization component is to be inferred subject to pre-and post-selection. The pointer is another photon's polarization which interferes (4th order quantum interference i.e., two-photon interference) with it and gives us a 'quantum non-demolition' measurement of photon's polarization. The interaction is effective as it is subject to detection of only 'one' signal and 'pointer' photon after the two photon interference. The unitary operator presented below is only considering the runs in which the non-demolition measurement protocol was successful.

The unitary operator that evolves the two photons when the observable is the projector $|\Pi_a = |a\rangle \langle a|$ is given by

$$U_M = (\mathbb{1} - \Pi_a) \otimes \mathbb{1} + \Pi_a \otimes R(g/2) \tag{A.1}$$

Here, $R(g/2)$ is the Jones matrix for the half-wave plate oriented at an angle θ from the Horizontal.

$$R(g/2) = \begin{pmatrix} \cos(g) & \sin(g) \\ \sin(g) & -\cos(g) \end{pmatrix} \tag{A.2}$$

In the limiting case of $g \to 0$, we have $R(g/2) = \sigma_z$ and with $g \to \pi/2$, $R(g/2) = \sigma_x$.

Strong Interaction Limit

Let's consider the limiting case when $R(g/2) = \sigma_x$. We have the measurement interaction operator as the CNOT gate.

$$U_M = \Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_x \tag{A.3}$$

Here, $\bar{a}$ is the polarization component orthogonal to a. Let the initial state of the system photon be given by

$$|\psi\rangle_s = \alpha |a\rangle + \beta |\bar{a}\rangle \tag{A.4}$$

Assuming that the pointer photon's polarization state is $|H\rangle$, the joint evolution of the system-pointer two-photon state is given as follows:

$$U_M |\psi\rangle_s |H\rangle = (\Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_x)(\alpha |a\rangle + \beta |\bar{a}\rangle) |H\rangle \tag{A.5}$$

$$= \beta |\bar{a}\rangle |H\rangle + \alpha |a\rangle |V\rangle \tag{A.6}$$

The resultant state after the interaction is entangled. If there is no post-selection, by measuring the pointer state to be $|V\rangle$ one can infer the state of the system to be $|a\rangle$ and if the pointer turns out to be $|H\rangle$, the system is inferred to be $|\bar{a}\rangle$. In either case the system reduces to $|a\rangle$ or $|\bar{a}\rangle$ from $|\psi\rangle_s$.

Zero Interaction Limit

Let's consider the case when $g \to 0$. Thus $R(g/2) \to \sigma_z$. The measurement interaction becomes controlled-Z (CZ) gate.

$$U_M = \Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_z \tag{A.7}$$

The joint evolution of the system-pointer state is given by

$$U_M |\psi\rangle_s |H\rangle = (\Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_z)(\alpha |a\rangle + \beta |\bar{a}\rangle) |H\rangle \tag{A.8}$$
$$= \beta |\bar{a}\rangle |H\rangle + \alpha |a\rangle |H\rangle \tag{A.9}$$
$$= (\alpha |a\rangle + \beta |\bar{a}\rangle) |H\rangle \tag{A.10}$$
$$= |\psi\rangle_s |H\rangle \tag{A.11}$$

The system and the pointer are separable. The pointer's state will always be found to be $|H\rangle$ and the system is undisturbed.

Note that the CZ gate do not disturb the system state only if the pointer is initialized in the state $|H\rangle$. If the initial state of the pointer were $|V\rangle$, we would have

$$U_M |\psi\rangle_s |V\rangle = (\Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_z)(\alpha |a\rangle + \beta |\bar{a}\rangle) |V\rangle \tag{A.12}$$
$$= \beta |\bar{a}\rangle |V\rangle - \alpha |a\rangle |V\rangle \tag{A.13}$$
$$= (-\alpha |a\rangle + \beta |\bar{a}\rangle) |V\rangle \tag{A.14}$$

The system state is changed from $|\psi_s\rangle$ to $(-\alpha |a\rangle + \beta |\bar{a}\rangle)$, even when the pointer state is intact. Thus, the disturbance on the system is not related to the shift of the pointer for arbitrary pointer states.

If the initial pointer state is $|+\rangle$, we have,

$$U_M |\psi\rangle_s |+\rangle = (\Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes \sigma_z)(\alpha |a\rangle + \beta |\bar{a}\rangle) |+\rangle \tag{A.15}$$
$$= \beta |\bar{a}\rangle |+\rangle - \alpha |a\rangle |-\rangle \tag{A.16}$$

This is an entangled state and knowing the pointer reduces the system state. In this case both the system is disturbed as well as the pointer shows the change as well. This is in contrast to the combination of Gaussian pointer states and birefringent crystal/ Stern Gerlach as interaction operations, where the strength of interaction do not depend on the initial pointer (centre of incident Gaussian). The relative changes in the pointer state (separation of the Gaussian) is independent of initial pointer state (centre of the gaussian). In this sense, U_M can be only *effectively* weak for the given initial pointer state $|H\rangle$ when we approach the limit $g \to 0$.

Weak Interaction limit

Here, we wish to consider only first order terms in the expansion of $R(g/2)$ matrix. Thus the appropriate limit is $g << 1$ or $g << \pi/2$. In this limit we have the rotation matrix as

$$R(g/2) = \begin{pmatrix} \cos(g) & \sin(g) \\ \sin(g) & -\cos(g) \end{pmatrix} \xrightarrow{g<<1} \begin{pmatrix} 1 & g \\ g & -1 \end{pmatrix} \tag{A.17}$$

$$= \sigma_z + g\sigma_x \tag{A.18}$$

Note that all terms involving g^2 and above have been ignored. The interaction operator now becomes,

$$U_M = \Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes (\sigma_z + g\sigma_x) \tag{A.19}$$

Next, we study the joint evolution of the two-photon state.

$$U_M |\psi\rangle_s |H\rangle = (\Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes (\sigma_z + g\sigma_x))(\alpha |a\rangle + \beta |\bar{a}\rangle) |H\rangle \tag{A.20}$$

$$= \beta |\bar{a}\rangle |H\rangle + \alpha |a\rangle (|H\rangle + g |V\rangle) \tag{A.21}$$

$$= |\psi\rangle_s |H\rangle + \alpha g |a\rangle |V\rangle \tag{A.22}$$

The system remain unaffected only in the limit $g \to 0$.

Note that the pointer state at the end of the interaction $|H\rangle + g|V\rangle$ is not normalized. This is because $\sigma_z + g\sigma_x$ is not unitary to begin with for non-zero g. In the limit, $g \to 0$, the pointer state becomes normalized and the interaction becomes unitary. This is similar to the case with the Gaussian pointer as well after we truncate the series expansion. If we truncate the evolution operator to the first order, it is not unitary for a non-zero interaction strength. Therefore, this should always be considered as a limiting case of zero interaction strength where the evolution is unitary. The state after the evolution, therefore would need normalization while computing the expectation value of observables.

A.2 Post-selection and Weak Value

Let the post selected state of the system be $|\phi\rangle_s$. We wish the now infer change in the state of the pointer in the weak interaction limit. Since, $U_M = \Pi_{\bar{a}} \otimes \mathbb{1} + \Pi_a \otimes (\sigma_z + g\sigma_c)$ the pointer shifts occurs only due to the projector Π_a. The pointer is unaffected by the system's polarization component along $\bar{a}$. The final pointer state after the post-selection is given by

$$\langle\phi|_s U_m |\psi\rangle_s |H\rangle = \langle\phi|_s \Pi_{\bar{a}} |\psi\rangle_s \mathbb{1} |H\rangle + \langle\phi|_s \Pi_a |\psi\rangle_s (\sigma_z + g\sigma_x) |H\rangle \tag{A.23}$$

$$= \langle\phi|_s (\mathbb{1} - \Pi_a) |\psi\rangle_s |H\rangle + \langle\phi|_s \Pi_a |\psi\rangle_s (|H\rangle + g|V\rangle) \tag{A.24}$$

$$= \langle\phi|\psi\rangle_s |H\rangle + \langle\phi|_s \Pi_a |\psi\rangle_s g|V\rangle \tag{A.25}$$

$$= \langle\phi|\psi\rangle_s (|H\rangle + g\frac{\langle\phi|_s \Pi_a |\psi\rangle_s}{\langle\phi|\psi\rangle_s}|V\rangle) \tag{A.26}$$

In the above expression, $\langle\phi|\psi\rangle_s$ denotes the probability amplitude of successful post-selection. Thus the change in pointer state is proportional to $g\frac{\langle\phi|_s \Pi_a |\psi\rangle_s}{\langle\phi|\psi\rangle_s}$ or $g\langle\Pi_a\rangle^{(w)}$. In this sense, one can claim that the pointer shift, in this case, the polarization shift of the pointer photon, is proportional to the interaction strength g and the weak value of Π_a. Since the change in pointer polarization is linear with respect to g, taking a derivative of the change in pointer state with respect to g gives us the weak value.

If we do not assume $g << 1$ to begin with, we have the final pointer state as

$$\langle\phi|\psi\rangle_s \left(1 + (R(g/2) - 1)\frac{\langle\phi|_s \Pi_a |\psi\rangle_s}{\langle\phi|\psi\rangle_s}\right)|H\rangle \tag{A.27}$$

$$\langle\phi|\psi\rangle_s \left(|H\rangle + ((\cos(g) - 1)|H\rangle + \sin(g)|V\rangle)\frac{\langle\phi|_s \Pi_a |\psi\rangle_s}{\langle\phi|\psi\rangle_s}\right) \tag{A.28}$$

Note that the amplitude of the change in pointer state is proportional to the factor in front of $|V\rangle$. By taking the differential wrt and taking the limit $g \to 0$, we shall arrive at the weak value of Π_a.

Quantifying Polarization pointer

So far we have talked about 'measuring' the polarization change of the pointer without explicitly quantifying 'which observable' to use as a quantifier. To begin with let's compute

the expectation value of the pauli operators in the final pointer state. The normalized state is given by the vector

$$|\phi\rangle_p^f = \frac{1}{\sqrt{1+g^2|\langle\Pi_a\rangle^{(w)}|^2}}(|H\rangle + g\langle\Pi_a\rangle^{(w)}|V\rangle) \tag{A.29}$$

The expectation value of the Pauli operators in the state $|\phi\rangle_p^f$ are as follows.

$$\langle\sigma_x\rangle = \frac{1}{(1+g^2|\langle\Pi_a\rangle^{(w)}|^2)}(\langle H| + g\langle\Pi_a\rangle^{(w)*}\langle V|)\sigma_x(|H\rangle + g\langle\Pi_a\rangle^{(w)}|V\rangle) \tag{A.30}$$

$$= \frac{1}{(1+g^2|\langle\Pi_a\rangle^{(w)}|^2)}(\langle H| + g\langle\Pi_a\rangle^{(w)*}\langle V|)(|V\rangle + g\langle\Pi_a\rangle^{(w)}|H\rangle) \tag{A.31}$$

$$= \frac{1}{(1+g^2|\langle\Pi_a\rangle^{(w)}|^2)}2g\,\mathrm{Re}\Big\{\langle\Pi_a\rangle^{(w)}\Big\} \tag{A.32}$$

One can ignore g^2 in the normalization term, given $g << 1$. This is because doing a series expansion of $1/(1+g^2)$, we obtain $1-g^2$ as an approximation. If we choose to keep only first order terms we can effectively approximate the normalization factor to 1. But notice that g^2 is multiplied with the magnitude of the weak value. Thus this approximation, would not work in the amplification regime. However, since many experiments including the cheshire cat experiment we are concerned with, the desired value of observable is within the range of eigenvalues, and in such cases using the approximation is consistent.

In Eqn. A.30 to Eqn. A.32 in the derivation above $*$ is used to denote the complex conjugate of the weak value. Note that $\langle\sigma_x\rangle$ for the pointer polarization state would be a real number. Thus, it gives us information about the real part of the weak value of the projector. Similarly, we have the expectation value of all the Pauli operators as

$$\langle\sigma_x\rangle = 2g\,\mathrm{Re}\Big\{\langle\Pi_a\rangle^{(w)}\Big\} \qquad \langle\sigma_y\rangle = 2g\,\mathrm{Im}\Big\{\langle\Pi_a\rangle^{(w)}\Big\} \qquad \langle\sigma_z\rangle = 1 \tag{A.33}$$

Thus, the real and imaginary part of the weak value of Π_a can be obtained from the slope of the expectation value of pauli operators on the pointer state as a function of g.

A.3 Set up

The measurement interaction U_M can be experimentally implemented if we have a controlled phase gate to begin with. We take the particular case of CZ gate and show that U_M can be achieved by placing half wave plates in the input and output ports of the CZ gate. Here, we would describe the experimental setup needed to perform the weak measurement of the desired observable and would assume the availability of the CZ gate as black-box. The practical realization of the CZ gate shall be discussed later in the subsection A.5. The CZ gate is given by

$$C_z = \Pi_V \otimes \mathbb{1} + \Pi_H \otimes \sigma_z \tag{A.34}$$

Now, instead of the control polarization being along horizontal or vertical we want it along any arbitrary polarization direction a. Thus, we operate HWP in the input port of the C_z gate at angle $(\theta_a/2)$ so that the polarization gets rotated by θ_a. The C_z gate splits the polarization in H V basis. At the designated system exit port of the C_z gate, we again apply $R(\theta_a)$ to change the polarization back along a. Thus we have,

$$(R(\theta_a)\otimes\mathbb{1})C_z(R(\theta_a)\otimes\mathbb{1}) = (R(\theta_a)\otimes\mathbb{1})(\Pi_V\otimes\mathbb{1}+\Pi_H\otimes\sigma_z)(R(\theta_a)\otimes\mathbb{1}) \tag{A.35}$$

$$= R(\theta_a)\Pi_V(R(\theta_a)\otimes\mathbb{1} + (R(\theta_a)\Pi_H(R(\theta_a)\otimes\sigma_z \tag{A.36}$$

Now, $R(\theta_a)\Pi_H R(\theta_a) = R(\theta_a)\left|H\right\rangle\left\langle H\right|R(\theta_a) = \left|a\right\rangle\left\langle a\right| = \Pi_a$. Thus, we have

$$(R(\theta_a)\otimes\mathbb{1})C_z(R(\theta_a)\otimes\mathbb{1}) = (\mathbb{1}-\Pi_a)\otimes\mathbb{1}+\Pi_a\otimes\sigma_z \tag{A.37}$$

As we have already seen this operator is the zero interaction limit as it does not change the pointer as long as the initial pointer state $\left|\psi\right\rangle_p = \left|H\right\rangle$. To have effective two-photon interaction (due to two photon interference) we can change the initial polarization of the pointer. Equivalently, we can rotate the basis by using two half wave plates at the input and output port for the C_z gate designated for the pointer. The half wave plate when oriented

at θ_g, the polarization gets rotated by angle $2\theta_g$ (Note that g and θ_g). This is because $R(\theta_g)\sigma_z R(\theta_g) = R(2\theta_g) = R(g/2)$. Hence, the interaction strength $g = 4\theta_g$. Thus, we have the net interaction operator as

$$U_M = (R(\theta_a) \otimes R(\theta_g))C_z(R(\theta_a) \otimes R(\theta_g)) = (\mathbb{1} - \Pi_a) \otimes \mathbb{1} + \Pi_a \otimes R(g/2) \tag{A.38}$$

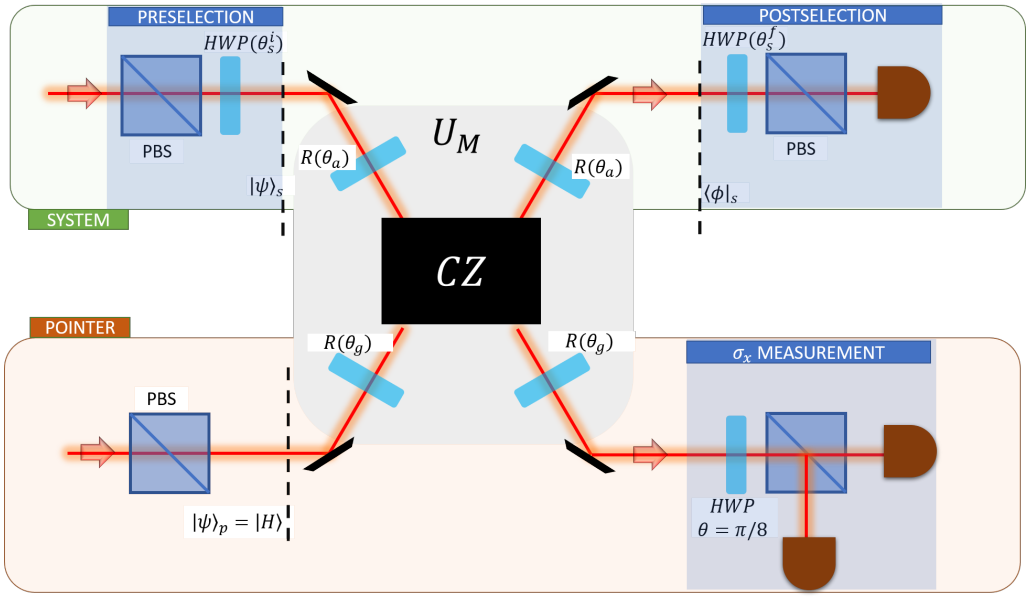


Figure A.1: Schematic of Polarization weak measurement using controlled phase gate

The system's polarization state can be prepared using a HWP at an angle θ_s^i after the PBS to have the pre-selected state $|\psi\rangle_s$. Simlarly, the system can be post-selected with the HWP at angle θ_s^f placed before the PBS of which only transmitted component is detected.

$$|\psi\rangle_s = \cos(2\theta_s^i)\,|H\rangle + \sin(2\theta_s^i)\,|V\rangle \tag{A.39}$$

$$|\phi\rangle_s = \cos(2\theta_s^i)\,|H\rangle + \sin(2\theta_s^i)\,|V\rangle \tag{A.40}$$

The pointer state is always prepared in the state $|H\rangle$. The expectation value σ_x in the final pointer state can be obtained by projecting it into the diagonal, off-diagonal basis. This can be done by placing a HWP at angle $\pi/8$ before the PBS and measuring the transmitted and reflected counts. The expectation value is simply $P_T - P_R$, where $P_{T/R}$ is the normalized

counts detected in the transmitted/reflected port.

A.4 Example

In this section, we shall compute the weak value of the operator σ_x as function of pre-selection angle θ_s^i and interaction strength $g = 4\theta_g$. We shall fix the post-selection to be $|H\rangle$ i.e. $\theta_s^f = 0$.

To infer weak value of σ_x for the system we need two runs of the experiment. In one, we obtain the $\langle\sigma_x\rangle_p^+$ for the projector $\Pi_+ = |+\rangle\langle+|$ by setting $\theta_a = \pi/8$ in the expression for the measurement interaction operator U_M. Next, we repeat the same, but with $\theta_a = \pi/8$, so that the projector becomes $\Pi_+ = |+\rangle\langle+|$ and we obtain $\langle\sigma_x\rangle_p^+$.

We have

$$\langle\sigma_x\rangle_p^+ = 2g\,\mathrm{Re}\{\langle\Pi_+\rangle_s^w\}\,\langle\sigma_x\rangle_p^- = 2g\,\mathrm{Re}\{\langle\Pi_-\rangle_s^w\} \tag{A.41}$$

Since , $\sigma_x = \Pi_+ - \Pi_-$, we have

$$\langle\sigma_x\rangle_p^+ - \langle\sigma_x\rangle_p^- = 2g\,\mathrm{Re}\Big\{\langle\sigma_x\rangle_s^{(w)}\Big\} \Rightarrow \frac{\partial(\langle\sigma_x\rangle_p^+ - \langle\sigma_x\rangle_p^-)}{2\partial g}\bigg|_{g=0} = \mathrm{Re}\Big\{\langle\sigma_x\rangle_s^{(w)}\Big\} \tag{A.42}$$

Thus, by taking slope of the different of the expectation value of σ_x in the final pointer state wrt. and evaluating it in the limit $g \to 0$ gives us the weak value of σ_x for the system.

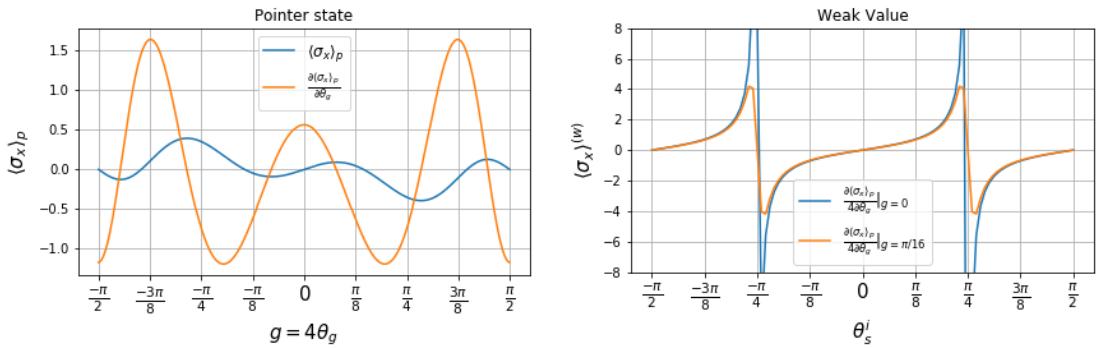


Figure A.2: [Left] Pointer variation as a function of g, [Right] Weak value derived from the pointer

Note that, if we evaluate the derivative at a non-zero g, we get reduced amplification.

The slopes inflexes beyond $g \geq \pi/8$ and hence for the interaction to be weak, we technically need $g << \pi/8$.

A.5 Realization of C_Z gate

In the previous section, we had discussed how we can use a C_Z gate to perform weak measurement of the polarization state of the photon using polarization state of another photon as pointer. However, the realization of C_Z gate was left as a blackbox. In this section, we discuss the realization of C_Z.

The partially polarizating beam splitter

The building block of the C_Z gate will be the partially polarization beam splitter (PPBS) also known as the polarization dependent beam splitter (PDBS).

A PPBS has different transmission probability for two orthogonal polarization. The standard PBS is the one where the transmission probability for incident $|H\rangle$ is 1 i.e., $T_H = 1$ and the transmission probability for the $|V\rangle$ is 0.Thus $T_V = 0$.

For a partially polarizing beam splitter one of the $0 < T_{H/R} < 1$. The PPBS of our interest is the case where $T_H = 1$ and $T_V = 1/3$. The operation is lossless and hence the reflectance for horizontally light is 0 $R_H = 0$ and that for vertically polarized light is $2/3$.

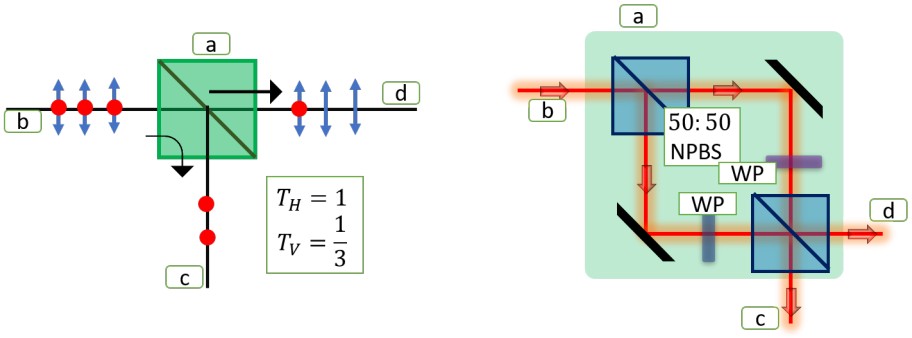


Figure A.3: [Left] Working of a PPBS [Right] Construction of a PPBS

The PPBS can be constructed from two 50:50 non-polarizing beam splitters in a Mach-Zehnder configuration with two waveplates, one in each arm. One of the waveplate (could be in principle, combination of waveplates) give 0 (relative) phase shift to $|H\rangle$ and some phase shift ϕ_V to $|V\rangle$. The other wave plate in the other arm creates a phase shift ϕ_H for

$|H\rangle$ and none for $|V\rangle$. In this way, we have independent control over how much of $|H\rangle$ component goes out at port d. For example, if we want effective T_H of the PPBS to be 1, we need $\phi_H = 0$ so that constructive interference occurs at port d for horizontally polarized light. To achieve $T_V = 1/3$, we would need $\phi_V = \arccos(-2/3)$. Next, without worrying about teh detailed construction of PPBS, we shall use the transmission and reflection amplitudes to study two-photon interference.

PPBS Transfer Matrix

Let's denote the transmission/reflection amplitudes for the input port x and output port y for the polarization component Z be given by $t/r^Z_{x\to y}$. Since, $T_H = 1$ and $T_V = 1/3$, we have the following amplitudes:

$$t^H_{a\to c} = 1 \qquad r^H_{a\to d} = 0 \qquad t^H_{b\to d} = 1 \qquad r^H_{b\to c} = 0 \tag{A.43}$$

$$t^V_{a\to c} = \sqrt{1/3} \qquad r^V_{a\to d} = -\sqrt{2/3} \qquad t^V_{b\to d} = \sqrt{1/3} \qquad r^V_{b\to c} = \sqrt{2/3} \tag{A.44}$$

One of the reflections $r^V_{a\to d} = -\sqrt{2/3}$ has a negative sign to imply the π phase shift in reflection. This helps in making the transfer matrix for the PPBS a unitary operator. The transformation of the creation operators is given as

$$\hat{a}^\dagger_H \to t^H_{a\to c}\hat{c}^\dagger_H + r^H_{a\to d}\hat{d}^\dagger_H = \hat{c}^\dagger_H \tag{A.45}$$

$$\hat{b}^\dagger_H \to r^H_{b\to c}\hat{c}^\dagger_H + t^H_{b\to d}\hat{d}^\dagger_H = \hat{d}^\dagger_H \tag{A.46}$$

$$\hat{a}^\dagger_V \to t^V_{a\to c}\hat{c}^\dagger_V + r^V_{a\to d}\hat{d}^\dagger_V = \sqrt{1/3}\,\hat{c}^\dagger_V - \sqrt{2/3}\,\hat{d}^\dagger_V \tag{A.47}$$

$$\hat{b}^\dagger_V \to r^V_{b\to c}\hat{c}^\dagger_V + t^V_{b\to d}\hat{d}^\dagger_V = \sqrt{2/3}\,\hat{c}^\dagger_V + \sqrt{1/3}\,\hat{d}^\dagger_V \tag{A.48}$$

Thus, the transfer matrix for the this PPBS can be written as:

$$\begin{pmatrix} \hat{a}^\dagger_H \\ \hat{b}^\dagger_H \\ \hat{a}^\dagger_V \\ \hat{b}^\dagger_V \end{pmatrix} \xrightarrow[T_H=1, T_V=1/3]{PPBS} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{1/3} & -\sqrt{2/3} \\ 0 & 0 & \sqrt{2/3} & \sqrt{1/3} \end{pmatrix} \begin{pmatrix} \hat{c}^\dagger_H \\ \hat{d}^\dagger_H \\ \hat{c}^\dagger_V \\ \hat{d}^\dagger_V \end{pmatrix} \tag{A.49}$$

The above transfer matrix is unitary and for that it was sufficient to have one of reflection

to have a π phase shift. Although, this may already seem like a controlled rotation gate, it is for the spatial port and polarization degree of freedom and not for the polarization degree of freedom for two photons.

Two-Photon Interference

We have four components possible cases for two photons incident on the PPBS from separate input ports. We write the transformations of the two-photon creation operators below:

$$\hat{a}_H^\dagger \hat{b}_H^\dagger \rightarrow \hat{c}_H^\dagger \hat{d}_H^\dagger \tag{A.50}$$

$$\hat{a}_H^\dagger \hat{b}_V^\dagger \rightarrow \hat{c}_H^\dagger (\sqrt{2/3}\,\hat{c}_V^\dagger + \sqrt{1/3}\,\hat{d}_V^\dagger) = \sqrt{2/3}\,\hat{c}_H^\dagger \hat{c}_V^\dagger + \sqrt{1/3}\,\hat{c}_H^\dagger \hat{d}_V^\dagger \tag{A.51}$$

$$\hat{a}_V^\dagger \hat{b}_H^\dagger \rightarrow (\sqrt{1/3}\,\hat{c}_V^\dagger - \sqrt{2/3}\,\hat{d}_V^\dagger)\hat{d}_H^\dagger = \sqrt{1/3}\,\hat{c}_V^\dagger \hat{d}_H^\dagger - \sqrt{2/3}\,\hat{d}_V^\dagger \hat{d}_H^\dagger \tag{A.52}$$

$$\hat{a}_V^\dagger \hat{b}_V^\dagger \rightarrow (\sqrt{1/3}\,\hat{c}_V^\dagger - \sqrt{2/3}\,\hat{d}_V^\dagger)(\sqrt{2/3}\,\hat{c}_V^\dagger + \sqrt{1/3}\,\hat{d}_V^\dagger) \tag{A.53}$$

$$= \frac{2}{3}\hat{c}_V^\dagger \hat{c}_V^\dagger + \frac{1}{3}\hat{c}_V^\dagger \hat{d}_V^\dagger - \frac{2}{3}\hat{d}_V^\dagger \hat{c}_V^\dagger + \frac{1}{3}\hat{d}_V^\dagger \hat{d}_V^\dagger \tag{A.54}$$

We now assume that the two photons are indistinguishable. Hence $\hat{c}_V^\dagger \hat{d}_V^\dagger = \hat{d}_V^\dagger \hat{c}_V^\dagger$ from the bosonic commutator being 0. Then, we have,

$$\hat{a}_V^\dagger \hat{b}_V^\dagger \rightarrow \frac{2}{3}\hat{c}_V^\dagger \hat{c}_V^\dagger - \frac{1}{3}\hat{c}_V^\dagger \hat{d}_V^\dagger \frac{1}{3}\hat{d}_V^\dagger \hat{d}_V^\dagger \tag{A.55}$$

Now, we consider the general polarization state of two qubits is incident on the two input ports of the PPBS.

$$|\psi\rangle_i = \gamma_1 |HH\rangle + \gamma_2 |HV\rangle + \gamma_3 |VH\rangle + \gamma_4 |VV\rangle \tag{A.56}$$

$$= (\gamma_1 \hat{a}_H^\dagger \hat{b}_H^\dagger + \gamma_2 \hat{a}_H^\dagger \hat{b}_V^\dagger + \gamma_3 \hat{a}_V^\dagger \hat{b}_V^\dagger + \gamma_4 \hat{a}_V^\dagger \hat{b}_V^\dagger)\,|00\rangle \tag{A.57}$$

After the beam splitter, the two photon state transfers to the following state

$$|\psi\rangle_f = (\gamma_1 \hat{c}_H^\dagger \hat{d}_H^\dagger + \gamma_2(\sqrt{2/3}\,\hat{c}_H^\dagger \hat{c}_V^\dagger + \sqrt{1/3}\,\hat{c}_H^\dagger \hat{d}_V^\dagger) + \gamma_3(\sqrt{1/3}\,\hat{c}_V^\dagger \hat{d}_H^\dagger - \sqrt{2/3}\,\hat{d}_H^\dagger \hat{d}_V^\dagger) + \gamma_4(\frac{2}{3}\hat{c}_V^\dagger \hat{c}_V^\dagger - \frac{1}{3}\hat{c}_V^\dagger \hat{d}_V^\dagger \frac{1}{3}\hat{d}_V^\dagger \hat{d}_V^\dagger))\,|00\rangle \tag{A.58}$$

Note that the terms $\hat{c}_H^\dagger \hat{c}_V^\dagger$, $\hat{d}_H^\dagger \hat{d}_V^\dagger$, $\hat{c}_V^\dagger \hat{c}_V^\dagger$ and $\hat{d}_V^\dagger \hat{d}_V^\dagger$ correspond to the two photons exiting in the same port.

Selection of coincident events

Our aim is to create a C_z gate. After the gate operation, control qubit and the system qubit must exit at different ports. Thus, if we restrict ourselves to consider only the events in which two photons exit in different ports, we can take a step towards constructing the C_Z gate. We can ensure that the two photons, exited at different ports we if consider only the events when we get a coincidence between the two output ports after further operations). Practically, we also need to be sure that only two-photons are incident on the PPBS and no more.

If we restrict ourselves to selection of coincident events only, i.e., consider the cases when the two photons exit at different ports, we have the following reduced two photon state after the PPBS.

$$|\psi\rangle_f \xrightarrow{coincidence} |\bar{\psi}\rangle_f = (\gamma_1 \hat{c}_H^\dagger \hat{d}_H^\dagger + \gamma_2 \sqrt{1/3}\, \hat{c}_H^\dagger \hat{d}_V^\dagger + \gamma_3 \sqrt{1/3}\, \hat{c}_V^\dagger \hat{d}_H^\dagger - \gamma_4 \frac{1}{3} \hat{c}_V^\dagger \hat{d}_V^\dagger)\, |00\rangle \tag{A.59}$$

$$= \gamma_1 |HH\rangle + \gamma_2 \sqrt{1/3}\ |HV\rangle + \gamma_3 \sqrt{1/3}\ |VH\rangle - \gamma_4 \frac{1}{3} |VV\rangle \tag{A.60}$$

For a ideal realization of C_Z gate, we should have $\gamma_{1,2,3} \xrightarrow{C_Z} \gamma_{1,2,3}$ and $\gamma_4 \xrightarrow{C_Z} -\gamma_4$. By selecting coincident events, we have been able to achieve the $-$ve sign in front of γ_4. However, it is scaled by a facctor $1/3$. To compensate, we must filter $|HH\rangle$ state by a factor of $1/3$ as well. The components HV and VH can needs to be filtered as well by a factor $\sqrt{1/3}$, Note that, this can be achieved, if at each of the output port of the PPBS we introduce another PPBS, but this time with $T_H = 1/3$ and $T_V = 1$ so that we can only consider the transmitted part.

Amplitude compensation

Alternatively, if we want to use the same PPBS, with $T_H = 1$ and $T_V = 1/3$, we can use two HWP's at 45°, one before and the other after the PPBS. The HWP acts as σ_x operation so

that $|H\rangle \leftrightarrow |V\rangle$. Then the PPBS filters the amplitude of $|H\rangle$. Finally, the σ_x again reverts $|H\rangle \leftrightarrow |V\rangle$ so that we can have the C_Z gate.

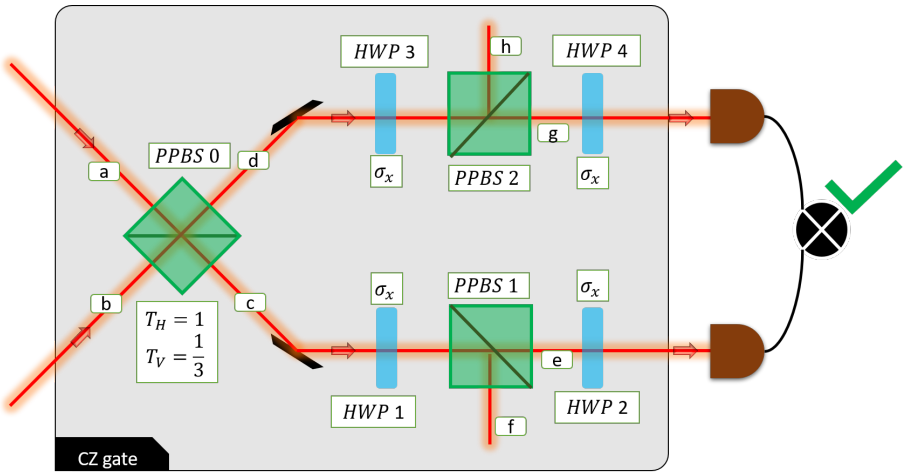


Figure A.4: Realization of a CZ using PPBS

The state after PPBS0, subject to future coincidence measurement, is $|\bar{\psi}_f\rangle$. The HWP1 and HWP3 forms the operator $\sigma_x \otimes \sigma_x$.

$$|\psi\rangle_{H1-H3} = (\sigma_x \otimes \sigma_x)\,|\bar{\psi}\rangle_f \tag{A.61}$$

$$= (\sigma_x \otimes \sigma_x)(\gamma_1 |HH\rangle + \gamma_2\sqrt{1/3}\ |HV\rangle + \gamma_3\sqrt{1/3}\ |VH\rangle - \gamma_4\frac{1}{3}|VV\rangle) \tag{A.62}$$

$$= \gamma_1 |VV\rangle + \gamma_2\sqrt{1/3}\ |VH\rangle + \gamma_3\sqrt{1/3}\ |HV\rangle - \gamma_4\frac{1}{3}|HH\rangle \tag{A.63}$$

We shall only consider the transmission and ignore all the photons that are reflected. We have $t^H_{c/d\to e/f} = 1$ and $t^V_{c/d\to e/f} = \sqrt{1/3}$. Thus, the state after PPBS1 and PPBS2 in the basis $|e\rangle, |g\rangle$ is given as follows:

$$|\psi\rangle_{P1-P2} = \gamma_1\frac{1}{3}|VV\rangle + \gamma_2\sqrt{1/3}\sqrt{1/3}\ |VH\rangle + \gamma_3\sqrt{1/3}\sqrt{1/3}\ |HV\rangle - \gamma_4\frac{1}{3}|HH\rangle \tag{A.64}$$

$$= \frac{1}{3}(\gamma_1 |VV\rangle + \gamma_2 |VH\rangle + \gamma_3 |HV\rangle - \gamma_4 |HH\rangle) \tag{A.65}$$

Now, we can renormalize the polarization state of the two qubits to discard the pre factor $1/3$.

$$|\bar{\psi}\rangle_{P1-P2} = \gamma_1 |VV\rangle + \gamma_2 |VH\rangle + \gamma_3 |HV\rangle - \gamma_4 |HH\rangle \tag{A.66}$$

Now, all the amplitudes of the polarization components are same and γ_4 has additionally acquired a $-$ve sign i.e. its associated polarization component gets a phase shift of π. The only thing remaining is to flip the basis once again by the action of HWP2 and HWP4. We have the final state as

$$(\sigma_x \otimes \sigma_x) |\bar{\psi}\rangle_{P1-P2} = \gamma_1 |HH\rangle + \gamma_2 |HV\rangle + \gamma_3 |VH\rangle - \gamma_4 |VV\rangle \tag{A.67}$$

Thus, effectively we have the C_Z transformation.

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ -\gamma_4 \end{pmatrix} \xleftarrow{C_Z} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} \tag{A.68}$$

This C_Z gate realization is the blackbox that we used while describing weak measurement in Fig. A.1.

A.6 Comparison with using spatial pointer

There are some advantages and disadvantages of using the transverse spatial degree of freedom of the photon as a pointer versus using the polarization degree of freedom in the two-photon interaction scheme.

- The two-photon interference do not have direct classical analogue and hence this method ubiquitously establishes the weak value obtained to have quantum origin. When the source is changed to a coherent source which is deemed classical, we would have multi-photon events at the two detectors. In principle, one would be able to detect such events either by using cascading or using PNR detector and post-select the events which are deemed to be quantum. On the other hand, the interaction used to entangle the spatial degree of freedom with the polarization of the same photon works irrespective of whether the state of light is classical or quantum. The

interaction Hamiltonian is agnostic to the photon statistics. For the two -photon interaction we skip the Hamiltonian[13] which is not very well defined and directly start with the effective evolution operator.

- The spatial pointer lives in an infinite dimensional Hilbert space where as the polarization degree of freedom is 2-dimensional. However, the pointer variable itself be it mean of the transverse profile or the expectation value of the σ_x observable is continuous. In the two-photon interference method, the strength of the interaction is dependent on the initial polarization state of the pointer photon, while the shift in the spatial pointer only depends on the interaction time and parameters in the Hamiltonian.
- To obtain the weak value, one needs to compute the expectation value of the Pauli observable of the polarization state of the photon as a function of the interaction strength and then take the derivative. Note that the same can be done with the spatial pointer as well But in practice this is not needed as the weak value can be directly read off as the shift in mean position of the transverse profile. Thus, nothing needs to be changed within the pre- and post-selection to obtain the weak value.
- The transverse spatial profile is 2-dimensional and hence we can obtain the complex weak value of an observable in an experiment without any changes to the setup. Also, if weak values are known to be real, we can perform joint measurements. With the two-photon interference method to perform joint observations of real weak values of two observables we need another photon which is correlated with the other two. The detection similarly has to post-select all the events in which triple coincidence is observed. Producing two correlated photon is relatively easy with SPDC process. However, generating multi photon time correlated states are not that efficient. Further, post-selection of triple coincidence would dramatically reduce the overall efficiency of the experiment. This is the primary reason why we used transverse spatial profile of the photon as the pointer in the experimental implementation of the Quantum Cheshire Cat state instead.

[13] If $U = \exp\left(-\frac{i}{\hbar}Ht\right)$, one may attempt to write the Hamiltonian as $i\hbar\frac{1}{t}\log(U)$. But it neither gives any further insight nor can be established independently from the first principles.

B Probability from ABL rule and weak values for observables in cheshire cat effect

B.1 'which-path' degree of freedom

'path A projector' observable $\Pi_A \otimes \mathbb{1}$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \lvert A\rangle\lvert H\rangle$	$P\left(\Pi_A \otimes \mathbb{1} = 1; \lvert A\rangle\lvert H\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{1}$
1	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \lvert A\rangle\lvert V\rangle$	$P\left(\Pi_A \otimes \mathbb{1} = 1; \lvert A\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
0	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \lvert B\rangle\lvert H\rangle$	$P\left(\Pi_A \otimes \mathbb{1} = 0; \lvert B\rangle\lvert H\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
0	$\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \lvert A\rangle\lvert H\rangle$	$P\left(\Pi_A \otimes \mathbb{1} = 0; \lvert B\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle \Pi_A \otimes \mathbb{1} \rangle^{(w)} = \dfrac{\langle\phi\rvert (\Pi_A \otimes \mathbb{1}) \lvert\psi\rangle}{\langle\phi\rvert\psi\rangle} = 1$	

Table A.1: Projector for arm A

'path B projector' observable $\Pi_B \otimes \mathbb{1}$		
Eigenvalue	**Eigenvector**	**ABL probability**
0	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \lvert A\rangle\lvert H\rangle$	$P\left(\Pi_B \otimes \mathbb{1} = 0; \lvert A\rangle\lvert H\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{1}$
0	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \lvert A\rangle\lvert V\rangle$	$P\left(\Pi_B \otimes \mathbb{1} = 0; \lvert A\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
1	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \lvert B\rangle\lvert H\rangle$	$P\left(\Pi_B \otimes \mathbb{1} = 1; \lvert B\rangle\lvert H\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
1	$\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \lvert A\rangle\lvert H\rangle$	$P\left(\Pi_B \otimes \mathbb{1} = 1; \lvert B\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert \,\middle\vert\, \lvert\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle \Pi_B \otimes \mathbb{1} \rangle^{(w)} = \dfrac{\langle\phi\rvert (\Pi_B \otimes \mathbb{1}) \lvert\psi\rangle}{\langle\phi\rvert\psi\rangle} = 0$	

Table A.2: Projector for arm B

'which-path' observable $\Sigma_z \otimes \mathbb{1}$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Sigma_A \otimes \mathbb{1} = 1; \|A\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{1}$
1	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \|A\rangle\|V\rangle$	$P\left(\Sigma_A \otimes \mathbb{1} = 1; \|A\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
-1	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \|B\rangle\|H\rangle$	$P\left(\Sigma_A \otimes \mathbb{1} = -1; \|B\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
-1	$\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Sigma_A \otimes \mathbb{1} = -1; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle\Sigma_z \otimes \mathbb{1}\rangle^{(w)} = \frac{\langle\phi\|(\Sigma_z \otimes \mathbb{1})\|\psi\rangle}{\langle\phi\|\psi\rangle} = 1$	

Table A.3: 'which-path' observable

B.2 Polarization in path A

σ_x polarization component in arm A observable $\Pi_A \otimes \sigma_x$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix} = \|A\rangle\|+\rangle$	$P\left(\Pi_A \otimes \sigma_x = 1; \|A\rangle\|+\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = 0.5$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix} = \|A\rangle\|-\rangle$	$P\left(\Pi_A \otimes \sigma_x = -1; \|A\rangle\|-\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = 0.5$
0	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \|B\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_x = 0; \|B\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
0	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \|B\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_x = 0; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle\Pi_A \otimes \sigma_x\rangle^{(w)} = \frac{\langle\phi\|(\Pi_A \otimes \sigma_x)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 0$	

Table A.4: σ_x polarization component in arm A

σ_y **polarization component in arm** A **observable** $\Pi_A \otimes \sigma_y$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\\ 0\\ 0\end{pmatrix} = \|A\rangle\|\circlearrowleft\rangle$	$P\left(\Pi_A \otimes \sigma_y = 1; \|A\rangle\|\circlearrowleft\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \frac{1}{2}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\\ 0\\ 0\end{pmatrix} = \|A\rangle\|\circlearrowright\rangle$	$P\left(\Pi_A \otimes \sigma_x = -1; \|A\rangle\|\circlearrowright\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \frac{1}{2}$
0	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 0\\ 1\\ 0\end{pmatrix} = \|B\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_y = 0; \|B\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{0}$
0	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 0\\ 0\\ 1\end{pmatrix} = \|B\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_y = 0; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle \Pi_A \otimes \sigma_y \rangle^{(w)} = \frac{\langle\phi\|(\Pi_A \otimes \sigma_y)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 0$	

Table A.5: σ_y polarization component in arm A

σ_z **polarization component in arm** A **observable** $\Pi_A \otimes \sigma_z$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\begin{pmatrix}1\\ 0\\ 0\\ 0\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_z = 1; \|A\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{1}$
-1	$\begin{pmatrix}0\\ 1\\ 0\\ 0\end{pmatrix} = \|A\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_z = -1; \|A\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{0}$
0	$\begin{pmatrix}0\\ 0\\ 1\\ 0\end{pmatrix} = \|B\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_z = 0; \|B\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{0}$
0	$\begin{pmatrix}0\\ 0\\ 0\\ 1\end{pmatrix} = \|B\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_z = 0; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\middle|\, \psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle \Pi_A \otimes \sigma_z \rangle^{(w)} = \frac{\langle\phi\|(\Pi_A \otimes \sigma_z)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 1$	

Table A.6: σ_z polarization component in arm A

B.3 Polarization in path B

σ_x polarization component in arm B observable $\Pi_B \otimes \sigma_x$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\1\end{pmatrix} = \|B\rangle\|+\rangle$	$P\left(\Pi_B \otimes \sigma_x = 1; \|B\rangle\|+\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{6}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix} = \|B\rangle\|-\rangle$	$P\left(\Pi_B \otimes \sigma_x = -1; \|B\rangle\|-\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{6}$
0	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Pi_B \otimes \sigma_x = 0; \|A\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{2}{3}$
0	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \|A\rangle\|V\rangle$	$P\left(\Pi_B \otimes \sigma_x = 0; \|A\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
Weak Value $\langle \Pi_B \otimes \sigma_x \rangle^{(w)} = \frac{\langle\phi\|(\Pi_B \otimes \sigma_x)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 1$		

Table A.7: σ_x polarization component in arm B

σ_y polarization component in arm B observable $\Pi_B \otimes \sigma_y$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \|B\rangle\|\circlearrowleft\rangle$	$P\left(\Pi_B \otimes \sigma_y = 1; \|B\rangle\|\circlearrowleft\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{6}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-i\end{pmatrix} = \|B\rangle\|\circlearrowright\rangle$	$P\left(\Pi_B \otimes \sigma_y = -1; \|A\rangle\|\circlearrowright\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{6}$
0	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Pi_B \otimes \sigma_y = 0; \|A\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{6}$
0	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \|B\rangle\|V\rangle$	$P\left(\Pi_B \otimes \sigma_y = 0; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
Weak Value $\langle \Pi_B \otimes \sigma_y \rangle^{(w)} = \frac{\langle\phi\|(\Pi_B \otimes \sigma_y)\|\psi\rangle}{\langle\phi\|\psi\rangle} = -i$		

Table A.8: σ_y polarization component in arm B

σ_z polarization component in arm B observable $\Pi_B \otimes \sigma_z$		
Eigenvalue	**Eigenvector**	**ABL probability**
0	$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \|A\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_z = 0; \|A\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{1}$
0	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \|A\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_z = 0; \|A\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
1	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \|B\rangle\|H\rangle$	$P\left(\Pi_A \otimes \sigma_z = 1; \|B\rangle\|H\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
-1	$\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \|B\rangle\|V\rangle$	$P\left(\Pi_A \otimes \sigma_z = -1; \|B\rangle\|V\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle \Pi_B \otimes \sigma_z \rangle^{(w)} = \frac{\langle\phi\|(\Pi_B \otimes \sigma_z)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 0$	

Table A.9: σ_z polarization component in arm B

B.4 Polarization degree of freedom (without 'which-path' consideration)

σ_x polarization component irrespective of the path $\mathbb{1} \otimes \sigma_x$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix} = \|A\rangle\|+\rangle$	$P\left(\mathbb{1} \otimes \sigma_x = 1; \|A\rangle\|+\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{4}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix} = \|A\rangle\|-\rangle$	$P\left(\mathbb{1} \otimes \sigma_x = -1; \|A\rangle\|-\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{4}$
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\1\end{pmatrix} = \|B\rangle\|+\rangle$	$P\left(\mathbb{1} \otimes \sigma_x = 1; \|B\rangle\|+\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{4}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix} = \|B\rangle\|-\rangle$	$P\left(\mathbb{1} \otimes \sigma_x = -1; \|B\rangle\|-\rangle \,\middle|\, \langle\phi\| \,\|\psi\rangle\right) = \frac{1}{4}$
Weak Value	$\langle \mathbb{1} \otimes \sigma_x \rangle^{(w)} = \frac{\langle\phi\|(\mathbb{1} \otimes \sigma_x)\|\psi\rangle}{\langle\phi\|\psi\rangle} = 1$	

Table A.10: σ_x polarization component in the MZI

σ_y **polarization component irrespective of the path** $\mathbb{1}\otimes\sigma_y$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\\ 0\\ 0\end{pmatrix} = \lvert A\rangle\lvert\circlearrowleft\rangle$	$P\left(\mathbb{1}\otimes\sigma_y = 1; \lvert A\rangle\lvert\circlearrowleft\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \frac{1}{4}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\\ 0\\ 0\end{pmatrix} = \lvert A\rangle\lvert\circlearrowright\rangle$	$P\left(\mathbb{1}\otimes\sigma_y = -1; \lvert A\rangle\lvert\circlearrowright\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \frac{1}{4}$
1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 0\\ 1\\ i\end{pmatrix} = \lvert B\rangle\lvert\circlearrowleft\rangle$	$P\left(\mathbb{1}\otimes\sigma_y = 1; \lvert B\rangle\lvert\circlearrowleft\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \frac{1}{4}$
-1	$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 0\\ 1\\ -i\end{pmatrix} = \lvert B\rangle\lvert\circlearrowright\rangle$	$P\left(\mathbb{1}\otimes\sigma_y = -1; \lvert B\rangle\lvert\circlearrowright\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \frac{1}{4}$
Weak Value	$\langle\mathbb{1}\otimes\sigma_y\rangle^{(w)} = \frac{\langle\phi\rvert(\mathbb{1}\otimes\sigma_y)\lvert\psi\rangle}{\langle\phi\vert\psi\rangle} = -i$	

Table A.11: σ_y polarization component in the MZI

σ_z **polarization component irrespective of the path** $\mathbb{1}\otimes\sigma_z$		
Eigenvalue	**Eigenvector**	**ABL probability**
1	$\begin{pmatrix}1\\ 0\\ 0\\ 0\end{pmatrix} = \lvert A\rangle\lvert H\rangle$	$P\left(\mathbb{1}\otimes\sigma_z = 1; \lvert A\rangle\lvert H\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \mathbf{1}$
-1	$\begin{pmatrix}0\\ 1\\ 0\\ 0\end{pmatrix} = \lvert A\rangle\lvert V\rangle$	$P\left(\mathbb{1}\otimes\sigma_z = -1; \lvert A\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \mathbf{0}$
1	$\begin{pmatrix}0\\ 0\\ 1\\ 0\end{pmatrix} = \lvert B\rangle\lvert H\rangle$	$P\left(\mathbb{1}\otimes\sigma_z = 1; \lvert B\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \mathbf{0}$
-1	$\begin{pmatrix}0\\ 0\\ 0\\ 1\end{pmatrix} = \lvert B\rangle\lvert V\rangle$	$P\left(\mathbb{1}\otimes\sigma_z = -1; \lvert B\rangle\lvert V\rangle \,\middle\vert\, \langle\phi\rvert\,\rvert\psi\rangle\right) = \mathbf{0}$
Weak Value	$\langle\mathbb{1}\otimes\sigma_z\rangle^{(w)} = \frac{\langle\phi\rvert(\mathbb{1}\otimes\sigma_z)\lvert\psi\rangle}{\langle\phi\vert\psi\rangle} = 1$	

Table A.12: σ_z polarization component in the MZI

C Comment of Cheshire Cat using two photon interference

In a very recent work [arXiv:2004.07451], Kim et al claimed to have made the first genuine experimental observation of the Quantum Cheshire Cat effect. We dispute this claim on the ground that the setup employed is not adequate for making the weak measurements that define this interferometric effect. Half of the necessary weak values are not observed, and the other half is obtained indirectly by combining results measured with distinct setups.

In their work [A.1, A.2], henceforth designated by “P”, Kim et al. report what they term "*the first genuine experimental observation of the quantum Cheshire cat effect*". In this Comment, we argue that their claim does not withstand scrutiny. Our main argument is that the setup employed in P is inadequate to observe the quantum Cheshire cat (QCC) effect because the very nature of the effect requires weak measurements (WM) to be carried out on both arms of the interferometer. In P however WM are made on a single arm; moreover these WM are not the ones relevant for the QCC effect. Rather the WM required for observing the QCC are deduced from independent WM of a different observable.

The QCC effect [A.3] describes the spatial separation of a particle and one of its properties in the sense of WM. A single photon traveling in a Mach-Zehnder type interferometer has a weak value of 1 for the spatial projector Π_u on the upper arm u and 0 on the lower arm l; for a given property, say linear polarization represented by σ_z, the opposite holds. Hence the weak values (WV) defining the QCC are

$$\langle \Pi_u \rangle_w = 1 \qquad \langle \Pi_u \sigma_z \rangle_w = 0 \tag{A.69}$$

$$\langle \Pi_l \rangle_w = 0 \qquad \langle \Pi_l \sigma_z \rangle_w = 1. \tag{A.70}$$

All these WV can in principle be observed jointly by introducing weak couplings between the relevant probes on each arm and the photon prepared and detected in suitably chosen pre and post-selected states.

As remarked in P, early experiments on the QCC [A.4, A.5] were unable to implement the required weak measurements and resorted instead to strong couplings. Each WV in Eqs. (A.69)-(A.70) was obtained indirectly by combining several strong measurements,

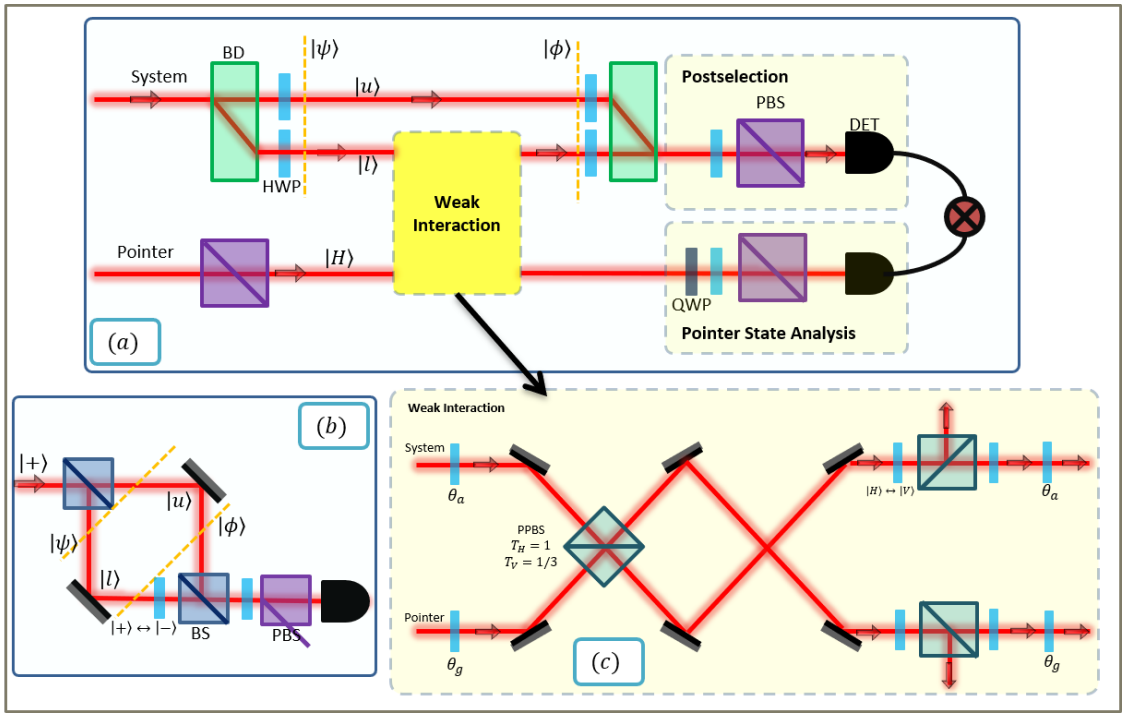


Figure A.5: $[a]$The experimental setup used in P. The two beam displacers (BD) form the Mach-Zehnder interferometer. The HWPs placed in the upper arm $|u\rangle$ and lower arm $|l\rangle$ makes the pre-selected state $|\psi\rangle = \frac{1}{\sqrt{2}}(|u\rangle|+\rangle + |l\rangle|-\rangle)$ and the post-selected state $\frac{1}{\sqrt{2}}(|u\rangle + |l\rangle)|+\rangle$. The pointer state is the polarization degree of freedom of another photon which has to be prepared in $|H\rangle$. $[c]$ The weak interaction is implemented by small rotation of C_z gate effectively achieved by two-photon interference at partially polarizing beam splitters (PPBS) and subsequent amplitude adjustments of the output system and pointer states. The implementation of weak interaction is subject to coincidence of system and pointer photons later which occurs with a probability of $1/9$. $[b]$ The Mach-Zehnder interferometer for QCC [A.3] with pre-selected state $|\psi\rangle = \frac{1}{2}(|u\rangle + |l\rangle)|+\rangle$ and post-selected state $|\phi\rangle = \frac{1}{2}(|u\rangle + |l\rangle)|+\rangle$.

employing distinct setups. This prevents the QCC from being observed because a strong measurement aimed at extracting a given WV disturbs the other ones, so that Eqs. (A.69)-(A.70) cannot be measured jointly even in principle [A.6]. Unfortunately, the setup used in P is also unable in principle to observe the QCC. There are several serious experimental shortcomings as well.

The main problem is that joint measurements of the observables are not performed. Since with parametric down conversion only one photon can be used as a pointer, the system and pointer can interact only in one arm in a given run. In P, measurements on only arm l are performed. No observation is done on arm u, so that the weak values in Eq. (A.69) cannot be determined directly from the experiment. How can one then observe spatial

separation? The authors argue (above Eq. (2) of P) that they can infer the WV on arm u from complementarity. But to the best of our knowledge, complementarity is not well defined for WM. For strong measurements, complementarity would be applicable: then the operators in each of the sets $\{\Pi_u, \Pi_l\}$ and $\{\Pi_u\sigma_z, \Pi_l\sigma_z\}$ are complementary, but both sets cannot be measured jointly with strong interactions. Crucially, for the measurements of $\{\Pi_u, \Pi_l\sigma_z\}$ that lie at the heart of the QCC effect (particle observed on arm u and polarization on l) there is no such complementarity: a joint measurement needs to be performed and this is only possible with WM. Hence the sole measurement of WV along l is in principle insufficient in order to observe the spatial separation of properties. Note that the "sum rule" for WV could in principle have been invoked but would have required an additional measurement: for example for the polarization the sum rule yields (See eg Eq. (37) of [A.7]) $\langle\Pi_u\sigma_z\rangle_w = \langle\sigma_z\rangle_w - \langle\Pi_l\sigma_z\rangle_w$ where $\langle\sigma_z\rangle_w$ is the polarization weak value that should be measured before the entrance or after the exit beam displacers.

A second major problem with the setup employed in P is that along arm l, none of the WV of Eq. (A.70) are directly measured. Instead $\langle\Pi_l\rangle_w$ and $\langle\Pi_l\sigma_z\rangle_w$ are reconstructed from independent WM of the polarization projectors Π_H and Π_V. The WV (A.70) are extracted by combining the pointer averages $\langle\sigma_x^H\rangle_p \pm \langle\sigma_x^V\rangle_p$ obtained employing different setups (the HWP angles θ_a and θ_g on Fig. 1(c) are changed) on different photons. Hence even if one wanted to rely on the sum rule (by performing the required additional measurement), it would be hard to claim that single photon spatial separation can be observed with the present setup.

Moreover genuine weak measurements are not performed. In the quantum non-demolition scheme employed in P the meter photon is not a von Neumann pointer but is coupled to the measured photon through a nondeterministic polarization-dependent coupling scheme [A.8, A.9]. In P, the required measurements are not done in the weak regime. Instead, the WV are obtained by interpolating, from measurements made at varying strengths g, a quantity that should in principle be measured in the neighbourhood of $g \approx 0$. This creates large uncertainties and estimated WV dependent to a large extent on the chosen polynomial function and data range included in the fit (this is particularly striking in Fig. 3(a) of P).

Note further that as a result of the nondeterministic coupling scheme employed, in which coincidences of system and pointer photons are needed, one is actually dealing with a

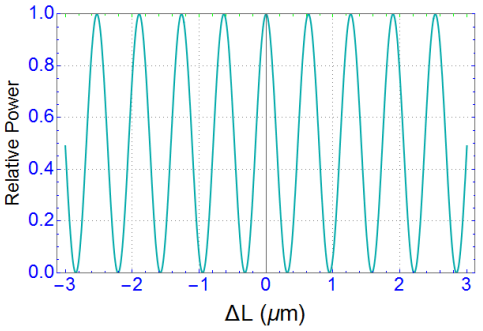
sub-ensemble of the original pre and post selection ensemble. Due to the probabilistic nature of system and pointer photons emerging in different ports after the two-photon interference and subsequent amplitude reduction (see Fig. 1(c)), the overall probability of successful implementation of weak interaction is $1/9$. Thus the experiment loses $8/9$ of the system photons that are part of the pre and post-selection ensemble.

Overall we have argued that strictly speaking the experiment reported in P deals with interpolating WV from orthogonal polarization projector measurements made in the sole lower arm of the interferometer. These results represent a progress relative to the early experiments on the QCC [A.4, A.5] in that weak values are obtained from quantum non-demolition measurements. But contrary to what is claimed in P, this is still far from representing a genuine and unambiguous observation of the Quantum Cheshire Cat effect.

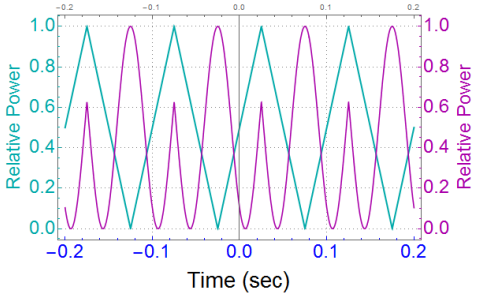
D Stabilization of the Mach Zehnder Interferometer

D.1 Scanning path difference of the interferometer for Laser based partial Cheshire Cat Experiment

We change the path difference of the interferometer by applying a voltage to the piezo. As a function of voltage, the piezo expands linearly which in turn changes the path length between interferometer. If we scan the piezo with a triangular ramp voltage signal, in the absence of noise, we would get power at one beam splitter port as a function of time as



(a) Intensity vs path difference



(b) Power vs time as the piezo is scanned (Without noise)

Using a DAQ card we can read voltages from the detector and infer the phase drift. Using LabView, we design a PID algorithm that maintains the phase shift at a desired set point.

As is the case with any control system, we do have residual fluctuation of the phase after the PID loop. We monitor the same and incorporate in the uncertainty calculation in the expected weak value for the experiment.

D.2 Finding set point for the single photon based experiment

We need an auxillary laser to monitor the phase difference between two arms of the interferometer when we have the single photon source. For this, as mentioned in Ch 4, we have used Helium Neon laser that goes through the interferomter and later filtered using dichroic mirrors and bandpass filters. To characterize, we use a 810 nm pulsed laser beam and obtain the voltage at separate detectors as the piezo is scanned.

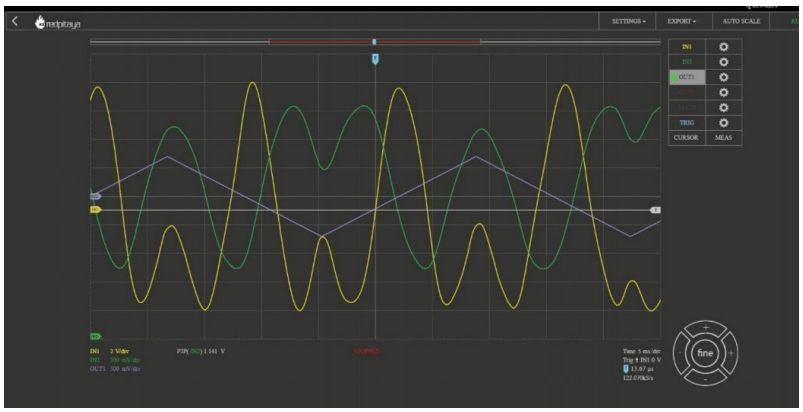


Figure A.7: IN1 = from Detector with Helium Neon beam incident, IN2= 810 nm beam (projection close to Diagonal) , OUT1= Scan signal before amplification

It is important that the path difference is exactly 0 as the two wavelengths are different. If the path difference is any integer multiple of the wavelength may give zero phase for the 632.8 nm beam it would not correspond zero phase difference for the 633 nm. We can ensure zero phase difference for 810 beam only when the path difference is scanned within one wavelength. The aim then is to take use the reference voltage for the 633.8 nm where the 810 nm path difference is exactly zero and lock the interferometer using that as a set point.

Adjusting the path difference to zero is achieved by maximizing the visibility with pulsed 810 nm beam as the piezo is scanned by moving the precision actuator on which it is mounted.

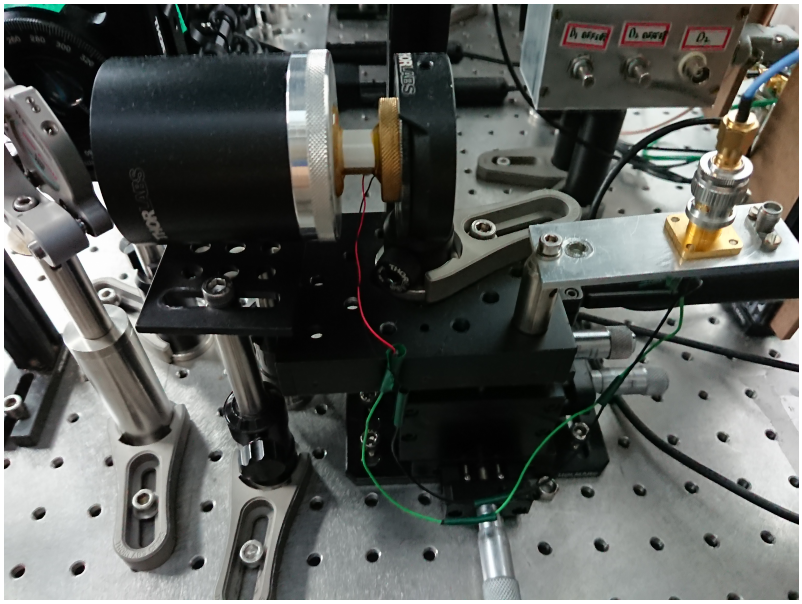


Figure A.8: In Partial Cheshire Cat experiment with single photon source (Ch 4), the corner cube retroreflector is attached to a piezo which recieves volatage from the DAQ card for scanning and stabilization. The corner cube is mounted on a rotation stage so that the incident and emergent beams can be adjusted to be in the same height. For further alignment the whole assembly is mounted on a 3-axis translation stage. To perform fine adjustment to the path length difference bewteen the two arms of the interferometer a motorized actuator is mounted on top of the 3-axis manual translation stage. This actuator is used to maximize visibility and ensure that the path difference is within one wavelength.

E Error Budget

E.1 Calculation of Pointer Shift

Here, we aim to compute the error propagated to the measurement of centroid due to experimental uncertainties in designing the experiment. Major contribution to the error is slight optical misalignment that results due to limitation of precision with which optomechanics can remain aligned for the duration of one experimental run.

The electric field in the transmitted arm of BS_i can be written as

$$E_{T_i} = A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{y^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z) \begin{pmatrix} \alpha_T \\ \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \tag{A.71}$$

Here, A_T is the amplitude of light transmitted through the beam splitter BS_i. We shall be

following a convention in which the local direction of propagation of the beam is along z axis. We consider the intensity profile to be fairly Gaussian and the deviations can be addressed with the spatial function η_S. The polarization along $\hat{x}$ and $\hat{y}$ are represented in Jones column vector. When the polarization is $\hat{x}$, we have $\alpha_T = 1$. Similarly, we have the electric field for the reflected arm as

$$E_{R_i} = A_R \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{y^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z) \begin{pmatrix} \alpha_R \\ \sqrt{1-|\alpha_R|^2}\exp(i\phi_R) \end{pmatrix} \tag{A.72}$$

Notice that $\hat{z}$ in expression A.71 and A.72 do not represent the same direction, in fact they are orthogonal in space after the first beam splitter $\mathrm{BS_i}$. But since, we wish not to complicate things with direction changes after each reflection, we shall denote locally all direction of propagation as $\hat{z}$, until we consider the beam past the second beam splitter $\mathrm{BS_f}$. The physical direction of the beam can always be inferred in the context of the schematic of the experimental set up A.9. Although, we ideally aim for $\mathrm{BS_i}$ to be a 50:50 beam splitter, in practice, the beam A_R need not be same as A_T. The reflected beam gets across components like the corner cube retroreflector which not only changes the intensity but introduces ellipticity in the polarization. The latter is compensated by the use of half wave plate $\mathrm{H_C}$ together with the quarter wave plate $\mathrm{Q_C}$. Since we want the pre-selected state to be

$$\left|\psi(t_{pre})\right\rangle = \frac{1}{\sqrt{2}}\left(|T\rangle\,|H\rangle + i\exp(i\varphi)\,|R\rangle\,|V\rangle\right). \tag{A.73}$$

, then we have $\alpha_R \to 0$.

Now, in the schematic Figure A.9, below we have placed both beam displacer and glass plate in both the arms simultaneously. In practice, we shall have only two pointers and hence can do pairwise permutations of the operators only. In this calculation however, we shall assume that all the four operators exist. Since the presence of the two half-wave plates $\mathrm{H_{T/R_1}}$ and $\mathrm{H_{T/R_2}}$ are desired to be oriented such that they make $\pi/8$ angle from $\hat{x}$, they together would make the identity operator and hence in theory can be assumed to be always present. The beam displacers presence and absence can be turned on and off by changing the displacement γ_T^{BD} that it causes between the two orthogonal polarizations

to be zero or non-zero. We shall, now, evolve the two fields ,transmitted and reflected, individually, assuming they pass though all the four optics and then finally superpose the them at the output port of BS_f after taking care of angular misalignments.

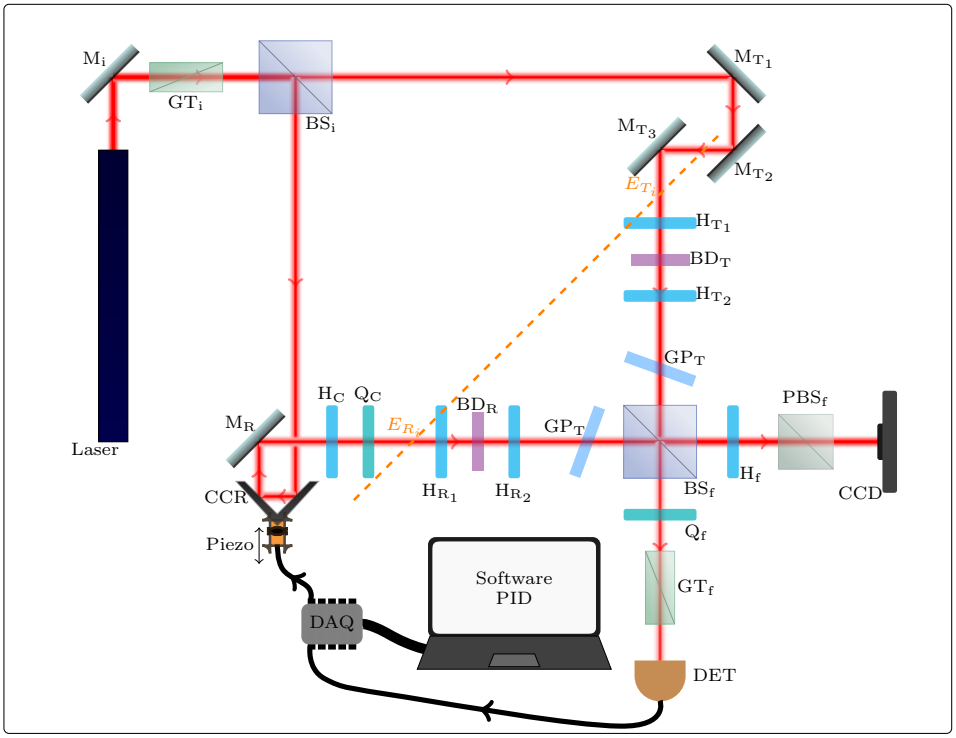


Figure A.9: Schematic for the calculation of pointer shift. We consider both the beam displacer and the glass plate in both the arms on the interferometer to derive the expression for pointer shifts. While doing the experiment, we have one glass plate and one beam displacer and depending on the case, we remove the terms concerned with components not present in the experiment.

After the waveplate H_{T_1}, the transmitted field would be

$$E_{T_1} = A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{y^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha_T \\ \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \tag{A.74}$$

$$= A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{y^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z) \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \tag{A.75}$$

Here, we have used the Jones matrix for a half wave plate oriented at an angle θ from $\hat{x}$.

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \tag{A.76}$$

, with $\theta = \pi/8$. Here, we have not accounted for any error in alignment of the half wave plate as the error in polarization due to misalignment can be taken into absorbed by α_T and ϕ_T. Next, we consider that the beam displacer causes the desired shift $\gamma_T^{BD_V}$ of the vertical polarization along $\hat{y}$ axis which includes the additional shift due to tilting. The horizontal component, due to the effect of tilt, gets displaced by $\gamma_T^{BD_H}$. The phase difference between the horizontal and vertical component emerging out of the polarizing beam displacer is denoted by φ_T^{BD}. Thus, we have the electric field as,

$$\begin{aligned} E_{T_2} = A_T &\left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_H})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z) \\ &\frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ 0 \end{pmatrix} \\ + A_T &\left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_V})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_T^{BD}\right) \\ &\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \end{aligned} \tag{A.77}$$

After the half wave plate $\mathrm{H}_{\mathrm{T}_2}$, we have the state as

$$E_{T_3} = A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_H})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z)$$
$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ 0 \end{pmatrix}$$
$$+A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_V})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_T^{BD}\right)$$
$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix}$$
(A.78)

$$= A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_H})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp(i\,k\,z)$$
$$\frac{1}{2} \begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix}$$
$$+A_T \left(\exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp\left(-\frac{(y-\gamma_T^{BD_V})^2}{4\sigma_y^2}\right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_T^{BD}\right)$$
$$\frac{1}{2} \begin{pmatrix} \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ -\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix}$$
(A.79)

The glass plate $\mathrm{GP_T}$ causes a displacement of γ_T^{GP} along the horizontal irrespective of the polarization and thus the field becomes

$$E_{T_3} = A_T \left(\exp\left(-\frac{(x-\gamma_T^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_T^{BD_H})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp(i\,k\,z) \tag{A.80}$$

$$\frac{1}{2} \begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \tag{A.81}$$

$$+A_T \left(\exp\left(-\frac{(x-\gamma_T^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_T^{BD_V})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_T^{BD}\right) \tag{A.82}$$

$$\frac{1}{2} \begin{pmatrix} \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ -\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \tag{A.83}$$

Similarly, the reflected arm after going through H_{R_1}, BD_R, H_{R_2} and GP_R can be represented by the electric field,

$$E_{R_3} = A_R \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_H})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp(i\,k\,z)$$
$$\frac{1}{2} \begin{pmatrix} \alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R) \\ \alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R) \end{pmatrix}$$
$$+A_R \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_V})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_R^{BD}\right)$$
$$\frac{1}{2} \begin{pmatrix} \alpha_R - \sqrt{1-|\alpha_R|^2}\exp(i\phi_R) \\ -\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R) \end{pmatrix} \tag{A.84}$$

Now that we have accounted for all the operators, we can consider misalignments. The beam reflected from BS_i is transmitted through BS_f to reach the beam profiler CCD and hence we can consider this as our reference (i.e. $E_{R_4} = E_{R_3}$) and the direction of its propagation can be considered along $\hat{z}$ axis and its point of incidence at BS_f can be considered as the origin in $x-y$ plane. The beam transmitted through BS_i gets reflected from BS_f to reach the beam profiler. Thus, we can add the effect of misalignment to this beam. We shall consider that the undesired relative displacement are given by $\gamma_{x/y}^{MA}$ along

horizontal/vertical direction. After reflection, the beam is slightly tilted from z axis and hence we can write the electric field with new variables $X, Y \& Z$ as given below

$$\begin{aligned} E_{T_4} = A_T &\left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_H}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\ &\exp(i\,k\,Z)\frac{1}{2}\begin{pmatrix} \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ \alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \\ + A_T &\left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_V}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\ &\exp\left(i\,k\,Z + i\varphi_T^{BD}\right)\frac{1}{2}\begin{pmatrix} \alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \\ -\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T) \end{pmatrix} \end{aligned} \tag{A.85}$$

Here, the coordinate transformation is made using the following sequential transformations,

$$\begin{pmatrix} X' \\ Z' \end{pmatrix} = \begin{pmatrix} \cos(\theta_{xz}) & -\sin(\theta_{xz}) \\ \sin(\theta_{xz}) & \cos(\theta_{xz}) \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \tag{A.86}$$

$$\begin{pmatrix} Y' \\ Z \end{pmatrix} = \begin{pmatrix} \cos(\theta_{yz}) & -\sin(\theta_{yz}) \\ \sin(\theta_{yz}) & \cos(\theta_{yz}) \end{pmatrix} \begin{pmatrix} y \\ Z' \end{pmatrix} \tag{A.87}$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos(\theta_{xy}) & -\sin(\theta_{xy}) \\ \sin(\theta_{xy}) & \cos(\theta_{xy}) \end{pmatrix} \begin{pmatrix} X' \\ Y' \end{pmatrix} \tag{A.88}$$

Here, θ_{xz}, θ_{yz} and θ_{xy} are the angles by which the reflected beam from $\mathrm{BS_f}$ is misaligned wrt. the transmitted beam.

$$\begin{aligned} X = z\cos(\theta_{xy})(\cos(\theta_{xz}) - \sin(\theta_{xz})) + \sin(\theta_{xy})(-y\cos(\theta_{yz}) \\ + z\sin(\theta_{yz})(\cos(\theta_{xz}) + \sin(\theta_{xz}))) \end{aligned} \tag{A.89}$$

$$Y = z\sin(\theta_{xy})(\cos(\theta_{xz}) - \sin(\theta_{xz})) + \cos(\theta_{xy})(y\cos(\theta_{yz}) \\ -z\sin(\theta_{yz})(\cos(\theta_{xz}) + \sin(\theta_{xz}))) \tag{A.90}$$

$$Z = z\cos(\theta_{yz})(\cos(\theta_{xz}) + \sin(\theta_{xz})) + y\sin(\theta_{yz}) \tag{A.91}$$

The electric field after the second beam splitter in the port that leads to the beam profiler is given by

$$E_5 = \frac{1}{\sqrt{2}}(E_{R_4} + \exp(i\varphi_{\Delta L})E_{T_4}) \tag{A.92}$$

Here, $\varphi_{\Delta L}$ is the phase difference between the two arms of the beam splitter accounting for any polarization independent optical path difference.
After passing through, the half wave plate H_f oriented at an angle θ, we have the electric field

$$E_6 = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} (E_{R_4} + \exp(i\varphi_{\Delta L})E_{T_4} \tag{A.93}$$

Absorbing all the constant factors in $A_{T/R}$, we can write E_6 as

$$
\begin{aligned}
E_6 = {} & A_R \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_H})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp(i\,k\,z) \\
& \begin{pmatrix} \cos(2\theta)(\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) + \sin(2\theta)(\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \\ \sin(2\theta)(\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) - \cos(2\theta)(\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \end{pmatrix} \\
& + A_R \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_V})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_R^{BD}\right) \\
& \begin{pmatrix} \cos(2\theta)(\alpha_R - \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) + \sin(2\theta)(-\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \\ \sin(2\theta)(\alpha_R - \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) - \cos(2\theta)(-\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \end{pmatrix} \\
& + \exp(i\varphi_{\Delta L}) A_T \left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_H}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\
& \exp(i\,k\,Z) \\
& \begin{pmatrix} \cos(2\theta)(\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) + \sin(2\theta)(\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) \\ \sin(2\theta)(\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) - \cos(2\theta)(\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) \end{pmatrix} \\
& + \exp(i\varphi_{\Delta L}) A_T \left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_V}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\
& \exp\left(i\,k\,Z + i\varphi_T^{BD}\right) \\
& \begin{pmatrix} \cos(2\theta)(\alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) + \sin(2\theta)(-\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) \\ \sin(2\theta)(\alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) - \cos(2\theta)(-\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) \end{pmatrix}
\end{aligned}
\tag{A.94}
$$

After the final polarizing beam splitter $\mathrm{PBS_f}$, we have

$$
\begin{aligned}
E_f = A_R & \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_H})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp(i\,k\,z) \\
& (\cos(2\theta) + \sin(2\theta))(\alpha_R + \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \\
& + A_R \left(\exp\left(-\frac{(x-\gamma_R^{GP})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(y-\gamma_R^{BD_V})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \exp\left(i\,k\,z + i\varphi_R^{BD}\right) \\
& (\cos(2\theta) - \sin(2\theta))(\alpha_R - \sqrt{1-|\alpha_R|^2}\exp(i\phi_R)) \\
& + \exp(i\varphi_{\Delta L}) A_T \left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_H}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\
& \exp(i\,k\,Z)(\cos(2\theta) + \sin(2\theta)(\alpha_T + \sqrt{1-|\alpha_T|^2}\exp(i\phi_T)) \\
& + \exp(i\varphi_{\Delta L}) A_T \left(\exp\left(-\frac{(X-\gamma_T^{GP}-\gamma_x^{MA})^2}{4\sigma_x^2} \right) \exp\left(-\frac{(Y-\gamma_T^{BD_V}-\gamma_y^{MA})^2}{4\sigma_y^2} \right) + \eta_S(x,y) \right) \\
& \exp\left(i\,k\,Z + i\varphi_T^{BD}\right)(\cos(2\theta) - \sin(2\theta))(\alpha_T - \sqrt{1-|\alpha_T|^2}\exp(i\phi_T))
\end{aligned}
\tag{A.95}
$$

The power density is computed as

$$
\rho = |E_f|^2 \tag{A.96}
$$

The x and y centroid at a given z, are given by

$$
c_x = \frac{\int x\rho \; dxdy}{\int \rho \; dxdy} \tag{A.97}
$$

$$
c_y = \frac{\int y\rho \; dxdy}{\int \rho \; dxdy} \tag{A.98}
$$

E.2 Error Estimation

The centroid in Eqn. A.97 and Eqn. A.98 depends not only on the accuracy with which we determine the intended displacement caused operators but also due to misalignment. The tilt angles can be about $60\mu rad$ inferred by uncertainty of fluctuations of the centroid. The overlap of the beams coming from two arms can be claimed up to $5\mu m$. Considering that different optical component cause different loses, the power ratio in the two arms was about 1.1:1. Another major source of error is phase difference $\varphi_{\Delta L}$ which can be stabilized

only up to 8 deg.

All the above range of values were plugged to obtain Eqn. A.97 and Eqn. A.98 the theoretical band that accounts for experimental uncertainties.

F Alignment with Pulsed Laser

F.1 Determining the coherence length of the pulsed laser

The coherence length of single photon source is of the order of µm. The typical coherence length of interferometric grade single longitudinal mode lasers are of the order of few metres and above. Thus, it is impractical to adjust the path difference between two arms of the Mach-Zehnder interferomter up to few µm with a laser as the change in visibility would be negligible. Therefore, we need a source with coherence length of the same order as that of the single photon source. For this, a femto-second pulsed source is suitable.

We use the Ti-Saphire pulsed laser source in femto-second mode to make fine adjustment to the path difference between the two arms of the interferometer. But first, we need to determine the coherence length of the pulsed source itself. One can compute the coherence length from the quoted pulse length itself but whether the laser is behaving as per the specifications needs to be verified. The standard method to determine the coherence length of a femto-second pulsed laser is to find the auto-correlation experiment in which the power of higher(usually second) harmonic generated beam is computed as a function of path delay between the two incident beams. However, here we aim to to determine the coherence length without using the $\chi^{(2)}$ crystal meant for SHG. In principle, we can know the the coherence length by plotting visibility vs. path delay in any two-path interferometer such as Michelson interferometer and Mach-Zehnder interferometer. The problem is, since the coherence length is less than 100µm, the absolute path difference between the two arms need to be precisely adjusted within that tolerance. This is cumbersome to achieve. Therefore we use the displaced Sagnac interferometer in which naturally the path difference between the two arms is almost zero (tolerance is about 200 nm).

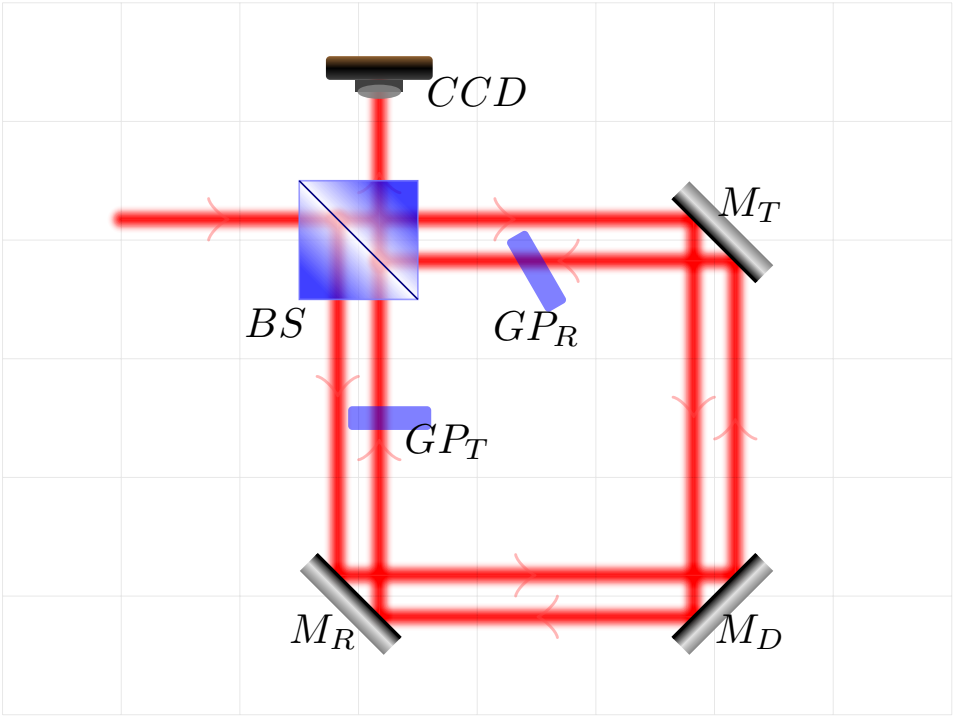


Figure A.10: Experimental set up to determine the coherence length of the pulsed laser

- First, the displaced sagnac interferometer is aligned to obtain the collinear interference.
- the beam splitter *BS* is slightly tilted and the mirrors are slightly tweaked so that we have the non-collinear interference pattern (double-slit like). The beam profiler's position needs to be adjusted to ensure the maximum overlap between the two beams.
- Two glass plates of of same thickness (3mm) are inserted in the clockwise(GP_T, transmitted arm of the beam splitter) and anti-clockwise (GP_R, reflected arm of the beam splitter) arms of the interferometer. One of them is mounted on abase-rotation mount.
- The glass plate on base rotated mount in steps of 1 deg. This tilts the glass plate and causes a longitudinal path delay. Note that , it also causes a lateral displacement. But in our case the lateral displacement will be in the order of few tens of microns where as the beam size is more than 1mm. Thus the loss in visibility due the lateral displacement can be ignored.
- As the longitudinal optical path delay changes the visibility also changes. Thus the

beam profile can be recorded and visibility can be inferred.

- From the plot of visibility vs. path delay, we can infer the coherence length of the pulsed laser.

.

The above method of inferring the coherence length of the pulsed laser forms an alternative to the usual auto-correlation method. The advantage here is that we do not need to deal with crystal dispersions and obtain the coherence length simply using an interferometric technique.

F.2 Extra path difference due to the Corner Cube Retro-reflector

The aim is to compute the additional optical path difference due to the corner cube retro reflector which shall provide the ball park estimate of path compensation to be done with the extra mirror loop.

The corner cube retro-reflector mainly consists of three reflecting surfaces. For the purpose of the calculation, we shall assume that the incident beam is along $-\hat{z}$. Two of the reflecting planes intersects along $\hat{y}$ axis. We shall calculate the path traversed by the beam in the glass material and then multiply with the refractive index of the material to find the optical path difference. The first task is to have orientation of the three planes so that the point of reflections could be computed. Let the common point where the three reflecting surfaces meet be at origin. They make 120 degree with each other in their projection on the $x-y$ plane and 35 deg in the their projection on the $y-z$ plane. So the the normals to the respective planes can be computed as follows:

- Rotation of $\hat{y}$ about x axis by -35 deg
- Rotation of $\hat{y}$ about x axis by -35 deg, followed by 120 deg rotation about z axis
- Rotation of $\hat{y}$ about x axis by - 35 deg, followed by -120 deg rotation about z axis

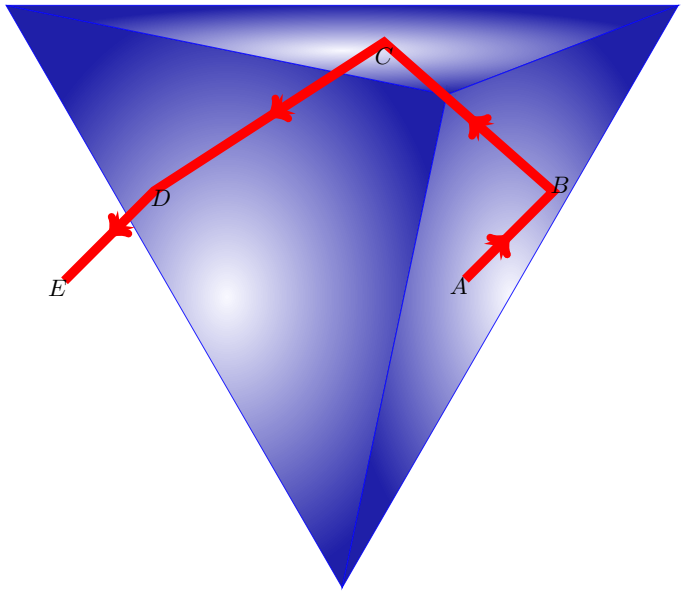


Figure A.11: Calculation of extra optical path difference due to CCR

The direction vectors te three normals to the respective reflective surfaces are obtained by the above mentioned rotation operation.

$$\hat{n}_x = \begin{pmatrix} \frac{\sqrt{3}}{2}\cos 35^\circ \\ -\frac{1}{2}\cos 35^\circ \\ \sin 35^\circ \end{pmatrix} \quad \hat{n}_y = \begin{pmatrix} -\frac{\sqrt{3}}{2}\cos 35^\circ \\ -\frac{1}{2}\cos 35^\circ \\ \sin 35^\circ \end{pmatrix} \quad \hat{n}_z = \begin{pmatrix} 0 \\ \cos 35^\circ \\ \sin 35^\circ \end{pmatrix} \tag{A.99}$$

The equation of the plane with which the incident beam intersects is given by the equation of the plane normal to $\hat{n}_x$ passing through origin.

$$\hat{n}_x \cdot \{x, y, z\} = 0 \tag{A.100}$$

$$\Rightarrow \frac{\sqrt{3}}{2}\cos(35^\circ)x - \frac{1}{2}\cos(35^\circ)y + \sin(35^\circ)z = 0 \tag{A.101}$$

$$\Rightarrow z = \frac{\frac{\sqrt{3}}{2}\cos(35^\circ)x - \frac{1}{2}\cos(35^\circ)y}{\sin(35^\circ)} \tag{A.102}$$

If the incident beam is moving along $-\hat{z}$ with horizontal and vertical shifts of x_0 and y_0 respectively, we obtain the following point of incidence marked (B) in Fig. A.10.

$$B = \left(x_0, y_0, \frac{\frac{\sqrt{3}}{2}\cos(35^\circ)x_0 - \frac{1}{2}\cos(35^\circ)y_0}{\sin(35^\circ)} \right) \tag{A.103}$$

In the experiment, we aim to make $y_0 = 0$ i.e. to have no vertical displacement caused due to the corner cube retro-reflector. In practice, a graph paper was placed on the front mount of the CCR (without touching the glass surface as there is about 6.6mm gap between the front edge of the mount and the glass surface.) and it was ensured that the vertical offset is less than 1mm. Similarly, horizontal offset was found to be 15 ± 1mm.
Now, we know the point of incidence (B), the incident direction $\vec{AB}$ and hence we can find out the direction vector $\vec{BC}$ for reflection. The angle of incidence equals angle of reflection and hence

$$\hat{n}_x \cdot \vec{AB} = \hat{n}_x \cdot \vec{BC}. \tag{A.104}$$

But the above equation alone does not constraint the reflected vector to be in the plane of incidence. For that we have to ensure that the angle between the incident and the reflected ray is twice the angle of incidence.

$$\arccos(\vec{AB} \cdot \cdot \vec{BC}) = 2\arccos(\hat{n}_x . \vec{AB}). \tag{A.105}$$

Simultaneously solving the above equations involving arcosines might be cumbersome and hence we use the alternate approach of using the reflection transform matrix. The reflection transform matrix of a plane normal to the vector $\hat{v} = (v_x, y_y, v_z)$ is given as follows.

$$R_v = \begin{pmatrix} 1 - \frac{2v_x^2}{\|\vec{v}^2\|} & -\frac{2v_x v_y}{\|\vec{v}^2\|} & -\frac{2v_x v_z}{\|\vec{v}^2\|} \\ -\frac{2v_y v_x}{\|\vec{v}^2\|} & 1 - \frac{2v_y^2}{\|\vec{v}^2\|} - \frac{2v_y v_z}{\|\vec{v}^2\|} & \\ -\frac{2v_z v_x}{\|\vec{v}^2\|} & -\frac{2v_z v_y}{\|\vec{v}^2\|} & 1 - \frac{2v_z^2}{\|\vec{v}^2\|} \end{pmatrix} \tag{A.106}$$

In our case, R_v, with $\vec{v} = \hat{n_x}$ acting on the incident direction vector $\hat{AB} = (0, 0, -1)$ gives the direction of the reflected direction vector $\hat{BC}$.

$$\hat{BC} = \begin{pmatrix} \frac{\sqrt{3}}{2}\cos(20^\circ) \\ -\frac{1}{2}\cos(20^\circ) \\ -\sin(20^\circ) \end{pmatrix} \tag{A.107}$$

The equation of the line BC, with the known point $B$is given by

$$\frac{x-x_0}{\frac{\sqrt{3}}{2}\cos(20^\circ)} = \frac{y-y_0}{-\frac{1}{2}\cos(20^\circ)} = \frac{z-\left(\frac{\frac{\sqrt{3}}{2}\cos(35^\circ)x_0-\frac{1}{2}\cos(35^\circ)y_0}{\sin(35^\circ)}\right)}{-\sin(20^\circ)} \tag{A.108}$$

Now the aim is find the point of incidence (C) on the next surface which would lie on the the line given by the equation above and intersect with the plane normal to $\hat{n_z}$ and passing through origin. We simultaneously solve the equation of the line BC with the equation of the plane given by

$$\hat{n_z}\cdot\{x,y,z\} = 0 \tag{A.109}$$

$$\Rightarrow \cos(35^\circ)y + \sin(35^\circ)z = 0 \tag{A.110}$$

We obtain the point (C) as

$$C = \begin{pmatrix} \frac{\left(5x_0+\sqrt{3}y_0\right)\cot(35^\circ)\csc(20^\circ)+4x_0\sec(20^\circ)}{4\sec(20^\circ)+2\cot(35^\circ)\csc(20^\circ)} \\ y_0 - \frac{\cot\left(\frac{7\pi}{36}\right)\csc\left(\frac{\pi}{9}\right)\left(\sqrt{3}x_0+y_0\right)}{4\sec\left(\frac{\pi}{9}\right)+2\cot\left(\frac{7\pi}{36}\right)\csc\left(\frac{\pi}{9}\right)} \\ \frac{\cot(35^\circ)\left(\left(\sqrt{3}x_0-y_0\right)\cot(35^\circ)\csc(20^\circ)-4y0\sec(20^\circ)\right)}{4\sec(20^\circ)+2\cot(35^\circ)\csc(20^\circ)} \end{pmatrix} \tag{A.111}$$

Next, we need to find the point (D). For that, we need to find the equation of the line CD. The reflected ray's direction vector is found by acting the matrix R_v, with $v = \hat{n}_y$ on the vector BC. Thus, we obtain the equation of the line as

$$\frac{x-\left(\frac{\left(5x_0+\sqrt{3}y_0\right)\cot(35^\circ)\csc(20^\circ)+4x_0\sec(20^\circ)}{4\sec(20^\circ)+2\cot(35^\circ)\csc(20^\circ)}\right)}{\frac{1}{2}\sqrt{3}\cos(20^\circ)} \tag{A.112}$$

$$=\frac{y-\left(y_0-\frac{\cot\left(\frac{7\pi}{36}\right)\csc\left(\frac{\pi}{9}\right)\left(\sqrt{3}x_0+y_0\right)}{4\sec\left(\frac{\pi}{9}\right)+2\cot\left(\frac{7\pi}{36}\right)\csc\left(\frac{\pi}{9}\right)}\right)}{\frac{3}{2}\sin(20^\circ)\cos(20^\circ)} \tag{A.113}$$

$$=\frac{z-\left(\frac{\cot(35^\circ)\left(\left(\sqrt{3}x_0-y_0\right)\cot(35^\circ)\csc(20^\circ)-4y0\sec(20^\circ)\right)}{4\sec(20^\circ)+2\cot(35^\circ)\csc(20^\circ)}\right)}{\frac{1}{2}\cos^2(20^\circ)-\sin^2(20^\circ)} \tag{A.114}$$

We simultaneously solve the equation of the line with the plane with normal $\hat{n_y}$ and passing through origin.

$$\hat{n}_y\cdot\{x,y,z\}=0 \tag{A.115}$$

$$\Rightarrow -\frac{\sqrt{3}}{2}\cos(35^\circ)x-\frac{1}{2}\cos(35^\circ)y+\sin(35^\circ)z=0 \tag{A.116}$$

Using the above equations, we find the intersection point (D), with the assumption that $y_0=0$.

Since, the device is a retro-reflector the beam DE is along $\hat{z}$. The total distance traveled by the beam inside the glass material is given by

$$L_g=\overline{AB}+\overline{BC}+\overline{CD}+\overline{DE} \tag{A.117}$$

We have found the coordinates of B, C and D depending on x_0. Since the distance from origin to the front surface of the glass is 35.4mm, we assign the z-coordinate of (A) and (E). For calculation purpose, we use $x_0=15$mm and obtain L_g to be about 70.5mm. Thus the optical path, given the refractive index of the glass to be 1.51 is about 106.5mm. Given that the mount has a gap of about about 6.5mm between its outer ring and the glass the additional longitudinal path delay is 13mm. From the point of reflection from the mirror to the mount, there is a gap of about 1cm. Thus over and above 20mm is added to the extra path and thus, the total path traversed is calculated to be 139.5mm. Note, in the transmitted arm of the Mach Zehnder interferomter, we would already account for the lateral path delay which is about 30mm. Thus, the two mirrors in the transmitted arm of the MZI,

should compensate for about 110mm and thus each needs to be placed at least 2 or 3 holes behind. If placed 3 holes behind, The difference of about 40mm need to be accounted for by moving the CCR back by about 2cm. This can be achieved since the CCR is placed on a translation stage. On the other hand, if we place the two mirrors 2 holes behind, we have to move the corner cube about 5mm closer.

F.3 Alignment of the Mach-Zehnder interferometer with pulsed laser

In the experimental set up 4.12, we have the CCR attached to the piezo mounted on a translation stage. The two extra mirrors in the transmitted arm are are placed 3 inches behind to compensate for the extra path difference caused by the CCR.

- First, the MZI was aligned in non collinear configuration using CW mode in the femto-cavity. Then we mode-locked the laser to obtain femto-second pulses. The fringes on the the beam profiler now disappeared.
- The beam emerging from the other port of the beam splitter BS_2 is made incident on a fast photo diode. Since the fringes are non-collinear a pin-hole is used to allow only a small portion of the fringe to fall on the detector. The piezo connected to the CCR is now scanned and the change in intensity measured by the photo-diode is seeked for. If the path difference is larger than the coherence length there is no interference and hence no oscillation in path difference.
- To begin with, we directly do not aim to see the interference pattern in the beam profiler. This is because, even when the piezo is not scanned, as we manually move the translation stage the vibrations are enough to wash out the interference pattern unless we have exactly compensated for the path difference (within say 10 microns tolerance).
- The CCR, now mounted on an actuated translation stage was moved till we notice the oscillations in oscilloscope connected to the voltage output of the fast photo diode. There are few positions of the stage where we notice tiny oscillations but the visibility never increases beyond 1 or 2 %. These are most probably due to multiple reflections of the beam. But the true zero path delay position is nearby them and is

identified with almost 100% visibility.

- Once we are near the zero path delay position of the actuator, we switch to the beam profiler to record the fringes. We scan about 250 microns and recorded fringes with various filtering conditions: 810-10 Bandpass filter (Thorlabs), 808-3 Bandpass filter(Semrock). From the visibility plot, we can compute the coherence length at various filtering conditions. This gives us an experimental reference of what to expect from the singles when we go to the single photon source.
- When we have no filter, the peak visibility reaches only about 60%. With the filter with 3nm bandpass width, the visibility obtained is greater than 99%.

G Uncertainty and Error Propagation in Pointer Shift

The centroid of the beam in a weak measurement experiment can have random errors due to the pointing fluctuations in the beam due to which the centroid wanders over time. The physical cause of fast beam (Hz - kHz) wander is mainly mechanical vibrations of opto-mechanical components. For example, any mechanical vibration in the fiber tip with respect to the collimating lens would cause a transverse beam wander. Nevertheless, there would also be systematic errors creeping in because of non-idealness of many experimental parameters. Some are due to non-idealness of the state preparation and others are due to non-idealness in experimental realization of the desired operator. For the sake of rigour, one need to analyze the pointer shift in the experiment involving weak measurement as a result of interference between several spatial or polarization modes of a photon. However, if we ignore the contribution from multi-photon events, then the leading contribution to the error or uncertainty in the pointer shift is the non-idealness of the spatial profile and the polarization of the beam. This can be better understood by treating the photon propagation as a pulse and describing its spatio-temporal mode as we do for the electric field of a light pulse. However, strictly speaking here we would be describing the probability amplitude of the photon.

G.1 Description of the beam

The electric field amplitude of the pulse can be described as having a transverse Gaussian profile. Due to divergence the size of Gaussian beam increases as it propagates and is given as the beam waist $w(z)$ as the beam propagates along $\hat{z}$.

$$w(z) = w_0\sqrt{1+\left(\frac{z}{z_R}\right)^2} \tag{A.118}$$

The transverse profile of the beam is ideally assumed to have azimuthal symmetry and is typically described as a Gaussian only depended on the radial distance $r = \sqrt{x^2+y^2}$ from the centre. However, in practice, the beam can have slight elliptic transverse profile. The ellipse profile is described by two axes named major(M) and minor(m) which are orthogonal to each other. However, they need not, in general, correspond to $\hat{x}$ and $\hat{y}$. But the entire transverse profile E_T can be written in terms of the components along $\hat{x}$ and $\hat{y}$ with different beam waist to account for the ellipticity.

$$E_T(x,y,Z) = \exp\left(-\frac{x^2}{w_x(Z)^2}\right)\exp\left(-\frac{y^2}{w_y(Z)^2}\right) \tag{A.119}$$

We have used $Z = z - \frac{c}{n}t$ to signify that the propagation of the beam in time is implicitly considered in the medium with refractive index n. c is the speed of light in vaccuum.
The longitudinal profile depends on design of the source from which single photons were generated. In general, the longitudinal profile is determined by the narrowest bandpass filter used in the experiment. However, if the design of the source is such that the linewidth of photons is narrower than that of the bandpass filter, then the longitudinal profile is determined by the spectral properties of the source. For the SPDC source, we can consider the longitudinal profile E_L as a Gaussian[14] of appropriate temporal width described by w_z.

$$E_L(Z) = \exp\left(-\frac{z^2}{w_z^2}\right)\exp\left(-\imath(kZ + k\frac{x^2+y^2}{2R(z)}) - \phi_g(z)\right) \tag{A.120}$$

Here,$k = 2\pi n/\lambda$ is the angular wave-number corresponding to the wavelength λ and $R(z)$ is the radius of curvature of the wavefront of the diverging Gaussian beam, which is given

[14] We can consider it as Lorentzian as well

by

$$R(z) = z\left(1+\left(\frac{z_R}{z}\right)^2\right) \tag{A.121}$$

Here, z_R is Rayleigh range described as $z_R = \frac{\pi w_0^2 n}{\lambda}$. The phase $\phi_g(z) = \arctan(z/z_R)$ is known as the Gouy phase. In the experimental set up, the beam waist would be located close to the collimating lens and as the beam traverses it is only diverging. Given that we are already in the paraxial approximation, the beam would interact with the components at $z >> z_R$ and hence the Gouy phase can be assumed to be almost constant. Therefore, Gouy phase would not affect the pointer shift in our setup. However, in weak measurement experiments where lenses are used to focus the beam on the interaction apparatus proper consideration of the Gouy phase may be required. In this experiment, we have used cylindrical lenses after the post-selection and the discussion on the Gouy phase after the lens would be discussed in that context.

Since, the light is not perfectly monochromatic, we need to integrate the wavelength over the bandwidth of the source as well. However, since we have already included the longitudinal temporal profile, the uncertainty in the wavelength or frequency would emerge from the Fourier transform of the temporal profile. Thus, the electric field can be written as

$$E(x,y,z,t) = \mathcal{A}\; E_T(x,y,z)E_L(x,y,z,t) \tag{A.122}$$

$$= \mathcal{A}\; \exp\left(-\frac{x^2}{w_x(Z)^2}\right)\exp\left(-\frac{y^2}{w_y(Z)^2}\right)\exp\left(-\frac{z^2}{w_z^2}\right)\exp\left(-\imath(kZ+k\frac{x^2+y^2}{2R(z)})\right) \tag{A.123}$$

The electric field is a real quantity and hence the complex conjugate is assumed to be added to the above term. In other words, we would be only considering the real part to describe physical object and square complex magnitude to describe probabilities.

Here, $\mathcal{A}$ is the amplitude of the electric field. However, when we consider the case of single photons this can be used as constant which normalizes the probability of the photon being somewhere to be 1. The above description of electric field is in time domain. The measurement process needs to be described as a time integral of the square magnitude

of the field over a time scale in which the photon interacts with the detector (typically about ps-ns). However, this often can be cumbersome to compute. Instead, it is sometimes simpler to consider the frequency/wavelength domain in which we can treat the beam to be broadband. Here, we need to perform the integration over wavelength instead.

$$E(x,y,z,\lambda) = \mathcal{A}\int_{\lambda}\rho_f(\lambda)\,\exp\left(-\frac{x^2}{w_x(z)^2}\right)\exp\left(-\frac{y^2}{w_y(z)^2}\right)\exp\left(-\iota\frac{2\pi}{\lambda}(z+\frac{x^2+y^2}{2R(z)})\right)d\lambda \tag{A.124}$$

Note that here $Z = z - ct$ is replaced with z everywhere as we are not in time domain. Here, $\rho(\lambda)$ is spectral profile of the source. This is a useful description to conceptualize single photons. Since the arrival time from a stochastic source is randomly distributed on an average the temporal profile is uniform. The temporal profile only emerges when we introduces a path delay between two arms of an interferometer. Nevertheless, many a times, it is easier to consider the ensemble of photons with the given distribution of wavelength for the analysis.

The complete electric field vector is described by mentioning the direction of electric field as well. Thus, we have the electric field vector for the initially prepared state as

$$\vec{E} = \mathcal{A}\int_{\lambda}\rho_f(\lambda)\exp\left(-\frac{x^2}{w_x(z)^2}\right)\exp\left(-\frac{y^2}{w_y(z)^2}\right)\exp\left(-\iota\frac{2\pi}{\lambda}(z+\frac{x^2+y^2}{2R(z)})\right)d\lambda\hat{x} \tag{A.125}$$

Here, $\hat{x}$ corresponds to the Horizontal polarization. We can also describe the state in the Jones vector formalism as

$$\vec{E} = \begin{pmatrix} E(x,y,z,\lambda) \\ 0 \end{pmatrix} \tag{A.126}$$

However, care should be taken that in the Jones vector formalism the time evolution of the electric field $\exp(\iota\omega t)$ is usually implicitly assumed and hence the term $\exp\left(-\iota\frac{2\pi}{\lambda}(z+\frac{x^2+y^2}{2R(z)})\right)$ should carefully be used to interpret the divergence of the beam and not temporal oscillation.

With the above description, we have the following parameters which need to be experimentally determined and the errors in the measurement of which would eventually propagate

to the pointer shift in an experiment involving weak measurement.

- $w_x(z), w_y(z)$: The transverse beam waist as a function of z. This can be obtained by capturing the beam profile at few different values of z. This would give us the beam divergence and using λ we can compute the function. The R(z) can also be computed from this data.
- $\rho(\lambda)$: The exact spectrum is usually determined from spectrograph obtained using a spectrometer. In this experiment however, we indirectly determine the same by varying the path length in the interferometry to obtain visibility of the interference pattern. The longitudinal/temporal coherence length can be obtained in this way from which the spectral width can be inferred.
- Non-Gaussian beam: So far we have considered the transverse mode to be Gaussian. But in practice, there is always some deviation from that. In general, there is slight contribution from higher order Laguerre Gaussian modes. Also, dust in various optics adds noise and speckles which makes the wavefront undulated.

The approach here would be to consider the effect of non-idealness in the above parameters on the pointer shift one by one i.e., by considering all other parameters to be ideal. We shall begin by considering the measurement of σ_x weak value in the reflected arm of the interferomter when the transmitted arm is blocked. To recall, the pre-selected polarization state would be vertical. The post-selected state is obtained by rotating the half-wave plate placed before a polarizing beam splitter. Thus, we have the pre- and post-selected state as $|\psi\rangle = |V\rangle$ and $|\phi\rangle = S(\theta)|H\rangle$ respectively. Here, $S_H(\theta)$ is the Jones matrix for the half-wave plate.

To begin with, we shall consider the weak interaction with the σ_x operator is ideally realized by the experimental components. In practice, the realization also depends on the beam divergence and wavelength. These shall be included in the subsequent sections. But for now, we shall presume that the σ_x component shifts the diagonal $|+\rangle$ state by an amount a along the horizontal direction and it shifts the anti-diagonal $|-\rangle$ polarization component by an amount $-a$.

Effect of beam size

We shall assume that the divergence of the beam is minimal i.e. the propagation distance from the pre-selection and post-selection and subsequent measurement is small enough so that the size of the beam can be assumed to be constant.

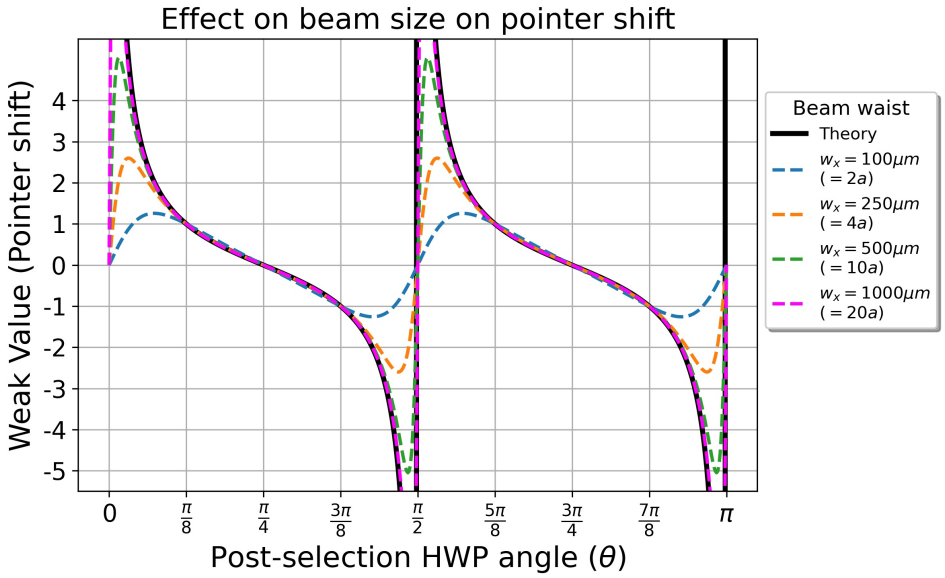


Figure A.12: The displacement caused by the σ_x interaction is set to $a = 50\mu m$. When beam size w_x is comparable to a, we barely have any weak value amplification as the weak value is almost within the eigenvalue range i.e., between -1 and 1. As, we increase the beam size, the ratio a/w_x approaches zero. Thus, the pointer shift obtained becomes closer to the weak value which is the ideal case in the limit $\lim a/w_x \to 0$.

Typically, when the beam size is more than 20 times, the displacement, the pointer shift is almost identical to the weak value if we consider the eigen spectrum in the range $(-5,5)$. Of course, the mismatch would be noticeable on a larger range simply because the weak value for this setup tends towards infinity when pre and post- selection becomes orthogonal whereas the pointer shift would always be finite.

Effect of beam divergence

Having beam divergence also implies that the beam size also increases with distance. We assume that the weak measurement setup is small enough that the beam size at the detection place is same as the beam size in the plane of weak interaction. In practice, this is not the case, but can be a good approximation if the divergence of the beam is small and

the weak measurement is made at a distance z from the waist of the beam. In the paraxial approximation, the divergence of the beam depends on the beam size at the waist $\theta_d = \frac{\lambda}{\pi w_0}$.

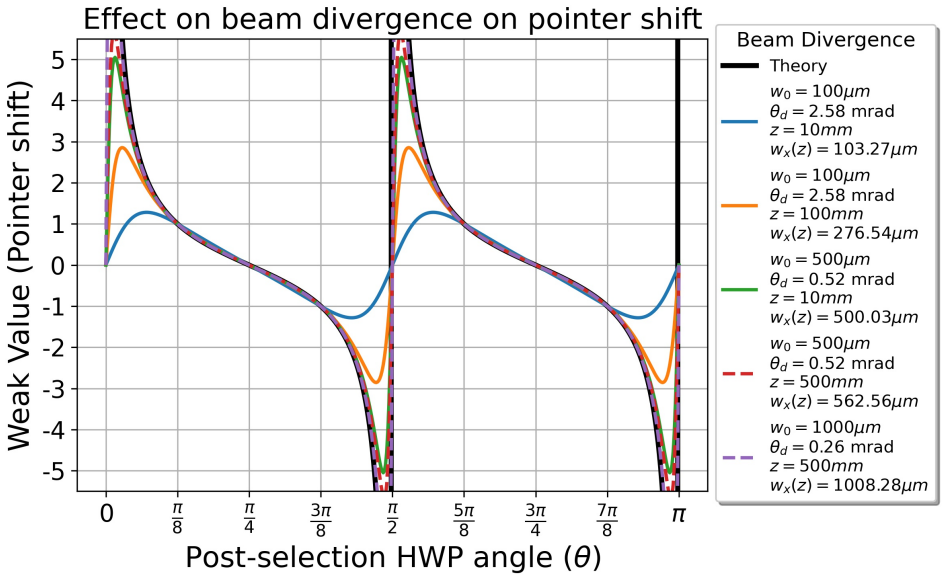


Figure A.13: When the divergence is large around 2.5 mrad for the given waist of $100\mu m$, the longer the distance of the setup from the waist makes the pointer shift a better indicator of weak value. Even, when the divergence is same, for a higher z with lower $a/w_x(z)$, we have a better match of the pointer shift with the weak value.

Thus, it seems, the effect of divergence only affects the pointer shift by indirectly changing the beam size. It is important to note that although we have considered divergence, in the paraxial approximation, we have not considered any distribution of k-vectors. As long as the distribution is symmetric about the considered propagation direction the centroid of the beam should not be affected due to the symmetry of k vector distribution. However, if there is a small angle between two interfering beams, the symmetry would be broken and the distribution of k vectors may play a role.

H Mechanical Stability of the Mach Zehnder Interferometer

While claiming the centroid shift as a signature of the Quantum Cheshire Cat effect, it is important to establish that centroid shifts due to the intended displacements. For example, in order to claim the centroid shift as the signature of the photon being in arm A, it is

imperative to establish that the shift is only obtained when the glass plate is tilted and not otherwise. However, due to drift of beam over time, the centroid of the resultant beam shifts also shifts over time. These shifts should be much less than 50μm in so that noise is much below the desired shift.

The second beam-splitter BS2 in the Fig. A.14 below is mounted on glued to the post instead of it being on a kinematic mount. Kinematic mounts are often sensitive to temperature and the screws relax over time. Thus avoiding the tip-tilt degree of freedom reduces the possibility of drift in those degrees of freedom. But, the alignment procedure becomes cumbersome. All the 4 mirrors used in the Mach-Zehnder interferometer are of Polaris and hence thermal drift is expected to be minimal. The corner cube, however, is mounted on a 3-axis stage which might drift. This will make the reflected beam after the corner cube to have translation drift. The first beam splitter BS1 however is mounted on a kinematic platform mount and hence the reflected arm might have angular drifts as well. The Mach Zehnder interferomenter is enclosed from sides (not top) to avoid air currents and thus make the temperature variation much slower.

H.1 Drift of Centroid of Individual Beams

The centroid of the beams in each arm are measured and we aim to estimate the centroid shift of the resultant interference that would be caused due to drifts of the individual beams.

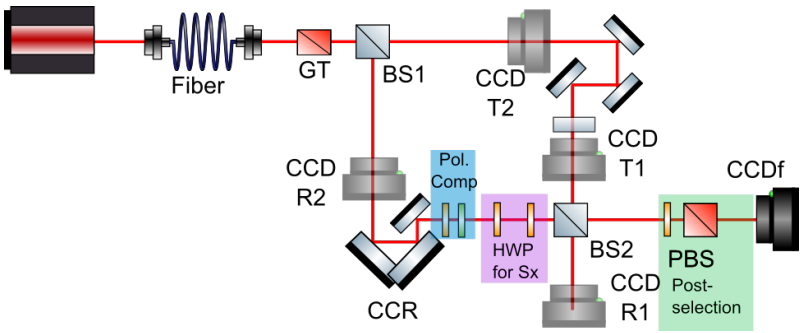


Figure A.14: Two CCD beam profilers were used to compare the beam drifts in the transmmitted and reflected arm respectively. In the first run, the CCD in the transmitted arm was placed in the position marked T1. Since the beam from the transmitted arm is blocked by the CCDT1, the reflected arm could be monitored by placing the CCD at R1. The intensity is lower there and hence the exposure is appropriately adjusted. This experiment gives us the drift of the centroid approximately at the second beam splitter BS2 which could be due to mirrors and corner cube along with the deviation caused due to the first beam splitter BS1. To isolate the effect due to the first beam splitter BS1, two CCDs were placed at position marked T2 and R2 respectively. Based on the observed drift, the shift in due to interference at the CCDf is estimated.

Beam Drift measured near the second beam splitter

The CCD in the transmitted arm was placed in the position marked T1. Since the beam from the transmitted arm is blocked by the CCDT1, the reflected arm could be monitored by placing the CCD at R1. The centroid here captures the the drift due to the stage as well as pointing fluctuations due to various reflections. The temperature logger was placed near the second beam splitter in order to monitor local temperature fluctuations.

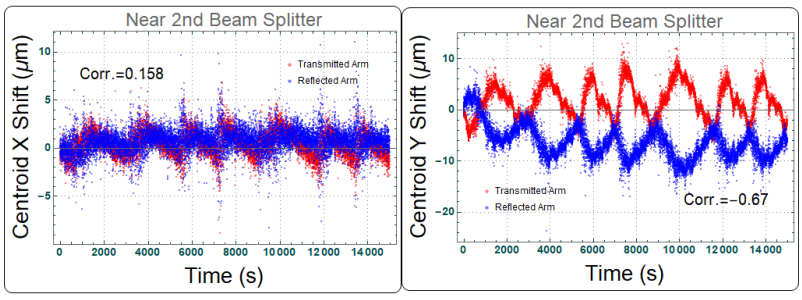


Figure A.15: [Left] Horizontal centroid as a function of time. The drifts are small but the shifts are weakly correlated. But because of reflection in the final beam splitter for one CCD this practically translated to weak anti-correlation. [Right] The vertical shifts are clearly anti-correlated. Since these are from independent arms and pass through different optical components the source of drift has to be either temperature or first beam splitter itself.

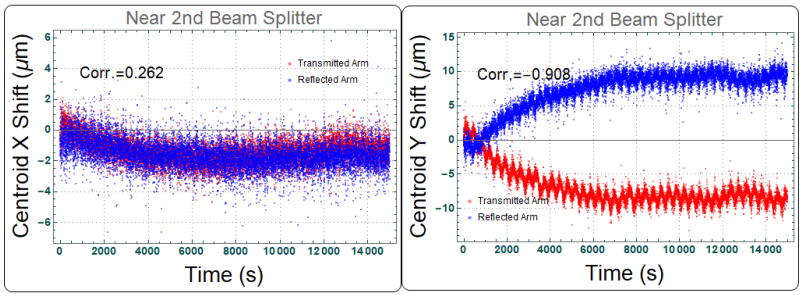


Figure A.16: Another run of the experiment. This is late night (2-6 am). We do not see the oscillations.

Beam Drift measured near the Mirror/CCR

The beam profiler was placed in the position marked T2 and R2 respectively. The only components that could affect the beam drift is the BS1. The transmitted beam is expected to be less affected compared to the reflecting beam if the BS1 mount is relaxing.

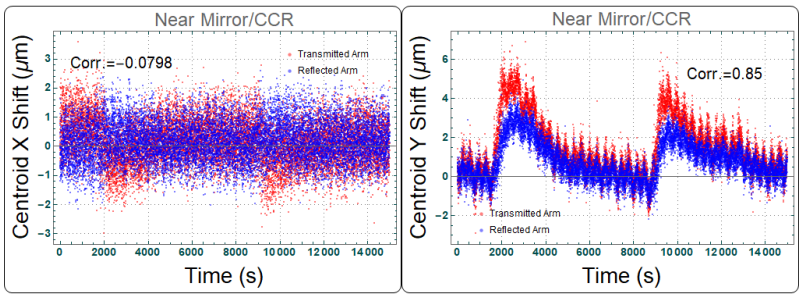


Figure A.17: [Left] The horizontal is pretty stable with very tiny anti-correlation [Right] The vertical showed occasional jumps but this time the drifts in both the arms are correlated.

Since the vertical drift in both the arms are correlated, the cause cannot be attributed to the beam splitter mount. The drift therefore must be occurring before that. Thus, the short Polarization maintaining fiber and Glan Thompson must be the source of the drift. Switching to the longer fiber or manually changing the stress or temperature of the fiber may help establish the cause.

H.2 Calculation of the effect on the centroid of the resultant beam

In the following calculation, we shall compute the effect of beam drift on the centroid of the resultant interference pattern. For simplicity, it is assumed that the drift is entirely due to transverse displacement and not due to any angular deviation. The effect of angular deviation shall be captured in calculation in the next section. In this section, we shall

assume that the two beams are collinear i.e., have same propagation vector after the second beam splitter of the Mach Zehnder Interferometer.
Let the two beams be denoted by the Gaussians g_A and g_B.

$$g_A(x,y) = \mathcal{A}_A \exp\left(-\frac{(x-x_{A_0})^2}{4\sigma_{x_A}^2}\right)\exp\left(-\frac{(y-y_{A_0})^2}{4\sigma_{y_A}^2}\right) \tag{A.127}$$

$$g_B(x,y) = \mathcal{A}_B \exp\left(-\frac{(x-x_{B_0})^2}{4\sigma_{x_B}^2}\right)\exp\left(-\frac{(y-y_{A_0})^2}{4\sigma_{y_A}^2}\right) \tag{A.128}$$

We assume that the phase difference between the two arms of the interferometer is maintained at 0. Initially, the polarization of Gaussian beam g_A in arm A is $|H\rangle$ and that of the be beam in arm B is $|V\rangle$. After the half-wave plate oriented at an angle α and the polarizing beam splitter, the amplitudes of the beams g_A and g_B that are successfully post-selected would be weighed by $\cos(2\alpha)$ and $\sin(2\alpha)$ respectively. Therefore, the resultant field at one of the output port of the second 50:50 beam-splitter is given as follows:

$$I_d = \left|\frac{1}{\sqrt{2}}\left(\cos(2\alpha)g_A(x,y) + \sin(2\alpha)g_B(x,y)\right)\right|^2 \tag{A.129}$$

The horizontal and vertical centroid of the resultant intensity distribution is obtained by the following expressions:

$$c_x(x_{A_0}, x_{B_0}, y_{A_0}, y_{B_0}) = \frac{\iint_{-\infty}^{\infty} dx\, dy\, x\, I_d(x,y)}{\iint_{-\infty}^{\infty} dx\, dy\, I_d(x,y)} \tag{A.130}$$

$$c_y(\alpha; x_{A_0}, x_{B_0}, y_{A_0}, y_{B_0}) = \frac{\iint_{-\infty}^{\infty} dx\, dy\, y\, I_d(x,y)}{\alpha; \iint_{-\infty}^{\infty} dx\, dy\, I_d(x,y)} \tag{A.131}$$

After the cylindrical lens, we can only observe the horizontal centroid c_x. Because of size of the multi-mode fiber, we in principle would need perform the integral piecewise first to obtain the discreet intensity values and then to obtain the centroid. Nevertheless, in this section we shall assume a perfect detector and simulate how the horizontal centroid would change as a function of time.

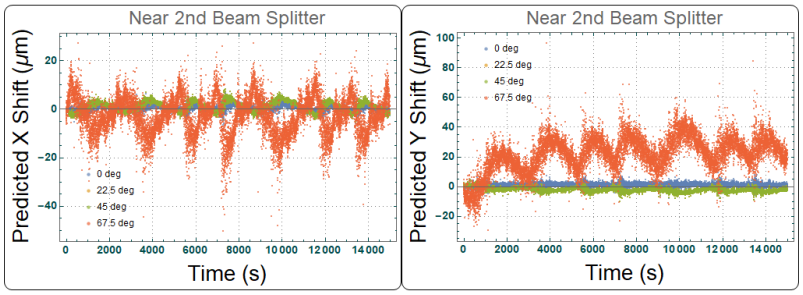


Figure A.18: The centroid of the beams in individual arm was used to calculate the expected change in centroid of the post-selected beam for the four (post-selection) HWP angles. The Horizontal and vertical shifts are about 20 and 50 μm respectively near destructive interference.

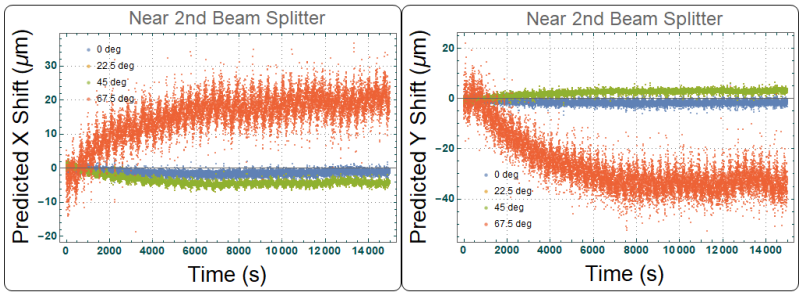


Figure A.19: Predicted centroid for another data set.

Therefore, over long period of time (few hours) it is important to ensure that the beam drift is not substantial. Since the drift and periodic cycles have correlations with the temperature of the lab, the data acquisition was done during the periods where the temperature changes are minimal.

I Effect of incoherence on weak values and cheshire cat effect

I.1 Effect of Incoherence on weak values

Let's consider the path difference between the two arm of the Mach-Zehnder Interferomter $2\delta L$ is much larger than the coherence length of the source so that we can treat incoherent combination at the second beam splitter. Here, we shall compute the centroid shift as a function of θ - the angle of the half-wave plate that determines the post-selected state $|\phi\rangle(\theta)$. We shall assume the condition that the glass plate, causing Horizontal displacement

is placed in the transmitted arm and the beam displacer, causing vertical displacement, is placed in the reflected arm.

The component beam emerging at the concerned port of the second beam splitter can be written as the following gaussian beam.

$$\frac{1}{2}g_T(x-x_{(GP)},y)\,|H\rangle \tag{A.132}$$

The component of the beam after in the reflected arm after the first half-wave plate that forms the σ_x interaction is given by

$$\frac{1}{\sqrt{2}}g_R(x,y)\,|A\rangle \tag{A.133}$$

The beam displacer vertically displaces the $|V\rangle$ polarization component and thus we have

$$\frac{1}{2}(g_R(x,y)\,|H\rangle - g_R(x,y-y_{(BD)})\,|V\rangle) \tag{A.134}$$

After the last half-wave plate, the component of the beam that emerges at the concerned port of the beam splitter is given by

$$\frac{1}{2\sqrt{2}}(g_R(x,y)\,|D\rangle - g_R(x,y-y_{(BD)})\,|A\rangle) \tag{A.135}$$

Since the source is considered incoherent, there is no interference at the second beam splitter or the polarizing beam splitter used for post-selection. Therefore, we shall propagate the component of the beams till the detector and then incoherently combine them.

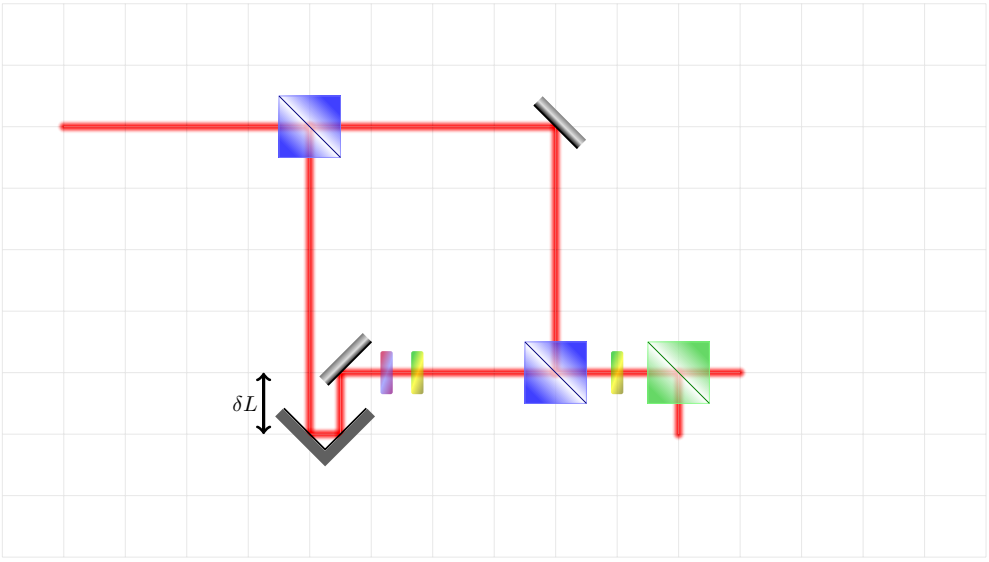


Figure A.20: Macroscopic Path difference in Mach Zehnder interferometer. If δL compares with coherence length of the source, the visibility would decrease and weak values would not get amplified as much

After the post-selection half-wave plate, the component from transmitted arm of the MZI that gets transmitted through the PBS is given by

$$\frac{1}{2} g_T(x - x_{(GP)}, y)\cos(2\theta) \tag{A.136}$$

Similarly, the component from the reflected arm that reaches the detector is given by

$$\frac{1}{4}(g_R(x,y)(\cos(2\theta)+\sin(2\theta)) + g_R(x, y - y_{(BD)})(\cos(2\theta)-\sin(2\theta))) \tag{A.137}$$

The resultant centroid shift is due to the incoherent combination of the above two expressions A.136 and A.137. This is compared with the coherent combination assuming the zero phase difference between the two arms of the MZI.

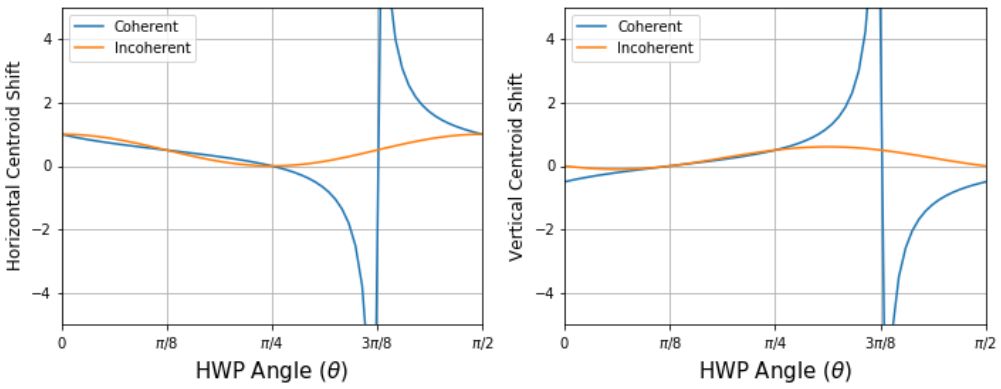


Figure A.21: Comparision of shift of horizontal/vertical centroid in presence of both glass plate and beam displacer. Here, $x_{(GP)} = y_{(BD)} = 1$

- For a considerable range of the half-wave plate angles (from 0 to $\pi/4$) away from the amplification zone, the centroid shift obtained with an incoherent source is almost similar in values to what one would obtain with the coherent source. Therefore, if one performs the experiment with incoherent source for the half wave plate angles in the range between 0 to $\pi/4$ one would not be able to infer based on the centroid shift whether the shift is due to interference or incoherent combination. Thus, one could infer the partial Cheshire cat effect with incoherent source if one considers only centroid shifts in the region with no amplification. Note that the centroid shift with the incoherent source is between 0 and 1 and thus is bounded between the eigenvalue.
- The distinction between coherent and incoherent source however can be made when we consider the normalized power integrated over all detector positions. With coherent source, we have sinosoidal variation where as with the incoherent source, we have a flat line.

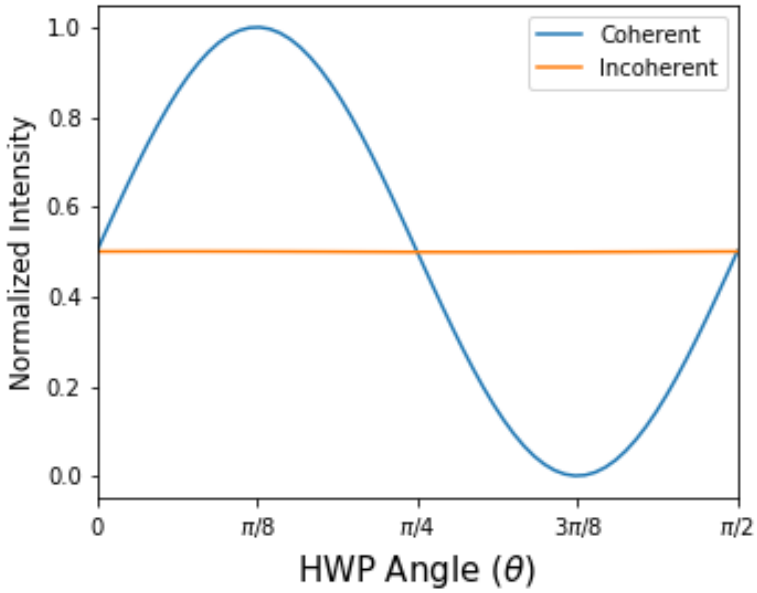


Figure A.22: Comparison of intensity at the detector for coherent and incoherent source

Therefore, it is important to verify that the path difference between the two arms of the Mach Zehnder interferometer is less than the coherence length of the single photon source.

J Stability of photon source and Statistics of centre over Datasets

J.1 Stability of the source

For the partial chehsire cat experiment with sngle photons as a function of post-selection angle we have two data sets. The duration over which the experiment was performed was few hours. As the detector scans through the profile of the beam, it is important that source remains stable so that no bias is introduced in computing the centre of the beam profile. For this, we monitor the count rate reference heralding arm of the single photons and normalize the signal photons that went through the partial cheshire cat setup.

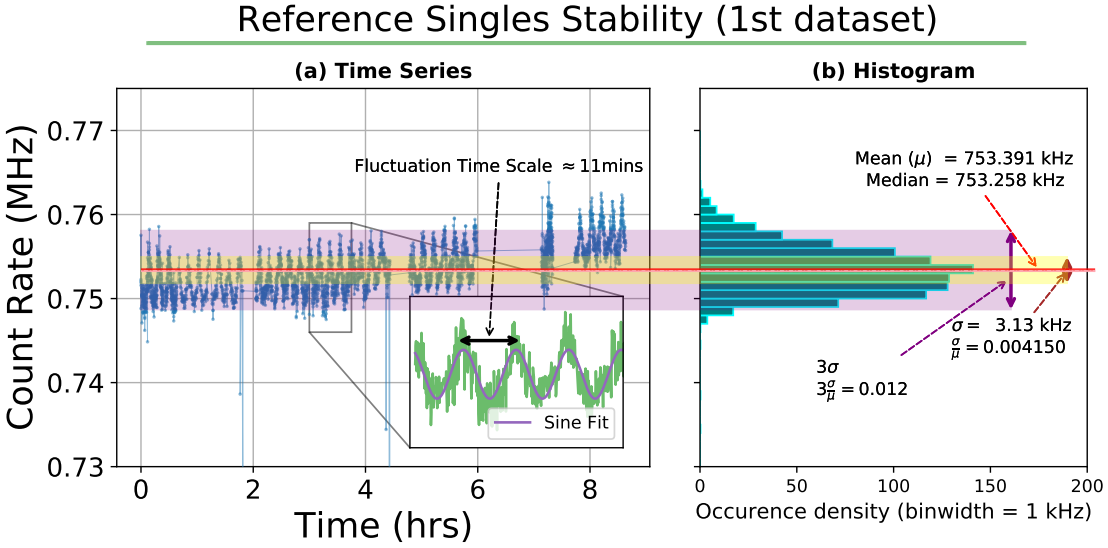


Figure A.23

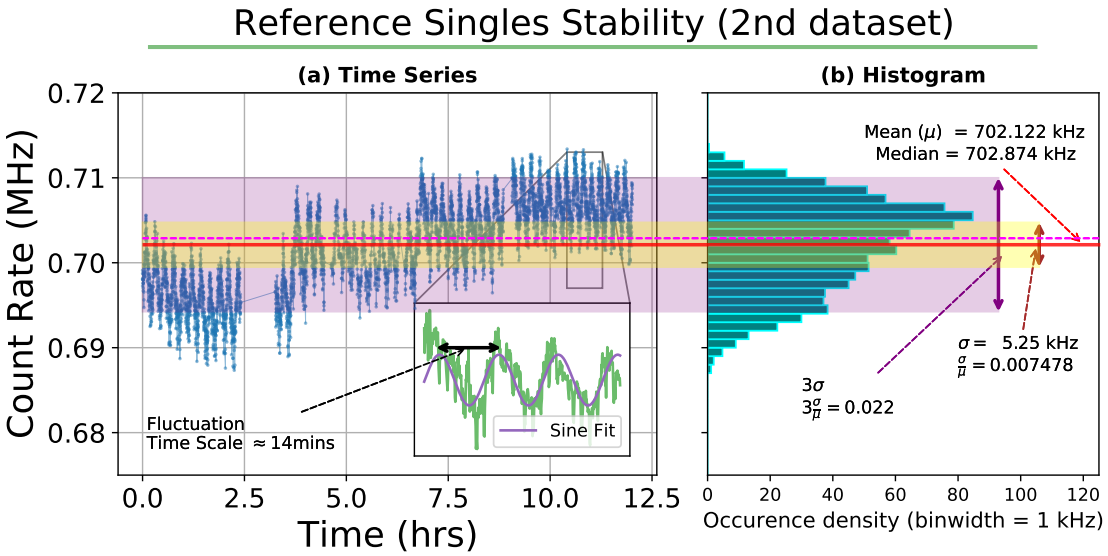


Figure A.24

The count rate fluctuates with periodicity of 11-14 mins, which correlates with temperature fluctuation due to the air-conditioning unit periodically turning on and off. The relative magnitude (σ/μ) of the fluctuation however is small in comparision to the statistical fluctuation of signal photons photon numbers. Nevertheless, we correct for these fluctuations while computing the centroid.

J.2 Centroid Comparison

For gaussian beam profiles, the second order centroid defined as

$$C_n = \frac{\int x^n \rho(x) dx}{\int \rho(x) dx} \quad n = 2 \tag{A.138}$$

Similarly, for $n = 1$ or 3, we have the first and third order centroid for the intensity distribution $\rho(x)$.

If we generate an array of random numbers with normal distribution and fit the histogram with a Gaussian, the obtained gaussian centre is the same as the computed second order centroid.

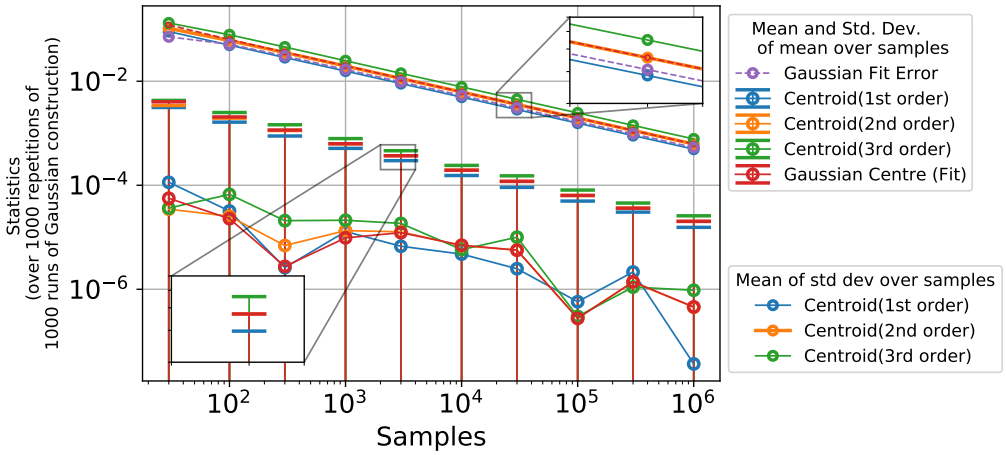


Figure A.25: **Simulation** The number of photons (samples) that generate the Gaussian is varied. It is repeated over R = 1000 times. We take mean of std dev over samples and compare it with std dev of means.

Thus, for all cases, we report the centre obtained for the Gaussian fit as the centre and use it interchangeably with the second order centroid. However, when the beam shape is not Gaussian, i.e., in the amplification region, we these two may defer, We report the Gaussian centre for those regions nevertheless.

J.3 Statistics of centre over various datasets

During data acquisition, if we collect one reading of the photon counts for a position of the detector and move it so that we reconstruct the beam profile, we call it as one scan of the

detector. However, when the detector is made to sit at a poistion collecting few readings and then move, the errors in the statistics that one would obtain would be different.

We compare the both cases in the two figurs below:

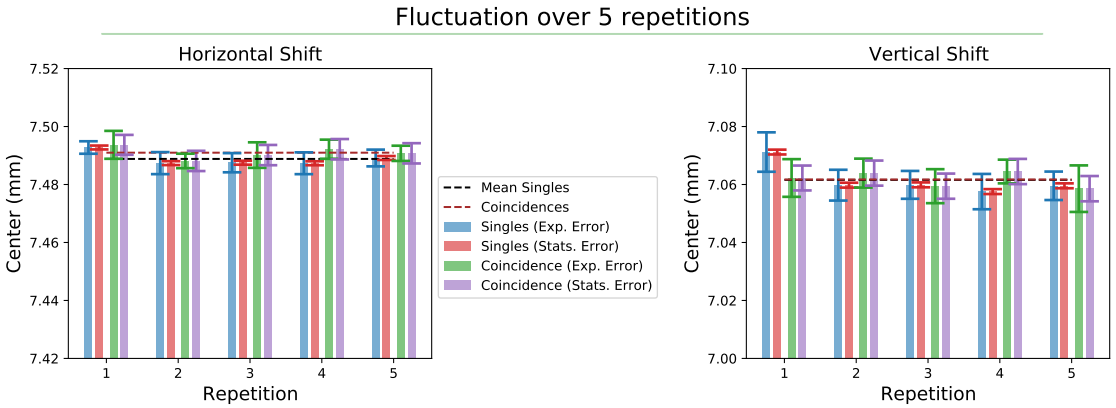


Figure A.26: **PCC Region Data** With a single scan, 5 times readings are taken. The statistical uncertainty is compared with the experimentally obtained error in the Gaussian fit. Most values within the error bar. For singles, uncertainty is not limited by statistics. But for coincidences they are comparable.

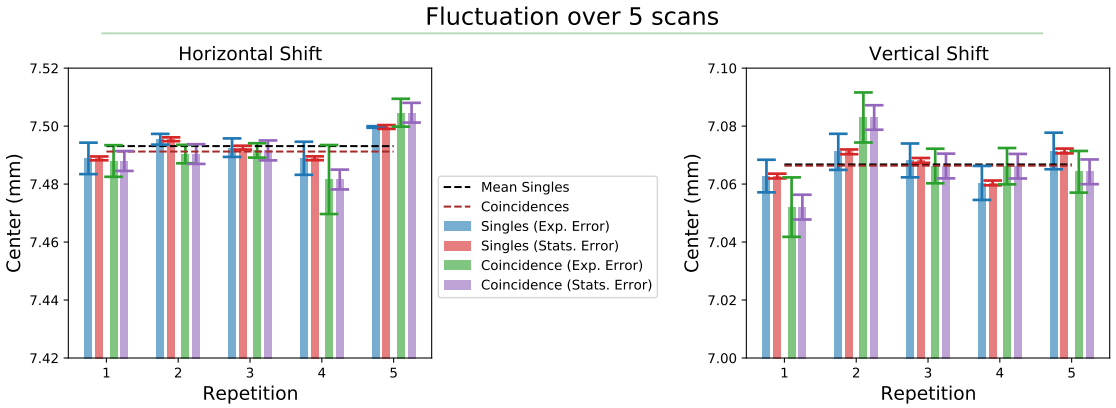


Figure A.27: **PCC Region Data** 5 times we scan the profile each time taking only one reading . The statistical uncertainty is compared with the experimentally obtained error in the Gaussian fit

Whether, we can combine the datasets over various scans or repetitions depends on whether the systematic errors, if present, gets randomized over different scans or repetitions. The following combination of statistics indicates that upon repeatiion or scan the readings for centroid computed from singles or coincidences are often within one standard deviation.

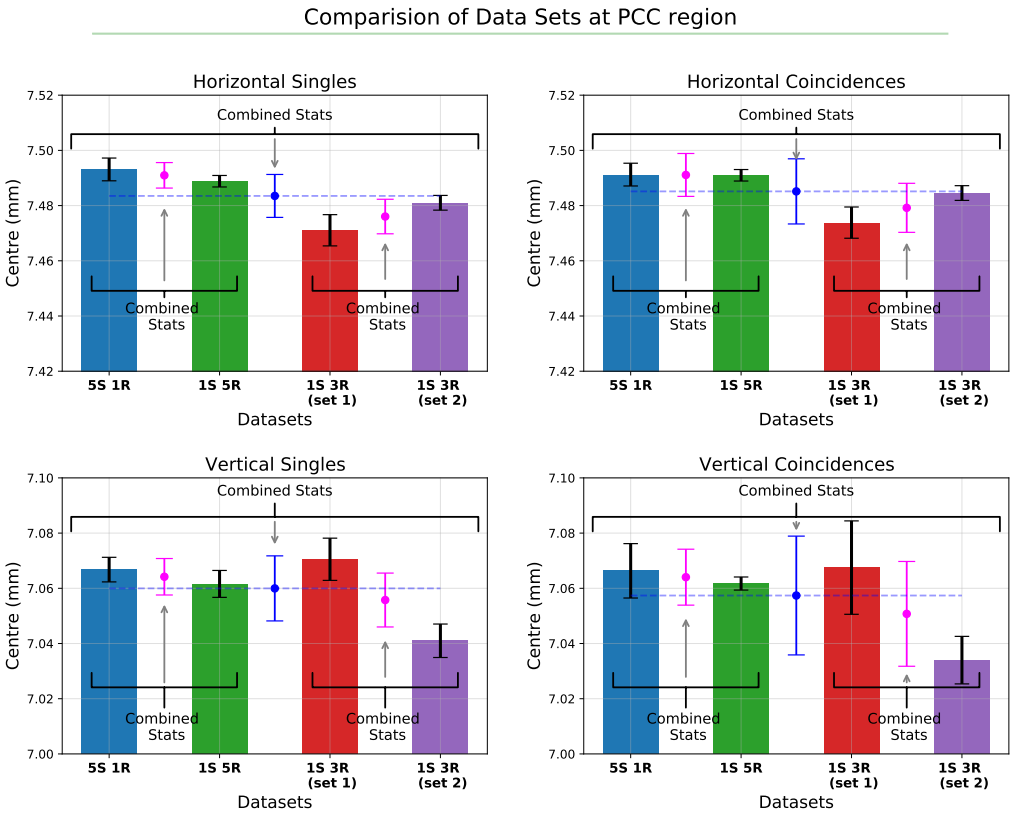


Figure A.28: Comparison of the stats in PCC region. Additionally, two scans with 3 repetitions are considered. The combined mean and RMS error (after finding the dominant contribution) is shown.

K Comment on other experiments on superposition principle

Around the same time when the microwave experiment was completed, another [A.10] group had reported the observation of non-zero κ in a single-photon version of the experiment. The parameters used in that experiment is as follows:

Wavelength (λ)	810 nm	680 nm
Slit Width (w)	200 nm	215 nm
Inter Slit distance (p)	4.6 μm	3.8 μm
Slit Thickness (t)	110 nm	100 nm
Slit Height (h)	100 μm	not mentioned (assumed ∞)
Beam Spot Size	400 nm[15] 15 μm[16]	400 nm

Summary of the Claims

- *Intensity distribution of far field intensity distribution is wavelength dependent(intensity oscillates vs wavelength) in presence of exotic loop paths and the dependence on wavelength is smooth when loop trajectories are absent*

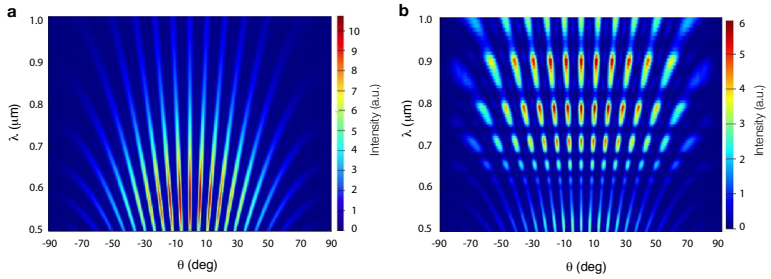


Figure A.29: Source: [A.10] The left figure (a) is when the incident polarization is along the long axis and the right figure (b) is for when we have polarization along the short axis. The simulation parameters involve plane linearly polarized wave incident on slits(w = 150 nm and p = 4.6 μm, $h = \infty$). The interference pattern likely have been obtained after 2D near to far field transformation

  - They claim that if polarization is along the width(short axis), this increases the contribution of surface plasmon enhanced loop trajectories.

- *Polarization along the long axis decreases the visibility where as polarization along the short axis increases the visibility*

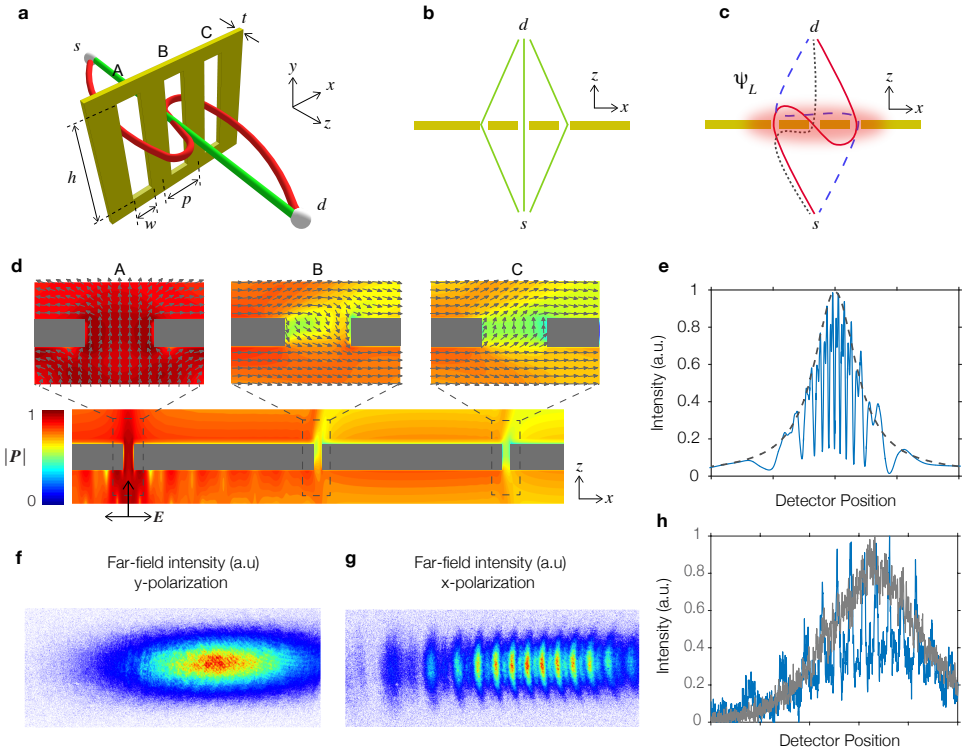


Figure A.30: Source [A.10] : It was claimed that "(f) Experimental evidence that shows the far-field pattern for a situation in which only one slit is illuminated with y-polarized light from an attenuated laser. (g) The presence of exotic looped trajectories leads to an increase in the visibility of the far-field pattern. This effect is observed when x-polarized light illuminates one of the slits"

– *Modelling with plasmons gives a better match with experiment*

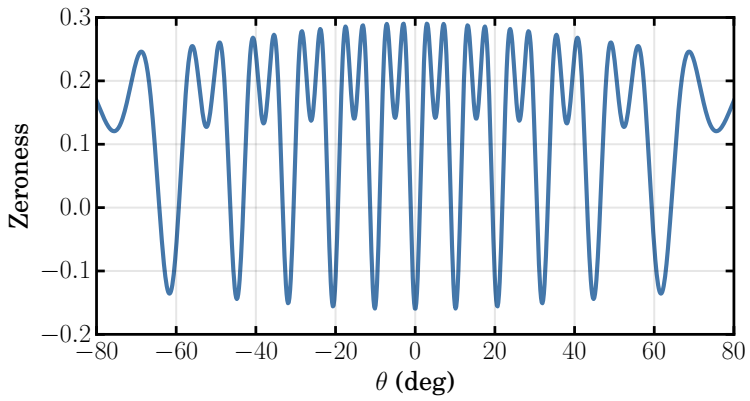
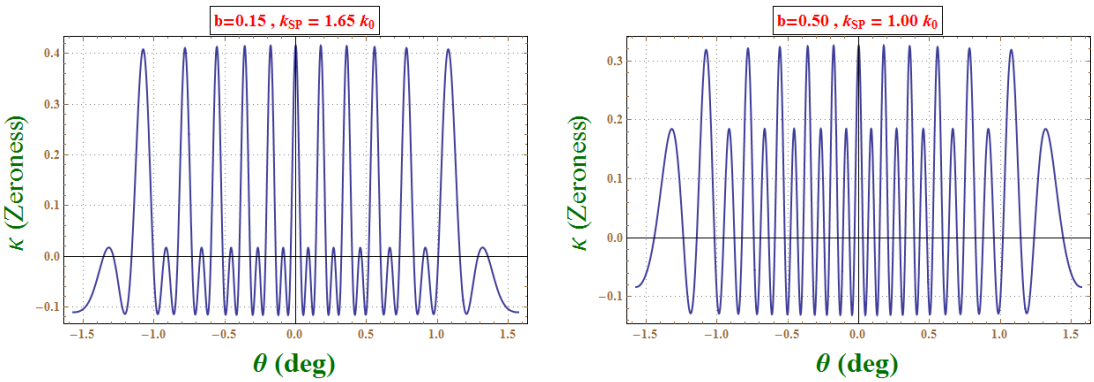


Figure A.31: Source [A.10] κ from the model which includes plasmons in path integrals

The effect of polarization on interference pattern of slits is understood. The origin of non-zero κ comes from geometry effects although polarization and plasmons may

enhance the same. Next, we reproduce their result and show that the modelling with plasmons does not appreciable changes κ. Further, we show that the experimentally reported κ which are reported on the peak of the oscillations can occur due to experimental imperfections in the various slits.

- To model the plasmons they have had a different weight multiplied to the hugging paths (via the parameter b) and also have a different wave vector for surface plasmons k_SP.
- In their work, they have used parameters $k_SP = 1.65k_0$ and $b = 0.15$. So to remove the effect of surface plasmons, we would have to use $k_SP = 1.65k_0$ and $b = 0.5$. We compare the results below



- We find that modeling of plasmons does not appreciably changes the order of kappa.

K.1 Reproducing the FDTD simulations results

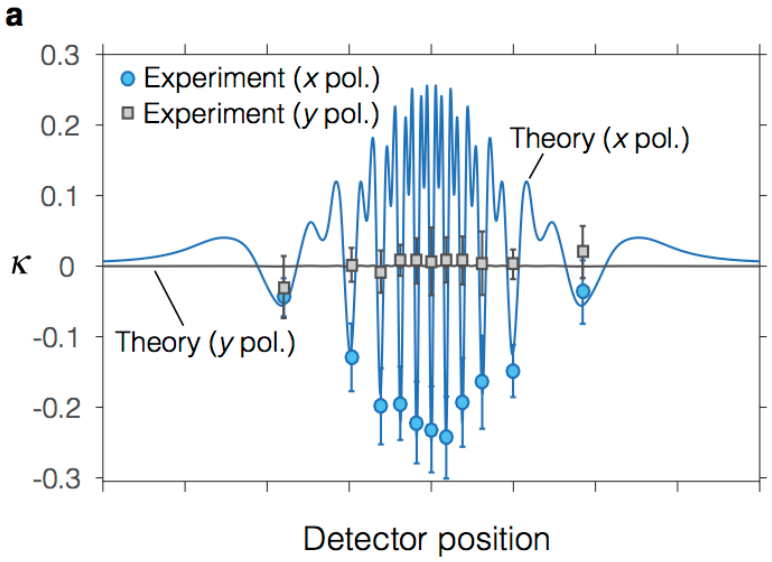
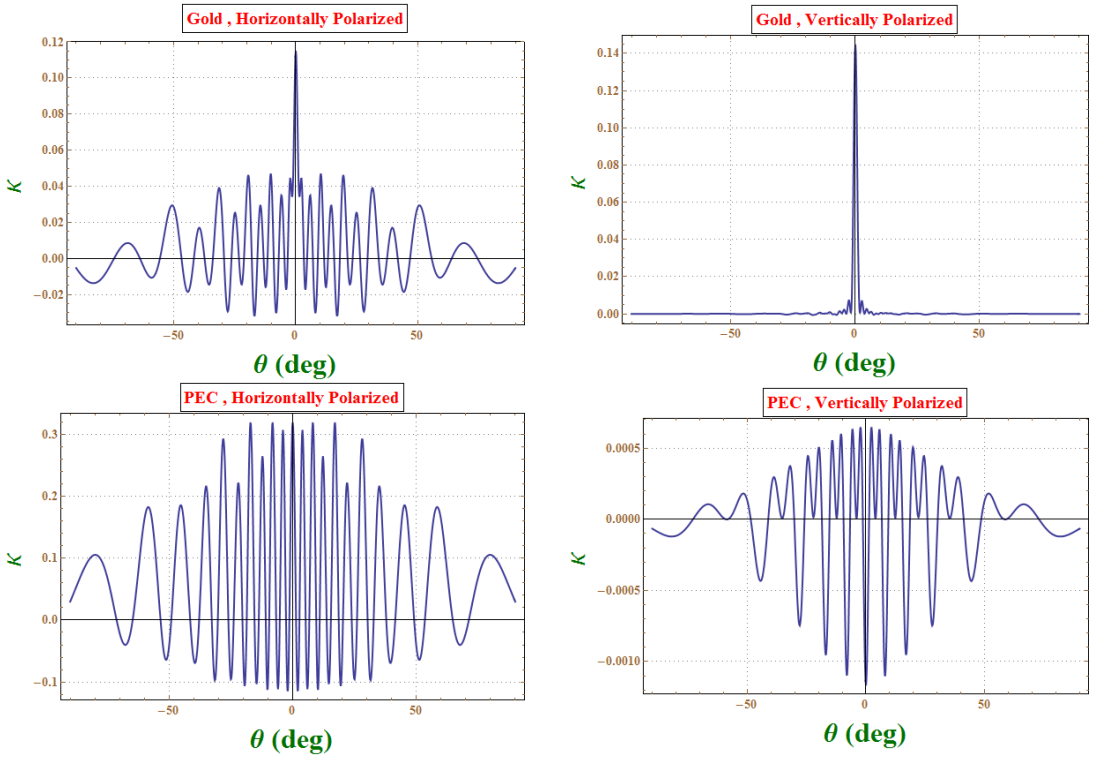


Figure A.32: Source [A.10] : Expt and Theory comparision.



- The vertically polarized PEC FDTD simulation matches the modulation provided. But its magnitude is less by two orders.

- The PEC horizontally polarized matches the order with [A.10], but modulation is different, we have a central peak.
- Gold with Horizontally polarized light also shows a central peak.
- In any case, PEC with horizontal polarization produces kappa of about 0.3 and hence the fact that "Gold enhances kappa" does not seem to be a necessary argument when compared with metals. But when compared with slits are made with absorbing materials, Gold or any metal would enhance the effect.

K.2 Estimation of Classical “error" Kappa from Slit Images

- In experiment in [A.10] they seem to have all the seven combinations on a plate that slides to have different slit combinations. It is unclear what exactly are the error associated with the beam spot size. Thus, here we will assume that the source remains fixed for cases, which is the best case scenario.
- First of all to rule out a false kappa due to a false alignment
  - One need to make sure that as the plate slides the slit A in combination ABC is exactly at the same place as in combination A. From the images of slit they have provided it does not seem so (see the images below). However, let us even assume that this is an imaging error (in translation).
  - Assuming alignment is perfect now we need that all the respective slits needs to be identical in various combinations. That is, the interslit distance between C and A in combination CA and in combination ABC should be same. Also the slit width of A should be same for all combinations A, AB , ABC and CA.



Figure A.33: ABC slit combination



Figure A.34: CA slit combination



Figure A.35: BC slit combination



Figure A.36: CA slit combination subtracted from ABC slit combination

- In the bottom right image, we first binarize the image and then translate CA image on top of ABC image such that A combination matches the best in both images. Still, we see there is a spot left (This is the best case i.e. translating any pixels left or right introduces more difference in area). We can see that this is due to the slight notches we see in the slits (i.e they are not perfect rectangles). Also, note that C is mismatched stating that the inter-slit distance is not precise.
- To estimate how much the above two non-idealness would contribute to kappa, we did the following
  - All the images were translated such that they have best match with ABC. For AB combination BC was translated and for A and C combination image B was translated. Thus, this is a better situation than what would be in the experiment.
  - Using a horizontal slice of the above images we compute far field interference pattern and compute kappa from there.
  - We see that different slices give different modulation to kappa and the order of 0.1 (greater than the order of far field kappa using path integral only) is typical.

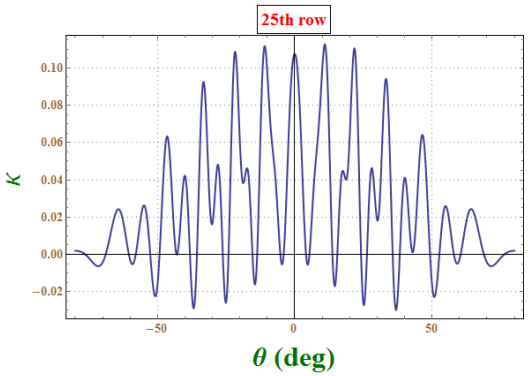


Figure A.37: from 25th row of image

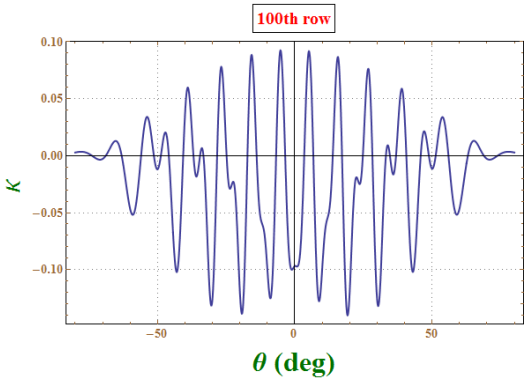


Figure A.38: from 100th row of image

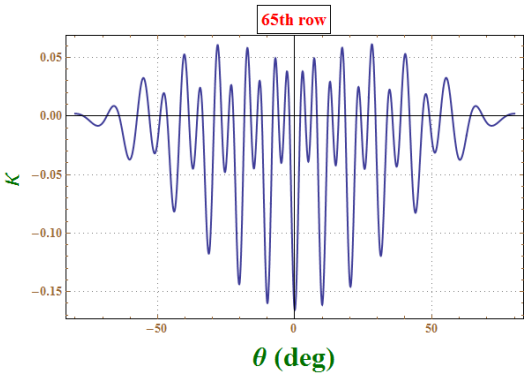


Figure A.39: from 65th row of image

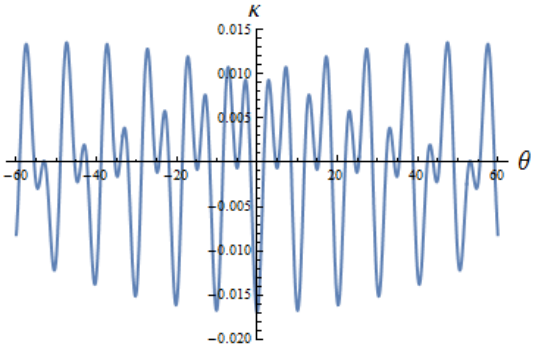


Figure A.40: Analytical Kappa

- From the above analysis, kappa of the order of 0.1 can be easily explained by systematic experimental errors.

Also, although the value of kappa as a function of position could be computed for all positions, only the peaks are reported. This, along with the possibility of systematic errors, do not rule out that the observation could be due to systematic errors.

References

[A.1] Yosep Kim et al. “Observing the “quantum Cheshire cat” effect with noninvasive weak measurement”. In: (2020). DOI: 10.1038/s41534-020-00350-6.

[A.2] Yosep Kim et al. “Observing the quantum Cheshire cat effect with noninvasive weak measurement”. In: *npj Quantum Information* 7.1 (2021), p. 13. ISSN: 2056-6387. DOI: 10.1038/s41534-020-00350-6.

[A.3] Yakir Aharonov et al. “Quantum Cheshire Cats”. In: *New Journal of Physics* 15.11 (2013), p. 113015. DOI: https://doi.org/10.1088/1367-2630/15/11/113015.

[A.4] Tobias Denkmayr et al. “Observation of a quantum Cheshire Cat in a matter-wave interferometer experiment”. In: *Nature Communications* 5.4492 (2014).

[A.5] James M. Ashby, Peter D. Schwarz, and Maximilian Schlosshauer. “Observation of the quantum paradox of separation of a single photon from one of its properties”. In: *Phys. Rev. A* 94 (1 2016), p. 012102. DOI: 10.1103/PhysRevA.94.012102.

[A.6] Q. Duprey et al. “The Quantum Cheshire Cat effect: Theoretical basis and observational implications”. In: *Annals of Physics* 391 (2018), pp. 1–15. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2018.01.011.

[A.7] A Matzkin and A K Pan. “Three-box paradox and ‘Cheshire cat grin’: the case of spin-1 atoms”. In: *Journal of Physics A: Mathematical and Theoretical* 46.31 (2013), p. 315307. DOI: 10.1088/1751-8113/46/31/315307.

[A.8] G. J. Pryde et al. “Measuring a Photonic Qubit without Destroying It”. In: *Phys. Rev. Lett.* 92 (19 2004), p. 190402. DOI: 10.1103/PhysRevLett.92.190402.

[A.9] G. J. Pryde et al. “Measurement of Quantum Weak Values of Photon Polarization”. In: *Phys. Rev. Lett.* 94 (22 2005), p. 220405. DOI: 10.1103/PhysRevLett.94.220405.

[A.10] Omar S Magaña-Loaiza et al. “Exotic looped trajectories of photons in three-slit interference”. In: *Nature Communications* 7.1 (2016). DOI: 10.1038/ncomms13987.