2013-01-15

Probing the Completeness of Quantum Theory with Entangled Photons

Stuart, Terence Everett

UNIVERSITY OF CALGARY

Probing the Completeness of Quantum Theory

with Entangled Photons

by

Terence Stuart

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PHYSICS AND ASTRONOMY

CALGARY, ALBERTA

January, 2013

Abstract

Quantum theory provides a way to describe the behaviour of matter and energy on the nano-scale, but its predictions are sometimes very surprising. Intuition, based on experience from our daily lives, suggests there are a set of assumptions that should be applicable even when describing quantum systems. In this thesis we conduct a sequence of experiments that test some alternative theories to quantum theory and, in doing so, show that we must abandon some assumptions we made based on intuition if we are to accurately describe nature. In this thesis we also conduct an experiment that bounds the maximum predictive power that theories may have if they are to describe nature at the quantum level accurately. These experiments were conducted using a novel source of polarization entangled photon pairs, whose construction and characterization are also discussed.

Acknowledgements

First, I would like to thank the professors, teaching assistants, and my fellow students who have helped me through my course-work, those who TA'd alongside me, and all those who have been good friends. You are too numerous to list here, but I thank you. I thank the Department of Physics and IQIS (now IQST) support staff who have helped me. In particular, I would like to thank Lucia Wang, Nancy Lu, Leslie Holmes, and Tracy Korsgaard. Although our lab is well equipped and funded, we have also received extremely helpful loans of equipment from Pat Irwin, who runs the senior undergraduate labs, and Peter Gimby of the junior undergraduate labs. I also thank Renato Renner and Roger Colbeck, who collaborated with us on the work discussed in Chapter 5.

I would like to thank the past and present QC2 students and post-docs for all the help, patience, and friendship they've given me: Ahdiyeh Delfan, Chris Healey, Jeongwan Jin, Cecilia La Mela, Mike Lamont, Itzel Lucio Martinez, Hassan Mallahzadeh, Xiaofan Mo, Allison Rubenok, Erhan Saglamyurek, and Raju Valivarthi. I would like to thank the QC2 group's administrators, Hyejeong Hwang and Catherine Kosior, who, amongst other things, helped me avoid the reefs and krakens on the savage seas of peoplesoft. Vladimir Kiselyov constructed some of the electronics used in this experiment and generally kept the QC2 lab from degenerating into complete chaos, for which I am very grateful! I would like to offer special thanks to the QC2 members who helped proof-read my thesis: Philip Chan, Neil Sinclair, Daniel Oblak, and Morgan Hedges. In particular, I would like to thank Joshua A. Slater, who collaborated on several experiments in this thesis and was, by far, my most thorough proof reader. Without him this thesis would be considerably more unpleasant to read! I would also like to give special thanks to Félix Bussières, a former PHD student with the QC2 group, for helping design the entanglement source used in these experiments, showing me the ropes of working in an optics lab, and teaching

me that, "you can if you believe you can!".

I also thank my supervisor, Dr. Wolfgang Tittel, who is one of those rare individuals who is both brilliant and able to avoid making those around him feel dim by comparison. He has given me the guidance I needed to stay on track when I needed it and the freedom to follow my interests. His energy and enthusiasm are both inspiring and contagious. The QC2 group he has assembled has been a true pleasure to be a part of.

Finally, I would like to thank my family, who have cheered me on and supported me throughout my life. My sisters, Laurie and Diana have always inspired me with their generosity. My parents, Jim and Sharon, possess almost boundless patience and perseverance. My brothers-in-law, Dan and Ally, have been brothers in more than law to me. Finally, I would like to thank Morgan, my niece, and Andrew, my nephew, who taught me that things always look better after you've spent some time playing in the snow.

Table of Contents

List of Tables

List of Figures

Glossary

APD	Avalanche Photo-Diode
BS	50/50 Beam Splitter
BSC	Babinet-Soleil phase Compensator
CCW	Counterclockwise
CEMT	Classical Electromagnetic Theory
CHSH	Clauser Horne Shimony and Holt
CM	Classical Mechanics
CW	Clockwise
EPR	Einstein Podolsky Rosen
$g^{(2)}(0)$	Second-order correlation function for $\tau = 0$
HWP	Half ($\frac{\lambda}{2}$) waveplate
InGaAs	Indium Gallium Arsenide
LHV	Local Hidden Variable
ML	Maximum Likelihood
NLHV	Non-Local Hidden Variable
PBS	Polarizing Beam Splitter
PPLN	Periodically Poled Lithium Niobate
QC2	Quantum Cryptography and Communications Lab
QKD	Quantum Key Distribution
QST	Quantum State Tomography
QT	Quantum Theory
QWP	Quarter ($\frac{\lambda}{4}$) waveplate
Si	Silicon
SPDC	Spontaneous Parametric Down-Conversion
SV	Spacetime Variable
TDC	Time-to-Digital Converter

Chapter 1

Introduction

The universe is written in a language that the discipline of physics strives to discover. When we observe something unfamiliar in nature, the laws of physics must often be changed or extended in order to remain consistent with our observations. Discovering where physical laws do not correctly or completely describe nature is crucial if we wish to speak nature's language correctly. Only recently have we begun to appreciate how strangely nature behaves on the scale of a single quantum. This behaviour is unlike anything that exists in the macroscopic world that our own biology has equipped us to observe. Quantum theory allows us to speak about matter on the quantum scale.

The remarkable accuracy of predictions made by quantum theory has allowed it to serve as a solid foundation for tremendous advances in both our understanding of the universe and in practical applications. Some consider this foundation beyond question, similar to how one of Max Planck's professors considered the study of physics to be unworthy of Planck's pursuit since "almost everything is already discovered". No matter how good our theories may be, they must always defer to what has experimentally been observed. Experiments are our most reliable teachers. Even experiments that do not contradict quantum theory often reveal nuances that we would not otherwise think to investigate.

Entanglement is one of the stranger properties predicted by quantum theory; so much so that Einstein, in the absence of experimental data, proposed that quantum theory was incomplete. Yet, in experiments, we have observed behaviour that is consistent with the existence of entangled states. It is still worth asking if there might an alternative to quantum theory that agrees with our observations but more completely describes nature.

One way of doing this is to propose alternative theories or extensions to quantum theory. Some of these alternative theories and extensions make predictions that are sufficiently different from those of quantum theory that we can devise experiments to test them. These tests are, of course, also tests of quantum theory itself. Even if QT itself is not invalidated by our observations, invalidations of alternative theories can enhance our understanding of QT and of nature.

In this thesis we will discuss an experimental source of entangled photon pairs and experiments conducted with it that were designed to test aspects of the foundations of quantum theory.

1.1 Random Events: Deterministic vs Probabilistic

Probabilistic events are events whose outcomes we cannot predict with certainty. However, if we repeatedly measure outcomes of such an event a sufficiently large number of times, we may be able to predict the distribution of these outcomes. We often think of a coin toss as a probabilistic event. For a fair coin we assume that heads and tails are both equally likely outcomes. Thus, if we repeatedly toss a fair coin many times we will obtain outcomes that are evenly split between heads and tails. Any prediction we make for an individual toss will be correct with probability $P = 0.5$ (i.e. half of the time). Yet, according to classical mechanics, the outcome of a coin toss is a completely deterministic event. We could correctly predict the outcome with probability $P = 1$ if we had sufficient knowledge of the coin's properties, its environment, and the forces applied to toss the coin. We call the combination of these components the coin's "system". The difficulty of measuring the state of the coin's system before the toss and complexity of calculating the outcome are what makes a coin toss appear probabilistic.

Truly probabilistic events do not exist according to classical mechanics. We are forced

to treat rolls of dice, the weather, etc. as probabilistic only because the states of the pertinent systems are difficult or possibly impossible to measure completely and, provided the initial state of the system under test is fully known, the outcomes require excessively intensive calculations to predict.
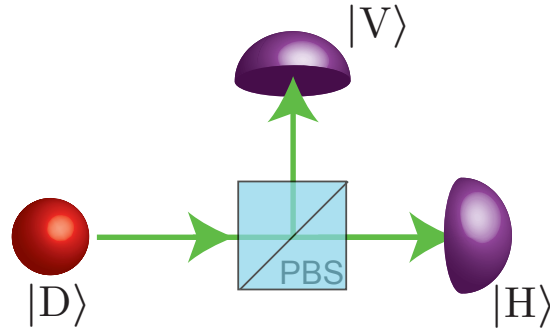
Figure 1.1: **Polarizing Beam Splitter.** A polarizing beam splitter (PBS) is an optical component that selectively transmits the horizontal component and reflects the vertical component of light entering it. QT states that a single diagonally polarized photon, upon encountering a PBS, will enter a superposition of being both transmitted and reflected by the PBS. It is then detected by one of two detectors.

QT states that there are indivisible energy quanta, such as single photons. This is one of QT's more radical departures from classical EM theory (CEMT). One implication of this is that, according to QT, some outcomes are completely unpredictable even if we possess complete knowledge of the system and have infinite computational resources. In Figure 1.1 we see one example of a system that, according to QT, will behave in a probabilistic fashion. Diagonally polarized light, which can be viewed as a coherent superposition of horizontally and vertically polarized light in equal proportions, is sent to a polarizing beam splitter (PBS). CEMT predicts that any diagonal light entering the PBS, regardless of intensity, will be split into two equal halves and both detectors will detect light, or "click". However, if there is only a single photon entering the PBS, QT predicts that a single diagonally polarized photon cannot divide and will therefore exit the PBS from both paths in superposition. Conservation of energy dictates that

this single quanta can only cause a detection in one of these detectors. The photon is in an equal superposition of reaching both detectors so QT offers no way to know which detector will actually click. Hence, even if we have complete knowledge of this system, QT says that can correctly predict which detector will click with a probability of 0.5.

Unlike CEMT, classical mechanics (CM) does offer predictions for single particle behaviour. The key difference between CM and QT in the above example is that CM would predict that the photon would exit the PBS from just one path rather than exiting in a superposition of both. Thus, CM predicts that the question of which detector will click is settled well before the moment of detection. CM can predict which path the photon will exit the PBS from deterministically. Provided the complete state of the photon and PBS are known, CM states that the photon's exit path could have been predicted even before the photon encountered the PBS. The superposition state of the photon after the PBS and resulting probabilistic choice of detector predicted by QT is like nothing in the macroscopic world we are used to observing.

The notion of truly probabilistic quantum outcomes, as proposed by Born, deeply disturbed Einstein and prompted him to say that God "is not playing at dice" [1]. This is part of what caused Einstein to question the completeness of quantum theory. If QT could be shown to be incomplete, there would be room for extra parameters or variables to be added to QT that might allow determinism to be recovered. It would not be until Bell's work on hidden variable models that the conflict between quantum physics and determinism would start to become clear.

1.1.1 Entanglement: A Puzzling Resource

Just as Einstein did not accept the notion of a probabilistic universe, he believed that QT's prediction of entanglement showed that QT itself was incomplete [2]. The famous Einstein Podolsky Rosen (EPR) Paradox paper proposed and analyzed a thought exper-

iment. They first make the assumption that QT is both correct and complete. They then look for a paradox resulting from their thought experiment that shows one of these assumptions cannot hold. EPR states that in order for a theory to be considered a complete description of reality, "every element of the physical reality must have a counterpart in the physical theory". They define elements of physical reality as any physical quantity that we can predict with certainty.

In EPR's thought experiment, two particles, A and B, are allowed to interact in such a way that we would now call them "entangled", a term coined by Schrödinger [3]. These particles are then widely separated so that they do not interact any further. We may then choose a measurement to conduct on particle A and use the result to predict what state particle B must be in. We could measure the position of particle A and, in the process, reduce the state of particle B to a position state: φ_r. We can predict with certainty that particle B will be found in this state if particle B's position were measured. Thus, position must be an element of the physical reality of particle B's system that existed after particles A and B stopped interacting. However, we might also choose to measure the momentum of particle A, which would place particle B in a momentum state: ψ_k. In this case we can also predict particle B's state with certainty, so momentum must also be an element of particle B's physical reality. Based on this logic, position and momentum must both be simultaneous elements of particle B's reality. Yet, since the associated operators do not commute, QT says that particle B cannot exist in both of these states simultaneously, and position and momentum cannot simultaneously be elements of reality as defined by EPR. This paradox implies that the assumptions made previously cannot hold. Einstein wrote about this paradox further in a subsequent paper and stated:

"Since there can be only one physical condition of B after the interaction and which can reasonably not be considered as dependent on the particular measurement we perform on the system A separated from B it may be con-

cluded that the function is not unambiguously coordinated with the physical condition. This coordination of several ψ functions with the same physical condition of system B shows again that the ψ function cannot be interpreted as a (complete) description of a physical condition of a unit system." [4]

In this quote, the "several ψ functions" describing the same physical condition of system B corresponds to φ_r and ψ_k above.

If QT's description of systems A and B was complete, then measurements of these systems could only reveal what was already contained in that description. The above thought experiment shows an example where this clearly is not the case, so EPR interprets the above paradox as proof that QT is not a complete theory, but this is not the only possible interpretation. We may take issue with EPR's notion that measurements reveal elements of reality that must have a counterpart in the expression of the entangled state prior to measurement. Instead, we could argue that a measurement on particle A defines the reality of particle B. For example, if we measure the momentum of particle A, only then is particle B placed in a momentum state. This resolves the paradox because only one measurement on particle A can be performed and particle B can be in just one corresponding state. However, particles A and B are separated such that this interpretation requires non-local interaction between particles A and B. Needless to say, non-local action did not sit well with Einstein! A more palatable way out of the paradox was simply to assume that Quantum theory does not completely describe reality, as EPR argued. If this was the case, then it might be possible to create a complete description of reality by adding local "hidden" variables (LHVs) to quantum states.

1.1.2 Tests of Locality and Beyond

If quantum theory provides a description of reality that is incomplete, then there could exist an alternative theory to QT or extension of QT that is more complete. One approach

to this question is to propose model theories that differ from QT in specific ways and then look at where their predictions agree or conflict with those of QT and what we observe in nature. LHV models provided one possible way to address the EPR paradox, and John S. Bell proposed an inequality in 1964 [5] that showed QT and LHV models differ from each other in terms of their predictions.

Bell considered a LHV model in which QT was extended by the addition of a local hidden variable: λ. This variable could be discrete or continuous, real or complex, a function, or even a list of values and functions. Bell placed no restrictions on what λ could be. If we could measure λ we could use this additional information to improve upon the predictions of QT. Figure 1.2 shows an experiment that could be conducted on entangled particles, for which LHV models and QT offer different predictions about the outcomes. Two entangled particles are separated and sent to two analyzers. Each analyzer takes as input a measurement setting $(\hat{a}, \hat{b})$ and outputs a result (A, B) after measurement. Note that while measurement outcome A may rely on $\hat{a}$ and λ, it must be independent of B and $\hat{b}$ for Bell's locality assumption to be satisfied i.e. A and B are functions of the form: $A(\hat{a}, \lambda)$ and $B(\hat{b}, \lambda)$. With this formalism, Bell showed that the measurement predictions made by deterministic LHV models are incompatible with those made by QT.

If deterministic models are found to be inconsistent with our observations of nature, then it is worth considering models that, while probabilistic, offer better predictions than QT. For example, in Figure 1.1 QT was able to correctly predict which detector would click with a probability of P=0.5. Could an alternative theory correctly predict which detector would click with a higher probability? The Leggett model (see section 4.3 - The Leggett Inequality) was developed to help answer this question. This thesis will also discuss further work in the same vein that provides a general way to evaluate models based on their predictive power (see Chapter 5). The key resource needed for all these
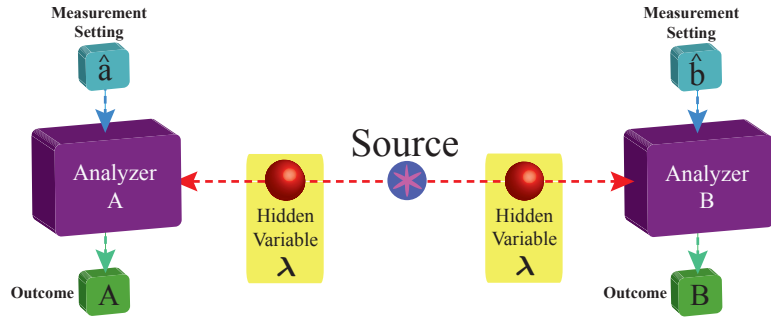
Figure 1.2: **LHV Models.** Bell considered deterministic LHV models that offer predictions for the outcome of measurements performed on an EPR pair of particles, that are split and sent to two qubit analyzers. Each of these analyzers measures its respective particle according to measurement settings $\hat{a}$ and $\hat{b}$, producing output results A and B respectively. The hidden variable λ is accessible by both analyzers when evaluating their particles.

tests is, of course, entanglement. Although it will not be discussed further in this thesis, it is interesting to note that there are alternative theories to quantum theory that seem to produce identical predictions, such as that proposed by Bohm [6]. At present, we are not aware of testable differences between QT and Bohm's model, but it is important to keep in mind that QT may not be the only valid explanation of our observations.

1.2 Entanglement as a Resource

Entanglement is a uniquely quantum phenomena that vividly highlights where classical theory breaks down. As such, entanglement is an essential resource for probing the quantum aspects of nature. While QT's prediction of entangled states was once controversial, is has since been established and we now use entangled states in a variety of applications.

1.2.1 Quantum States and Entanglement

Qubits

In quantum physics we are concerned with the states of single quanta, such as atoms or photons. Similar to how a computer's smallest unit of storage is a bit that can be either a 0 or a 1, the smallest unit of information in quantum systems is a qubit. Unlike a classical bit which must be either 0 or 1, a qubit may exist in a coherent superposition of two orthogonal basis states: $|0\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $|1\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. This basis may correspond to, for example, the up and down spins of a spin 1/2 electron $\{|\uparrow\rangle, |\downarrow\rangle\}$, an atom in a system that has just two energy levels of interest $\{|E_1\rangle, |E_2\rangle\}$, or orthogonal polarizations of a photon. In this thesis we will chose $|0\rangle = |H\rangle$ and $|1\rangle = |V\rangle$ where $|H\rangle$ is a horizontally polarized photon and $|V\rangle$ is a vertically polarized photon. Note that this choice is arbitrary, and we could chose other definitions for the basis states such as, for example, right and left circular polarized light: $|0\rangle = |R\rangle$ and $|1\rangle = |L\rangle$ (see equation 1.3). We will use polarized photon states from here on since the experiments we conduct use polarization states. A polarization qubit may be generally expressed as:

$$|\psi\rangle = \alpha|H\rangle + \beta|V\rangle. \tag{1.1}$$

Here, α and β are complex probability amplitudes that are normalized such that $\alpha^*\alpha + \beta^*\beta = 1$. If we discard any global phase, qubits can also be generally expressed as:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|H\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|V\rangle, \tag{1.2}$$

where $\sin\left(\frac{\theta}{2}\right)$ and $\cos\left(\frac{\theta}{2}\right)$ determine the $|H\rangle$ and $|V\rangle$ components of the state and ϕ is the relative phase between those components.

For ease of reference, we define notation for some specific states:

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle + |V\rangle\right) \qquad\qquad |R\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle + i|V\rangle\right)$$

$$|-\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle - |V\rangle\right) \qquad\qquad |L\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle - i|V\rangle\right). \qquad (1.3)$$

The Bloch Sphere

These states are often depicted graphically as vectors on the Bloch sphere, as shown in Figure 1.3. Arbitrary states can be specified using θ and ϕ. It is also common to look at just one great circle on the Bloch sphere, such as one where $\phi = 0$. States on this great circle are sometimes specified solely by giving θ. For example, $|\theta = \frac{\pi}{2}\rangle$ would correspond to $|\theta = \frac{\pi}{2}, \phi = 0\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle + |V\rangle\right) = |+\rangle$. Linearly polarized states on the $\phi = 0$ plane are also sometimes expressed in shorthand using $\frac{\theta}{2}$ expressed in degrees. For example, in this shorthand representation $|45°\rangle = |+\rangle$.
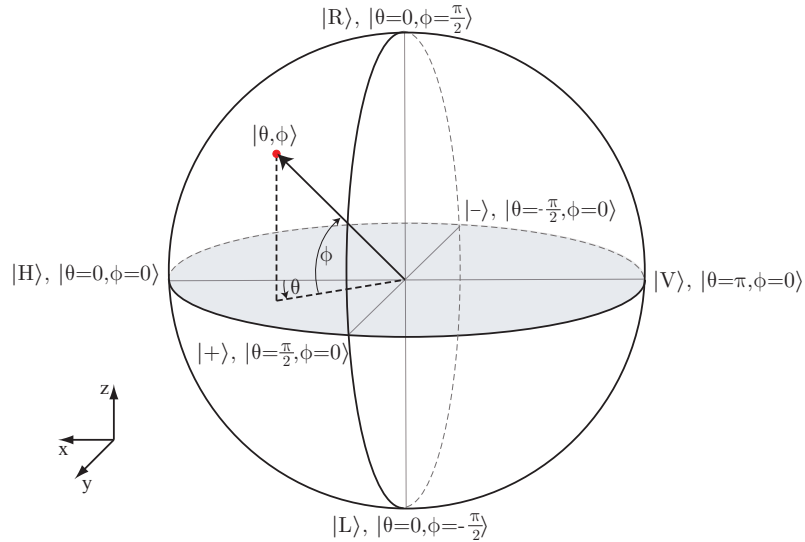
Figure 1.3: **Bloch Sphere.** The Bloch sphere allows us to graphically represent qubit states. Arbitrary states may be expressed in terms of θ and ϕ. The great-circle corresponding to $\phi = 0$ is shaded in.

The Density Matrix and Projective Measurements

So far we have dealt with pure states only, but we must also be able to express mixtures of pure states. Density matrices allow us to describe both pure states and mixed states. We can express a single pure state, $|\theta, \phi\rangle$ in the form of a density matrix:

$$
\begin{aligned}
\rho &= |\psi\rangle\langle\psi| \\
&= \sin^2 \frac{\theta}{2}|H\rangle\langle H| + e^{-\iota\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2}|H\rangle\langle V| + e^{\iota\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2}|V\rangle\langle H| + \cos^2\frac{\theta}{2}|V\rangle\langle V| \\
&= \begin{pmatrix} \sin^2(\frac{\theta}{2}) & e^{-\iota\phi}\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) \\ e^{\iota\phi}\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) & \cos^2(\frac{\theta}{2}) \end{pmatrix}.
\end{aligned} \tag{1.4}
$$

We can then express mixed states as a proportional sum of orthogonal pure states:

$$
\rho_M = \sum_i P_i|\psi_i\rangle\langle\psi_i| \tag{1.5}
$$

where P_i is the probabilistic weighting of pure state $|\psi_i\rangle$ such that $\sum_i P_i = 1$.

Say we have an unknown state described by a density matrix, ρ, and we projectively measure this state onto a pure state, $|\psi\rangle$. The outcome probability will be:

$$
P(|\psi\rangle) = Tr\left(|\psi\rangle\langle\psi|\rho\right) \tag{1.6}
$$

If we were to projectively measure this qubit on two states forming an orthogonal basis, such as $|H\rangle$ and $|V\rangle$, we require for normalization that the sum of the two resulting probabilities must be 1. The trace of a density matrix corresponding to a physically valid state must therefore be 1. The off-diagonal terms capture the relative phase between the two components of the state.

If, in an experiment, we use a qubit analyzer to make a projective measurement onto state $|\psi\rangle$, we will register a number of detection events, $C(|\psi\rangle)$. To experimentally find $P(|\psi\rangle)$ we must also projectively measure onto $|\psi^\perp\rangle$, which is the orthogonal state to $|\psi\rangle$, so that we can calculate:

$$P(|\psi\rangle) = \frac{C(|\psi\rangle)}{C(|\psi\rangle) + C(|\psi^{\perp}\rangle)} \tag{1.7}$$

Bipartite States

We can extend this formalism to deal with multi-qubit systems. A pure 2-qubit, or bipartite system can be expressed as:

$$|\psi\rangle = A|HH\rangle + B|HV\rangle + C|VH\rangle + D|VV\rangle, \tag{1.8}$$

where A, B, C and D may be complex numbers that include phase information, as α and β did above for a single qubit state.

Bipartite Projection Measurements

For a bipartite state defined by a density matrix, ρ, the outcome probability of projecting the two qubits of this state onto the pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively will be:

$$P(|\psi_1\psi_2\rangle) = Tr\left(|\psi_1\psi_2\rangle\langle\psi_1\psi_2|\rho\right) \tag{1.9}$$

If, in an experiment, we used two qubit analyzers to projectively measure onto these states, we would obtain a number of coincidence detections between the two qubit analyzers: $C(|\psi_1\psi_2\rangle)$. We would need to conduct a total of four such measurements to find:

$$P(|\psi_1\psi_2\rangle) = \frac{C(|\psi_1\psi_2\rangle)}{C(|\psi_1\psi_2\rangle) + C(|\psi_1^{\perp}\psi_2\rangle) + C(|\psi_1\psi_2^{\perp}\rangle) + C(|\psi_1^{\perp}\psi_2^{\perp}\rangle)} \tag{1.10}$$

where $|\psi_i^{\perp}\rangle$ indicates the orthogonal state to $|\psi\rangle_i$.

Entanglement

An entangled system is a special case of bipartite states. Entangled states are non-separable in that the full information associated with such states cannot be fully expressed

by describing each of its constituent components separately. For example, consider a bipartite entangled state, such as the $|\Phi^+\rangle$ Bell state:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} \left(|H\rangle_a |H\rangle_b + |V\rangle_a |V\rangle_b \right), \tag{1.11}$$

Here the subscripts a and b refer to each of the two photons in the entangled pair. We will now drop this notation and assume that the order of states specifies which particle has which state. We can express the density matrix for this state as:

$$\begin{aligned}
\rho_{\Phi^+} &= |\Phi^+\rangle\langle\Phi^+| \\
&= \frac{1}{2} \left(|HH\rangle\langle HH| + |HH\rangle\langle VV| + |VV\rangle\langle HH| + |VV\rangle\langle VV| \right). \\
&= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned} \tag{1.12}$$

An entangled state need not be limited to two particles. It could involve any number of particles or even different degrees of freedom of a single particle. In practical experiments these particles are usually pairs of single quanta such as atoms or photons. They may be entangled in any available degree of freedom. For example, a $|\phi^+\rangle$ Bell state could be formed by two photons, a and b, entangled in polarization ($|\phi^+\rangle = \frac{1}{\sqrt{2}} \left(|H_a H_b\rangle + |V_a V_b\rangle \right)$). Time-bin entangled states ($|\phi^+\rangle = \frac{1}{\sqrt{2}} \left(|E_a E_b\rangle + |L_a L_b\rangle \right)$) are also possible where $|E\rangle$ indicates a photon that arrives early with respect to some reference time and $|L\rangle$ indicates a photon that arrives late. We will limit further discussion to bipartite polarization entangled photons since this corresponds to what has been used in our own experiments.

1.2.2   Polarization Entangled Photon Pair Source Design Requirements

The entanglement source used in our experiments was designed with laboratory tests of quantum theory in mind. This application requires several properties including a

high degree of entanglement (see Section 3.2.1), ease of projections onto arbitrary bases, and high efficiency distribution/measurement that may include links between spatially separated locations for future experiments. Many of these requirements coincide with those of quantum key distribution (QKD), so this source might have applications in that field as well.

Efficiency of Distribution and Measurement

For optimal signal-to-noise ratios we need to maximize the probability that single particles both reach their destinations and are detected. Our experiments require us to distribute entangled particles to several devices in our own lab and future experiments may require distribution of entanglement over fiber links and free-space between separated labs.

Photons are the natural choice of qubits for this application because they have a low probability of being absorbed by the media they travel through and limited decoherence, unlike, for example, atoms. Photons have a reasonable probability of travelling long distances through air or fiber before loss becomes an insurmountable problem. For example, QKD has been experimentally conducted at distances of up to 144 km [7]. However, even for experiments where photons need only travel a few meters, the wrong choice of wavelength can introduce substantial loss. Figure 1.4 shows the transmission of photons through air at different wavelengths. A wavelength of 810 nm is one of several possible choices yielding high transmission through air.

Fiber optic cables are another frequently used conduit for photon-based communications because they simplify alignment and eliminate the need for direct lines-of-sight between sender and receiver. They are highly useful for providing an optical path between spatially separated labs. As with air, loss is also a problem in fiber. Figure 1.5 shows the loss versus wavelength in telecommunications fiber.

For flexibility of operation in both free-space and fiber, it is advantageous for a source
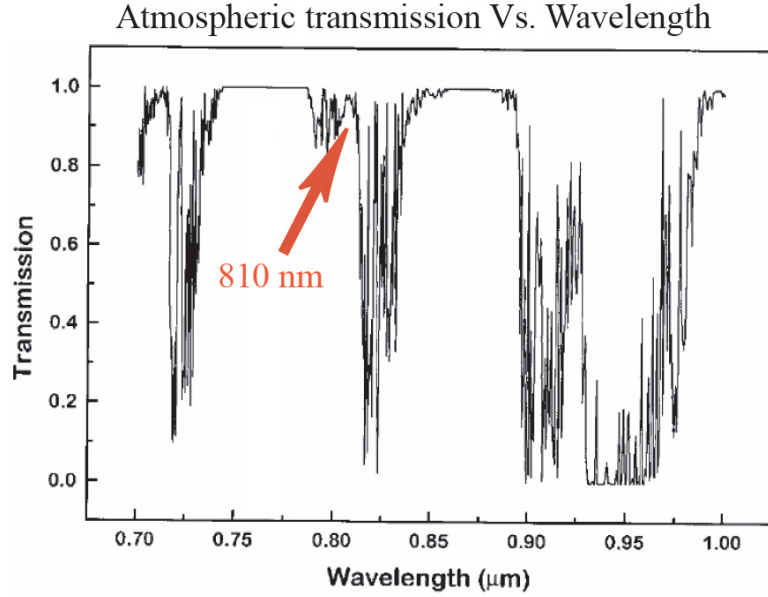
Figure 1.4: **Loss in Atmosphere.** On this plot of atmospheric transmission vs wavelength, we observe that 810 nm corresponds to a relatively high transmission. From ref. [9].

of entangled photon pairs to produce pairs such that one photon is at 1550 nm and the other is at 810 nm. As we will see in section 2.3.2, the detectors available to us that operate at 1550 nm have some disadvantages that can be overcome if we are using one 810 nm detector as well, rather than two detectors designed for detecting 1550 nm photons.

Visibility and Projection

It is currently easier to obtain high quality entangled states with polarization entangled photons than it is with some other commonly used degrees of freedom, such as time. Additionally, projecting polarization states onto arbitrary bases is relatively easy. A $\frac{\lambda}{4}$ waveplate, a $\frac{\lambda}{2}$ waveplate, a polarizing beam splitter, and detectors are all that is needed to construct a polarization qubit analyzer capable of making projective measurements along any vector on the Bloch sphere. While transmission through air has relatively little effect on polarization states, time varying birefringence in fiber links can alter polarization
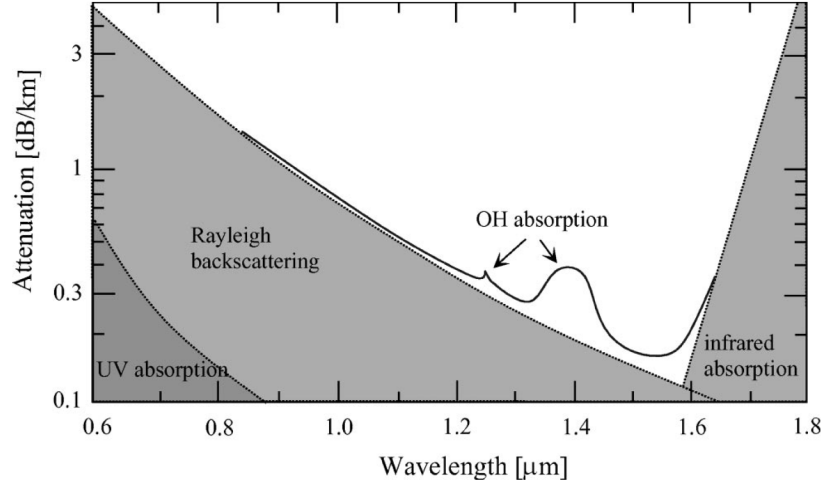
Figure 1.5: **Loss in Fiber.** On this plot of fiber attenuation vs wavelength, we observe that 1550 nm corresponds to relatively low loss and hence high transmission. From: [9].

states, making active stabilization necessary [11]. Time-bin entangled photon states are robust against this effect, which is why they are frequently used instead of polarization encoded states for communications.

1.3 This Thesis

1.3.1 Motivation

We live in a universe that cannot be adequately described by classical EM and mechanical theories alone. QT was proposed to successfully describe matter on the single quantum scale where other theories break down, but some aspects of it have proved to be surprising. The probabilistic nature of measurement outcomes and the bizarre properties of entanglement both conflict with the macroscopic picture of the world we are familiar with. Many have asked and continue to question to what extent quantum physics truly describes reality [10]. Alternative models or possible extensions to quantum physics have been proposed with the goal of either recovering some properties of classical physics or revealing weaknesses in quantum theory itself. Since entanglement is at the heart of the

disagreement between classical and quantum theory, experiments based on entanglement are crucial to our understanding of nature. In this thesis we will discuss tests of several alternative models to quantum theory as well as an experiment that uses measurements of an entangled system to bound the predictive power of any alternative theory or extension to quantum mechanics, provided certain plausible conditions are met (see Chapter 5). This experiment is significant in that it both provides an alternative method of ruling out previously tested models, such as LHV models or Leggett's model, while also providing a criteria for judging previously untested models and even models that are yet to be proposed.

1.3.2 Organization

In Chapter 2 we will discuss the design and construction of a novel experimental source of high quality entangled photons pairs at adjustable, non-degenerate wavelengths. This chapter will cover how photon pairs are generated, review other sources of entanglement, and describe how the source itself works. Chapter 3 will provide the results of characterizing this source. Chapter 4 will discuss direct tests of alternative models to quantum theory including violations of the CHSH Bell inquality, the Beautiful Bell inequality, and the Leggett inequality. Chapter 5 is a discussion of the experiment that bounds the predictive power of alternative theories to quantum mechanics. Chapter 6 is a summary of these results and a discussion of potential work that may be pursued in the future.

1.3.3 Collaborations

The work presented here would not have been possible without the collaboration of others. The design, construction and some characterization of the entanglement source was carried out in collaboration with Félix Bussières who was then a PhD student in our group. Software design for experiment-control/data-collection were carried in conjunction

with Félix Bussières and Joshua A. Slater, another PhD student. Vladimir Kiselyov, the QC2 group's engineer, constructed some of the electronic logic used to gather statistics. Joshua A. Slater was involved with some of the source characterization as well as the test of the Beautiful Bell and Leggett inequalities and the experiment that bounded the predictive power of alternative theories to quantum physics. This experiment was done in collaboration with Roger Colbeck and Renato Renner, both at ETH Zurich in Switzerland, the theorists who laid the foundation for this work.

Chapter 2

Sources of Entangled Photons

Spontaneous Parametric Down-Conversion (SPDC) in non-linear crystals is a widely used method for producing pairs of photons for quantum level experiments and applications. In this section we will first cover how SPDC lets us create photon pairs and then explain how the entanglement source we have built use SPDC crystals to create entangled pairs of photons. We will then discuss the results of characterizing this source of polarization entangled photon pairs.

2.1   Photon Pair Production via Spontaneous Parametric Down-Conversion

2.1.1   Spontaneous Parametric Down-Conversion

Generating many photons at a time is a task our ancestors were adept at before they evolved into Homo sapiens. Producing exactly two photons on command remains a challenging problem! SPDC does not allow us to generate exactly two photons on command, but it does allow us to create pairs of photons in a probabilistic manner such that we can conduct quantum level experiments.

SPDC is a process in which a pump photon is split into two daughter photons. Depending on the SPDC crystal used, photon pairs may be produced such that their polarizations are both the same as that of the pump, both perpendicular to that of the pump, or the pairs might be perpendicularly polarized to each other with one photon sharing the same polarization as the pump. We use SPDC crystals that produce co-polarized pairs at the same polarization as the pump.

As we discussed in Section 1.2.2, we would like to produce pairs of photons such

that one photon is at 810 nm and the other at 1550 nm. This is accomplished by sending pump photons into a nonlinear crystal, such as lithium niobate, that is designed to mediate energy transfer from the pump mode to two daughter photon modes. This happens on a probabilistic basis such that the number of pairs emitted follows a Poisson distribution (see Section 2.1.2). We can choose the pump intensity such that the mean probability of producing pairs in a specific period of time is very low. In this regime, during a specific period of time there is a high probability of producing no pairs, a low probability of producing just one pair, and a very low probability of producing multiple pairs. We can therefore produce approximately a single pair at a time although we cannot predict the precise moment that each pair will be produced.

Each time a pair is generated, a single pump photon is split into two daughter photons such that energy and momentum are conserved. In this process the following equations, derived from conservation of energy and momentum, must be obeyed:

$$\hbar\omega_p = \hbar\omega_s + \hbar\omega_i$$

$$\hbar\mathbf{k}_p = \hbar\mathbf{k_s} + \hbar\mathbf{k}_i \tag{2.1}$$

The first equation expresses conservation of energy and the second expresses conservation of momentum. As the momentum is associated with spatial phase change, this is often referred to as the phase matching condition, but in this thesis we will refer to both equations in combination as the phase matching conditions. Here, ω refers to frequencies and $\mathbf{k}$ refers to wave vectors. In these equations the subscript p refers to pump photons entering the crystal while s and i refer to the two daughter photons, which we call signal and idler photons for ease of reference (and consistency with literature). Rearranging the phase matching condition derived from conservation of energy, cancelling out the reduced

Planck's constant ($\hbar$), and converting frequencies to wavelengths, we obtain:

$$\frac{n_p}{\lambda_p} = \frac{n_s}{\lambda_s} + \frac{n_i}{\lambda_i}. \tag{2.2}$$

Here, the n terms are the refractive indices of the medium at the three wavelengths. In order for natural phase matching to occur a material must satisfy both of the phase matching conditions simultaneously. If the material used is birefringent, these indices of refraction will also depend on polarization, so a material that meets phase matching conditions may only do so at one combination of polarizations for the pump, signal, and idler photons. Unfortunately, if we require a specific set of pump, signal and idler wavelengths and polarizations, there may exist no known material able to satisfy these requirements.
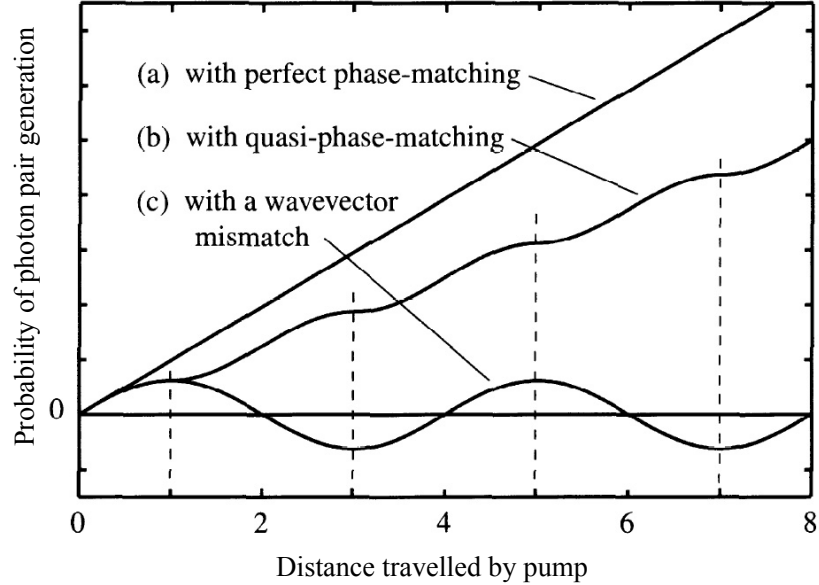
Figure 2.1: **SPDC: Probability of Photon Pair Production vs distance pump travels in medium.** Curve (a) corresponds to perfect, or natural, phase matching that occurs when conservation of energy and phase matching conditions are met naturally by a material. Curve (b) shows quasi phase matching that occurs when a non-linear crystal is poled to periodically reverse its electric dipole moment. Curve (c) shows the lack of phase matching. Reproduced from: [12].

Quasi phase matching is a way to modify non-linear crystals to meet the phase match-

ing conditions at wavelengths that are different from those that satisfy the natural phase matching conditions of the crystal material. One way to achieve this is to periodically pole a non-linear crystal by exposing it to an intense electric field during the manufacturing process to periodically reverse its electric dipole moments. Figure 2.1 provides a comparison of how the probability of producing photon pairs grows as pump photons travel through a crystal with phase matching, quasi phase matching, and without phase matching. Without phase matching, the probability of generating photon pairs oscillates around zero and does not grow [12]. If the electric dipole moments of the crystal are periodically reversed with a period that is the same as the period of the oscillation in the curve without phase matching, a roughly linear growth in pair generation probability can be obtained. The quasi-phase matching condition can then be modified to include this poling period:

$$\frac{n_p}{\lambda_p} - \frac{1}{\Lambda} = \frac{n_s}{\lambda_s} + \frac{n_i}{\lambda_i}. \tag{2.3}$$

Here, Λ is the poling period of the crystal. While crystals can be poled with different periods, it is also possible to tune the quasi-phase matching conditions by varying the temperature of the crystal, causing it to expand or contract, thus changing Λ. However, the indices of refraction are also temperature dependent, which must be taken into account:

$$\frac{n_p(T)}{\lambda_p} - \frac{1}{\Lambda(T)} = \frac{n_s(T)}{\lambda_s} + \frac{n_i(T)}{\lambda_i}. \tag{2.4}$$

Figure 2.2 shows how signal and idler wavelengths may be manipulated via temperature for three different poling periods and a 532 nm pump wavelength. Changing the poling period of the crystal changes the phase matching conditions such that different signal and idler wavelengths will be produced at a given crystal temperature. We calculated the indices of refraction for this plot using a Sellmeier equation with coefficients

23
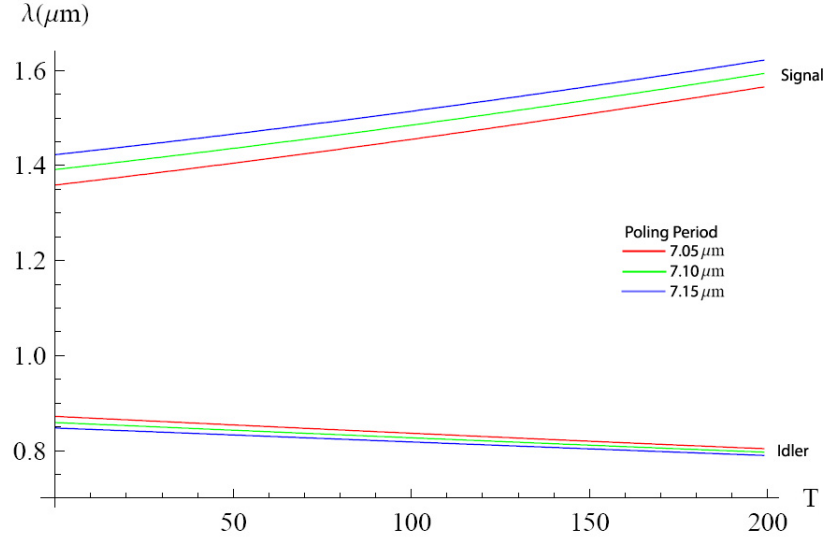
Figure 2.2: **Wavelengths satisfying phase matching conditions in PPLN.** Signal and idler wavelengths plotted versus temperature T ($°C$) for a 532 nm pump and three different poling periods.

from [13]. The Sellmeier equation offers a way to calculate the index of refraction for a material corresponding to a specific wavelength and temperature, and is based on empirically measured coefficients. Based on these calculations, we expect signal and idler wavelengths of 810 nm and 1550 nm to be obtained at approximately 178 $°C$ for a poling period of 7.05 μm. PPLN crystals with this poling period are commercially available from Covesion, as shown in figure 2.3. When one considers that the pump wavelength may also be changed, this method of creating photon pairs provides considerable flexibility for producing different signal and idler wavelengths.
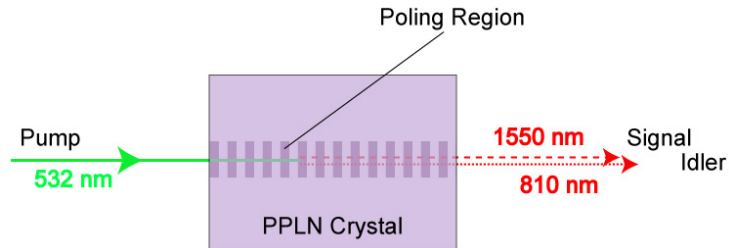
Figure 2.3: **PPLN SPDC Crystal.** A periodically poled lithium niobate (PPLN) crystal with a poling period of $\Lambda = 7.05$ μm generates photon pairs via non-degenerate SPDC at 810 nm and 1550 nm.

24

2.1.2 Photon Pair Statistics: $g^{(2)}(0)$ Measurement

Note: The measurements in this subsection were completed as a part of the author's 598 project in undergraduate studies, but are included here to provide a complete description of the entanglement source.

The SPDC crystals we use produce photon pairs probabilistically according to a Poisson distribution:

$$P_i = \frac{e^{-\mu}\mu^i}{i!}, \tag{2.5}$$

where i is the number of pairs produced in a time period of interest and μ is the mean number of photon pairs produced per time period. The time period being considered here is dependent on the smallest time period that can be resolved by the experimental apparatus. For example, the limiting factor might be the time resolution of the detectors.

If $\mu \ll 1$ then the probability of producing zero pairs in an event, $P_0 = e^{-\mu}$, will be quite close to 1, the probability of producing just one pair per event, $P_1 = \mu P_0$, will be low, and the probability of producing two pairs in a single event, $P_2 = \frac{\mu^2}{2}P_0$ will be very low.

When multiple pairs are produced within a single time period it is possible for photons from different pairs to be detected and assumed to belong to the same pair. The speed of detectors and other electronic components of the experimental apparatus determine how close two pair events may occur in time before we cannot distinguish them. It is also possible for detector imperfections, such as dark counts or after-pulsing in one or both detectors, to produce detection events that may be interpreted as corresponding to parts of a single pair. We must therefore characterize the performance of a SPDC crystal based source of photon pairs as a part of a system that includes detectors.

For the experiments conducted in this thesis, it is important to minimize the probability that we detect photons from two different pairs and then interpret them as having

come from a single pair. The lower the mean photon number used, the lower this probability should be. It was also necessary to minimize the probability of interpreting dark counts and after-pulsing as parts of a single pair. The dark count rate of a detector is usually a constant, so decreasing the mean photon number will actually increase the relative probability of dark counts. We must therefore choose the pump power such that μ is neither too high nor too low. Figure 2.4 shows an experiment designed to allow us to do just that.

Pump light from a 532 nm laser is sent through one PPLN crystal such that photon pairs are generated at non-degenerate wavelengths of 810 nm and 1550 nm. The pump is filtered out and the remaining photon pairs are coupled into fiber and split according to wavelength. The 810 nm photons are detected by a silicon avalanche photo-diode (Si APD) that is used to trigger two indium gallium arsenide (InGaAs) APDs in the 1550 nm portion of the setup. The 1550 nm photons encounter a 50/50 beam splitter (BS) and enter a superposition of travelling to each of two InGaAs APDs, A and B. Detection signals from all three detectors are sent to a time-to-digital converter (TDC) that records when each detection event occurs relative to the others.

Say that the photon pair source and experiment were ideal such that there are no dark-counts or after-pulsing, the detectors are 100 % efficient, and no more than one pair is generated closer together than the experiment is able to resolve. In this ideal case, when we detect an 810 nm photon at the Si APD we would know that exactly one 1550 nm photon is in the 1550 nm side of the apparatus. We should therefore see exactly one detection at one of the two InGaAs APDs. Detection events at the two InGaAs APDs would be perfectly anti-correlated. A measure of the correlation between detections at InGaAs APDs A and B is therefore a measure of how far a real experiment deviates from the ideal mentioned above.

We can quantify the degree of correlation between two events using the second-order

correlation function, denoted $g^{(2)}(\tau)$, which is a measure of the probability of measuring event B at time $t = \tau$ assuming we first measured event A at time $t = 0$. This probability is normalized with respect to the probability of A and B occurring on their own. For $\tau = 0$, $g^{(2)}(0)$ is the normalized correlation function for two simultaneous events. We can write this as:

$$g^{(2)}(0) = \frac{P(A \cap B)}{P(A)P(B)}, \tag{2.6}$$

where $P(A \cap B)$ is the probability of A and B happening simultaneously and $P(A)$ and $P(B)$ are the probabilities of events A and B happening at all, respectively.
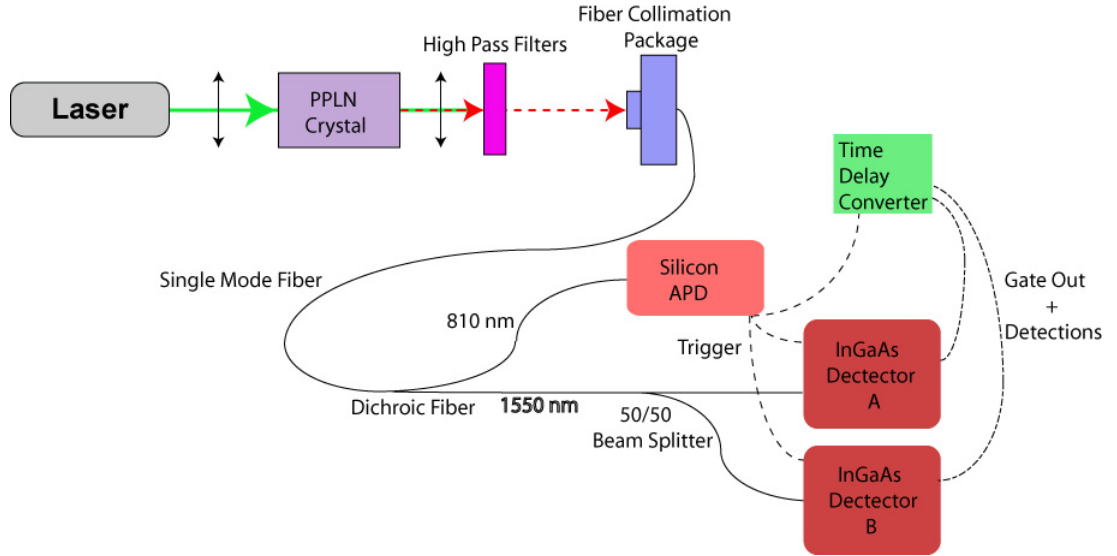
Figure 2.4: **Measurement of $\mathbf{g^{(2)}(0)}$.** 1550 nm photons are heralded by the presence of 810 nm photons at the silicon APD. The 1550 nm light is split in two by a BS and sent to two different InGaAs detectors. Detection events are sent to a TDC so that coincidence detections may be recorded.

We interpret an event, A, as corresponding to a detection in InGaAs APD A and B as a detection in InGaAs APD B. In this context, A and B occurring "simultaneously" means that the TDC received detection signals from the two InGaAs APDs closer together than the time resolution of the InGaAs APDs, which is the limiting factor on time

resolution in our experiments. This resolution is approximately half a nanosecond.

Since we are looking for detections at detectors A and B whenever the Si APD clicks, the equation for $g^{(2)}(0)$ will look slightly different because all of the events are now conditional on the detection of a 810 nm photon in the Si APD. Equation 2.6 becomes:

$$g^{(2)}(0) = \frac{P(A \cap B | Si)}{P(A|Si)P(B|Si)}. \tag{2.7}$$

Here, detections in the Si APD are abbreviated as Si. Using Bayes' theorem to rewrite the conditional probabilities, this reduces to:

$$g^{(2)}(0) = \frac{P(A \cap B \cap Si)}{P(A \cap Si)P(B \cap Si)} P(Si). \tag{2.8}$$

For this experiment, a $g^{(2)}(0)$ of 1 would indicate that $P(A \cap B) = P(A)P(B)$ and events A and B therefore both happen without any dependence on each other. If $g^{(2)}(0)$ were > 1 this would indicate that A and B are correlated such that if InGaAs APD A clicks, there is an increased probability that InGaAs B clicks. A $g^{(2)}(0)$ less than one indicates anti-correlation, with zero being the minimum value possible. At $g^{(2)}(0) = 0$ event A never happens at the same time as event B. We therefore want to choose a mean photon number such that the $g^{(2)}(0)$ value is as close to zero as possible since, when this is the case, it shows that we are producing single 1550 nm photons, which in turn implies we are working in the single pair regime.

For a pump power of 73.1 μW we measured $g^{(2)}(0) = 0.04 \pm 0.01$. The pump intensity was chosen to produce a $g^{(2)}(0)$ well within the single photon pair regime. In this regime the probability of multiple pair generation events or dark counts happening should be very low. Data for this this $g^{(2)}(0)$ measurement was taken over ~ 25 minutes and contains approximately 22 million 810 nm signal clicks and a hundred thousand 1550 nm InGaAs clicks, but there were just 10 coincidence detections. One more or one less coincidence

would have changed the values for $g^{(2)}(0)$ noticeably, which is why the uncertainty is high relative to the value for $g^{(2)}(0)$.

The Si APD detection rate, which is proportional to the mean photon number, used in this experiment is approximately the same as the rate the entanglement source was operated at in the other experiments discussed in this thesis, so this value for the $g^{(2)}(0)$ is representative of the impact non single-pair events might have had on the entanglement source's operation.

2.2 Review of Designs used for Producing Polarization Entangled Pairs

2.2.1 Sequential crystal designs

There have been many designs for producing entangled photon pairs using SPDC crystals. One of the conceptually simplest is shown in Figure 2.5. In this design two periodically poled Potassium Titanyl Phosphate, or PPKTP, crystals are placed sequentially in the path of diagonally polarized pump photons [14]. The first crystal is oriented to down-convert pump photons into $|HH\rangle$ pairs while the second crystal is oriented to create $|VV\rangle$ pairs. Since the diagonally polarized pump is in a superposition of being horizontally and vertically polarized, each pump photon that is down-converted will enter into a superposition of having down-converted in either the first or second crystal. This creates a $|\psi\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{\imath\phi}|VV\rangle\right)$ entangled state where ϕ is a phase dependent on the lengths of the two PPKTP crystals and their indices of refraction.

While elegantly simple, this design does have some drawbacks. In general, if we want to create maximal entanglement in one degree of freedom, we want other degrees of freedom to contain no information about which crystal a pair was created in. In this design it may become possible to distinguish which of the two crystals a given pair of photons was generated in, which creates a partially mixed state rather than a maximally

29
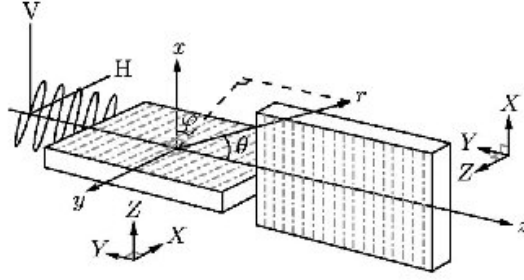
Figure 2.5: **Sequential crystal entanglement source.** Two sequential PPKTP crystals in a configuration for creating polarization entanglement. $|+\rangle$ polarized pump photons enter a superposition of down-converting in the first crystal to $|HH\rangle$ or in the the second crystal to $|VV\rangle$, yielding a $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$ state. Reprinted figure with permission from: [15].

entangled one. For example, say that a sequential crystal source is designed to create pairs of photons where one photon has a significantly different wavelength than the other. The index of refraction will be different at these two wavelengths so, after down-conversion, the two photons belonging to a pair will start to accumulate a relative phase delay. This delay is proportional to the length of crystal they travel through before leaving the source. If the crystals are long enough we may be able to distinguish which crystal a given pair was created in just by measuring the delay between the two photons in a pair. This chromatic dispersion limits the length of the crystals that can be used in sequential crystal sources, and makes some materials difficult to work with. Specifically, lithium niobate exhibits a significantly higher degree of chromatic dispersion than KTP [15]. We chose to use periodically poled lithium niobate (PPLN) crystals because they satisfy the requirements in Section 1.2.2 and are easily obtainable.

2.2.2 Interferometer based designs

Entanglement sources may also be based upon an interferometer in which light is split into two paths that differ in at least one degree of freedom (in addition to spatial mode) and are then recombined, producing a superposition state. These designs can be based on

30

several different types of interferometers. In most interferometer based designs, such as in the Mach-Zehnder design shown in Figure 2.6, there are two spatially distinct paths that pump photons travel through before recombining into the same mode. Entanglement can be created if each path creates pairs of photons, but the two paths create states that are orthogonal in one degree of freedom. The quality of entangled states produced by such sources is reduced if the paths can be distinguished in degrees of freedom other that chosen to encode the state. An unstable phase may also arise between the two paths of such an interferometer if the path length difference is not constant. In a typical interferometer, the phase between components of the state will change by a large amount if the path length difference changes by a fraction of wavelength. This can happen in relatively short time periods due to temperature fluctuations and air flow in the interferometer.
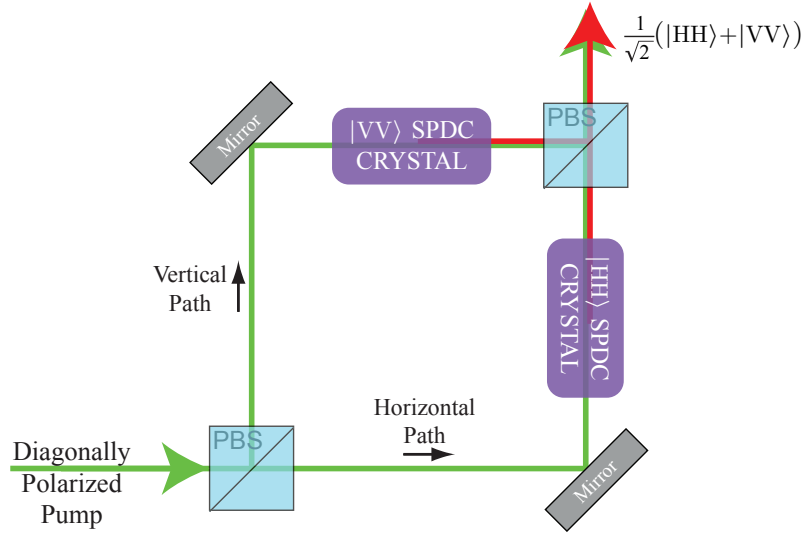
Figure 2.6: **Mach-Zehnder Interferometer based entanglement source.** This entanglement source is based on a Mach-Zehnder interferometer. Diagonally polarized pump light first encounters a PBS, which places it in a superposition of being transmitted into the horizontal path of the interferometer or reflected into the vertical path. In the horizontal path, pump photons encounter a SPDC crystal oriented to produce $|HH\rangle$ pairs which are transmitted through the second PBS. Similarly, in the vertical path pump photons are down-converted to $|VV\rangle$ photon pairs that are reflected by the second PBS. The two paths are recombined on the second PBS, resulting in an entangled state. Pump light is shown in green and down-converted photon pairs are shown in red.

While a controllable path length difference is often the goal for many interferometer designs, Sagnac interferometers are designed to have zero path length difference. In a Sagnac interferometer there are two counter-propagating paths that overlap in the same spatial mode. Any phase variations arising from changes in path length that might be caused by temperature fluctuations or air-flow will occur in both paths simultaneously and cancel out.

Figure 2.7 shows a polarization entanglement source based on a Sagnac interferometer that inspired our own design. In this design, 405 nm diagonally polarized pump photons are sent into the interferometer where they first encounter a polarizing beam splitter (PBS) that places them in a superposition of travelling both clockwise (CW) and counterclockwise (CCW) through the interferometer. In the CCW path, horizontal light is down-converted to wavelength-degenerate 810 nm $|HV\rangle$ pairs that are then rotated by a $\frac{\lambda}{2}$ waveplate to $|VH\rangle$ and split by the PBS into paths 1 and 2. In the CW path vertically polarized photons are first rotated to horizontal polarization before being down-converted by the same crystal to $|HV\rangle$ pairs that are split by the PBS into paths 1 and 2. The superposition of both paths results in a entangled state $\frac{1}{\sqrt{2}}\left(|H_1V_2\rangle + e^{\imath\phi}|V_1H_2\rangle\right)$. Here, the subscripts 1 and 2 indicate which path a given photon leaves the PBS by. The relative phase ϕ is introduced by the waveplate since, in the CW path, pump photons are rotated, while in the other path, down-converted pairs are rotated. The phase imparted by a waveplate is dependent on the wavelength of light passing through it, and the wavelengths of the pump and down-converted photon pairs are quite different. The waveplate is chosen such that it corresponds to a Nth-order $\frac{\lambda}{2}$ waveplate for the desired wavelengths.

This design does not require any active stabilization. However, for this design to work N must be low at at the pump, signal, and idler wavelengths simultaneously. Depending on the wavelengths chosen, this may be a difficult condition to satisfy, particularly if
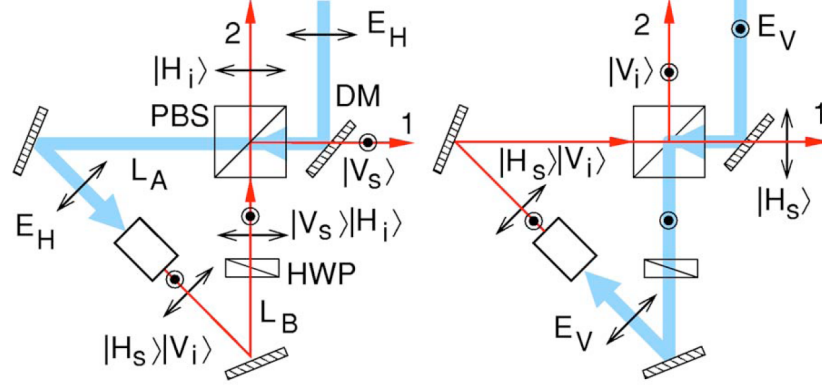
Figure 2.7: **Degenerate wavelength Sagnac interferometer based entanglement source.** This entanglement source is based on a Sagnac interferometer and produces signal and idler photons that are degenerate in wavelength. In (a) H-polarized pump (E_H) propagates counter-clockwise and (b) V-polarized pump (E_V) propagates counter–clockwise [16]. Pump light is shown in light blue and down-converted photon pairs are shown in red. See text for details. Reprinted figure with permission from [16].

non-degenerate signal and idler wavelengths are desired. As a result, fine-tuning the signal and idler wavelengths produced by this design may require changing the pump wavelength or the optics.

2.3 Our Experimental Design

2.3.1 Optics

One way to avoid the constraints imposed by placing a $\frac{\lambda}{2}$ waveplate inside the interferometer is to use two down-conversion crystals, one for each polarization. Figure 2.8 shows the optical setup for our entanglement source. It is a Sagnac interferometer with two orthogonally oriented PPLN crystals. Again, the Sagnac design eliminates fluctuating phases that arise from changes in path length difference due to lab temperature and air fluctuations. However, the use of two crystals does introduce relative phase components of the entangled state that are generated by each crystal. This phase depends on the difference in temperature between the two crystals. This phase is typically constant since

the temperature of the two crystals is controlled to meet the quasi phase matching conditions. A Babinet-Soleil phase compensator (BSC) is included to add a phase between the horizontal and vertical components of the pump light, making it elliptically polarized. This phase is chosen to cancel out the phase arising from the difference in temperature between the two PPLN crystals. A BSC is composed of two wedges of birefringent material that slide over each other to create a waveplate of controllable thickness. It imparts a finely controllable relative phase between the horizontally and vertically polarized components of light travelling through it.
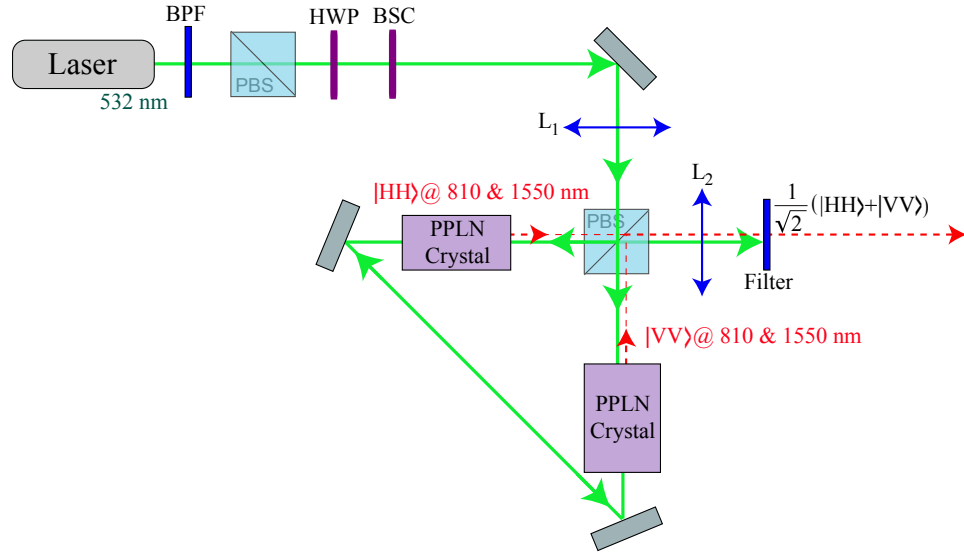
Figure 2.8: **Polarization entanglement source optical setup.** Here BPF is a band–pass filter, HWP is a $\frac{\lambda}{2}$ waveplate, BSC is a Babinet-Soleil phase compensator, L_1 and L_2 are achromatic lenses, Filter is a high(frequency)-pass filter, and PBSs are polarizing beam splitters. See text for details.

Light from an inexpensive 532 nm laser pointer is first attenuated to the desired power using neutral density filters (not shown), wavelength filtered using a broad band-pass filter to remove unwanted spectral lines, and then polarization filtered using a polarizing beam splitter (PBS). A $\frac{\lambda}{2}$ waveplate is placed after the PBS to allow us to rotate the pump to be diagonally polarized while the BSC applies a relative phase between the

horizontal and vertical components. The pump beam is then focused by an achromatic lens L_1 (focal length: 250 mm) to converge at a point that is close to the position of the second crystal encountered in each of the counter-propagating paths. Prior to entering the interferometer, pump light is elliptically polarized such that it is an equal superposition of being horizontally or vertically polarized. When this light encounters the PBS at the entrance of the interferometer, it enters into a superposition of travelling both clockwise (CW) and counter-clockwise (CCW) paths of the interferometer until both paths recombine on the same PBS.

In the CW branch of the interferometer, horizontally polarized pump light first encounters a PPLN crystal that is oriented to meet phase matching conditions for SPDC with vertically polarized light. The pump light will pass through this crystal without significant interaction because the phase matching conditions are not met at this polarization. The second PPLN crystal encountered by pump light in this path is oriented to allow SPDC to occur such that some of the horizontally polarized pump is down-converted to approximately one pair of horizontally polarized photons at non-degenerate wavelengths of 810 nm and 1550 nm. This pair is transmitted through the PBS and exits the source. The CCW path is similar, except that vertically polarized pairs are produced in the second crystal encountered and then reflected into the same output mode as the horizontal pairs from the CW path. After these two paths are recombined by the PBS, an entangled $|\Phi^\phi\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\phi}|VV\rangle\right)$ state results. The phase, ϕ, arises from different indices of refraction in the two crystals (see Section 3.3). After exiting the interferometer through the PBS, a lens, L_2 (focal length: 100 mm), is used to collimate the output of the interferometer and a filter is used to block the 532 nm pump, leaving only the entangled photon pairs.

These pairs are then separated according to wavelength by a dichroic mirror and sent to wavelength specific qubit analyzers consisting of a $\frac{\lambda}{4}$ waveplate, a $\frac{\lambda}{2}$ waveplate, a
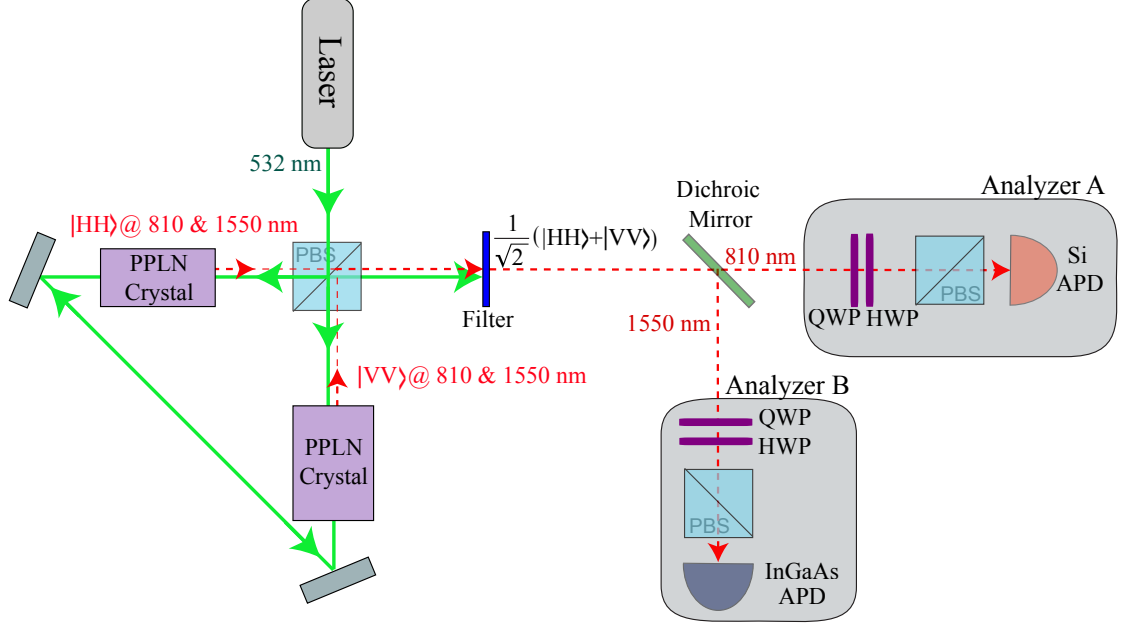
Figure 2.9: **Polarization entanglement source with qubit analyzers.** Entangled states produced by the source are split according to wavelength on a dichroic mirror and distributed to qubit analyzers A and B, which are each composed of a $\frac{\lambda}{4}$ waveplate (QWP), a $\frac{\lambda}{2}$ waveplate (HWP), polarizing beam splitter (PBS) and wavelength specific single photon detectors (Si APD and InGaAs APD). See text for details.

PBS, and wavelength specific detectors, as shown in Figure 2.9. These analyzers allow arbitrary projection measurements to be made on each of the photons. A free running silicon avalanche photo-diode (Si APD) is used in the 810 nm analyzer and a triggered Indium Galium Arsenide (InGaAs) APD is used in the 1550 nm analyzer.

The current lower limit on the size of this interferometer is set by the physical dimensions of the PPLN crystal ovens. This constraint is what made it necessary to use a lens with a focal length of 250 mm for L_1. As a result, the pump beam has a Rayleigh range that is longer than has been shown to be optimal [17]. This suggests that a custom oven design, which would allow for a smaller interferometer, may increase the number of down-conversion events this source produces for a given pump power. Given that the laser pointer provided ample power for the experiments in this thesis, this suboptimal efficiency was not a limiting factor. However, it may be worth addressing in future

designs.

2.3.2 Coincidence Detection

Our experiment uses avalanche photo-diodes (APDs) to detect photons. These APDs are prone to producing spurious "dark counts" at random times that do not correspond to the detection of photons. They may also experience "after pulsing" if charge carriers from a prior detection event are trapped and subsequently released, potentially causing multiple detection signals to arise from a single photon. The probabilities of these happening together determine if an APD can be used in a "free running" mode or must be actively gated.

The Si APD used in this experiment has relatively low probabilities of producing dark counts and after-pulsing, so it can be operated in a free running mode such that it is almost always ready to detect photons. For a short time after a detection the APD must be quenched, resulting in a short period of "dead time" during which the detector is unable to detect photons.

Si APDs have a very low single photon detection efficiency at 1550 nm, so we use an InGaAs APD for this wavelength. The InGaAs APD used in this experiment has a relatively high probability of both producing dark counts and after pulsing, so it had to be actively gated. A trigger signal is sent telling the gating electronics when the APD should be biased above breakdown voltage and thereby become ready to detect photons. Dark counts and after pulsing can be greatly mitigated if the InGaAS APD is biased below its breakdown voltage whenever it does not need to be ready to detect photons.

A "coincidence" detection occurs when the detectors in both qubit analyzers register clicks at times corresponding to the detection of simultaneously created photons. In our experiment the Si APD is used to trigger the InGaAs APD, so a fiber "delay line" must be inserted into the 1550 nm portion of the setup to give this trigger time to propagate

through to the InGaAs detector before the photon reaches it, as shown in Figure 2.10. In addition to showing the delay line, this figure also shows how photons are coupled into fiber after passing through the analyzers' waveplates and PBSs.
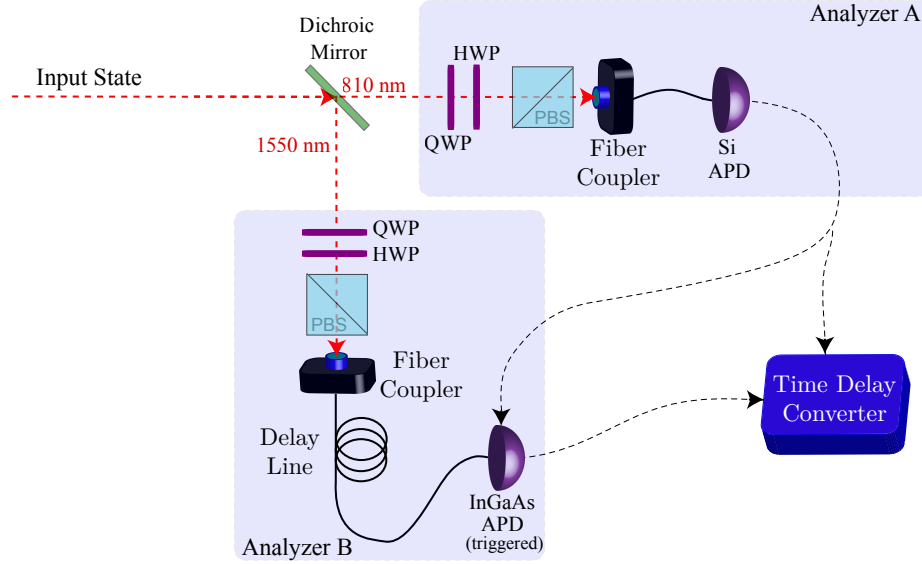
Figure 2.10: **Entanglement source delay, fiber couplings, and electronics.** Figure 2.9 was simplified by leaving out the fiber couplings, delay line, and electronic connections shown here. Photons from each analyzer, after passing through waveplates and PBSs, are coupled into fiber and routed to detectors. The 1550 nm line contains a delay line so that the InGaAs APD may be triggered by detection signals from the Si APD shortly before 1550 nm photons arrive. Both detectors send signals to a TDC that measures the relative time of detection events.

We use a TDC to determine when coincidence detections occur. The TDC operates similarly to a stop-watch. The trigger signal sent by the Si APD to the InGaAs APD is also sent to the TDC, which prompts it to start timing. The InGaAs APD sends two signals to the TDC: "detection" and "gate-out". The gate-out signal is only sent when the InGaAs APD is not recovering from a previous detection event. The TDC records this signal so that the InGaAs APD's dead time won't alter the statistics of our measurements. If the InGaAs APD does detect a photon, the TDC records how much time elapsed between the detection signals from the two detectors. This information is then gathered and collated on a computer that is interfaced with the TDC.

The period of time that elapses between signal and idler detections stemming from the same pair is constant and depends on the length of the optical delay fiber in the 1550 nm analyzer and the electronic delay of the detectors and delay circuits. If the delay between two detections is substantially longer or shorter than this constant delay, then the detections must correspond to events where either a dark count or after pulsing occurred in at least one detector, or photons from two different pairs were detected. This system is able to resolve distinct events that are separated in time by as little as ~ 0.5 ns, so we can disregard events where the period between signal and idler detections differs from the constant delay by more than this.

Chapter 3

Source Characterization

3.1 Visibility

Consider an experiment using a source of polarization entangled photons and qubit analyzers as described in Section 2.3, except, to make things simpler, we assume the entangled state is a perfect $|\phi^+\rangle$ state and the projections made by the analyzers are also perfect. Analyzer A is set to project its part of the entangled system onto a state, $|a_0\rangle$, and analyzer B projects onto states, $|b_i\rangle$, that lay on the great circle around the Bloch sphere that includes $|b_0\rangle = |a_0\rangle$ and $|b_0^\perp\rangle = |a_0^\perp\rangle$, where $|a_0^\perp\rangle$ is orthogonal to $|a_0\rangle$.

We should observe maximal coincidence detections when Analyzer B projectively measures on $|b_0\rangle$ and minimal coincidence detections when Analyzer B projectively measures on $|b_0^\perp\rangle$. We can calculate the visibility from these measurements as:

$$V = \frac{C(|a_0\rangle, |b_0\rangle) - C(|a_0\rangle, |b_0^\perp\rangle)}{C(|a_0\rangle, |b_0\rangle) + C(|a_0\rangle, |b_0^\perp\rangle)}, \tag{3.1}$$

where $C(|a_j\rangle, |b_i\rangle)$ is the number of coincidence detections observed when analyzers A and B project onto states $|a_j\rangle$ and $|b_i\rangle$.

This method of measuring visibility assumes that we can make perfect projections onto the above states. In a practical experiment with imperfect projections, this value for visibility is sensitive to how those projections differ from ideal. The state itself will also be imperfect. Visibility for imperfect states and projections can be more reliably determined if we set one analyzer to project onto a single state, $|a_0\rangle$, and vary the projections made by the other analyzer onto many states, $|b_i\rangle$, that are in a great circle on the Bloch sphere that includes $|b_0\rangle = |a_0\rangle$. We can then calculate visibility by fitting

40

the resulting curve.

For example, say that $|a_0\rangle = |H\rangle$ and $|b_i\rangle = |\theta_B\rangle$ where $|\theta_B\rangle = \cos(\frac{\theta_B}{2})|H\rangle + \sin(\frac{\theta_B}{2})|V\rangle$ as discussed in Section 1.2.1. We can express the combination of states that the qubit analyzers project onto as:

$$
|H\rangle \otimes |\theta_B\rangle = |H\theta_B\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} \cos\frac{\theta_B}{2} \\ \sin\frac{\theta_B}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta_B}{2} \\ \sin\frac{\theta_B}{2} \\ 0 \\ 0 \end{pmatrix}. \tag{3.2}
$$

If we project $|\Phi^+\rangle$ entangled photon pairs onto this combination of states, the probability of coincidences between the qubit analyzers is:

$$
P(|H\rangle, |\theta_B\rangle) = Tr\left(|H\theta_B\rangle\langle H\theta_B|\rho_{\Phi^+}\right)
$$

$$
= Tr\left(\begin{pmatrix} \cos\frac{\theta_B}{2} \\ \sin\frac{\theta_B}{2} \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \cos\frac{\theta_B}{2} & \sin\frac{\theta_B}{2} & 0 & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \right)
$$

$$
= \frac{1}{2}\cos^2\frac{\theta_B}{2}. \tag{3.3}
$$

If we vary θ_B and plot the number of coincidence detections observed, the result will be a sinusoidal curve. We can thereforen fit a theoretical expression to this curve as follows:

$$
C(|H\rangle, |\theta_B\rangle) = V\cos^2(P_2\theta_B - P_2) + P_3, \tag{3.4}
$$

where $C(|H\rangle, |\theta_B\rangle)$ is the coincidence rate observed as θ_B is varied, P_i are fit parameters, and V is visibility. The visibility obtained using this method will generally be more

accurate than deriving a value for V based on just two points corresponding to what we think are orthogonal projections in each analyzer. In general, a maximally entangled state will produce a visibility of one for ideal measurements on any circle on the Bloch sphere.

Even if we find that the visibility resulting from one such measurement is as high as $V = 1$, the presence of entanglement is not conclusively demonstrated. A product state, if properly chosen, can produce identical results for this measurement. For example, say that we perform the same measurements on a product state: $|\psi_P\rangle = |R\rangle \otimes |H\rangle$. We will obtain:

$$P(|H\rangle, |\theta_B\rangle) = Tr\left(|H\theta_B\rangle\langle H\theta_B|\rho_P\right)$$

$$= Tr\left(\begin{pmatrix} \cos\frac{\theta_B}{2} \\ \sin\frac{\theta_B}{2} \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \cos\frac{\theta_B}{2} & \sin\frac{\theta_B}{2} & 0 & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 0 & -\imath & 0 \\ 0 & 0 & 0 & 0 \\ \imath & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right)$$

$$= \frac{1}{2}\cos^2\frac{\theta_B}{2}, \tag{3.5}$$

where $\rho_P = |\psi_P\rangle\langle\psi_P|$ is the density matrix of the above product state. This result corresponds to a visibility of 1 and is identical to what we obtained for a maximally entangled state!

If we choose a set of states on a different great circle on the Bloch the results will be quite different. For example, say that we chose to project onto $|a_0\rangle = |L\rangle$ and $|b_i\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle + e^{\imath\phi}|V\rangle\right)$. The result for a $|\Phi^+\rangle$ state will be a curve with a visibility of one, but the resulting curve for ρ_P would be a constant for all ϕ, resulting in a visibility of zero.

We measured two visibilities from projective measurements onto states that lie in two orthogonal great circles on the Bloch sphere: one that includes $|H\rangle$, $|V\rangle$, $|+\rangle$, and $|-\rangle$
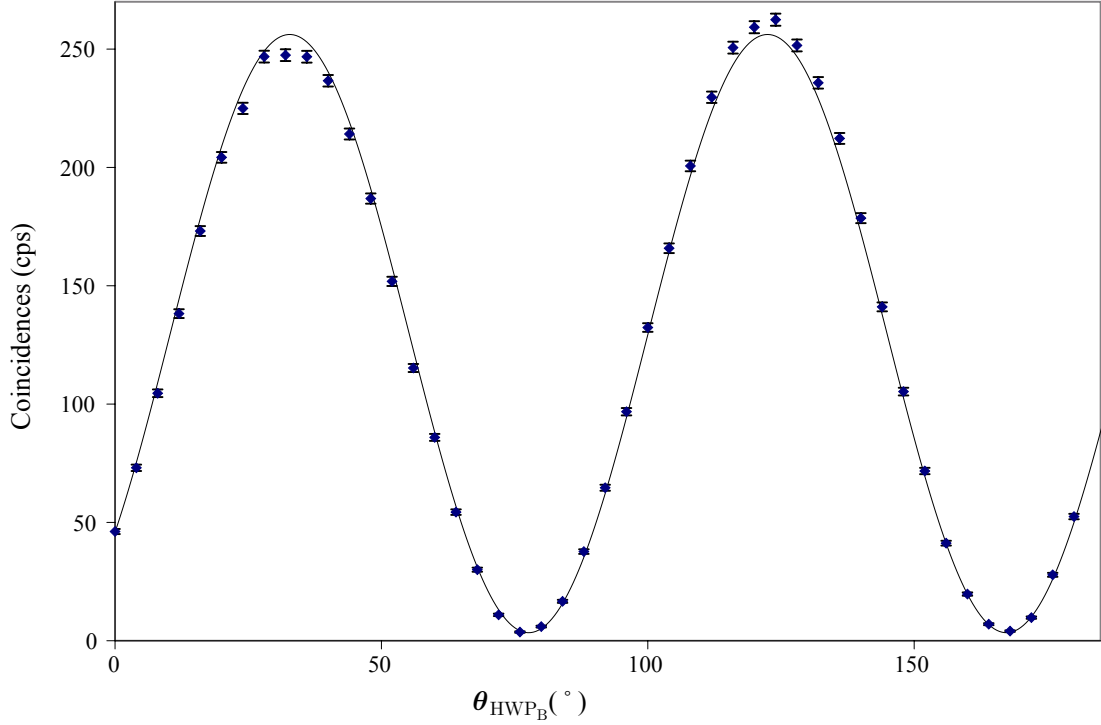
Figure 3.1: **Measurement of Visibility.** Coincidence detections collected with analyzer A set to project on $|R\rangle$ and analyzer B set to project on states on the great circle that includes $|+\rangle$, $|-\rangle$, $|R\rangle$, and $|L\rangle$. The visibility measured by fitting this curve is 97.4 %.

and one that includes $|R\rangle$, $|L\rangle$, $|+\rangle$, and $|-\rangle$. The first set of measurements yielded a visibility of $V = 99.1$ %. The second set of measurements yielded a curve (shown in Figure 3.1) with visibility $V = 97.4$ %.

Even two visibility measurements, as we have just described, are insufficient to prove that the system is entangled. A minimal set of measurements to prove the presence of entanglement would require the first analyzer to project onto four states that are comprised of two orthogonal pairs of states in two bases that are mutually perpendicular on the Bloch sphere. The other analyzer would need to project onto the same pairs of orthogonal states [18]. We will describe a method for quantifying a state's degree of entanglement in Section 3.2.1).

3.2 Quantum State Tomography

Tomography is a process by which an accurate model of an object an be constructed from a set of measurements that individually are not sufficient to describe the object. For example, a 3D model of a statue's surface can be generated by photographing it from several known angles and then constructing the model from those photographs. Similar to a single photograph of the statue, a single projective measurement of a quantum state is insufficient to fully determine the state. Quantum states are altered when they are measured, so Quantum State Tomography (QST) [19] requires many copies of the same state to accurately determine the state. Many different projection measurements are made on these states, similar to photographing the statue from different angles. With the right choice of projection measurements an accurate representation of the original state can then be reconstructed.

To illustrate how QST can be performed, let's start with a single qubit. A single qubit can be expressed in terms of the Stokes parameters:

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{S}\vec{\sigma}\right), \tag{3.6}$$

where:

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \hat{\sigma}_1 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \hat{\sigma}_2 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\sigma}_3 \equiv \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix}, \qquad \tag{3.7}$$

Here, $\mathbb{1}$ is the identity matrix, $\hat{\sigma}_i$ are the Pauli matrices, and the S_i values are Stokes parameters, which are given by $S_i = Tr(\hat{\sigma}_i\rho)$. The Stokes parameters can be obtained by making projective measurements onto the pairs of states $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$, and $\{|L\rangle, |R\rangle\}$ respectively:

$$S_x = P(|H\rangle) - P(|V\rangle),$$

$$S_y = P(|+\rangle) - P(|-\rangle)$$

$$S_z = P(|R\rangle) - P(|L\rangle) \tag{3.8}$$

$$\tag{3.9}$$

where $P(a_i)$ is the probability that the state being measured is projected onto state a_i.

We can find all of the Stokes parameters by making six projective measurements on the Bloch sphere, as shown in Figure 3.2. Once we have the Stokes parameters we can write the density matrix of the state as in Equation 3.6.
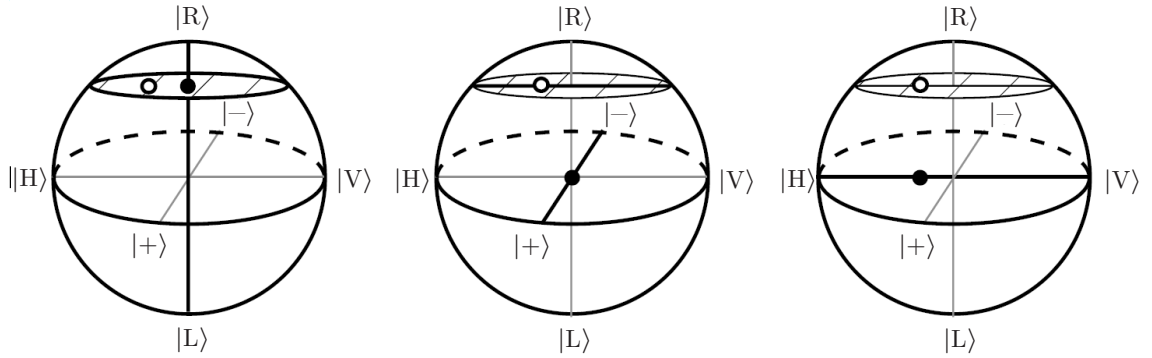
Figure 3.2: **Quantum State Tomography performed on a single qubit.** A sequence of three sets of two orthogonal projection measurements on identical qubits (state to be determined indicated by ○) is sufficient to determine the state of those qubits. The first two projection measurements on $\{|R\rangle$ and $|L\rangle\}$ gives S_2 and restricts the possible states of the qubit to a plane. A second pair of projection measurements on $\{|+\rangle$ and $|-\rangle\}$ S_1 and restricts the possible states to those on a line in that plane. A third pair of projection measurements on $\{|H\rangle$ and $|V\rangle\}$ give S_3 and restricts the possible states of the qubit to a single point in the Bloch sphere that corresponds to the state being measured. Reproduced from [19].

This procedure can be extended to perform tomography on bipartite states. Say that qubits A and B are members of a bipartite state. We can write their state in terms of Stokes parameters as:

$$\rho = \frac{1}{4} \sum_{i_A=0,i_B=0}^{3} S_{i_A,i_B} \left(\hat{\sigma}_{i_A} \otimes \hat{\sigma}_{i_B} \right), \tag{3.10}$$

where the tensor product, $(\hat{\sigma}_{i_A} \otimes \hat{\sigma}_{i_B})$ is a 4x4 matrix formed from the combination of any two of the identity and Pauli matrices from above. There are sixteen Stokes parameters corresponding to S_{i_A,i_B}. Each of these parameters can be calculated from four probabilities, so the most naieve possible way to construct the full density matrix would be to measure 64 probabilities. In practice, many of these measurements would be redundant. It is straightforward to find all of the Stokes parameters for a bipartite system by projectively measuring all sixteen combinations of the states, $|H\rangle$, $|V\rangle$, $|+\rangle$ and $|R\rangle$.

A minimal set of measurements do not necessarily provide the most accurate results. Adding $|-\rangle$ and $|L\rangle$ to the set of states above can improve the accuracy of results considerably. This corresponds to making 36 projective measurements [19]. For our purposes, accuracy was more important than speed, so we performed 36 projective measurements when performing QST.

Measurement uncertainties in real experiments can cause problems if the density matrix is formed directly from the measured Stokes parameters, as above. The result can be a density matrix that does not correspond to a possible physical state. For example, this method could easily result in a density matrix where $Tr(\rho) \neq 1$. Fortunately, there are more advanced methods for analyzing the same experimental data that do not have this problem. One example is Maximum Likelihood [20] (ML), which is a computational method that can be used to find the physically valid density matrix that is most likely to produce the observed projective measurement probabilities. All density matrices derived from QST shown in this thesis were calculated using ML.

Table 3.1 shows a complete set of 36 measurements conducted on our source to measure the density matrix shown in Table 3.2.

Basis States		HWP_A	QWP_A	HWP_B	QWP_B	$C(a_i, b_j)$	$\Delta C(a_i, b_j)$
a_i	b_j	(°)	(°)	(°)	(°)	(cps)	(cps)
$|H\rangle$	$|H\rangle$	0	0	0	0	899.0	4.7
$|H\rangle$	$|V\rangle$	0	0	45	0	10.8	0.5
$|H\rangle$	$|+\rangle$	0	0	22.5	45	474.9	3.4
$|H\rangle$	$|-\rangle$	0	0	-22.5	45	463.0	3.4
$|H\rangle$	$|R\rangle$	0	0	0	45	464.5	3.4
$|H\rangle$	$|L\rangle$	0	0	0	-45	479.6	3.5
$|V\rangle$	$|H\rangle$	45	0	0	0	9.8	0.5
$|V\rangle$	$|V\rangle$	45	0	45	0	919.1	4.8
$|V\rangle$	$|+\rangle$	45	0	22.5	45	454.2	3.4
$|V\rangle$	$|-\rangle$	45	0	-22.5	45	451.9	3.4
$|V\rangle$	$|R\rangle$	45	0	0	45	461.1	3.4
$|V\rangle$	$|L\rangle$	45	0	0	-45	458.6	3.4
$|+\rangle$	$|H\rangle$	22.5	45	0	0	421.2	3.2
$|+\rangle$	$|V\rangle$	22.5	45	45	0	499.7	3.5
$|+\rangle$	$|+\rangle$	22.5	45	22.5	45	906.8	4.8
$|+\rangle$	$|-\rangle$	22.5	45	-22.5	45	17.7	0.7
$|+\rangle$	$|R\rangle$	22.5	45	0	45	443.0	3.3
$|+\rangle$	$|L\rangle$	22.5	45	0	-45	437.8	3.3
$|-\rangle$	$|H\rangle$	-22.5	45	0	0	507.5	3.6
$|-\rangle$	$|V\rangle$	-22.5	45	45	0	410.1	3.2
$|-\rangle$	$|+\rangle$	-22.5	45	22.5	45	22.2	0.7
$|-\rangle$	$|-\rangle$	-22.5	45	-22.5	45	902.3	4.7
$|-\rangle$	$|R\rangle$	-22.5	45	0	45	483.7	3.5
$|-\rangle$	$|L\rangle$	-22.5	45	0	-45	485.1	3.5
$|R\rangle$	$|H\rangle$	0	45	0	0	472.4	3.4
$|R\rangle$	$|V\rangle$	0	45	45	0	455.1	3.4
$|R\rangle$	$|+\rangle$	0	45	22.5	45	438.8	3.3
$|R\rangle$	$|-\rangle$	0	45	-22.5	45	469.9	3.4
$|R\rangle$	$|R\rangle$	0	45	0	45	19.1	0.7
$|R\rangle$	$|L\rangle$	0	45	0	-45	920.1	4.8
$|L\rangle$	$|H\rangle$	0	-45	0	0	484.3	3.5
$|L\rangle$	$|V\rangle$	0	-45	45	0	446.9	3.3
$|L\rangle$	$|+\rangle$	0	-45	22.5	45	456.0	3.4
$|L\rangle$	$|-\rangle$	0	-45	-22.5	45	491.3	3.5
$|L\rangle$	$|R\rangle$	0	-45	0	45	935.4	4.8
$|L\rangle$	$|L\rangle$	0	-45	0	-45	21.4	0.7

Table 3.1: **Tomographic Data.** This table shows raw data collected to find the density matrix shown in Table 3.2. The coincidence rates between the Si avalanche photodiode (APD) and the triggered 1550 nm InGaAs APD ($C(a_i, b_j)$) for each set of qubit analyzer settings are given in average counts per second (cps), as are their one-standard-deviation uncertainties ($\Delta C(a_i, b_j)$). Projections onto basis states a_i and b_j were implemented using one quarter wave plate followed by one half wave plate in each analyzer. These waveplates were set at angles HWP_A, QWP_A, HWP_A, and QWP_A. Data collection time for each point was 30 seconds.
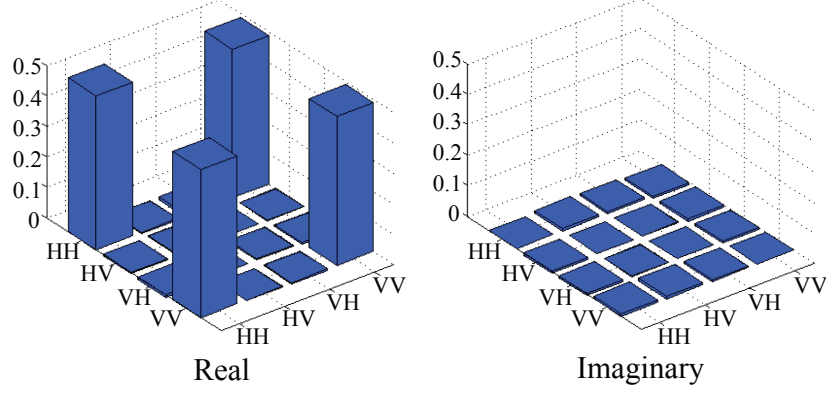
Figure 3.3: **Graphical Density matrix.** Graphical representation of density matrix shown in Table 3.2.

(a) ρ_{Re} (b) ρ_{Im}

	$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$			$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$
$|HH\rangle$	0.5031	0.0056	-0.0196	0.4828		$|HH\rangle$	0.0000	0.0020	0.0046	-0.0007
$|HV\rangle$	0.0056	0.0033	0.0006	0.0113		$|HV\rangle$	-0.0020	0.0000	0.0002	-0.0012
$|VH\rangle$	-0.0196	0.0006	0.0032	-0.0115		$|VH\rangle$	-0.0046	-0.0002	0.0000	-0.0036
$|VV\rangle$	0.4828	0.0113	-0.0115	0.4904		$|VV\rangle$	0.0007	0.0012	0.0036	0.0000

Table 3.2: **Density matrix derived from QST.** Real and imaginary parts of the density matrix generated by maximum likelihood QST for data shown in Table 3.1.

3.2.1 Quantifying Entanglement Quality

There are several ways of quantifying how entangled a state is. We have already discussed "visibility", but it is important to note that the choice of great-circles on the Bloch sphere used to measure visibility can result in different values. "Fidelity" to a known state is another measure of the quality of a state. For example, if we measure a density matrix, ρ, we can say that it's fidelity to a specific pure state, $|\psi\rangle$ is:

$$F = Tr\left(|\psi\rangle\langle\psi|\rho\right). \tag{3.11}$$

The fidelity of a state says nothing about the quantity of entanglement present. A maximally entangled state may have zero fidelity to a different entangled state.

"Tangle" was proposed as a way to quantify the degree of entanglement of a state

without reference to specific bases or states [21]. For a bipartite state, the tangle can be calculated from the state's density matrix, ρ, as follows. First, we calculate $R = \rho \Sigma \rho^T \Sigma$ where ρ^T is the transpose of ρ and

$$\Sigma = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{3.12}$$

We then order the eigenvalues of R in decreasing order such that $r_1 \geq r_2 \geq r_3 \geq r_4$ and calculate the Tangle as:

$$T = (\text{Max}\{0, \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3} - \sqrt{r_4}\})^2 \tag{3.13}$$

Tangle is bounded by $0 \leq T \leq 1$. A state with $T = 0$ is a mixed state with no entanglement and a state with $T = 1$ is maximally entangled.

Tangle is useful for predicting the kind of outcomes we can expect. For example, in Section 4.1.2, we will perform a consistency check by comparing a measured value for the CHSH S-parameter with a prediction based on tangle from the results of QST. It is not completely straightforward to calculate this from fidelity or visibility unless certain conditions are met.

3.3   Phase Stability

The crystals used for SPDC in our entanglement source impart a phase both to pump photons and down-converted signal and idler photon pairs travelling through them. For entanglement sources that use a single SPDC crystal, the phase affects the state globally and is, hence, not observable, i.e. $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ cannot be distinguished from each other and are effectively equivalent states. In our Sagnac interferometer there is one

significant asymmetry in that the $|HH\rangle$ and $|VV\rangle$ photon pairs travel through different birefringent crystals after down-conversion, picking up a relative phase. These crystals were not manufactured at the same time and have slightly different poling periods when kept at the same temperature. They must therefore be maintained at slightly different temperatures for the spectra of down-converted photons from each of the two crystals to maximally overlap (see Section 3.4). As a consequence, the two crystal's have slightly different indices of refraction for a specific wavelength. The state produced by our source of entangled photon pairs must therefore be written as:

$$|\Phi^\phi\rangle = \frac{1}{\sqrt{2}} \left(e^{i\phi_1} |HH\rangle + e^{i\phi_2} |VV\rangle \right) = \frac{1}{\sqrt{2}} \left(|HH\rangle + e^{i\phi} |VV\rangle \right), \qquad (3.14)$$

where $\phi = \phi_2 - \phi_1$ is the relative phase between the $|HH\rangle$ and $|VV\rangle$ pairs. Figure 3.4 shows how this phase changes if we vary the temperature of one of the PPLN crystals. We expect $C(|R\rangle, |L\rangle)$ to correspond to a maximum for a $|\Phi^+\rangle$ state, but it oscillates between being a maximum and a minimum depending on crystal temperature. When $C(|R\rangle, |L\rangle)$ is a minimum on this graph, the relative phase, ϕ is approximately equal to π such that the state produced by the source is actually close to being a $|\Phi^-\rangle = \frac{1}{\sqrt{2}} (|HH\rangle - |VV\rangle)$ state. Note that the tangle of the state is also changing as we vary temperature (see section 3.2.1), resulting in an envelope inside of which the coincidence detections oscillate.

We can control this relative phase by controlling crystal temperature, but spectral overlap (and tangle) also depend on crystal temperature. There is a specific temperature difference between the two PPLN crystals that results in maximum overlap between the spectra of photons produced by the two crystals, yielding a state with the highest tangle. However, this corresponds to a relative phase that is not a convenient value to work with. We would prefer that the phase be zero or π, since these phases would yield a $|\Phi^+\rangle$ or $|\Phi^-\rangle$ state respectively.

Once we observed this effect we added the Babinet-Soleil phase compensator to the
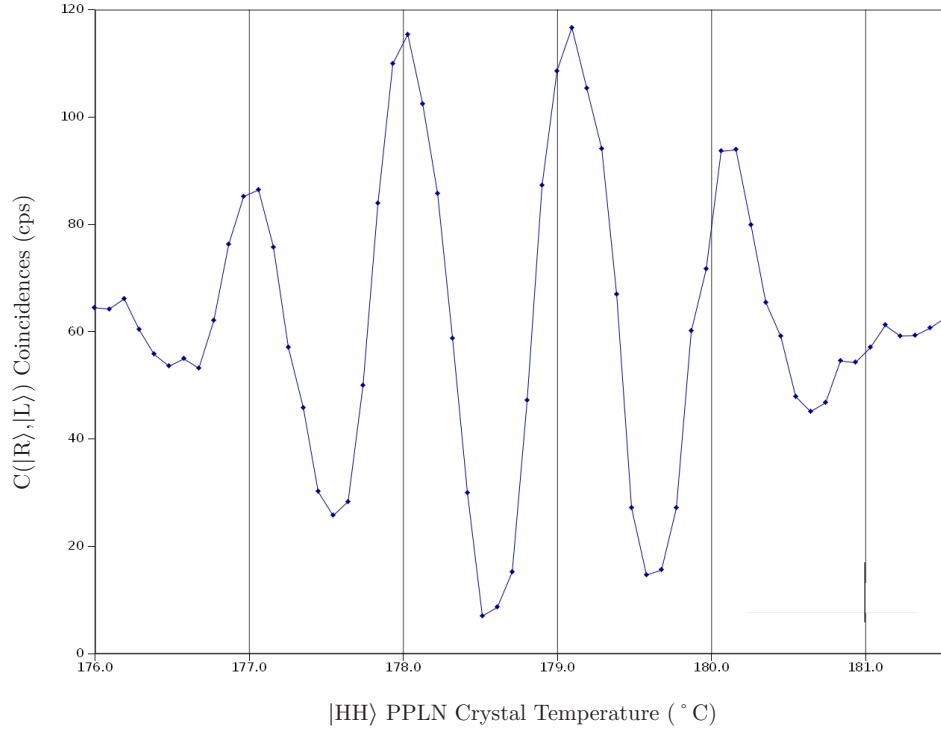
Figure 3.4: **Phase dependence on temperature.** This plot shows coincidence measurements with analyzer A set to project onto $|R\rangle$ and analyzer B set to project onto $|L\rangle$. For a $|\Phi^+\rangle$ entangled state $C(|R\rangle, |L\rangle)$ should be a maximum. The temperature of the PPLN crystal in the CCW path (which produces $|VV\rangle$ pairs) was kept constant while the other PPLN crystal's temperature was varied.

source, as discussed in Section 2.3, which allows us to set the relative phase to be anything we want independently of crystal temperature. For the experiments conducted in this thesis we set the relative phase to $\phi = 0$ in order to obtain a $|\Phi^+\rangle$ state. This is a convenient state to work with, but we could just as easily have set the phase to $\phi = \pi$ to generate a $|\Phi^-\rangle$ state.

3.4   Effects of Spectral Distinguishability

One interesting feature of our entanglement source is its ability to produce a state with a degree of spectral distinguishability that can be manipulated. Adjusting the temper-

ature of one PPLN crystal relative to the other allows us to change the phase-matching conditions in one crystal and alter the spectrum of photons it produces, resulting in a bipartite state with an adjustable degree of overlap between the spectra of the $|HH\rangle$ and $|VV\rangle$ photon pairs. When the spectra do not perfectly overlap, there is additional information available that reveals what crystal a given pair of photons was created in, thus reducing the tangle of the state. We can use this to examine the relationship between entanglement quality and the spectral overlap of photons produced by the two crystals.

Figure 3.5 shows examples of two $\sim 810\ nm$ spectra, one gathered from the $|HH\rangle$ PPLN crystal at $T = 165.2\ °C$ and the other gathered from the $|VV\rangle$ PPLN crystal at $T = 165.70\ °C$. For these temperatures the two spectra have a relatively small overlap, and the value for C produced will be small, but non-zero. The shape of these spectra differs substantially from theoretical models of SPDC, some of which suggest these should be a sinc-squared function [22]. This might be partly due to the shape of the pump spectrum, as shown in Figure 3.6.

To see how tangle is related to spectral overlap, we varied the temperature of the PPLN crystal in the CW path of our entanglement source while keeping the other crystal's temperature constant. This shifted the spectrum of the $|HH\rangle$ component of the state relative to the $|VV\rangle$ component, resulting in different degrees of spectral overlap, O, which we calculate as:

$$O = \int S_{VV}(\lambda)S_{HH}(\lambda)d\lambda. \tag{3.15}$$

where $S_{VV}(\lambda)$ is the the $\sim 810\ nm$ signal spectral density as a function of wavelength, λ, for the SPDC crystal that produces $|VV\rangle$ photons pairs and $S_{HH}(\lambda)$ is the spectral density of the SPDC crystal that produces $|HH\rangle$ photon pairs. We measured the signal spectrum of the $|VV\rangle$ SPDC crystal, which was kept at a constant temperature of $T = 165.70\ °C$ using a temperature controlled oven that is stable to $\pm\ 0.01\ °C$. We also measured spectra of signal photons from the $|HH\rangle$ SPDC crystal at several different temperatures.
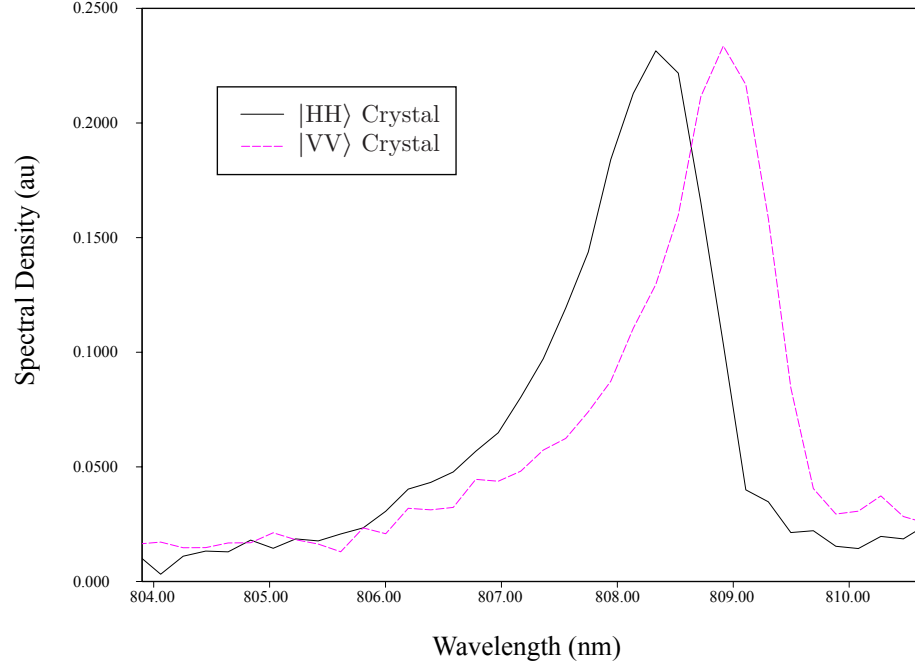
Figure 3.5: **Single photon spectra for two crystals at different temperatures.** This plot shows single photon spectra gathered for $\sim 810\ nm$ photons from the entanglement source's $|VV\rangle$ PPLN crystal at $T = 165.70\ °C$ and from the $|VV\rangle$ PPLN crystal at $T = 165.2\ °C$. Amplitude is in arbitrary units.
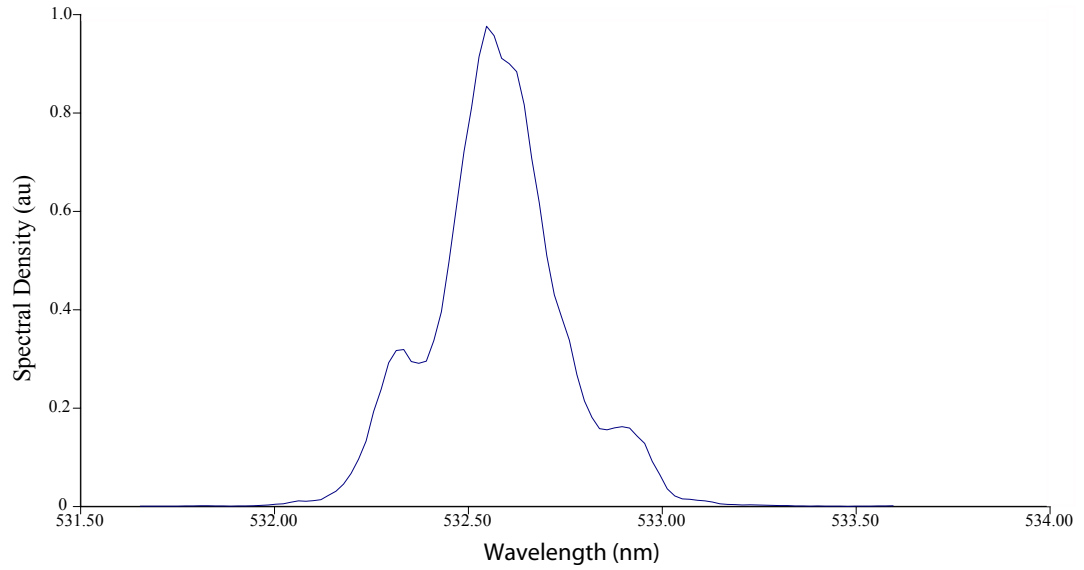
Figure 3.6: **Pump Spectrum.** This plot shows the spectrum measured for the $\sim 532\ nm$ laser pointer used as the pump in our entanglement source.

At each of these temperatures we also performed QST on the resulting bipartite states to find density matrices and associated tangles for each temperature. The results are shown in Figure 3.7 and Table 3.3 and Figure 3.8. In Figure 3.8, note that the off-diagonal term amplitudes increase with spectral overlap.
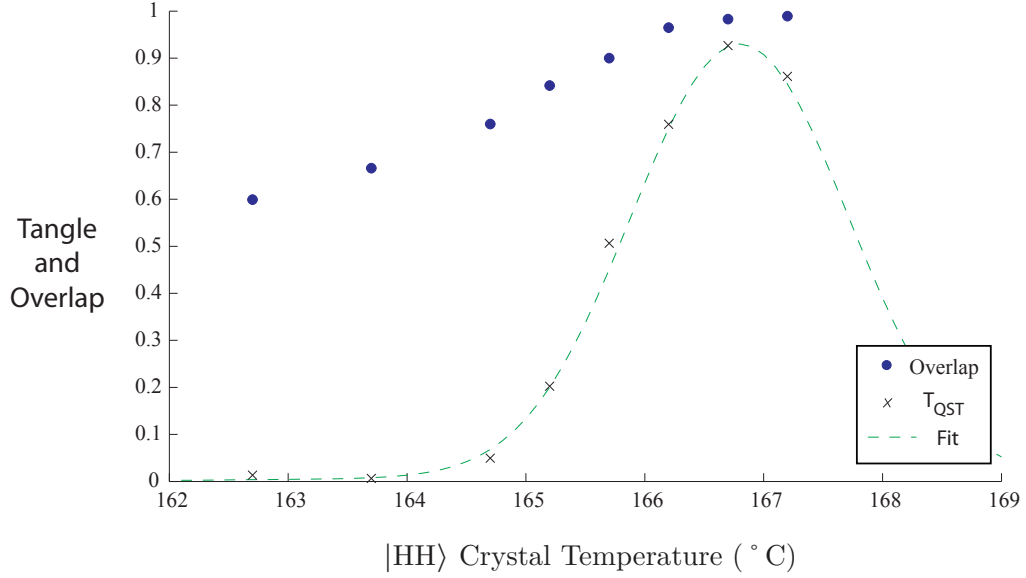
Figure 3.7: **Tangle vs spectral overlap.** This plot shows tangles derived from density matrices (shown in Table 3.3) measured via QST, T_{QST}, as the spectral overlap was changed by varying the temperature of the $|HH\rangle$ PPLN crystal. The $|VV\rangle$ crystal's temperature was kept constant. Also shown is the overlap, O, of the measured spectra.

The spectra for the down-converted signal photons is relatively broad, which is why the overlap is well above zero at $T = 162\ °C$ and climbs slowly to a maximum of nearly one at $T = 167\ °C$. Tangle, on the other hand, is close to zero except where spectral overlap is high.

162.70 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5037	0.0040	-0.0219	0.0458	0.0000	-0.0038	-0.0080	-0.0410
$\|HV\rangle$	0.0040	0.0044	0.0009	0.0154	0.0038	0.0000	0.0028	0.0083
$\|VH\rangle$	-0.0219	0.0009	0.0046	-0.0063	0.0080	-0.0028	0.0000	-0.0129
$\|VV\rangle$	0.0458	0.0154	-0.0063	0.4873	0.0410	-0.0083	0.0129	0.0000

163.70 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.4680	0.0036	-0.0162	0.0241	0.0000	-0.0075	-0.0105	-0.0493
$\|HV\rangle$	0.0036	0.0045	0.0002	0.0111	0.0075	0.0000	-0.0007	0.0064
$\|VH\rangle$	-0.0162	0.0002	0.0535	0.0005	0.0105	0.0007	0.0000	-0.0131
$\|VV\rangle$	0.0241	0.0111	0.0005	0.4741	0.0493	-0.0064	0.0131	0.0000

164.70 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5004	0.0035	-0.0246	0.0808	0.0000	-0.0074	-0.0079	0.0803
$\|HV\rangle$	0.0035	0.0035	-0.0015	0.0199	0.0074	0.0000	0.0004	0.0075
$\|VH\rangle$	-0.0246	-0.0015	0.0044	0.0009	0.0079	-0.0004	0.0000	-0.0118
$\|VV\rangle$	0.0808	0.0199	0.0009	0.4917	-0.0803	-0.0075	0.0118	0.0000

165.20 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5010	0.0056	-0.0171	-0.1925	0.0000	-0.0057	-0.0184	-0.1248
$\|HV\rangle$	0.0056	0.0050	0.0021	0.0151	0.0057	0.0000	-0.0023	0.0157
$\|VH\rangle$	-0.0171	0.0021	0.0048	-0.0045	0.0184	0.0023	0.0000	-0.0048
$\|VV\rangle$	-0.1925	0.0151	-0.0045	0.4892	0.1248	-0.0157	0.0048	0.0000

165.70 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.4970	-0.0059	-0.0234	0.3588	0.0000	0.0067	0.0032	0.0068
$\|HV\rangle$	-0.0059	0.0034	0.0024	0.0173	-0.0067	0.0000	0.0018	-0.0014
$\|VH\rangle$	-0.0234	0.0024	0.0037	0.0046	-0.0032	-0.0018	0.0000	-0.0064
$\|VV\rangle$	0.3588	0.0173	0.0046	0.4958	-0.0068	0.0014	0.0064	0.0000

166.20 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5108	0.0028	-0.0130	-0.4400	0.0000	0.0007	-0.0243	0.0098
$\|HV\rangle$	0.0028	0.0047	-0.0013	0.0143	-0.0007	0.0000	-0.0044	0.0144
$\|VH\rangle$	-0.0130	-0.0013	0.0057	-0.0060	0.0243	0.0044	0.0000	-0.0093
$\|VV\rangle$	-0.4400	0.0143	-0.0060	0.4788	-0.0098	-0.0144	0.0093	0.0000

166.70 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5017	-0.0069	-0.0051	0.3879	0.0000	0.0089	0.0068	-0.2917
$\|HV\rangle$	-0.0069	0.0041	0.0024	-0.0018	-0.0089	0.0000	0.0033	-0.0009
$\|VH\rangle$	-0.0051	0.0024	0.0042	-0.0013	-0.0068	-0.0033	0.0000	-0.0085
$\|VV\rangle$	0.3879	-0.0018	-0.0013	0.4900	0.2917	0.0009	0.0085	0.0000

167.20 °C

	ρ_{Re}					ρ_{Im}		
	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5072	-0.0010	-0.0176	-0.3646	0.0000	0.0038	-0.0226	0.2922
$\|HV\rangle$	-0.0010	0.0054	-0.0003	0.0191	-0.0038	0.0000	-0.0035	0.0064
$\|VH\rangle$	-0.0176	-0.0003	0.0042	-0.0039	0.0226	0.0035	0.0000	-0.0154
$\|VV\rangle$	-0.3646	0.0191	-0.0039	0.4832	-0.2922	-0.0064	0.0154	0.0000

Table 3.3: **Tangle versus Spectral Overlap Density Matrices.** Density matrices measured as $|HH\rangle$ SPDC crystal temperature (shown above) was varied. Phase was adjusted for maximal fidelity to a $|\Phi^+\rangle$ (positive values for off-diagonal terms $|HH\rangle\langle VV|$ and $|VV\rangle\langle HH|$) or $|\Phi^-\rangle$ (negative values for off-diagonal terms) state. The $|VV\rangle$ SPDC crystal temperature was kept at a constant 165.70 °C.
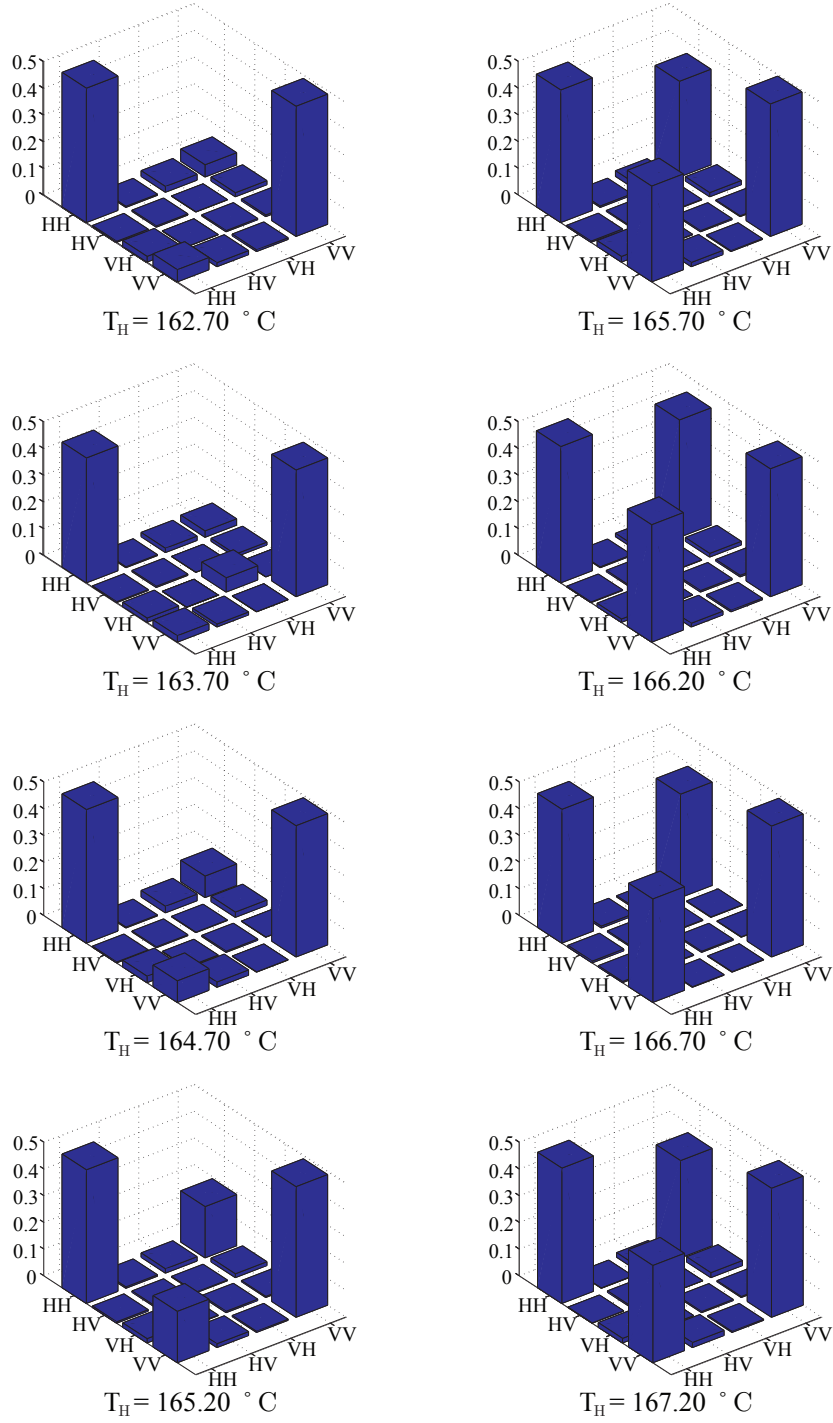
Figure 3.8: **Tangle versus Spectral Overlap Graphical Density Matrices.** This plot shows magnitudes of the real components of the density matrices shown in Table 3.3.

Chapter 4

Tests of CHSH Bell, Beautiful Bell, and Leggett Models

As we have discussed, entanglement is a purely quantum phenomenon without a classical analogue. A source of high quality entanglement and a pair of qubit analyzers allow us to perform a wide range of experiments that provide insight into where classical theories of electromagnetism and mechanics, or models sharing some of their properties, break down and quantum theory becomes necessary to explain our observations of nature. It is in these situations where we can both test the predictions of quantum theory and look for clues about properties that alternative theories must have if they are to be valid. One way of doing this is to construct alternative theories that make predictions that are not in full agreement with those of quantum theory. We can then perform experimental tests of these predictions to reveal if those models, or QT itself, are in agreement with observations. In this chapter we will discuss two different experimental tests of Bell inequalites and a test of the Leggett model.

4.1   The CHSH Bell Inequality

4.1.1   Background

In thermodynamics, we can describe the pressure a gas exerts on the walls of a box containing it as the result of gas molecules continually bouncing off the inside of the box. Meanwhile, the temperature of the gas is a measure of how fast the molecules are moving. Thus, we can make a conceptual link between temperature and pressure. This link is not obvious without the concept of gas molecules. After the EPR paper was published it was thought that, similar to the concept of gas molecules, there might exist some underlying

mechanic of nature that was not a part of QT. A more complete theory, such as one based on LHVs, might allow us to recover the ability to make deterministic predictions, an ability we are used to having when dealing with matter on a macroscopic scale. In 1964 John S. Bell proposed an inequality that showed there was a conflict between predictions made by LHV models and those made by QT [5]. Thus, experimentally testing and violating Bell's inequality would show that LHV models are not valid.
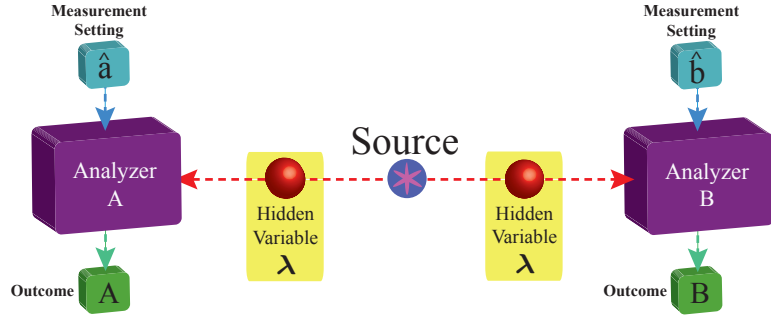
Figure 4.1: **LHV Models.** Bell considered deterministic LHV models that offer predictions for the outcome of measurements performed on an EPR pair of particles that are split and sent to two qubit analyzers. Each of these analyzers measures its respective particle according to measurement settings $\hat{a}$ and $\hat{b}$, producing output results A and B respectively. The hidden variable λ is accessible by both analyzers when evaluating their particles.

Figure 4.1 shows a way to construct such a test of Bell's inequality. Entangled pairs of particles are split between two analyzers, Alice (A) and Bob (B), who each projectively measure their particles in bases $\hat{a}$ and $\hat{b}$ respectively, producing outputs A and B. We assume that Alice and Bob are separated such that their measurement events are outside of each other's light cones. In no reference frame could Alice's measurement have an effect on Bob's without violating causality. This leads to the assumption that Alice's result cannot depend on Bob's measurement setting or result, and vice-versa: $A(\hat{a}, \hat{b}, \lambda : B) = A(\hat{a}, \lambda)$ and $B(\hat{b}, \hat{a}, \lambda : A) = B(\hat{b}, \lambda)$. The Bell inequality does not restrict what form λ can take except to assert that it is strictly local and cannot be used to introduce non-local effects into the system that would violate the above conditions.

The CHSH Bell inequality was proposed in 1969 by John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt as a way to make experimental tests of LHV models [23] feasible. It describes a parameter, S, that can be experimentally determined. If an experimental value for the S parameter were to exceed the bound that CHSH derives for local hidden variable (LHV) theories, then all LHV theories would be invalidated. This inequality was first violated experimentally in 1972 by Stuart Freedman and John Clauser [24], but only marginally. The first strong violation of the CHSH inequality was in 1981 by Alain Aspect, Philippe Grangier, and Gérard Roger in an experiment using polarized entangled photons generated by cascaded atomic transitions [26].

In order to find an experimental value for the CHSH Bell S-Parameter, Alice and Bob must perform the experiment shown in Figure 4.1 in several different measurement bases. After an entangled pair is distributed between them, Alice and Bob each project their halves onto one of two orthogonal states. Figure 4.2 shows one possible set of measurement bases that can be used to test the CHSH Bell inequality. For example, if Alice projectively measures in basis $\hat{a}_1$, her result would be one of the two orthogonal basis vectors in $\hat{a}_1$: a_1 or $a_1^\perp$. She would assign a value to A of $+1$ or -1, respectively, for each of these results.

In the CHSH inequality, Alice and Bob each measure in one of two bases, chosen uniformly and at random, so there are four possible combinations of bases. For each combination of bases, $\hat{a}_i = \{a_i, a_i^\perp\}$ and $\hat{b}_j = \{b_j, b_j^\perp\}$, Alice and Bob must measure the correlation coefficient, which is a measure of how strongly correlated their results are. In our experiment we use one detector in each of Alice and Bob's analyzers, so we make projection measurements onto each basis vector separately. We can then form the correlation coefficient as:

$$E(\hat{a}_i, \hat{b}_j) = P(a_i, b_j) + P(a_i^\perp, b_j^\perp) - P(a_i^\perp, b_j) - P(a_i, b_j^\perp) \tag{4.1}$$
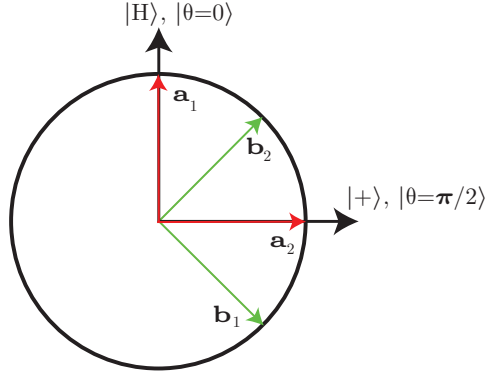
Figure 4.2: **CHSH Measurement Bases.** One set of optimal measurement bases for the CHSH Bell inequality when using $|\Phi^+\rangle$ is shown here on the equator of the Bloch sphere. Only one vector for each basis is shown. The orthogonal vector associated with each basis is rotated by π from the vector shown. For example, if Alice measures in basis $\hat{a}_1$ she would project onto one of two states: $\{a_1, a_1^\perp\}$ where $a_1^\perp = a_1 + \pi$.

where $P(a_i, b_j) = \frac{C(a_i,b_j)}{\sum_{i,j} C(a_i,b_j)}$ and $C(a_i, b_j)$ is the number of "coincidence" detections, or number of times where the detectors held by Alice and Bob registered a click simultaneously, observed when Alice projectively measures along the a_i basis vector and Bob along the b_j vector.

Once all four correlation coefficients have been measured by Alice and Bob, we can calculate the Bell S parameter as:

$$S = E(\hat{a}_1, \hat{b}_1) - E(\hat{a}_1, \hat{b}_2) + E(\hat{a}_2, \hat{b}_1) + E(\hat{a}_2, \hat{b}_2). \tag{4.2}$$

The CHSH bell inequality predicts that S must fall within the range: $-2 \le S \le 2$. QT, on the other hand, predicts that we obtain a result of $-2\sqrt{2} \le S \le 2\sqrt{2}$. We say that we have violated the CHSH Bell inequality when we experimentally measure the S parameter and find it to be outside the bounds predicted by the CHSH inequality. Note that the bounds predicted by QT are a maximal violation that only occurs for a perfect experiment and an optimal set of measurement bases, which vary depending on the entangled state being used. It is possible to measure a value for the CHSH S parameter

that is less than $2\sqrt{2}$ even with a perfect experiment if sub-optimal measurement settings are chosen. For a $|\Phi^+\rangle$ state one set of optimal measurement settings (also shown in Figure 4.2, is:

$$a_1 \rightarrow |\theta = 0\rangle$$
$$a_2 \rightarrow |\theta = \frac{\pi}{2}\rangle$$
$$b_1 \rightarrow |\theta = \frac{3\pi}{4}\rangle$$
$$b_2 \rightarrow |\theta = \frac{\pi}{4}\rangle$$

A minimal violation of the CHSH Bell inequality requires an entanglement visibility of roughly 71 %.

4.1.2 Results

In our measurements, coincidence detections were collected over a period of 40 seconds for each combination of basis states. Our raw data is shown in Table 4.1. We measured a value of $S = 2.757 \pm 0.008$. The uncertainty is based on Poissonian statistics. Using QST (see section 3.2) we measured a density matrix with a tangle of $T = 0.884$ immediately before this measurement. Based on this we would expect a maximum S parameter value of $S_{max} = 2\sqrt{1+T} = 2.75$, which is consistent with the measured value. We can say with high confidence that our measured value for the CHSH Bell S parameter violates the classical bound of 2 and invalidates all LHV models.

Table 4.2 shows a density matrix determined by performing QST shortly before the CHSH Bell inequality was tested.

Bases	Basis Vectors		HWP_A	HWP_B	$C(a_i, b_j)$	$E(\hat{a}_i, \hat{b}_j)$	$\Delta E(\hat{a}_i, \hat{b}_j)$
	a_i	b_j	(°)	(°)	(cps)	()	()
$\{\hat{a}_1, \hat{b}_1\}$	a_1	b_1	0	11.25	432.38		
	a_1	$b_1^\perp$	0	56.25	80.38	0.6813	0.0018
	$a_1^\perp$	b_1	45	11.25	84.40		
	$a_1^\perp$	$b_1^\perp$	45	56.25	436.77		
$\{\hat{a}_1, \hat{b}_2\}$	a_1	b_2	0	33.75	75.42		
	a_1	$b_2^\perp$	0	78.75	441.00	-0.7042	0.0017
	$a_1^\perp$	b_2	45	33.75	448.27		
	$a_1^\perp$	$b_2^\perp$	45	78.75	78.95		
$\{\hat{a}_2, \hat{b}_1\}$	a_2	b_1	22.5	11.25	441.43		
	a_2	$b_1^\perp$	22.5	56.25	71.80	0.7088	0.0017
	$a_2^\perp$	b_1	67.5	11.25	79.70		
	$a_2^\perp$	$b_1^\perp$	67.5	56.25	447.45		
$\{\hat{a}_2, \hat{b}_2\}$	a_2	b_2	22.5	33.75	434.82		
	a_2	$b_2^\perp$	22.5	78.75	87.78	0.6632	0.0018
	$a_2^\perp$	b_2	67.5	33.75	88.30		
	$a_2^\perp$	$b_2^\perp$	67.5	78.75	434.65		

Table 4.1: **CHSH Inequality Measurement Settings and Data.** This table shows raw data collected to find $S = 2.757 \pm 0.008 > 2$. $HWP_{A/B}$ are the half waveplate settings that realize the measurements corresponding to basis vectors a_i and b_j. $C(a_i, b_j)$ is the coincidence rate measured with settings a_i and b_j. $E(\hat{a}_i, \hat{b}_j)$ is the correlation coefficient with settings $\hat{a}_i$ and $\hat{b}_j$. Data collection time for each point was 40 seconds. Uncertainties are derived from Poissonian statistics.

(a) ρ_{Re} (b) ρ_{Im}

	$\langle HH \vert$	$\langle HV \vert$	$\langle VH \vert$	$\langle VV \vert$		$\langle HH \vert$	$\langle HV \vert$	$\langle VH \vert$	$\langle VV \vert$
$\vert HH \rangle$	0.5031	0.0056	-0.0196	0.4828	$\vert HH \rangle$	0.0000	0.0020	0.0046	-0.0007
$\vert HV \rangle$	0.0056	0.0033	0.0006	0.0113	$\vert HV \rangle$	-0.0020	0.0000	0.0002	-0.0012
$\vert VH \rangle$	-0.0196	0.0006	0.0032	-0.0115	$\vert VH \rangle$	-0.0046	-0.0002	0.0000	-0.0036
$\vert VV \rangle$	0.4828	0.0113	-0.0115	0.4904	$\vert VV \rangle$	0.0007	0.0012	0.0036	0.0000

Table 4.2: **CHSH Inequality Test Density matrix.** Real and imaginary parts of the density matrix generated by maximum likelihood QST performed shortly before testing the CHSH Bell inequality.

4.1.3 Discussion

The violation of the CHSH Bell inequality seems to settle the debate over whether or not there exists a model using LHVs that accurately describes nature. However, there exist loopholes which are not closed in this experiment. For example, the single photon detectors used in this experiment are far from 100 % efficient. We are unable to measure all photons emitted by our source. One could argue that a conspiratorial agent is selecting which photons are detected in such a way as to produce misleading results [27]. Detectors with an efficiency greater than 82.8 % are needed to close this loophole [28] for tests of Bell inequalities that do not require supplementary assumptions. This loophole was closed in an experiment using trapped ions [29], but detection efficiencies this high are challenging to achieve using photonic states. However, a heralded source of entangled photons with a system detection efficiency of 83 % was recently demonstrated [30], which suggests closing this loophole with photonic states is now feasible.

Another loophole is the Locality loophole [31]. We assume that the measurement settings of one analyzer cannot effect the generation of entangled pairs of particles or the outcome of measurements in the other analyzer, but we cannot be sure of this unless measurements are chosen randomly, the two analyzers are widely separated, and the timing of the measurements is precise enough for us to conclude that faster-than-light interactions would be necessary for such an influence to occur. The first attempt to experimentally close this loophole [32] was in 1982, but this experiment is not considered definitive because the measurements settings were varied in a predictable rather than random way. The first experiment that definitively closed the locality loophole by widely separating the two qubit analyzers and using a true random number generator to select measurements settings was conducted in 1998 [33].

Violating a CHSH Bell inequality with all loopholes simultaneously closed remains a challenging problem but, since all loopholes have been closed individually, it is reasonable

to expect that we should not be misled by results from experiments that do not close all possible loopholes simultaneously.

The CHSH inequality has been violated experimentally many times before this, but it remains a useful test to perform when characterizing an entanglement source. A CHSH Bell inequality test is a indisputable witness for the presence of entanglement.

4.2 The Beautiful Bell Inequality

4.2.1 Background

The beautiful Bell inequality [34] was proposed by H. Bechmann-Pasquinucci and Nicolas Gisin in 2003 [35] and expanded upon by Gisin in 2008 [36]. These papers describe a family of inequalities that provide a way to test LHV models in arbitrary dimension Hilbert spaces with varied numbers of measurement bases per analyzer. The CHSH Bell inequality is a specific case of the beautiful Bell inequality for two dimensions and two measurement bases.

The beautiful Bell inequality is based on a game shown in Figure 4.3. This game is parameterized by two numbers: m and n. Bob receives as input a list of n values: $\{y_0, y_1, ...y_{n-1}\}$ where $y_i \in \{0, 1, ..., m-1\}$. Alice receives as input one number: x where $x \in \{0, 1, ..., n-1\}$. Alice must guess what y_x Bob was given and output her guess as a. Bob outputs b which is a boolean value that, if True, indicates the current round of the game should counted toward Alice and Bob's score. If b is False this indicates the result should be ignored. If Alice guesses correctly and the round counts, Alice and Bob both gain a point. If Alice's guess is not correct and the round counts, they both lose a point. The goal is for Alice and Bob to cooperatively maximize their score. This game can be played either using the resources provided by LHV models or, alternatively, using the resources of QT.

In the QT strategy Alice and Bob share entangled states whose halves are described in an m-dimensional Hilbert space. Optimal strategies exist for both LHV models and QT, so we can place upper bounds on the maximum score that Alice and Bob can achieve over a large number of trials for Bell's model and for QT. If we design an experiment to implement the strategy that takes advantage of QT and the result exceeds the bound for the LVH models, we will have invalidated LHV models.
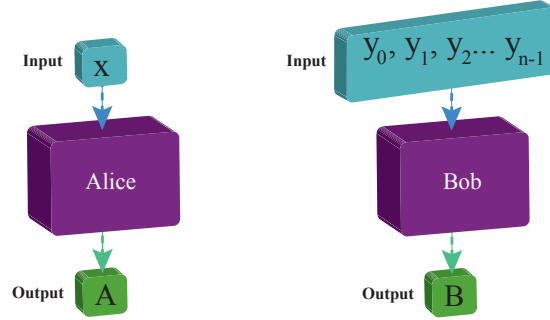
Figure 4.3: **Beautiful Bell inequality.** The Beautiful Bell inequality is derived from a game in which Alice receives an input, x, and must try to guess the xth element in the list of y_i values that Bob receives as input. Bob's chooses output B to specify whether or not the result of the round will be counted towards the total score for Alice and Bob. See text for details.

The CHSH Bell inequality is the $m = 2$, $n = 2$ case of the beautiful Bell inequality. Our source limits us to testing only the $m = 2$ cases, but we can increase n to $n = 3$ by adding measurement bases to Alice and Bob. Figure 4.4 shows the measurement basis required for optimal violation of the $m = 2$, $n = 3$ beautiful Bell inequality. Alice must measure in three mutually orthogonal bases while Bob must measure in four bases such that the average overlap between each of his measurements and each of Alice's measurements is minimized. Note that settings for an optimal violation are more challenging to measure than those for the CHSH Bell inequality since optimal violation requires projective measurements onto states that are not all in a single great circle on the Bloch sphere, as they are for an optimal violation of the CHSH Bell inequality. For this par-

ticular case of the beautiful Bell inequality, a suboptimal violation can still be obtained if the projective measurement states are chosen to be on a single great circle.
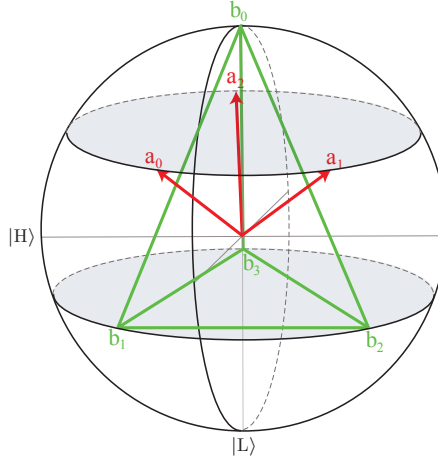
Figure 4.4: **Beautiful Bell measurement bases.** Alices measurement settings $\{\hat{a}_0, \hat{a}_1, \hat{a}_2\}$ correspond to mutually orthogonal vectors on the Bloch sphere. Bob's bases $\{\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3\}$ form a tetrahedron such that the average distance between Bob's settings and Alice's settings are maximized [36]. Only one basis vector (e.g. a_1 from $\hat{a}_1 = \{a_1, a_1^\perp\}$) from each basis is shown.

Similar to how we measure the CHSH inequality, we measure correlation coefficients for each combination of bases, which we then use to calculate the beautiful Bell S-parameter, S_{BB}, as:

$$S_{BB} = E(\hat{a}_0, \hat{b}_0) + E(\hat{a}_0, \hat{b}_1) - E(\hat{a}_0, \hat{b}_2) - E(\hat{a}_0, \hat{b}_3) + \tag{4.3}$$

$$E(\hat{a}_1, \hat{b}_0) - E(\hat{a}_1, \hat{b}_1) + E(\hat{a}_1, \hat{b}_2) - E(\hat{a}_1, \hat{b}_3) + \tag{4.4}$$

$$E(\hat{a}_2, \hat{b}_0) - E(\hat{a}_2, \hat{b}_1) - E(\hat{a}_2, \hat{b}_2) + E(\hat{a}_2, \hat{b}_3) \tag{4.5}$$

The beautiful Bell inequality predicts a bound of $S_{BB-Bell} \leq 6$ for LHV models while QT predicts a maximal value of $max(S_{BB-QT}) \leq 4\sqrt{3} = 6.928$. For a set of projective measurements onto states that are in a single great circle we expect $S_{BB-QT} \leq 2\sqrt{5} = 6.472$. A minimal violation of the beautiful Bell inequality requires an entanglement visibility of roughly 87 %.

4.2.2 Results

Table 4.3 shows correlation coefficients gathered that allowed us to produce a value of $S_{BB} = 6.67 \pm 0.08$. This value is greater than the Bell limit of 6, so we consider this to be a violation of LHV models. Note that the measured S_{BB} is also greater than the limit of 6.472, which corresponds to the maximum violation if the projective measurements are onto states that are on a single great circle of the Bloch sphere.

Bases	$E(\hat{a}_i, \hat{b}_j)$	$\Delta E(\hat{a}_i, \hat{b}_j)$	Bases	$E(\hat{a}_i, \hat{b}_j)$	$\Delta E(\hat{a}_i, \hat{b}_j)$
$\{\hat{a}_0, \hat{b}_0\}$	0.5742	0.0061	$\{\hat{a}_1, \hat{b}_2\}$	0.5763	0.0060
$\{\hat{a}_0, \hat{b}_1\}$	0.5247	0.0062	$\{\hat{a}_1, \hat{b}_3\}$	-0.5833	0.0061
$\{\hat{a}_0, \hat{b}_2\}$	-0.5641	0.0062	$\{\hat{a}_2, \hat{b}_0\}$	0.6124	0.0061
$\{\hat{a}_0, \hat{b}_3\}$	-0.5678	0.0061	$\{\hat{a}_2, \hat{b}_1\}$	-0.6255	0.0061
$\{\hat{a}_1, \hat{b}_0\}$	0.5446	0.0061	$\{\hat{a}_2, \hat{b}_2\}$	-0.5039	0.0061
$\{\hat{a}_1, \hat{b}_1\}$	-0.5307	0.0061	$\{\hat{a}_2, \hat{b}_3\}$	0.4645	0.0061

Table 4.3: **Beautiful Bell Measurement Settings and Data.** This table shows raw data collected to find $S_{BB} = 6.67 \pm 0.08 > 6$. $E(\hat{a}_i, \hat{b}_j)$ is the correlation coefficient measured with settings $\hat{a}_i$ and $\hat{b}_j$. Four coincidence measurements (not shown) consisting of 40 second samples were recorded for each correlation coefficient. Uncertainties are derived from Poissonian statistics.

(a) ρ_{Re}

	$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$
$|HH\rangle$	0.5085	0.0085	-0.0151	0.4773
$|HV\rangle$	0.0085	0.0028	-0.0006	0.0145
$|VH\rangle$	-0.0151	-0.0006	0.0038	-0.0075
$|VV\rangle$	0.4773	0.0145	-0.0075	0.4848

(b) ρ_{Im}

	$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$
$|HH\rangle$	0.0000	0.0028	-0.0027	-0.0337
$|HV\rangle$	-0.0028	0.0000	0.0028	0.0036
$|VH\rangle$	0.0027	-0.0028	0.0000	-0.0045
$|VV\rangle$	0.0337	-0.0036	0.0045	0.0000

Table 4.4: **Beautiful Bell Density matrix.** Real and imaginary parts of the density matrix generated by maximum likelihood QST performed shortly after the beautiful Bell inequality was tested.

Table 4.4 shows a density matrix determined by performing QST just after the test of the beautiful Bell inequality.

4.2.3 Discussion

We are not aware of any published violation of the $n = 3$, $m = 2$ (or higher dimension) beautiful Bell inequality, so this appears to be the first time it has been experimentally violated. Our value for S_{BB} also exceeds the quantum limit for projections onto states that are all on a single great circle of the Bloch sphere, which confirms that our qubit analyzers are indeed projecting onto states that are not all in a single great circle on the Bloch sphere.

4.3 The Leggett Inequality

4.3.1 Background

Experimental violations of the CHSH and beautiful Bell inequalities showed that LHV models are not consistent with experimental observations of nature and a local, fully deterministic, alternative to quantum theory based on hidden variables does not exist. If this is the case, then could there be a alternative probabilistic theory to QT that provides better predictions than quantum theory? In 2003, Anthony J. Leggett proposed a non-local hidden variable model [37] that differs from LHV models in two ways. First, it permits some non-local interactions. Second, unlike the deterministic LHV models we have considered previously, the Leggett model makes probabilistic predictions. These predictions can have a higher probability of correctness than those of QT. The Leggett model is interesting because it permits non-local hidden variables(NLHV) and has a predictive power (see Section 5) that lies between that of deterministic LHV models and QT. Also, experiments that rule out the LHV models do not automatically rule out the Leggett model.

Say that Alice and Bob are two parties who perform projective measurements in bases $\hat{a}$ and $\hat{b}$ and obtain outcomes A and B. As previously discussed, LHV models assume

that each party's outcomes may depend on a LHV, λ, but not on the other party's measurement setting nor outcome: $A(\hat{a}, \hat{b}, \lambda : B) = A(\hat{a}, \lambda)$ and $B(\hat{b}, \hat{a}, \lambda : A) = B(\hat{b}, \lambda)$. The Leggett model keeps the assumption that Alice's and Bob's measurement outcomes do not depend on the other party's measurement setting, since faster than light signalling would be permitted if Alice could change her measurement setting and instantly alter Bob's outcomes. However, the Leggett model does not assume that the outcomes are independent of each other: $A(\hat{a}, \hat{b}, \lambda : B) = A(\hat{a}, \lambda : B)$ and $B(\hat{b}, \hat{a}, \lambda : A) = B(\hat{b}, \lambda : A)$. This is why we say the Leggett model is non-local but non-signalling.

The Leggett model also restricts what λ can be since "allowing λ to be totally arbitrary in character" would permit a non-local hidden variable model to reproduce the predictions of quantum mechanics since, without restrictions on λ, the Leggett model would cover a general class of non-local theories that QT would be a subset of. The Leggett model restricts λ so that it carries information equivalent to two hidden spin vectors, one for each photon in the pair distributed to Alice and Bob. i.e. $\lambda = |\mathbf{u}\rangle_A \otimes |\mathbf{v}\rangle_B$. Alice and bob may projectively measure the hidden spin states associated with their particles. This results in a model that makes predictions with a higher probability of correctness than QT, but less than those of LHV models.

The Leggett inequality was first violated in 2007 [38]. This experiment assumed rotational invariance of their bases. That is to say, they measured in a set of bases and assumed that, if they were to rotate all bases by the same arbitrary amount, they would obtain the same results. QT says this would be true for an ideal state with perfect arbitrary-basis analyzers, but this may not be true in a real-world experiment that may have an imperfect state or biases in the measurements. This form of Leggett inequality also requires an entanglement visibility of at least 97.4 %. In 2007 and 2008 there were more violations of the Leggett inequality [39, 40, 41] that did not require this assumption and an alternate derivation of the Leggett inequality was derived and tested in 2008 by

Branciard, et al. [42]. The entanglement visibility required for a minimal violation of this inequality is 94.3 %, but it does require projective measurements onto states that lay outside of a single great circle of the Bloch sphere.
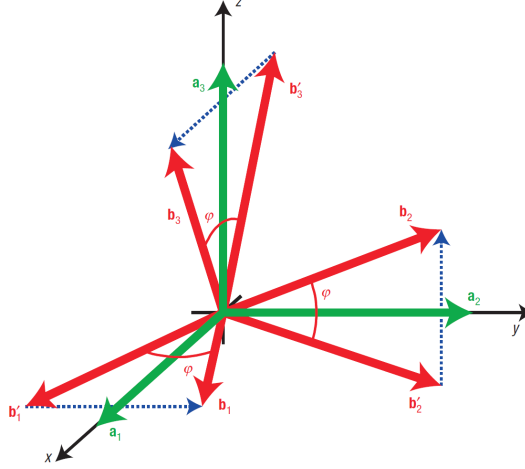
Figure 4.5: **Leggett Measurment Settings.** Settings used by Alice (green) and Bob (red) to test the Leggett inequality. b_1 and b_1' are each separated from a_1 by $\frac{\varphi}{2}$, and by φ from each other in the XY plane. Similarly, b_2 and b_2' lie in the YZ plane and b_3 and b_3' are in the XZ plane. Reprinted by permission from Macmillan Publishers from: [42].

We tested the 2008 version of Leggett inequality proposed by Branciard et al. [42]:

$$L_3(\varphi) \equiv \frac{1}{3} \sum_{i=1}^{3} |E(\hat{a}_i, \hat{b}_i) + E(\hat{a}_i, \hat{b}_i')| \tag{4.6}$$

Here, $E(\hat{a}, \hat{b})$ is the correlation function resulting when Alice and Bob measure in pairs of bases corresponding to those shown in Figure 4.5. The angle between each b_i and b_i' is φ . The value for φ that will produce an optimal violation of the inequality depends on the entangled state that Alice and Bob share. The bound provided by the Leggett model for L_3 is:

$$L_3 \leq 2 - \frac{2}{3} |\sin \frac{\varphi}{2}| \tag{4.7}$$

4.3.2 Results

Figure 4.6 shows the results we obtained for several different values of φ. Each measured point is above the solid red line, which corresponds to the bound of the Leggett model and is therefore a violation of the model. The maximal violation occurs at $\varphi = 40°$. At this setting, the measured value is $L_3 = 1.82 \pm 0.02$ while the Leggett model is bounded by 1.772. The data collected to find L_3 for $\varphi = 40°$ is shown in Table 4.5.

Bases	Basis Vectors a_i	b_j	HWP_A (°)	QWP_A (°)	HWP_B (°)	QWP_B (°)	$C(a_i, b_j)$ (cps)	$E(\hat{a}_i, \hat{b}_j)$ ()	$\Delta E(\hat{a}_i, \hat{b}_j)$ ()
$\{\hat{a}_1, \hat{b}_1\}$	a_1	b_1	22.5	45	17.5	45	346.08		
	a_1	$b_1^\perp$	22.5	45	-27.5	45	49.00	0.9083	0.0057
	$a_1^\perp$	b_1	-22.5	45	17.5	45	17.88		
	$a_1^\perp$	$b_1^\perp$	-22.5	45	-27.5	45	347.00		
$\{\hat{a}_2, \hat{b}_2'\}$	a_2	b_2'	22.5	45	27.5	45	341.80		
	a_2	$b_2'^\perp$	22.5	45	-17.5	45	49.00	0.8919	0.0057
	$a_2^\perp$	b_2'	-22.5	45	27.5	45	19.57		
	$a_2^\perp$	$b_2'^\perp$	-22.5	45	-17.5	45	354.02		
$\{\hat{a}_2, \hat{b}_2\}$	a_2	b_2	-22.5	0	-17.5	0	18.07		
	a_2	$b_2^\perp$	-22.5	0	27.5	0	4.00	-0.9081	0.0038
	$a_2^\perp$	b_2	22.5	0	-17.5	0	340.07		
	$a_2^\perp$	$b_2^\perp$	22.5	0	27.5	0	14.43		
$\{\hat{a}_2, \hat{b}_2'\}$	a_2	b_2'	-22.5	0	-27.5	0	17.32		
	a_2	$b_2'^\perp$	-22.5	0	17.5	0	334.30	-0.8972	0.0059
	$a_2^\perp$	b_2'	22.5	0	-27.5	0	341.30		
	$a_2^\perp$	$b_2'^\perp$	22.5	0	17.5	0	19.30		
$\{\hat{a}_3, \hat{b}_3\}$	a_3	b_3	0	0	5	10	340.52		
	a_3	$b_3^\perp$	0	0	50	10	13.43	0.9199	0.0059
	$a_3^\perp$	b_3	45	0	5	10	15.13		
	$a_3^\perp$	$b_3^\perp$	45	0	50	10	343.55		
$\{\hat{a}_3, \hat{b}_3'\}$	a_3	b_3'	0	0	-5	-10	337.38		
	a_3	$b_3'^\perp$	0	0	40	-10	11.00	0.9391	0.0060
	$a_3^\perp$	b_3'	45	0	-5	-10	10.43		
	$a_3^\perp$	$b_3'^\perp$	45	0	40	-10	344.50		

Table 4.5: **Leggett Inequality Measurerment Settings and Data ($\varphi = 45°$).** This table shows measurement settings used and data collected to find $L_3 = 1.82 \pm 0.02 > 1.772$ for $\varphi = 45°$. $HWP_{A/B}$ are the $\frac{\lambda}{2}$ waveplate settings and $QWP_{A/B}$ are the $\frac{\lambda}{2}$ waveplate settings that realize the measurements corresponding to basis vectors a_i and b_j. $C(a_i, b_j)$ is the coincidence rate detected measured with settings a_i and b_j. $E(\hat{a}_i, \hat{b}_j)$ is the correlation coefficient measured with settings $\hat{a}_i$ and $\hat{b}_j$. Data collection time for each point was 40 seconds. Uncertainties are derived from Poissonian statistics.
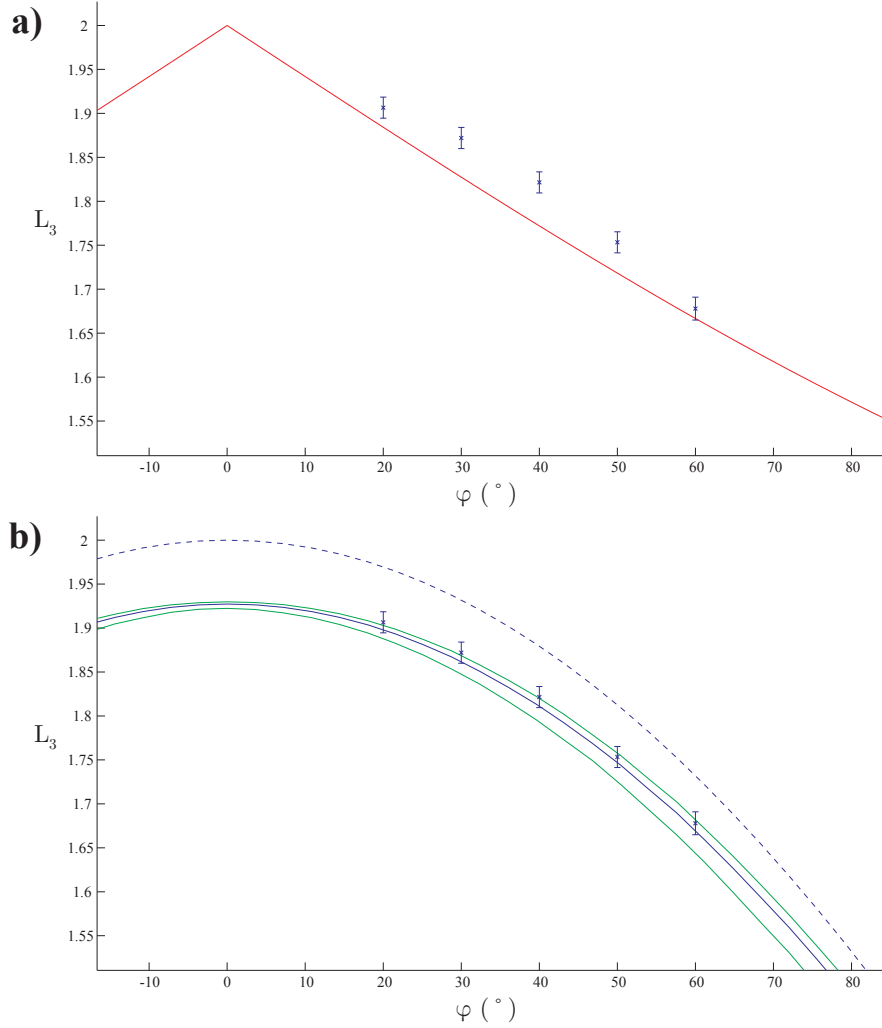
Figure 4.6: **Leggett inequality measurement results.** Experimentally measured values for $L_3(\varphi)$ are shown versus φ. The points with error bars on both graphs are the experimentally measured values for $L_3(\varphi)$. In **a)** the solid red line is the upper bound for the Leggett Model. All experimental data points above this line are violations of the Leggett inequality. In **b)** the blue solid line shows predicted L_3 values based on a measured density matrix (tangle $T = 0.905$) shown in Table 4.6. The green solid lines show one standard deviation. Both solid lines are obtained from a Monte-Carlo simulation that includes a $0.5°$ uncertainty in waveplate settings. The dashed line is the expected L_3 value for a perfect $|\Phi^+\rangle$ state.

	$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$
(a) ρ_{Re}				
$|HH\rangle$	0.5058	0.0083	-0.0079	0.4784
$|HV\rangle$	0.0083	0.0026	-0.0011	0.0133
$|VH\rangle$	-0.0079	-0.0011	0.0030	-0.0032
$|VV\rangle$	0.4784	0.0133	-0.0032	0.4886

	$\langle HH|$	$\langle HV|$	$\langle VH|$	$\langle VV|$
(b) ρ_{Im}				
$|HH\rangle$	0.0000	0.0014	0.0014	-0.0403
$|HV\rangle$	-0.0014	0.0000	0.0022	0.0051
$|VH\rangle$	-0.0014	-0.0022	0.0000	-0.0082
$|VV\rangle$	0.0403	-0.0051	0.0082	0.0000

Table 4.6: **Leggett Test Density matrix.** Real and imaginary parts of the density matrix generated by maximum likelihood QST performed shortly after the Leggett inequality was tested.

Table 4.6 shows a density matrix determined by performing QST just after the test of the Leggett inequality.

4.3.3 Discussion

To our knowledge, this is the second time that the Leggett inequality of the form in [42] has been violated. The challenging visibility requirements for this experiment made it a good test of the quality of our entanglement source and measurement capabilities. The violation of the Leggett inequality does not rule out all non-local hidden variable models for indeed, some may reproduce the predictions of quantum theory. However, this violation does show that the form assumed for the hidden variable, λ, disagrees with what we observe in nature. Leggett proposed this specific form for λ simply because it appeared reasonable, but there are many other forms of λ that could be tested.

Chapter 5

Bounding the Predictive Power of Alternative Theories to Quantum Mechanics

Devising alternative theories to QT, constructing corresponding testable inequalities, and experimentally testing them is, as we have seen, a labour intensive process. Some of these experiments are also challenging in their own right. This suggests that there might be a more efficient strategy than continuing to propose and test NLHV models. We should look for ways to experimentally test the general properties that non-local hidden variable theories must have in order to make predictions that are consistent with nature. This will allow us to begin developing a set of criteria that new theories proposed must meet in order to be worth considering. In this chapter we will conduct an experiment that does this.

Note: This remainder of this chapter is reproduced from [43]. This work was done in conjunction with Joshua A. Slater[1], Roger Colbeck[2], Renato Renner[2], and Wolfgang Tittel[1]

Bounding the Predictive Power of Alternative Theories to Quantum Mechanics

The question of whether the probabilistic nature of quantum mechanical predictions can be alleviated by supplementing the wave function with additional information has received a lot of attention during the past century. A few specific models have been

[1]Department of Physics and Astronomy, University of Calgary.
[2]Institute for Theoretical Physics, ETH Zürich

suggested, and subsequently falsified. Here we give a more general answer to this question: We provide experimental data that, as well as falsifying these models, cannot be explained within any alternative theory that could predict the outcomes of measurements on maximally entangled particles with significantly higher probability than quantum theory. Our conclusion is based on the assumptions that all measurement settings have been chosen freely (within a causal structure compatible with relativity theory), and that the presence of the detection loophole did not affect the measurement outcomes.

5.1   Background

Many of the predictions we make in everyday life are probabilistic. Usually this is caused by having incomplete information, as is the case when making weather forecasts. On the other hand, even with all the information available within quantum mechanics, the outcomes of certain experiments, e.g., the path taken by a member of a maximally entangled pair of spin-half particles that passes through a Stern-Gerlach apparatus, are generally not predictable before the start of the experiment (see Fig. 5.1). This lack of predictive power has prompted a long debate, going back to the paper by Einstein, Podolsky and Rosen [2], of whether quantum mechanics is the optimal way to predict measurement outcomes. In turn, these discussions have led to important fundamental insights. In particular, Kochen and Specker, and independently Bell, proved that there cannot exist any *noncontextual* theory that predicts observations with certainty [44, 45]. In a similar vein, Bell showed [5] that, in general, there cannot exist any additional local property (a local hidden variable) that completely determines the outcome of any measurement on the particle. Bell's argument relies on the fact that entangled particles give rise to correlations that cannot be reproduced in a local hidden variable theory. The existence of such correlations has been confirmed in a series of increasingly sophisticated

experiments [24, 46, 63, 25, 47, 48], and local hidden variable models have thus been ruled out.
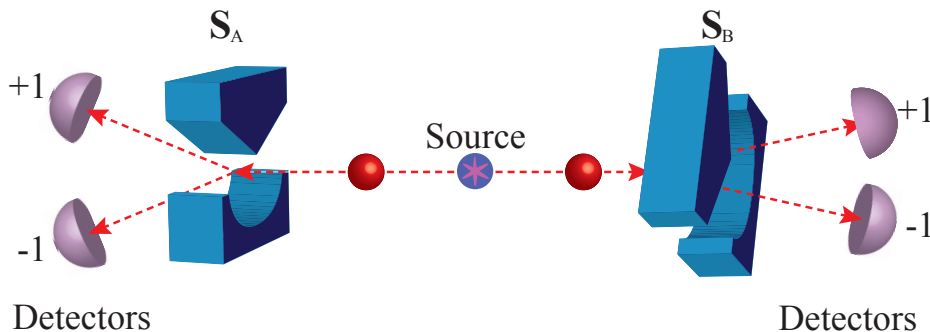

Figure 5.1: **Leggett Model.** A source emits two spin-half particles travelling to two distant sites where each particle's spin is measured along directions $\boldsymbol{S}_A$ and $\boldsymbol{S}_B$, respectively, using Stern-Gerlach apparatuses. If the particles are initially maximally entangled, then the probability of correctly predicting the result of the measurement on the particle on the left ($X = \pm 1$) is, according to quantum mechanics, given by $p_{\mathrm{QM}} = 0.5$.

The purpose of the above arguments was to refute theories in which access to hidden parameters would, in principle, allow perfect predictions of the outcomes of any experiment. However, these arguments do not rule out possible theories that have more predictive power than quantum mechanics, while remaining probabilistic [49]. Consider again the Stern-Gerlach apparatus in Fig. 5.1, in which, according to quantum mechanics, a member of a maximally entangled particle pair may be deviated in one of two directions, each with probability 0.5. One may now conceive of a theory that, depending on a hidden vector, **z** which may be seen as a "classical spin," would allow us to predict the direction of deviation with a larger probability, say 0.75, thereby improving the quantum mechanical prediction by 0.25. This corresponds to a proposal put forward by Leggett [37]. (We note that the given value of 0.75 assumes the most natural Leggett-type model, in which the direction of the hidden spin vector **z** is uniformly distributed [51]. Furthermore, we emphasize that the essence of Leggett's model is the existence of the

hidden local spin vector and not the additional nonlocal variables.) As in the case of local hidden variable models, Leggett-type hidden spin models have been shown to be incompatible with quantum theory [37] and falsified experimentally [38, 40, 39, 41, 42].

5.2 Theory

In this chapter we present experimental data that bounds the probability, δ, by which any alternative theory could improve upon predictions made by quantum theory about measurements on members of maximally entangled particles while being consistent with the assumption that measurement settings can be chosen freely, in the sense described later. We find that quantum theory is close to optimal in terms of its predictive power.

Our work develops a recent theoretical argument [37] that refutes alternative theories with increased predictive power based on the assumption that quantum theory is correct [52] (similar to Bell's and Leggett's arguments [5, 51]) and is itself based on a sequence of work [55, 56, 57]. Here we experimentally investigate this assumption for the case of maximally entangled particles. (In this sense, our work is related to [37] in the same way as experimental tests of the Bell inequality relate to Bell's theoretical work [5].) Furthermore, we provide a significantly strengthened relation between experimentally measurable quantities and the maximum increase of predictive power any alternative theory could have for these quantities. This allows us to obtain nontrivial bounds on the increased predictive power from experimental data obtained using present technology. In particular, we can falsify all local hidden variable models as well as all (including so far not considered) Leggett-type models.

Before describing the experiment, we briefly review the main features of the theory (see Appendix A for more details). Crucially, the framework used is operational; i.e., it refers only to directly observable quantities, such as measurement outcomes. For example,

the Stern-Gerlach experiment with entangled particles mentioned above outputs a binary value, X (Y), indicating in which direction particle one (two) is deviated. We associate with X (Y) a time coordinate t and three spatial coordinates (r_1, r_2, r_3), corresponding to a point in spacetime where the value X (Y) can be observed.

We call such observable values with spacetime coordinates spacetime variables (SVs). In the same manner, any parameters that are needed to specify the experiment (e.g., the orientations of the Stern-Gerlach apparatuses) can be modeled as SVs.

According to quantum theory, the outcome, X, of the measurement on particle one is random, even given a complete description of the measurement apparatus, A. However, an alternative theory may provide us with additional information, Ξ (which can also be modeled in terms of SVs [58]). We can then ask whether this additional information can be used to improve the predictions that quantum mechanics makes about X, which depend on the measurement setting A and the initial state (which we assume to be fixed). This question has a negative answer if the distribution of X, conditioned on A, is unchanged when we learn Ξ. This can be expressed in terms of the Markov chain condition [59],

$$X \leftrightarrow A \leftrightarrow \Xi. \tag{5.1}$$

The aim of this work is to place a bound on the maximum probability, δ, by which this condition can be violated. In other words, a bound of δ implies that the predictions obtained from quantum theory are optimal except with probability (at most) δ.

For the described experiment, our considerations rely only on the natural (and often implicit) assumption that measurement parameters can be chosen freely, i.e., independently of the other parameters of the theory. This assumption can be expressed in the above framework as the requirement that the SV corresponding to a measurement parameter, A, can be chosen such that it is statistically independent of all SVs whose co-

ordinates lie outside the future lightcone of A (Bell also used such a notion, see, e.g., [60]). When interpreted within the usual relativistic spacetime structure, this is equivalent to demanding that A is uncorrelated with any preexisting values in any frame. We note that any alternative theory that satisfies the free choice assumption automatically obeys the nonsignalling conditions, as shown in the Appendix A.
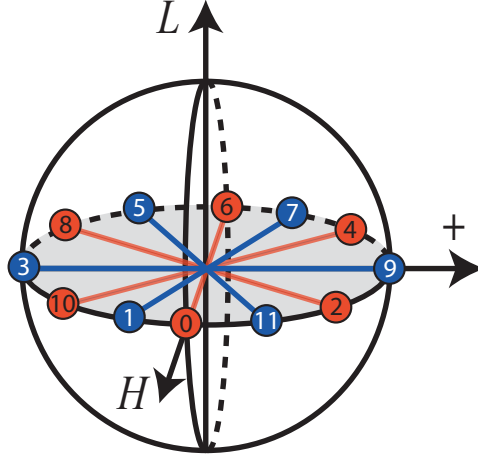
Figure 5.2: **Measurement settings.** Graphical depiction of the polarization measurements along $\boldsymbol{S}_A$ (red, labeled using index A, i.e., even numbers) and $\boldsymbol{S}_B$ (blue, labeled using index B, i.e., odd numbers) for $N = 3$.

As is the case in all falsifications of models that would improve the predictions given by standard quantum theory [5, 37], the argument leading to our bound on δ is based on the strength of correlations between measurement outcomes on entangled particles [61], and, in our case, on pairs of entangled qubits. We denote the projectors describing measurements on qubit one by $|a\rangle\langle a| = \frac{1}{2}(\mathbb{1} + \boldsymbol{S}_A(a)\boldsymbol{\sigma})$ and for qubit two by $|b\rangle\langle b| = \frac{1}{2}(\mathbb{1} + \boldsymbol{S}_B(b)\boldsymbol{\sigma})$ with

$$\boldsymbol{S}_A(a) = (\cos(a\pi/2N),\, \sin(a\pi/2N),\, 0)^{\mathbf{T}}$$

$$\boldsymbol{S}_B(b) = (\cos(b\pi/2N),\, \sin(b\pi/2N),\, 0)^{\mathbf{T}},$$

where $a \in \{0, 2, ..., (4N - 2)\}$, $b \in \{1, 3, \ldots, (4N - 1)\}$, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^{\mathbf{T}}$, and $^{\mathbf{T}}$

denotes "transpose". (The spin vectors $\boldsymbol{S}_A(a)$ and $\boldsymbol{S}_B(b)$ are conveniently depicted on the Bloch sphere; Fig. 5.2 shows the possible vectors for $N = 3$.) We note that projectors described by values of a (or b) that differ by $2N$ correspond to measurements of spin along opposite directions. Hence, each set of projectors describes N pairs of orthogonal measurements. This allows us to calculate, for each value of $a \in \{0, 2, ..., 2N - 2\}$ and $b \in \{1, 3, ..., 2N - 1\}$, the probability of detecting the two photons from a pair along the spin directions $\boldsymbol{S}_A(a)$ and $\boldsymbol{S}_B(b)$ (for which we assign $X, Y = +1$), and along the orthogonal directions $-\boldsymbol{S}_A(a)$ and $-\boldsymbol{S}_B(b)$ (for which we assign $X, Y = -1$). We denote this probability $P(X = Y | a, b)$. In turn, this allows us to establish the correlation strength

$$I_N \ := \ P(X = Y | 0, 2N - 1) + \sum_{\substack{a,b \\ |a-b|=1}} (1 - P(X = Y | a, b)). \tag{5.2}$$

We note that measuring I_N involves the same measurements as those required for testing a chained Bell inequality, first violated for $N \geq 3$ in [62].

Furthermore, deriving a bound on δ requires knowledge of the bias of the individual outcomes

$$\nu_N := \max_a \ D(P_{X|a}, P_{\bar{X}}),$$

where D denotes the variational distance, $D(P_X, Q_X) := \frac{1}{2} \sum_x |P_X(x) - Q_X(x)|$, and $P_{\bar{X}}$ denotes the uniform distribution on X. (In an experiment, due to imperfections in the generated bipartite state that lead to the local (single particle) states not being completely mixed, $P_X(x) \neq 1/2$, which implies a nonzero bias.)

As we show in detail in Appendix A, for each N, the maximum increase of predictive power, δ_N, of any alternative theory is bounded by

$$\delta_N = \frac{I_N}{2} + \nu_N. \tag{5.3}$$

Repeating the correlation and bias measurements for many N, we can obtain the best bound on δ via $\delta \leq \min_N \{\delta_N\}$.

Assuming a perfect experimental setup, quantum theory predicts that δ_N will approach 0 as N tends to ∞. For a realistic (imperfect) setup, however, δ_N reaches a minimum at some finite N, above which it is increasing in N.

5.3   Experimental Design
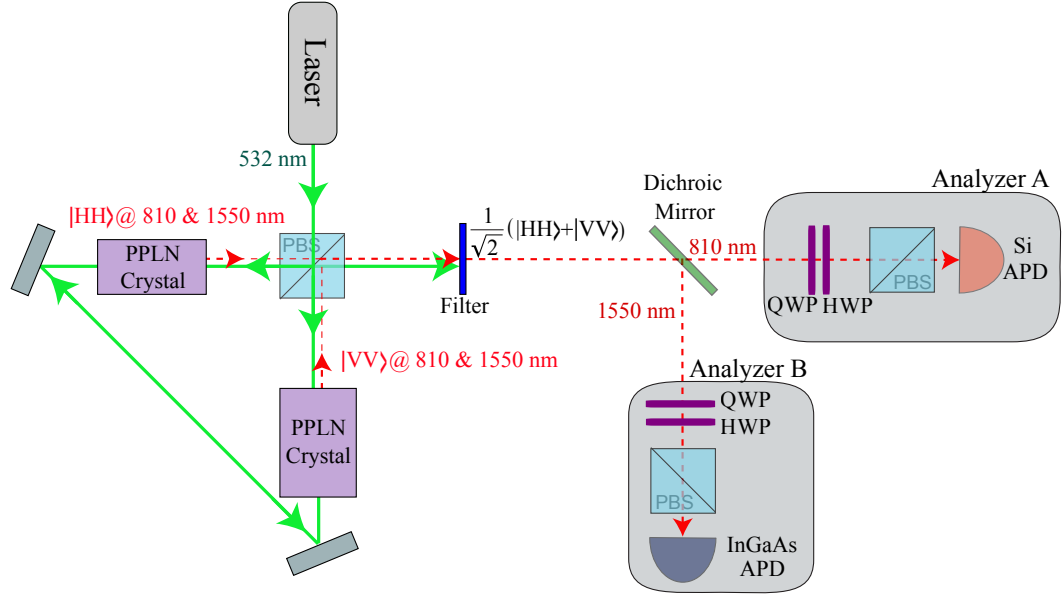
Figure 5.3: **Experimental Design.**(a) Experimental setup, see text for details. (b) Density matrix ρ_{real} of the biphoton state produced by our source as calculated via maximum-likelihood quantum state tomography [19](see Appendix A for actual values). The fidelity, $F = \langle \phi^+ | \rho_{\mathrm{real}} | \phi^+ \rangle$, between the detected state, ρ_{real}, and the ideal state, $|\phi^+\rangle$, given by Eq. 5.4, is (98.0 ± 0.1) %.

A schematic of our experimental setup, which is inspired by the source described in [16], is depicted in Fig. 5.3. A diagonally polarized, continuous wave, 532 nm wavelength laser beam is split by a polarizing beam splitter (PBS) and travels both clockwise and counterclockwise through a polarization Sagnac interferometer. The interferome-

ter contains two type-I, periodically poled lithium niobate (PPLN) crystals configured to produce collinear, nondegenerate, 810/1550 nm-wavelength photon pairs via spontaneous parametric down-conversion. As the optical axes of the two crystals are perpendicular to each other and photon-pair generation is polarization dependent, the clockwise-travelling, vertically polarized (counterclockwise travelling, horizontally polarized) pump light passes through the first crystal without interaction and may down-convert in the second crystal to produce two horizontally (vertically) polarized photons. For small pump power, recombination of the two biphoton modes on the PBS yields photon pairs with high fidelity to the maximally entangled state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle), \tag{5.4}$$

where $|H\rangle$ and $|V\rangle$ represent horizontal and vertical polarization states, respectively, and replace the usual spin-up and spin-down notation for spin-half particles. Behind the interferometer, the remaining pump light is removed using a high-pass filter. The entangled photons are separated on a dichroic mirror and sent to polarization analyzers that can be adjusted to measure the polarization of an incoming photon along any desired direction $\boldsymbol{S} = (S_H, S_+, S_L)^T$, where $\boldsymbol{S}$ is expressed in terms of its projections onto horizontal (H), diagonal ($+45°$), and left-circular (L) polarized components. The polarization analyzers consist of quarter wave plates (QWP), half wave plates (HWP), and PBSs. Finally, the 810 nm photons are detected using a free-running silicon avalanche photodiode (Si APD), and 1550 nm photons are detected using an InGaAs APD triggered by detection events from the Si APD.

For each setting $\boldsymbol{S}_A(a)$ (with a as described above), we establish the number of detected photons, $M(a)$, over 80 sec, from which we can calculate the bias

$$\nu_N = \frac{1}{2} \max_{a \in \{0,2,...,(2N-2)\}} \left\{ \frac{|M(a) - M(a+2N)|}{M(a) + M(a+2N)} \right\}.$$

84

Furthermore, for the joint measurements described by Eq. 5.2, we register the number of detected photon pairs over 40 sec to calculate

$$P(X = Y|a, b) = \frac{M(a, b) + M(a + 2N, b + 2N)}{M},$$

where, e.g., $M(a, b)$ is the number of joint photon detections for measurements along $\boldsymbol{S}_A(a)$ and $\boldsymbol{S}_B(b)$, and the normalization factor $M = M(a, b) + M(a, b + 2N) + M(a + 2N, b) + M(a + 2N, b + 2N)$. This allows us to establish δ_N via Eqs. 5.2 and 5.3.

5.4 Results

Our experimental results are depicted in Fig. 5.4 and summarized in Table 5.1. We measured δ_N for $N = 2$ to $N = 7$ and found the minimum, $\delta_7 = 0.1644 \pm 0.0014$, for $N = 7$. Using the above considerations, these data lead to our main conclusion that the maximum probability by which any alternative theory can improve the predictions of quantum theory is at most ~ 0.165. To put this result into context, we note that a deterministic local hidden variable theory would allow for predictions of the outcomes with probability 1; similarly, it is easy to verify that the Leggett model (with a uniform distribution over the hidden spin, z) would correctly predict the outcome with probability $p_{Leggett} = 3/4$. Since these values exceed the probability of a correct prediction based on quantum theory, 0.5, by more than delta, both theories are directly falsified by our result. (We refer to Appendix A for a more detailed discussion of the Leggett model, including variants with different distributions of the hidden spin vector.)

5.5 Discussion

We remark that our conclusion is based on the assumption that measurement settings can be chosen freely (this removes the need to experimentally close the locality loophole),
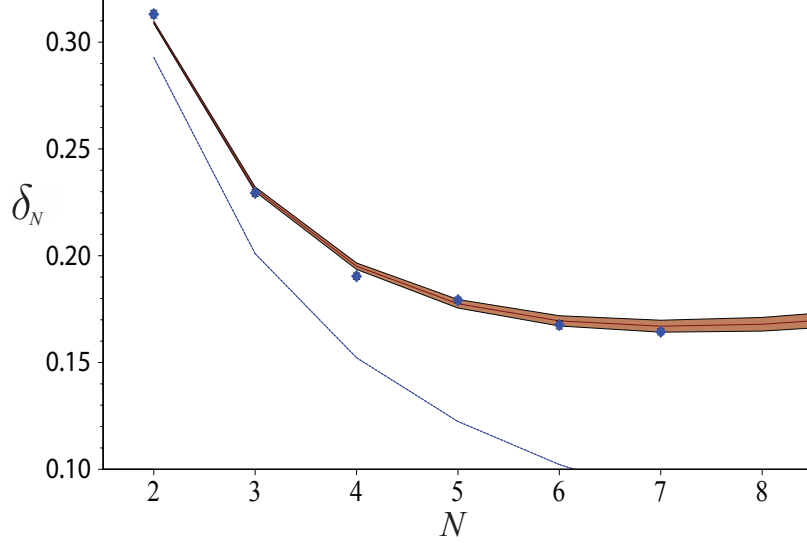
Figure 5.4: **Measured results for δ_N.** Experimentally obtained values δ_N (blue diamonds) with one-standard-deviation uncertainties (hidden by the size of the diamonds) calculated from measurement results assuming Poissonian statistics. Also shown is a curve joining the values predicted by quantum theory, including one-standard-deviation statistical uncertainties (solid red line and grey shaded area, respectively), calculated from the measured density matrix ρ_{real}. Note that the predicted value for δ increases for $N > 7$. The bounds of the shaded region are derived using Monte Carlo simulations and are consistent with the observed variations of the measured values. Finally, the dashed blue line is the theoretical curve, again calculated using quantum theory, that assumes the ideal $|\phi^+\rangle$ state, as in Eq. 5.4, and perfect experimental apparatus with zero noise. It asymptotically approaches zero as N tends to infinity. For instance, for $N = 7$ we find $\delta_7^{\text{ideal}} = 0.088$.

N	I_N	ν_N	δ_N
2	0.6213 ± 0.0035	0.0025 ± 0.0002	0.3131 ± 0.0018
3	0.4549 ± 0.0032	0.0020 ± 0.0002	0.2294 ± 0.0016
4	0.3757 ± 0.0029	0.0025 ± 0.0002	0.1904 ± 0.0015
5	0.3518 ± 0.0028	0.0033 ± 0.0002	0.1792 ± 0.0014
6	0.3290 ± 0.0028	0.0032 ± 0.0002	0.1677 ± 0.0014
7	0.3238 ± 0.0027	0.0025 ± 0.0002	0.1644 ± 0.0014

Table 5.1: **Summary of Results.** The table shows values for I_N, bias ν_N, as well as $\delta_N = I_N/2 + \nu_N$. Statistical uncertainties (one standard deviation) are calculated from measurement results assuming Poissonian statistics.

and our experiments do not close the detection loophole [63]. Hence, strictly, the above conclusion holds modulo the assumption that similar, loophole-free experiments would show the same results.

Further decreasing the experimentally established bound on δ would require photon pair sources and measurement apparatuses with rapidly increasing quality. For example, to decrease δ by more than a factor of two compared to our result, the fidelity of the source must exceed 99.6 % (assuming zero bias and perfect measurement apparatus) and N increases to 15 or beyond, resulting in 120 or more high-precision coincidence measurements. This is, to the best of our knowledge, unattainable with state-of-the-art sources [40, 39] (for more details see Appendix A).

In conclusion, under the assumption that measurements can be chosen freely, no theory can predict the outcomes of measurements on a member of a maximally entangled pair substantially better than quantum mechanics. In other words, any already-considered or yet-to-be-proposed theory that makes significantly better predictions would either be incompatible with our experimental observations, or be incompatible with our assumption that the measurement parameters can be chosen freely. While the former is true, for example, for local hidden variable theories (as already pointed out by Bell [5]) or for the Leggett model [37], the de Broglie-Bohm theory [64, 6] is an example of the second type – the theory cannot incorporate measurement parameters that satisfy our free-choice assumption (this is further discussed in Appendix A [51]).

Chapter 6

Summary and Outlook

The idea that matter is composed of indivisible building blocks is more than two and a half thousand years old, but it wasn't until relatively recently that we have found ways to experimentally test the implications of theories based on quantized matter, and the notion of quantized energy is comparatively new. We lacked a valid way to predict the outcomes of experiments involving a small number of quanta until the development of quantum theory in the last century. The predictions of quantum theory have, so far, been in remarkable agreement with experimental observations. However these predictions are probabilistic and, in some cases, completely random. This is in stark contrast to the macroscopic world we live in for which classical mechanics offers fully deterministic predictions. Quantum theory's inability to make such predictions has prompted a search for an alternative theory that is both correct and more complete in the sense that it would offer greater predictive power.

In this thesis we have conducted experiments that show that no deterministic local hidden variable model can be consistent with our observations of nature, provided the assumptions we used to avoid closing the locality and detection loopholes are valid. Similarly, we have tested the Leggett inequality and found that a broad class of non-local hidden variable models, which possess greater predictive power than QT, are also inconsistent with experimental observations. We have also experimentally determined an upper bound on the predictive power of alternative theories that, assuming certain reasonable conditions, allows us to rule out any alternative theory that can correctly predict the outcome of measurements with a probability greater than $P = 0.665$.

The key resource needed for all of these tests is a source of entangled particle pairs.

In this thesis we have constructed and characterized a source of polarization entangled photon pairs, which has very high visibility and tangle, making it suitable for the tests described above. In addition, this source is relatively inexpensive, compact, and designed with flexibility for future experiments in mind. The ability to produce photon pairs at non-degenerate wavelengths of 810 nm and 1550 nm allows us to consider experiments where there is either a long fiber or free-space link between two separated analyzers. The QC2 lab has a fiber link to SAIT that may be used for future long distance experiments. This may permit us to conduct experiments with the locality loophole closed. Single photon detectors that are efficient enough to permit closing the detection loophole have been developed recently [30], which suggests that it may soon be possible to repeat all of the experiments described in this thesis with the detection loophole closed.

The results of all experiments we have discussed in this thesis, as well as many yet to be proposed, can be improved upon using entangled states with a higher tangle. Although the source described in this thesis is of very high quality, it should be possible to improve it. For example, the laser pointer used as a pump for this source had a broader spectrum than expected and its spatial mode changed slowly with time. A mode-stabilized laser with a narrower spectrum could make producing states with higher tangles possible. Gains may also be realized by increasing the quality of PBSs used in the interferometer and qubit analyzers.

Over the last century we have begun to learn how to speak the language of nature in the context of a single quantum. This language has been full of surprising nuances. The experiments we have conduced here have allowed us to probe some of these nuances, and we have demonstrated that indeed, God does play at dice, and those dice are very close to being fair.

Bibliography

[1] Born, M. *The Born-Einstein Letters - Correspondence between Albert Einstein and Max and Hedwig Born* (Macmillan Press LTD, London and Basingstoke, 1971), 91 Irene Born, translator.

[2] Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Physical Review* **47**, 777–780 (1935).

[3] Schrödinger, E. Discussion of Probability Relations between Separated Systems. *Mathematical Proceedings of the Cambridge Philosophical Society* **31:04**, 555–563 (1935).

[4] Einstein, A. Physics and Reality *Journal of the Franklin Institute* **221**, 349–382 (1936) Jean Piccard, translator.

[5] Bell, J. S. On the Einstein-Podolsky-Rosen paradox. *Physics* **1**, 195–200 (1964).

[6] Bohm, D. A suggested interpretation of the quantum theory in terms of "hidden" variables. I. *Physical Review* **85**, 166–179 (1952).

[7] Manderbach, T. S., Weier, H., Furst, M., Ursin, R., Tiefenbacher, F., Scheidl, T., Perdigues, J., Sodnik, Z., Kurtsiefer, C., Rarity, J. G., Zeilinger, A., Weinfurter, H. Experimental Demonstration of Free-Space Decoy-State Quantum Key Distribution over 144 km. *Physical Review Letters* **98**, 010504 (2007).

[8] Hughes, R., Buttler, W., Kwiat, P., Lamoreaux, S., Morgan, G., Nordhold, J., Peterson, G. Free-space quantum key distribution in daylight. *Journal of Modern Optics* **47**, 549—562 (2000).

[9] Gisin, N., Ribordy, G., Tittel, W., Zbinden, H. Quantum cryptography. *Review of Modern Physics* **74**, 145–195 (2002).

[10] This is evidenced by the tests of Bell and Leggett inequalities discussed in this paper as well as the large number of publications on Bell and Leggett inequalities that have continued to appear even very recently.

[11] Ulrich, R. Polarization stabilization on single-mode fiber. *Applied Physics Letters* **35**, 840 (1979).

[12] Boyd, R. Nonlinear optics, 2nd Edition. Academic Press (2003).

[13] Jundt, D. H. Temperature-dependent Sellmeier equation for the index of refraction, n_e, in congruent lithium niobate. *Physical Review Letters* **22:20**, 1553—1555 (1997).

[14] Kwiat, P. G., Waks, E., White, A. G., Appelbaum, I., Eberhard, P. H. Ultrabright source of polarization-entangled photons. *Physical Review A* **60:2**, 773–776 (1999).

[15] Ljunggren, D. M., Tengner, M., Marsden, P., Pelton, M. Theory and experiment of entanglement in a quasi-phase-matched two-crystal source. *Physical Review A* **73**, 032326 (2006)

[16] Kim, T. Fiorentino, M.& Wong, F. N. C. Phase-stable source of polarization-entangled photons using a polarization Sagnac interferometer. *Physical Review A* **73**, 12316 (2006).

[17] Fedrizzi, A., Herbst, T., Poppe, A., Jennewein, A., Zeilinger, A. A wavelength-tunable fiber-coupled source of narrowband entangled photons. *Optics Express* **15:23**, 15377-15386 (2007).

[18] Eisenberg, H. S., Khoury, G., Durkin, G. A., Simon, C., Bouwmeester, D. Quantum

Entanglement of a Large Number of Photons. *Physical Review Letters* **93:19**, 193901 (2004).

[19] Altepeter, J. B., Jeffrey, E. R., Kwiat, P. J. Photonic State Tomography. *Advances In Atomic, Molecular, and Optical Physics* **52**, 105-159 (2005).

[20] Hradil, Z., Řeháček, M. Efficiency of Maximum-likelihood Reconstruction of Quantum States. *Fortschritte der Physik* **49, 10-11**, 1083-1088 (2001).

[21] Coffman, V., Kundu, J., Wooters, W. K. Distributed entanglement. *Physical Review A* **61**, 052306 (2000).

[22] Fiorentino, M., Spillane, S., Beausoleil, T., Roberts, M., Munro, M. Spontaneous parametric down-conversion in periodically poled KTP waveguides and bulk crystals. *Optics Express* **15:12**, 7479-7488 (2007).

[23] Clauser, J. F., Horne, M. A., Shimony, A., Holt, R. A. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters* **23**, 880-884 (1969).

[24] Freedman, S. J., Clauser, J. F. Experimental test of local hidden-variable theories. *Physical Review Letters* **28**, 938–941 (1972).

[25] Tittel, W. *et al.* Experimental demonstration of quantum correlations over more than 10 km. *Physical Review A* **57**, 3229–3232 (1998).

[26] Aspect, A., Grangier, P. & Roger, G. Experimental Tests of Realistic Local Theories via Bell's Theorem. *Physical Review Letters* **47**, 460–463 (1981).

[27] Pearle, P. M. Hidden-variable example based upon data rejection. *Physical Review D* **2**, 1418–1425 (1970).

[28] Eberhard, P. H. Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. *Physical Review A* **47**, R747 (1993).

[29] Rowe, M. A., Kielpinski, D., Meyer, V., Sackett, C. A., Itano, W. M., Monroe, C., Wineland, D. J., *Nature* **409**, 791–794 (2001).

[30] Ramelow, S., Mech, A., Giustina, M., Gröblacher, S., Wieczorek, W., Lita, A., Calkins, B., Gerrits, T., Nam, S. W., Zeilinger, A., Ursin, R. Highly efficient heralding of entangled single photons. arXiv:1211.5059 (2012).

[31] Zeilinger, A. Testing Bell's Inequalities With Periodic Switching. *Physics Letters A* **118**, 1 (1986).

[32] Aspect, A., Dalibard, J., Roger, G. Experimental Test of Bell's Inequalities Using Time-Varying Analyzers. *Physical Review Letters* **49:25**, 1804–1807 (1982).

[33] Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., Zeilinger, A. Violation of Bell's Inequality under Strict Einstein Locality Conditions. *Physical Review Letters* **81**, 5039–5043 (1998).

[34] The beautiful Bell inequality is called the "Bell inequality in arbitrary dimension" in Bechmann-Pasquinucci and Gisin's 2003 paper and the "elegant Bell" inequality in Gisin's 2008 paper.

[35] Bechmann-Pasquinucci, H., Gisin, N. Intermediate states in quantum cryptography and Bell inequalities. *Physical Review A* **67**, 062310 (2003).

[36] Gisin, N. Bell inequalities: many questions, a few answers. arXiv:0702021 (2008).

[37] Leggett, A. J. Nonlocal hidden-variable theories and quantum mechanics: An incompatibility theorem. *Foundations of Physics* **33**, 1469–1493 (2003).

[38] Gröblacher, S., Paterek,T., Kaltenbaek, R., Brukner, C., Zukowski, M., Aspelmeyer, M., Zeilinger, A. An expeimental test of non-local realism. *Nature* **446**, 871–875 (2007).

[39] Branciard, C., Ling, A., Gisin, N., Kurtsiefer, C., Lamas-Linares, A., Scarani, V. Experimental Falsification of Leggett's nonlocal variable model. *Physical Review A* **99**, 210407 (2007).

[40] Paterek, T., Fedrizzi, A., Gr oblacher, S., Jennewein, T., Zukowski, M., Aspelmeyer, M., Zeilinger, A. Experimental Test of Nonlocal Realistic Theories Without the Rotational Symmetry Assumption. *Physical Review Letters* **99**, 210406 (2007).

[41] Eisaman, M. D., Goldschmidt, E. A., Chen, J., Fan, J., Migdall, A. Experimental test of nonlocal realism using a fiber-based source of polarization entangled photon pairs. *Physical Review A* **77**, 032339 (2008).

[42] Branciard, C., Brunner, N., Gisin, N., Kurtsiefer, C., Lamas-Linares, A., Ling, A., Scarani, V. Testing quantum correlations versus single-particle properties within Leggett's model and beyond. *Nature Physics* **4**, 681–685 (2008).

[43] Stuart, T. E., Slater, J. A., Colbeck, R., Renner, R., Tittel, W. Experimental Bound on the Maximum Predictive Power of Physical Theories. *Physical Review Letters* **109**, 020402 (2012).

[44] Kochen, S. & Specker, E. P. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics* **17**, 59–87 (1967).

[45] Bell, J. S. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics* **38**, 447–452 (1966).

[46] Aspect, A., Grangier, P. & Roger, G. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Physical Review Letters* **49**, 91–94 (1982).

[47] Weihs, G., Jennewein, T., Simon, C., Weinfurter, H. & Zeilinger, A. Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters* **81**, 5039–5043 (1998).

[48] Rowe, M. A., Kielpinski, D., Meyer, V., Sackett, C. A., Itano, W. M., Monroe, C. & Wineland, D. J. Experimental violation of a Bell's inequality with efficient detection. *Nature* **409**, 791–794 (2001).

[49] In his later works, Bell uses definitions that potentially allow probabilistic models [50]. However, as explained in the Supplementary Information of [58], nondeterministic models are not compatible with Bell's other assumptions.

[50] Bell, J. S. La nouvelle cuisine. In *Between Science and Technology*, chap. 24, 97–115, Elsevier Science Publishers, (1990).

[51] More details, as well as other distributions are discussed in Appendix A.

[52] More precisely, the assumption that quantum theory is correct was divided into two parts, the first being that a particular experimental setup yields outcomes distributed as predicted by the theory, and the second that measurement processes can be considered as unitary operations if one takes into account the environment. Here we experimentally investigate the first assumption. Note that the second assumption is only required when bounding the predictive power for measurements on particles that are not members of maximally entangled pairs.

[53] Braunstein, S. L. & Caves, C. M. Wringing out better Bell inequalities. *Annals of Physics* **202**, 22–56 (1990).

[54] Pomarico, E., Bancal, J.-D., Sanguinetti, B., Rochdi, A. & Gisin, N. Various quantum nonlocality tests with a commercial two-photon entanglement source. *Physical Review A* **83**, 052104 (2011).

[55] Barrett, J., Hardy, L. & Kent, A. No signalling and quantum key distribution. *Physical Review Letters* **95**, 010503 (2005).

[56] Barrett, J., Kent, A. & Pironio, S. Maximally non-local and monogamous quantum correlations. *Physical Review Letters* **97**, 170409 (2006).

[57] Colbeck, R. & Renner, R. Hidden variable models for quantum theory cannot have any local part. *Physical Review Letters* **101**, 050403 (2008).

[58] Colbeck, R. & Renner, R. No extension of quantum theory can have improved predictive power. *Nature Communications* **2**, 411 (2011).

[59] Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley and Sons Inc., 2006), 2nd edn. Section 2.8.

[60] Bell, J. S. Free variables and local causality. *Epistemological Letters* **38**, (1977).

[61] Considering measurements restricted to non-entangled particles always leaves open the possibility of explaining the results using a local hidden variable model, see Appendix A.

[62] Pomarico, E. *et al.* Various quantum nonlocality tests with a commercial two-photon entanglement source. *Phys. Rev. A* **83**, 052104 (2011).

[63] Aspect, A. Bell's inequality test: More ideal than ever. *Nature* **398**, 189–190 (1999).

[64] de Broglie, L. La mécanique ondulatoire et la structure atomique de la matière et du rayonnement. *Journal de Physique, Serie VI* **VIII**, 225–241 (1927).

[65] Chen, Z. & Montina, A. Measurement contextuality is implied by macroscopic realism. *Physical Review A* **83**, 042110 (2011).

[66] Bell, J. S. Free variables and local causality. In *Speakable and unspeakable in quantum mechanics*, chap. 12 (Cambridge University Press, 1987).

[67] Paterek, T. *et al.* Experimental test of nonlocal realistic theories without the rotational symmetry assumption. *Physical Review Letters* **99**, 210406 (2007).

[68] Suarez, A. Why aren't quantum correlations maximally nonlocal? Biased local randomness as essential feature of quantum mechanics. e-print `arXiv:0902.2451` (2009).

Appendix A

Derivation of bound on the maximum predictive power of physical theories, and supplements.

A.1   Proof of the bound

In this section, we prove the bound given in Equation 5.1 in the main text, which is stated as Lemma 1 below. We use a bipartite scenario in which two spacelike separated measurements are performed on a maximally entangled state. We denote the choices of observable $A \in \{0, 2, \ldots, 2N - 2\}$ and $B \in \{1, 3, \ldots, 2N - 1\}$ and their outcomes $X \in \{+1, -1\}$ and $Y \in \{+1, -1\}$, respectively[1]. We additionally consider information that might be provided by an alternative theory (denoted Ξ), which is modelled as an additional system with input C and output Z [58]. We assume that this information is static, in the sense that its behaviour is independent of the coordinates associated with C and Z. If one makes the assumption that the measurements can be chosen freely, then the joint distribution $P_{XYZ|ABC}$ satisfies the non-signalling conditions

$$P_{XY|ABC} = P_{XY|AB} \tag{A.1}$$

$$P_{XZ|ABC} = P_{XZ|AC} \tag{A.2}$$

$$P_{YZ|ABC} = P_{YZ|BC}. \tag{A.3}$$

A proof of this was given in [58], which we now recap for completeness.

---

[1]Note that the measurements we speak of in this appendix have a slightly different form than those in the main text. Specifically, we now assume that measurements behave ideally, projecting onto one of two basis elements and leading to one of the two outcomes ± 1. In a real experiment, there is always the additional possibility of no photon detection (let us denote this outcome 0). The measurements discussed in the main text are configured to distinguish $+1$ from either -1 or 0, or to distinguish -1 from either $+1$ or 0. Both of these measurements are used in the experiment to infer the distribution of the ideal measurement with outcomes ± 1.

Recall that the free choice assumption states that for a parameter, A, to be free, it must be uncorrelated with all variables outside its future lightcone[2]. The setup is such that the measurements specified by A and B are spacelike separated and, since Ξ is static, we can consider its observation to also be spacelike separated from the other measurements.

We assume A is a free choice, which corresponds mathematically to the condition

$$P_{A|BCYZ} = P_A. \tag{A.4}$$

Furthermore, using the definition of conditional probability ($P_{Q|R} := P_{QR}/P_R$), we can write

$$P_{YZA|BC} = P_{YZ|BC} \times P_{A|BCYZ} = P_A \times P_{YZ|BC} \ ,$$

where we inserted (A.4) to obtain the second equality. Similarly, we have

$$P_{YZA|BC} = P_{A|BC} \times P_{YZ|ABC} = P_A \times P_{YZ|ABC} \ .$$

Comparing these two expressions for $P_{YZA|BC}$ yields the non-signalling condition (A.3). Repeating this argument symmetrically, the other non-signalling conditions can be similarly inferred.

Note that standard proofs of Bell's theorem and related results assume both no-signalling and that measurements are chosen freely (see, for example, [66] for a statement of Bell's notion of free choice, which is the same as ours). Although free choice implies no-signalling, the converse does not hold. Instead, no-signalling is needed for free choices to be possible, but does not imply that they are actually made.

Lemma 1 gives a bound on the increase in predictive power of any alternative theory in terms of the strength of correlations and the bias of the individual outcomes. The bound

---

[2]Note that this requirement can be seen as a prerequisite for non-contextuality, as pointed out in [65], where an alternative proof that quantum theory cannot be extended, based on the assumption of non-contextuality, is offered.

is expressed in terms of the variational distance $D(P_W, Q_W) := \frac{1}{2} \sum_w |P_W(w) - Q_W(w)|$, which has the following operational interpretation: if two distributions have variational distance at most δ, then the probability that we ever notice a difference between them is at most δ.

The bias is quantified by[3] $\nu_N := \max_a D(P_{X|a}, P_{\bar{X}})$, where $P_{\bar{X}}$ is the uniform distribution on X. To quantify the correlation strength, we define

$$I_N := P(X = Y | A = 0, B = 2N - 1) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y | A = a, B = b) . \tag{A.5}$$

This is equivalent to Equation 5.2 in the main text. We remark that $I_N \geq 1$ is a Bell inequality, i.e. is satisfied by any local hidden variable model.

Lemma 1. *For any non-signalling probability distribution, $P_{XYZ|ABC}$, we have*

$$D(P_{Z|abcx}, P_{Z|abc}) \leq \delta_N := \frac{I_N}{2} + \nu_N \tag{A.6}$$

for all a, b, c, and x.

To connect this back to the main text, we remark that the Markov chain condition $X \leftrightarrow A \leftrightarrow \Xi$ is equivalent to $P_{Z|abcx} = P_{Z|abc}$ (which corresponds to Ξ not being of use to predict X). Hence, from the operational meaning of the variational distance (given above), (A.6) can be interpreted that X and Z can be treated as uncorrelated, except with probability at most δ_N.

The proof is an extension of an argument given in [58] which is based on *chained Bell inequalities* [27, 53, 54] and generalizes results of [55, 56, 57]. Many steps of this proof mirror those in [58], which we repeat here for completeness. However, note that the bound derived in this Lemma is tighter than that of [58].

---

[3]A note on notation: we usually use lower case to denote particular instances of upper case random variables.

Proof. We first consider the quantity I_N evaluated for the conditional distribution $P_{XY|AB,cz} = P_{XY|ABCZ}(\cdot,\cdot|\cdot,\cdot,c,z)$, for any fixed c and z. The idea is to use this quantity to bound the variational distance between the conditional distribution $P_{X|acz}$ and its negation, $1 - P_{X|acz}$, which corresponds to the distribution of X if its values are interchanged. If this distance is small, it follows that the distribution $P_{X|acz}$ is roughly uniform.

For $a_0 := 0$, $b_0 := 2N - 1$, we have

$$
\begin{aligned}
I_N(P_{XY|AB,cz}) &= P(X = Y | A = a_0, B = b_0, C = c, Z = z) \\
&\quad + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y | A = a, B = b, C = c, Z = z) \\
&\geq D(1 - P_{X|a_0 b_0 cz}, P_{Y|a_0 b_0 cz}) + \sum_{\substack{a,b \\ |a-b|=1}} D(P_{X|abcz}, P_{Y|abcz}) \\
&= D(1 - P_{X|a_0 cz}, P_{Y|b_0 cz}) + \sum_{\substack{a,b \\ |a-b|=1}} D(P_{X|acz}, P_{Y|bcz}) \\
&\geq D(1 - P_{X|a_0 cz}, P_{X|a_0 cz}) \\
&= 2D(P_{X|a_0 b_0 cz}, P_{\bar{X}}) \; .
\end{aligned}
\tag{A.7}
$$

The first inequality follows from the fact that $D(P_{X|\Omega}, P_{Y|\Omega}) \leq P(X \neq Y|\Omega)$ for any event Ω (a short proof of this can be found in [57]). Furthermore, we have used the non-signalling conditions $P_{X|abcz} = P_{X|acz}$ (from (A.2)) and $P_{Y|abcz} = P_{Y|bcz}$ (from (A.3)), and the triangle inequality for D. By symmetry, this relation holds for all a and b. We hence obtain $D(P_{X|abcz}, P_{\bar{X}}) \leq \frac{1}{2} I_N(P_{XY|AB,cz})$ for all a, b, c and z.

We now take the average over z on both sides of (A.7). First, the left hand side gives

$$\sum_z P_{Z|abc}(z) I_N(P_{XY|AB,cz}) = \sum_z P_{Z|c}(z) I_N(P_{XY|AB,cz})$$

$$= \sum_z P_{Z|a_0 b_0 c}(z) P(X = Y|a_0, b_0, c, z)$$

$$+ \sum_{\substack{a,b \\ |a-b|=1}} \sum_z P_{Z|abc}(z) P(X \neq Y|a, b, c, z)$$

$$= P(X = Y|a_0, b_0, c) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y|a, b, c)$$

$$= I_N(P_{XY|AB,c}) \,, \tag{A.8}$$

where we used the non-signalling condition $P_{Z|abc} = P_{Z|c}$ (which is implied by (A.2) and (A.3)) several times. Next, taking the average on the right hand side of (A.7) yields $\sum_z P_{Z|abc}(z) D(P_{X|abcz}, P_{\bar{X}}) = D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc})$, so we have

$$2D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc}) \leq I_N(P_{XY|AB,c}) = I_N(P_{XY|AB}). \tag{A.9}$$

The last equality follows from the non-signalling condition (A.1) (if $P(X = Y|a, b, c)$ or $P(X \neq Y|a, b, c)$ depended on c, then there would be signalling from C to A and B).

Furthermore, note that

$$2D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc}) = \sum_z \left| P_{XZ|abc}(-1, z) - \frac{1}{2} P_{Z|abc}(z) \right| + \sum_z \left| P_{XZ|abc}(+1, z) - \frac{1}{2} P_{Z|abc}(z) \right|$$

and that both of the terms on the right hand side are equal (since $P_{Z|abc}(z) = P_{XZ|abc}(-1, z) + P_{XZ|abc}(+1, z)$) i.e. $\sum_z \left| P_{XZ|abc}(x, z) - \frac{1}{2} P_{Z|abc}(z) \right| \leq \frac{I_N}{2}$ for all a, b, c and x. Note also that $D(P_{X|a}, P_{\bar{X}}) = \left| P_{X|a}(x) - \frac{1}{2} \right|$ for all x.

Combining the above, we have

$$D(P_{Z|abcx}, P_{Z|abc}) = \sum_z \left| \frac{1}{2} P_{Z|abcx}(z) - \frac{1}{2} P_{Z|abc}(z) \right|$$

$$\leq \sum_z \left| \frac{1}{2} P_{Z|abcx}(z) - P_{X|abc}(x) P_{Z|abcx}(z) \right| + \sum_z \left| P_{X|abc}(x) P_{Z|abcx}(z) - \frac{1}{2} P_{Z|abc}(z) \right|$$

$$= \sum_z P_{Z|abcx}(z) \left| \frac{1}{2} - P_{X|abc}(x) \right| + \sum_z \left| P_{XZ|abc}(x, z) - \frac{1}{2} P_{Z|abc}(z) \right|$$

$$\leq D(P_{X|a}, P_{\bar{X}}) + \frac{I_N(P_{XY|AB})}{2}.$$

This establishes the relation (A.6). $\qquad \qquad \square$

Tightness

We can also establish that this bound is tight, as follows. Consider a classical model in which, with probability ε, we have $X = Y = Z = -1$, and otherwise $X = Y = Z = +1$ (independently of A, B and C). This distribution has $I_N(P_{XY|AB}) = 1$ and $\nu = \frac{1}{2} - \varepsilon$. It also satisfies $D(P_{Z|abcX=-1}, P_{Z|abc}) = 1 - \varepsilon$, which is equal to the bound implied by (A.6).

Use of bipartite correlations

The argument leading to the bound in Lemma 1 is based on a bipartite scenario. As mentioned in the main text, measurements on a single system can always be explained by a local hidden variable model. We give a simple argument for this here.

For a single system, we wish to recreate the correlations $P_{X|A}$. To do so, we introduce a hidden variable, Z_a for each possible choice of measurement, $A = a$, distributed according to $P_{Z_a} = P_{X|a}$ (i.e. this hidden variable is distributed exactly as the outcomes of the measurement $A = a$ would be and has the same alphabet). When measurement $A = a$ is chosen, the outcome is given by $X = Z_a$. This local hidden variable model recreates the correlations $P_{X|A}$ precisely. In other words, no experiment on a single system can rule out a local hidden variable model of this type.

Comment on the free choice assumption and the de Broglie-Bohm model

We now discuss our result in light of the de Broglie-Bohm model [64, 6]. There, C is not used, and the parameter Z includes both the wavefuction and the (hidden) particle trajectories that make up that model. Thus, according to the argument we give above, if A can be chosen freely, $P_{YZ|AB} = P_{YZ|B}$ (this is (A.3) in the case without C), and hence $P_{Y|ABZ} = P_{Y|BZ}$. However, the de Broglie-Bohm model does not, in general, satisfy this relation: the outcome Y is a deterministic function of the wavefunction, the particle positions and *both* A and B. The dependence on A is crucial in order that the model can recreate all quantum correlations. It hence follows that the de Broglie-Bohm model does not satisfy our free choice assumption, and hence it is not in contradiction with our main claim.

A.2   Application to Leggett models

In the Leggett model [37], one imagines that improved predictions about the outcomes for measurements on spin-half particles are available. More precisely, each particle has an associated vector (thought of as a hidden direction of its spin) and the outcome distribution is expressed via the inner product with the vector describing the measurement (see Figure 5.1 in the main text). Denoting the hidden vector for the first particle by $\mathbf{z}$, and its measurement vector $\boldsymbol{\alpha}$ (this is the Bloch vector associated with the chosen measurement direction; it was denoted $\mathbf{S}_A$ ($\mathbf{S}_B$) in the main text), its outcomes are distributed according to

$$P_{X|\boldsymbol{\alpha}\mathbf{z}}(\pm 1) = \frac{1}{2}(1 \pm \boldsymbol{\alpha} \cdot \mathbf{z}). \tag{A.10}$$

To relate this back to the discussion above, the Leggett model corresponds to the case that there is no C, and where the hidden vectors are contained in Z. Note that Leggett already showed his model to be incompatible with quantum theory [37] and experiments

have since falsified it using specific inequalities [38, 67, 39, 42]. Here we discuss the model in light of our experiment, which, it turns out, is sufficient to falsify it.

As presented above, the model is not fully specified since the distribution of the hidden vectors, $\mathbf{z}$, is not given. To discuss the implications of our experimental results we refer to four cases (corresponding to different distributions over $\mathbf{z}$). In order to agree with existing experimental observations, the distribution should be such that the uniform distribution is approximately recovered when $\mathbf{z}$ is unknown, i.e. $P_{X|\boldsymbol{\alpha}} = \sum_z P_{\mathbf{Z}}(\mathbf{z}) P_{X|\boldsymbol{\alpha}\mathbf{z}} \cong \frac{1}{2}$.

Before describing the four cases, we first note that adapting our previous analysis (starting from (A.9), for example) to the case of no C implies

$$\langle D(P_{X|\boldsymbol{\alpha}\mathbf{z}}, P_{\bar{X}})\rangle_{\mathbf{z}} \leq \delta_N, \tag{A.11}$$

for all $\boldsymbol{\alpha}$, where $\langle \cdot \rangle_{\mathbf{z}}$ denotes the expectation value over the vectors $\mathbf{z}$. In order to falsify a particular version of the Leggett model, we compute δ_N^{crit}, the smallest increase in predictive power under the assumption that a particular version of the Leggett model is correct (i.e. the smallest value of the left-hand-side of (A.11) over all $\boldsymbol{\alpha}$). We then show that δ_N^{crit} is above the maximum increase in predictive power compatible with the experimental data, δ_N, hence falsifying that version of Leggett's model.

First Case: We imagine that the vector $\mathbf{z}$ is uniformly distributed between two opposite vectors (i.e. $P_{\mathbf{Z}}(\mathbf{z_0}) = P_{\mathbf{Z}}(\bar{\mathbf{z}}_0) = \frac{1}{2}$ for some fixed vector $\mathbf{z_0} = -\bar{\mathbf{z}}_0$) in the same plane on the Bloch sphere as our measurements. From (A.10), we have $D(P_{X|\boldsymbol{\alpha}\mathbf{z_0}}, P_{\bar{X}}) = D(P_{X|\boldsymbol{\alpha}\bar{\mathbf{z}}_0}, P_{\bar{X}}) = \frac{|\boldsymbol{\alpha}\cdot\mathbf{z_0}|}{2}$. Hence, from (A.11) we require $\frac{|\boldsymbol{\alpha}\cdot\mathbf{z_0}|}{2} \leq \delta_N$ for all $\boldsymbol{\alpha}$. In order to make $\max_{\boldsymbol{\alpha}} |\boldsymbol{\alpha} \cdot \mathbf{z_0}|$ as small as possible, i.e. find $\delta_N^{\mathrm{crit1}}$, we require the vector $\mathbf{z_0}$ to be as far as possible from any of the possible $\boldsymbol{\alpha}$ vectors. If the fixed vector $\mathbf{z_0}$ is in the plane containing the measurements, this condition leads to $\max_{\boldsymbol{\alpha}} |\boldsymbol{\alpha} \cdot \mathbf{z_0}| = \cos \frac{\pi}{2N}$ (i.e. $\mathbf{z_0}$ is positioned exactly in between two settings for $\boldsymbol{\alpha}$). Hence, this specific version of the Leggett model is falsified if the measured $\delta_N < \delta_N^{\mathrm{crit1}} = \frac{1}{2}\cos\frac{\pi}{2N}$. As shown in Table A.1,

105

this is the case for all values of N assessed.

According to quantum theory, appropriately chosen measurements on a maximally entangled state lead to correlations for which $\delta_N = \frac{N}{2}(1 - \cos\frac{\pi}{2N})$. However, no experimental realization can be noise-free, and this affects the minimum δ_N attainable (see [68, 58]). One way to characterize the imperfection in the experiment is via the visibility. In an experiment with visibility V[4], we instead obtain $\delta_N = \frac{N}{2}(1 - V\cos\frac{\pi}{2N})$, which for fixed V has a minimum at finite N. In the case of this model, the minimum visibility required to falsify it is 0.821 (with such a visibility the model could be ruled out with $N = 3$).

N	δ_N^{crit1}	δ_N^{crit2}	δ_N^{crit3}	δ_N^{crit4}	δ_N^1	δ_N^2
2	**0.3536**	0.2	0.25	0.1768	0.3131±0.0018	0.3125±0.0025
3	**0.4330**	**0.3062**	**0.25**	0.2165	0.2294±0.0016	0.2437±0.0023
4	**0.4619**	**0.3266**	**0.25**	**0.2310**	0.1904±0.0015	0.2094±0.0023
5	**0.4755**	**0.3362**	**0.25**	**0.2378**	0.1792±0.0014	0.2015±0.0023
6	**0.4830**	**0.3415**	**0.25**	**0.2415**	0.1676±0.0019	0.1942±0.0021
7	**0.4875**	**0.3447**	**0.25**	**0.2437**	0.1644±0.0014	0.1948±0.0021
$V_{\min}$	0.821	0.906	0.946	0.951		

Table A.1: **Leggett models: critical values and experimental data.** This table shows the critical values of δ_N required to rule out each of the four Leggett-type models discussed in the text. Also shown are measured values for δ_N^1 and δ_N^2, where the superscript refers to measurements in the $|H\rangle - |+\rangle$ plane, and the $|+\rangle - |L\rangle$ plane of the Bloch sphere, respectively. Bold values have $\delta_N^1 < \delta_N^{\text{crit }i}$ and, if required $\delta_N^2 < \delta_N^{\text{crit }i}$, i.e. the Leggett model i is ruled out by the data for that N. The values of δ_N^2 are relevant for ruling out the second and fourth model. In the last row of the table, we note the minimum visibility required to rule out each of the four models.

Second case: We now suppose $\mathbf{z}$ is distributed as in the first case, but that $\mathbf{z_0}$ is no longer confined to the plane of measurements. In this case our basic measurements cannot

---

[4] The visibility is an alternative measure of the quality of the experiment (the fidelity was used in the main text). The visibility can be directly measured, while the fidelity (to the desired state) can be calculated from the state reconstructed tomographically. Assuming isotropic noise as the dominant source of imperfection (i.e. that we actually measure Werner states), fidelity and visibility are related through $V = (4F - 1)/3$.

strictly *rule out* this model: in principle, $\mathbf{z_0}$ could be close to orthogonal to the plane containing the measurement vectors. (We remark that if $\mathbf{z_0}$ is completely orthogonal to this plane, then it would not be useful for making predictions.) However, in order to rectify this we can include a second set of measurements in the set of random choices. This set should be identical to the first apart from being contained in an orthogonal plane. We denote the sets $\mathcal{A}_1$ and $\mathcal{A}_2$ and we separately measure the δ_N values for each plane, generating values denoted δ_N^1 and δ_N^2. Analogously to the first case discussed above, this version of the Leggett model is falsified unless for all $\boldsymbol{\alpha} \in \mathcal{A}_1 \cup \mathcal{A}_2$, $|\boldsymbol{\alpha} \cdot \mathbf{z_0}|/2 \leq \min(\delta_N^1, \delta_N^2)$. In order to make $\max_{\boldsymbol{\alpha}} |\boldsymbol{\alpha} \cdot \mathbf{z_0}|$ as small as possible, we require the vectors $\mathbf{z_0}$ to be as far as possible from any of the possible $\boldsymbol{\alpha}$ vectors. Consider now the four vectors $(0, \sin\phi, \cos\phi)$, $(0, -\sin\phi, \cos\phi)$, $(\cos\phi, \sin\phi, 0)$ and $(\cos\phi, -\sin\phi, 0)$ for $\phi \leq \frac{\pi}{4}$ (these represent two neighbouring pairs of measurement vectors (one in each plane), where we have chosen the coordinates such that they are symmetric). The vector equidistant from these (in their convex hull) is $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$. It is then not possible that for all $\boldsymbol{\alpha} \in \mathcal{A}_1 \cup \mathcal{A}_2$, $|\boldsymbol{\alpha} \cdot \mathbf{z_0}|/2 \leq \min(\delta_N^1, \delta_N^2)$ provided $\max(\delta_N^1, \delta_N^2) < \delta_N^{\mathrm{crit2}} = \frac{1}{2\sqrt{2}} \cos\frac{\pi}{2N}$. As shown in Table A.1, our experiment, which includes measurements of δ_N in an orthogonal plane, also rules out this version of the Leggett model. (The minimum visibility required to rule out this model is 0.906, which could do so using $N = 4$.)

Third case: We consider a slightly modified model in which $\mathbf{z}$ is distributed uniformly over the Bloch sphere. This model is arguably more natural since it is somewhat conspiratorial for $\mathbf{z}$ to be confined to a particular set of orientations with respect to the measurements we perform (particularly if that measurement is chosen freely), and is the model referred to in the main text. In this case, defining θ as the angle between $\boldsymbol{\alpha}$ and $\mathbf{z}$, we compute the left hand side of (A.11) as

$$\langle D(P_{X|\boldsymbol{\alpha}z}, P_{\bar{X}}) \rangle_{\mathbf{z}} = \int_{\theta=0}^{\pi} \mathrm{d}\theta \frac{|\cos\theta| \sin\theta}{4} = \frac{1}{4}.$$

This model is hence excluded if one finds $\delta_N < \delta_N^{\text{crit3}} = \frac{1}{4}$ (measurements are needed only in one plane). As shown in Table A.1, this is the case for $N \geq 3$. (The minimum visibility required to rule out this model is 0.946, which could do so for $N = 5$.)

Fourth case: Here we return to our measurements in two orthogonal planes and ask whether our data is sufficient to falsify the model for any distribution over $\mathbf{z}$. (We can think of this in terms of an adversarial picture. Suppose the set of possible measurement choices is known to an adversary, who can pick the vector $\mathbf{z}$ according to any distribution he likes. The aim is to show that our measurement results are not consistent with any such adversary.) For this model to be correct we need

$$
\frac{\langle |\boldsymbol{\alpha} \cdot \mathbf{z}| \rangle_{\mathbf{z}}}{2} \leq \delta_N^1 \text{ for all } \boldsymbol{\alpha} \in \mathcal{A}_1
$$
$$
\frac{\langle |\boldsymbol{\alpha} \cdot \mathbf{z}| \rangle_{\mathbf{z}}}{2} \leq \delta_N^2 \text{ for all } \boldsymbol{\alpha} \in \mathcal{A}_2.
$$

Again we can parameterize in terms of the four vectors introduced previously. When minimizing with respect to these four, we should take $P_{\mathbf{z}}$ to have support only on the set $(\sin \theta, 0, \cos \theta)$ (going off this line increases the inner product with measurement vectors in both sets). We thus have

$$
\langle |\boldsymbol{\alpha} \cdot \mathbf{z}| \rangle_{\mathbf{z}} = \begin{cases} \int_\theta \mathrm{d}\theta \rho(\theta) \cos \theta \cos \frac{\pi}{2N} & \text{for all } \boldsymbol{\alpha} \in \mathcal{A}_1 \\ \int_\theta \mathrm{d}\theta \rho(\theta) \sin \theta \cos \frac{\pi}{2N} & \text{for all } \boldsymbol{\alpha} \in \mathcal{A}_2 \end{cases}
$$

where $\rho(\theta)$ is the probability density over θ.

In other words, non-zero $\rho(\theta)$ gives contribution $\cos \theta \cos \frac{\pi}{2N}$ to the first integral, and $\sin \theta \cos \frac{\pi}{2N}$ to the second. In order that both integrals are equal, we should take $\rho(\theta)$ to be symmetric about $\theta = \frac{\pi}{8}$. For functions with this symmetry, non-zero $\rho(\theta)$ gives contribution $(\sin \theta + \cos \theta) \cos \frac{\pi}{2N}$ to both integrals. The minimum of this over $0 \leq \theta \leq \frac{\pi}{8}$ is $\cos \frac{\pi}{2N}$, which occurs for $\theta = 0$. It follows that the most experimentally challenging distribution to rule out is $\rho(\theta) = \frac{1}{2}(\delta_{\theta,0} + \delta_{\theta,\frac{\pi}{4}})$, where $\delta_{x,y}$ is the Kronecker delta (this being the distribution that requires the lowest measured δ_N to eliminate). For

this distribution, we have $\max_{\boldsymbol{\alpha}} \langle |\boldsymbol{\alpha} \cdot \mathbf{z}| \rangle_{\mathbf{z}}/2 = \frac{1}{4} \cos \frac{\pi}{2N}$, so this model is ruled out for $\max(\delta_N^1, \delta_N^2) < \delta_N^{\text{crit4}} = \frac{1}{4} \cos \frac{\pi}{2N}$. Again, as detailed in Table A.1, our experimental data is sufficient to do so. (The lowest visibility that could rule out this case is 0.951, which would do so for $N = 5$).

Note that, while discussing this case, we have so far ignored the requirement $\sum_z P_{\mathbf{Z}}(\mathbf{z}) P_{X|\boldsymbol{\alpha}\mathbf{z}} = \frac{1}{2}$. However, this condition can be ensured (without changing the critical value δ_N^{crit4}) by replacing the probability density $\rho(\theta)$ with $\frac{1}{2}(\rho(\theta) + \rho(\pi + \theta))$, i.e. by distributing the density of each vector evenly between itself and the vector orthogonal to it.

Comment on minimum visibilities required to rule out Leggett models

Here we briefly compare the visibilities required to rule out Leggett models using our approach with those needed in previously considered Leggett inequalities. We remind the reader that the technique used in the present work generates conclusions that apply to arbitrary theories and were not developed with Leggett's model in mind. Nevertheless, use of this new approach to rule out Leggett models requires comparable visibilities to those of previously discussed inequalities. More specifically, the claimed minimum visibilities are 0.974 in Gröblacher *et al.* [38] and 0.943 for the alternative inequality of Branciard *et al.* [39, 42], which is only slightly below the value we require to rule out all of the four models above.

We note that the average visibility for measurements in the plane used in the main text was 0.9781 ± 0.0008, while the average visibility in the orthogonal plane (measured for the purposes of ruling out the second and fourth cases) was 0.9706 ± 0.0014.

A.3   Visibility versus δ

As discussed in the main manuscript, assuming the quantum theoretical predictions to be optimum, the minimum measurable value for δ_N, hence the bound on δ, depends on

the quality of the generated bi-photon state and the measurement apparatus (character-ized, e.g., through the visibility). This is depicted in Figure A.1 where, for simplicity, we assume a bias of zero (i.e. $\nu_N = 0 \ \forall \ N$). In order to decrease δ by more than a factor of two compared to our result, the average visibility on the measurement plane must exceed 0.995 (assuming zero bias and perfect measurement settings), and the required value of N increases to 15 or beyond, resulting in 120 or more high-precision coincidence mea-surements. To decrease δ below 1 %, we require $V > 0.9999$ and $N > 111$. We emphasize that the precision required in the waveplate settings (that determine the spin measure-ments) increases with N, which rapidly constitutes another limitation to obtaining small values for δ, in addition to the need for a high-quality source.

A.4   Raw Data

The experimental settings as well as the associated measurement results that allow re-construction of the density matrix are given in Table 3.1. The most likely density matrix is detailed in Table A.2. Note that this density matrix is not used for the calculation of experimental values for δ_N, I_N or ν_N, but is included to characterize our source. The measurements settings used to experimentally determine δ_7 are depicted in Figure A.2, and Table A.3 lists the results used to calculate δ_7 from the bi-partite correlation I_7 and the bias ν_7.

(a) ρ_{Re}

	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.5031	0.0056	-0.0196	0.4828
$\|HV\rangle$	0.0056	0.0033	0.0006	0.0113
$\|VH\rangle$	-0.0196	0.0006	0.0032	-0.0115
$\|VV\rangle$	0.4828	0.0113	-0.0115	0.4904

(b) ρ_{Im}

	$\langle HH\|$	$\langle HV\|$	$\langle VH\|$	$\langle VV\|$
$\|HH\rangle$	0.0000	0.0020	0.0046	-0.0007
$\|HV\rangle$	-0.0020	0.0000	0.0002	-0.0012
$\|VH\rangle$	-0.0046	-0.0002	0.0000	-0.0036
$\|VV\rangle$	0.0007	0.0012	0.0036	0.0000

Table A.2: **Density matrix.** The real and imaginary parts of the density matrix gener-ated by maximum likelihood quantum state tomography.

Setting		HWP_A	HWP_B	R_{Si}	R_C	$1-P(m,n)$	$P(m,n)$	$\Delta P(m,n)$	ν	$\Delta\nu$
m	n	(°)	(°)	(cps)	(cps)					
0	13	0	41.79	41885	10.6					
0	27	0	86.79	41825	546.0	0.0207	0.9793	0.0007	0.0008	0.0002
14	27	45	86.79	41908	12.5					
14	13	45	41.79	42068	544.3					
0	1	0	3.21	41847	545.2					
0	15	0	48.21	41855	9.2	0.9836	0.0164	0.0006	0.0011	0.0002
14	15	45	48.21	41954	547.9					
14	1	45	3.21	42121	9.1					
2	1	6.43	3.21	41826	540.5					
2	15	6.43	48.21	41871	11.4	0.9798	0.0202	0.0007	0.0013	0.0002
16	15	51.43	48.21	42028	552.5					
16	1	51.43	3.21	42102	11.1					
2	3	6.43	9.64	41829	546.3					
2	17	6.43	54.64	41880	11.5	0.9781	0.0219	0.0007	0.0006	0.0002
16	17	51.43	54.64	42024	544.2					
16	3	51.43	9.64	41886	12.9					
4	3	12.86	9.64	41871	541.0					
4	17	12.86	54.64	41806	13.4	0.9745	0.0255	0.0007	0.0009	0.0002
18	17	57.86	54.64	41929	543.0					
18	3	57.86	9.64	42037	15.0					
4	5	12.86	16.07	41739	545.4					
4	19	12.86	61.07	41757	11.4	0.9779	0.0221	0.0007	0.0013	0.0002
18	19	57.86	61.07	41975	555.4					
18	5	57.86	16.07	41967	13.5					
6	5	19.29	16.07	41595	541.2					
6	19	19.29	61.07	41776	17.5	0.9717	0.0283	0.0008	0.0023	0.0002
20	19	64.29	61.07	42043	548.8					
20	5	64.29	16.07	42109	14.2					
6	7	19.29	22.5	41752	548.7					
6	21	19.29	67.5	41805	12.5	0.9760	0.0240	0.0007	0.0022	0.0002
20	21	64.29	67.5	42181	548.1					
20	7	64.29	22.5	42121	14.6					
8	7	25.71	22.5	41886	540.4					
8	21	25.71	67.5	41907	14.3	0.9727	0.0273	0.0008	0.0016	0.0002
22	21	70.71	67.5	42189	549.7					
22	7	70.71	22.5	42143	16.3					
8	9	25.71	28.93	41763	548.9					
8	23	25.71	73.93	41795	12.9	0.9737	0.0263	0.0008	0.0023	0.0002
22	23	70.71	73.93	42180	545.4					
22	9	70.71	28.93	42135	16.7					
10	9	32.14	28.93	42097	554.4					
10	23	32.14	73.93	42260	14.1	0.9744	0.0256	0.0007	0.0008	0.0002
24	23	77.14	73.93	42038	548.7					
24	9	77.14	28.93	42055	14.9					
10	11	32.14	35.36	42039	554.5					
10	25	32.14	80.36	42306	12.2	0.9768	0.0232	0.0007	0.0005	0.0002
24	25	77.14	80.36	42063	557.4					
24	11	77.14	35.36	42116	14.3					
12	11	38.57	35.36	42515	556.4					
12	25	38.57	80.36	42325	14.0	0.9753	0.0247	0.0007	0.0025	0.0002
26	25	83.57	80.36	41993	544.0					
26	11	83.57	35.36	42005	13.1					
12	13	38.57	41.79	42281	534.4					
12	27	38.57	86.79	42324	9.4	0.9825	0.0175	0.0007	0.0022	0.0002
26	27	83.57	86.79	41879	535.9					
26	13	83.57	41.79	41985	9.7					

Table A.3: **Raw Data used to calculate δ_7^1.** This table shows raw data collected to find $\delta_7^1 = 0.1644 \pm 0.0014$. $HWP_{A/B}$ are the half wave-plate settings that realize the measurements corresponding to m and n as shown in Figure A.2. The Si APD rates (R_{Si}) and the coincidence rates between the Si APD and the triggered InGaAs APD (R_C) are both given in average cps. $P(m,n)$ is the probability of correlated outcomes and ν is the bias for individual measurements as detailed in the Methods section. Data collection time for each point was 40 seconds. Uncertainties are one standard deviation.
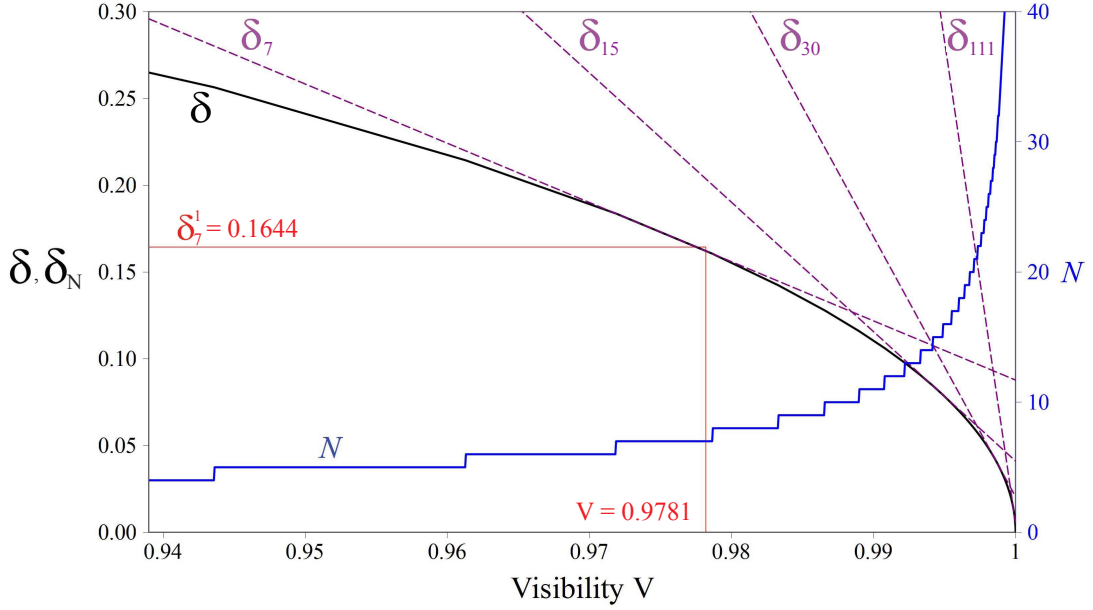
Figure A.1: δ (**minimum possible δ_N**) **and required number of bases per side N as a function of visibility V.** The stepped curve (N) uses the right y-axis. The curves are calculated using Werner states of varying visibility. The vertical and horizontal lines correspond to the average visibility for measurements in the plane used in the main text ($V = 0.9781$), and $\delta_7^1 = 0.1644$ respectively. The slight discrepancy between the intersection of these two lines and the curve showing $\delta(V)$ can be attributed to non-zero bias, imperfect measurements, and deviation of the experimental state from a Werner state. The dashed diagonal lines show δ_N as a function of visibility for $N = 7$, 15, 30, and 111. Note that, as $V \to 1$ and $\delta \to 0$, $N \to \infty$. Hence, significantly lowering δ requires not only higher visibilities than currently feasible [67, 39], but significantly more measurement settings (hence higher precision) also.
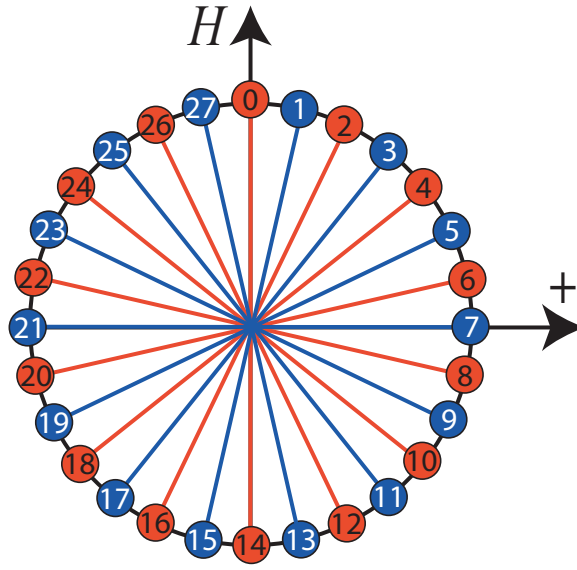
Figure A.2: **Measurement settings for** $N = 7$. All settings are in the $|H\rangle$-$|+\rangle$ plane in the Bloch sphere. The settings on one side of the bipartite setup are indicated in red (even numbers) and those on the other side are indicated in blue (odd numbers).

Appendix B

Copyright Permissions

B.0.1 Figures 2.5 and 2.7

APS physics

January 11, 2013

Terence Stuart
terence.stuart@gmail.com

Ref # 22231

Thank you for your permission request dated January 7, 2013. We are pleased to grant you a non-exclusive, non-transferable permission, English rights, limited to **print and electronic format**, provided you meet the criteria outlined below. Permission is for a one-time use and does not include permission for future editions, updates, databases, translations, or any other matters. Permission must be sought for each additional use. This permission does not include the right to modify APS material.

Please print the required copyright credit line on the first page that the material appears: "Reprinted (abstract/excerpt/figure) with permission from [FULL REFERENCE CITATION] as follows: authors names, journal title, volume number, page number and year of publication. Copyright (YEAR) by the American Physical Society.

The following language must appear somewhere on the website: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or part, without prior written permission from the American Physical Society."

Provide a hyperlink from the reprinted APS material (the hyperlink may be embedded in the copyright credit line). APS's link manager technology makes it convenient and easy to provide links to individual articles in APS journals. For information, see: http://link.aps.org/.

You must also obtain permission from at least one of the authors for each separate work, if you haven't done so already. The author's name and address can be found on the first page of the published Article.

Use of the APS material must not imply any endorsement by the American Physical Society.

Permission is granted for use of the following APS material only:
- Fig. 1, Phys. Rev. A 73, 032326 (2006)
- Fig. 1, Phys. Rev. A 73, 012316 (2006)
- Figs. 6, 8, Rev. Mod. Phys. 74, 145–195 (2002)

Permission is limited to the single title specified of the publication as follows:
A thesis entitled, "Probing the Completeness of Quantum Theory with Entangled Photons," to be published by the University of Calgary

If you have any questions, please refer to the Copyright FAQ at: http://publish.aps.org/copyrightFAQ.html or send an email to assocpub@aps.org.

Sincerely,

Melissa Overton
Publications Marketing Coordinator

GMail
by Google

Copyright Permission for Phys. Rev. A 73, 012316
2 messages

Terence Stuart   ████████████████   Fri, Jan 11, 2013 at 4:34 PM
To: ████████████

Dear Dr. Wong,
     I would like to reproduce, in my thesis, Figure 1 from a paper you published in Physical Review A in 2006:

http://pra.aps.org/abstract/PRA/v73/i3/e012316

I have received permission from APS, but also need permission from one of the authors.

This thesis will be submitted to the following online repositories:

University of Calgary Theses Repository http://theses.ucalgary.ca/
Library and Archives Canada http://collectionscanada.gc.ca/obj/s4/f2/frm-nl59-2-e.pdf

Please let me know if this is possible.

Cheers,

Terence Stuart

Franco Wong ████████████████   Mon, Jan 14, 2013 at 12:47 PM
To: Terence Stuart ████████████████████

Hi, Terence,

On behalf of my co-authors, you have the permission to reproduce Figure 1 of our PRA paper (vol. 73, article 012316).

After your thesis submission, I would be very interested in getting a pdf copy of your thesis if that's not too much trouble.

Regards,
Franco Wong
--
[Quoted text hidden]
--

_____
Dr. Franco N. C. Wong
Senior Research Scientist
MIT, ████████████████
████████████████
Phone: ████████████
email: ████████████
_____

████████████████████████████

Gmail by Google

Copyright Permission for Phys. Rev. A 73, 032326 (2006)

2 messages

Terence Stuart ████████████████████████   Fri, Jan 11, 2013 at 4:41 PM
To: ████████████████

Dear Dr. Ljunggren,
    I would like to reproduce, in my thesis, Figure 1 from a paper you published in Physical Review A in 2006:

http://pra.aps.org/abstract/PRA/v73/i3/e032326

I have received permission from APS, but also need permission from one of the authors.

This thesis will be submitted to the following online repositories:

University of Calgary Theses Repository http://theses.ucalgary.ca/
Library and Archives Canada http://collectionscanada.gc.ca/obj/s4/f2/frm-nl59-2-e.pdf

Please let me know if this is possible.

Cheers,

Terence Stuart

Daniel Ljunggren ████████████████████████   Mon, Jan 14, 2013 at 12:59 PM
To: Terence Stuart ████████████████

Dear Terence,

You are hereby granted to reproduce any of the figures in the named article as
long as they are properly cited.

I also attach the original eps file.

My own thesis also provides additional information on the subject if you
haven't found it already.
http://kth.diva-portal.org/smash/get/diva2:14623/FULLTEXT01

I'd like to get a electronic copy of your thesis once it is available.

Good luck!

/Daniel
[Quoted text hidden]

📄 **sourcecoordinate.eps**
45K

B.0.2 Figure 4.5

Note: All other reproduced figures are considered "fair dealing".