TESTING THE LIMITS OF NONLOCALITY

BY

JOSEPH BENJAMIN ALTEPETER

B.S., Washington University, 2000

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2006

Urbana, Illinois

Abstract

The nascent field of quantum information offers the promise of dramatic speed increases for certain types of computation and logical protocols which are classically impossible. The phenomenon of nonlocality is fundamental to this study, and polarization-entangled photons offer a clean, bright, and stable system for its experimental investigation. This thesis investigates, both experimentally and theoretically, the creation, manipulation, and measurement of polarization-entangled photon pairs.

Our entangled photon pairs are created through spontaneous parametric down-conversion within a pair of adjacent, orthogonally oriented nonlinear crystals. The quality and brightness of this source is then dramatically improved through the use of birefringent crystals which compensate for an angle-dependent phase shift. The use of these compensation crystals has allowed measured count rates of two million pairs per second with 97.7% fidelity (with a maximally entangled state) or alternately, ten thousand pairs per second with 99.5% fidelity.

By manipulating these entangled states, it is possible to study both how they change and how they can be used. We discuss the theory of state manipulation and the experimental implementation of extremely precise single-qubit operations. To study the operations themselves, we use quantum process tomography to characterize them and have successfully implemented the first experimental realization of ancilla-assisted process tomography. To study decoherence, we implemented the first experimental decoherence-free subspace. By using techniques developed during these investigations, we are able to transform our source of entangled photons into a source of any two-photon polarization state, mixed or pure.

Developing new techniques for state creation and manipulation is possible because of simultaneous development in state measurement. In addition to detailing both a theoretical analysis and experimental instructions for state tomography, we experimentally and theoretically compared state tomography with tests of local realism (Bell inequalities) and entanglement witnesses (an entanglement detection technique). In the process, we have measured the largest violations of local realism to date, both statistical (over 2400-σ) and absolute ($S = 2.826 \pm 0.005$—within 0.2% of a maximal violation).

To Mom and Dad.

Acknowledgments

This thesis is possible because of my teachers, and the many hours over many years that they spent making me a better person. Some of these teachers I met in elementary school: Miss Clark, Mrs. Langegrader, Miss Wank, Mrs. Marcinkiewicz, Miss Stecher, Miss Bockskopf, Sister John—who showed me my first equations and taught me that mathematics will be right more often than I will, Mrs. Winkelman, Miss Horn, Mrs. Foote, Mrs. Didde, Mrs. Graham, and Sister Jo Donna. Some I met in high school: Amy Applebee, Ed Bolhofner, Brenda Bollinger, Gerda Bosch, Ellen Bozich, Ellen Doweny, Larry Hester, Gail Johnson, Janice Mueth, Randy Naber, Earl Nold, Jean Pfeifer, Don Piekarski, Bob Robidoux, Carol Robinson, Mary Rodriguez, Cecil Sharp, Jeanne Weber, and Sherrie Wisdom. Some helped me through college: Andrea Heugatter, William Buhro, Brian Stratton, Mark Conradi, Carl Bender, Wim Dickhoff, Wai-Mo Suen, Malcolm Tobias, James Buckley, James Poag, Alan Wheeler, and Ken and Sally Goldman. Special thanks go to Ron Loui, who introduced me to research and taught me vi. In my graduate career, I've learned more than at any other time, due to the work of both professors—Andrew White, Daniel James, Paul Goldbart, and Dave Branning—and fellow students—Tzu-Chieh, Nick, Julio, Ria, Onur, and Matt. Special thanks go to my mentor, Paul Kwiat, who showed me how to learn and how to teach things that no one knows. Throughout my life, my friends have been some of my best teachers: Josh, Jeremy, Danny, Kevin, Rich, Jason, Brandon and Danielle, Phillip, Evan, Evan, Micah, Amy, Carla, Emily, Michelle, Aaron, and Daryl. And of course the most important teachers: Grandma and Grandpa, Mom and Dad, Greg, Stephanie, Corey, Tara, and the rest of my family.

In addition, this work described herein would not have been possible without the generous support of the National Science Foundation, the National Science Foundation Graduate Fellowship Program, the Australian-American Fulbright Association and the Fulbright Fellowship Program, the Disruptive Technologies Office (formerly the Advanced Research and Development Activity), a Multidisciplinary University Research Initiative, the National Security Agency, the Army Research Office, and the University of Illinois.

Special thanks to my thesis committee: Mike Weissman, Nigel Goldenfeld, Paul Goldbart, and Paul Kwiat.

Table of Contents

List of Tables

List of Figures

List of Examples

List of Abbreviations

A	anti-diagonal
AAPT	ancilla-assisted process tomography
BBO	β-Barium Borate (nonlinear down-conversion crystal)
BS	beam splitter
CG	Collins-Gisin
CHSH	Clauser, Horne, Shimony, and Holt (a type of Bell inequality)
CNOT	controlled NOT gate
CSIGN	controlled SIGN, equivalent to a controlled Z gate
D	diagonal
DFS	decoherence-free subspace
EAPT	entanglement-assisted process tomography
EPR	Einstein, Podolsky, and Rosen
FWHM	full-width half-maximum
H	horizontal
HWP	half waveplate
IF	interference filter
L	left-circular
MEMS	maximally entangled mixed state(s)
PBS	polarizing beam splitter
R	right-circular
SPCM	single-photon counting module
SQPT	standard quantum process tomogaphy
TTL	transistor-transistor logic
qubit, q-bit	quantum bit
qudit	d-level quantum system (generalization of the qubit)
QWP	quarter waveplate
V	vertical

List of Symbols

$\hat{}$	denotes a matrix (or, rarely, a unit vector)
$\hat{\rho}$	a density matrix used to represent a quantum state
$\|\psi\rangle$	a pure quantum state
$\|H\rangle$	horizontally polarized light ($\|H\rangle \equiv \|0\rangle$)
$\|V\rangle$	horizontally polarized light ($\|V\rangle \equiv \|1\rangle$)
$\|D\rangle$	diagonally polarized light $\left(\|D\rangle \equiv \frac{\|H\rangle+\|V\rangle}{\sqrt{2}}\right)$
$\|A\rangle$	anti-diagonally polarized light $\left(\|A\rangle \equiv \frac{\|H\rangle-\|V\rangle}{\sqrt{2}}\right)$
$\|R\rangle$	right-circularly polarized light $\left(\|R\rangle \equiv \frac{\|H\rangle+i\|V\rangle}{\sqrt{2}}\right)$
$\|L\rangle$	left-circularly polarized light $\left(\|L\rangle \equiv \frac{\|H\rangle-i\|V\rangle}{\sqrt{2}}\right)$
$\hat{\sigma}_i$	$\hat{\sigma}_0 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \hat{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \hat{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\hat{\rho}_W$	Werner state $\left(\hat{\rho}_W = P\|\gamma\rangle\langle\gamma\| + (1-P)\frac{1}{4}I\right)$
I	The Identity
$\|\phi^{\pm}\rangle$	Two of the four Bell states $\left(\frac{1}{\sqrt{2}}\left(\|HH\rangle \pm \|VV\rangle\right)\right)$
$\|\psi^{\pm}\rangle$	Two of the four Bell states $\left(\frac{1}{\sqrt{2}}\left(\|HV\rangle \pm \|VH\rangle\right)\right)$
F	Fidelity
T	Tangle
C	Concurrence
S_L	Linear entropy
S	Von Neuman entropy
$\|\gamma\rangle$	An arbitrary maximally entangled state
$\mathrm{Tr}\{\}$	Trace

1 Introduction

Until the early 20^{th} century, the language of physics described a world in which any phenomenon could—in principle—be measured and *predicted* by a nearby observer. The advent of quantum mechanics challenged part of this world view by allowing for the possibility of true randomness—true unpredictability. This thesis, however, concerns the experimental investigation of a more subtle and potentially more profound consequence of quantum theory: that a nearby quantum object can become inextricably linked to a distant system, in such a way that it is *impossible* to completely describe the near object without considering its far off partner. This linking between distant objects is called *nonlocality*, and is quite possibly the strangest and most fundamental of quantum phenomena.

This seemingly paradoxical behavior was first pointed out in 1935 by Einstein, Podolsky, and Rosen in their now famous EPR paper [25]. It wasn't until 1964 that Bell proposed a way to test this phenomenon [12], and in 1972 Freedman and Clauser experimentally observed the phenomenon of nonlocality for the first time [26]. Since then, many groups have corroborated these results using a variety of systems and techniques [75]. Interest in this fundamental phenomenon has been magnified by a confluence of additional experimental and theoretical results which have developed into the nascent field of quantum information [52]. These results include the realization that this type of strange linking can lead to quantum computers which are able to solve certain problems much faster than their classical counterparts (examples include factoring large numbers [62] and searching unordered databases [29]), and cryptographic schemes which base claims of absolute security on quantum mechanical law [28].

Much of the excitement surrounding quantum information stems from conjecture about what will be possible when and if we gain precise experimental control of large, complex quantum systems. However, at this early stage there is still a great deal that is not understood, even about very simple quantum systems. This thesis describes efforts to explore the limits of nonlocality using one of the simplest quantum systems: pairs of polarized photons. These efforts, more specifically, have pushed the limits of what is possible when creating, manipulating, and measuring quantum states. Our experiments have been carried out, not just to observe phenomena which had never before been seen, but in the hopes that the process might provide an intuitive insight into quantum mechanics itself.

This thesis is presented in a tutorial form (e.g., it includes a number of

subjects and derivations which could be found in textbooks) in order to allow those being introduced to the field, and particularly to experiments on entangled photons, to absorb the material without extensive reference to a number of different basic sources.

1.1 An abridged glossary of quantum information

1.1.1 The qubit

The fundamental unit of classical information is the *bit*, a single piece of information which can take exactly one of two values: 0 or 1. Quantum information's fundamental unit is analogous, a quantum system which has two levels, $|0\rangle$ and $|1\rangle$ and is referred to as the *qubit* or sometimes *q-bit*.

While quantum and classical bits share some characteristics (for example, when either type of bit is measured, there are exactly two possible results), there are important and irreconcilable differences between them. Classical states can be copied easily, while it is physically impossible to copy an unknown qubit. (It is this phenomenon which makes qubits useful for quantum cryptography). Classical bits always have exactly one value, either 0 or 1. Quantum bits can exist in superpositions of either $|0\rangle$ or $|1\rangle$, occupying a huge range of states somewhere "in between" the two extremes. In fact, quantum computers *require* superpositions of multiple-qubit states.

Finally, there are important disadvantages to quantum bits as opposed to classical bits. Because qubits are two-level quantum systems, they are generally more delicate than classical bits, being either difficult to transmit (electrons or atoms) or difficult to store (photons). Additionally, for most implementations of quantum bits, it is either difficult to avoid decoherence due to interaction with the environment (for electrons and atoms) or difficult to make the qubits interact with each other (for photons).

The experiments described in this thesis use polarization-encoded photons as the qubits, with a horizontally polarized photon representing a $|0\rangle$ and a vertically polarized photon representing a $|1\rangle$. (For more information on polarization-encoded qubits, see Example 2.1 in Section 2.1.1.)

1.1.2 Measurement

Measuring qubits is very different from measuring classical bits. There is only one way to measure a classical bit, and only one possible answer: 0 or 1. After it is measured, a classical bit is exactly the same as it was before. For quantum bits, there are many ways to make a measurement, and the act of making that measurement can nondeterministically *change* the state of the qubit being measured.

Consider one of the simplest qubits: a single electron's spin. The state of an electron's spin can be completely represented by an arrow pointing in space (for example, *straight up* along the positive z-axis). Measurements of an electron's spin *must be made* with respect to a single axis in space; for example, the z-axis or the x-axis. The result of this measurement will give one of two answers: the spin is pointing towards the *positive* direction of the measurement axis **or** the spin is pointing towards the *negative* direction of the measurement axis.

If the electron's spin is pointing along the measurement axis, the result of the measurement will be this direction, and the state of the spin will not be changed. However, if the measurement is made along a *different* axis, then the result of the measurement will be at least partially random. The further from the electron's spin the measurement axis is, the more random the result will be, until measurements made along axes which are perpendicular to the electron's spin will be totally random.

When measurements are made on an axis different than the electron's spin, the state of the electron's spin is changed. It thereafter points along the direction of the result of the measurement, changing *because* it was measured. This type of behavior exists for all quantum systems, though for some systems the "axes" of measurement are harder to visualize. In general, an "axis" of measurement is called a *basis* (plural bases).

1.1.3 Correlation and anticorrelation

The results of separate measurements on separate qubits (e.g., separate electron spins) can be related, either having a tendency to point in the same direction or in opposite directions[1]. These are called *correlated* results (sometimes "correlated" refers to results which always match and "anticorrelated" refers to results which never match). Correlation and anticorrelation can exist in one basis or in several. Two electrons with positive z-axis spin will exhibit perfect correlation if measured in the z basis (they will both always be measured as spin-up) but no correlation at all in the x or y bases. It is also possible for the measurement of some special states (see "Entanglement" below) to give *random*—yet perfectly correlated—results.

1.1.4 Entanglement

Two particles are *entangled* if their physical states cannot be described independently[2]. The quantum state of particles which *can* be described separately is said to be *separable*. In a classical world, all groups of objects are separable, and it is possible to describe them completely independently. They can interact

[1] Here, "direction" refers to one-half of a measurement basis. "Opposite directions" refers to the two different outcomes that can arise from a measurement in a single basis. For a 2-level electron's spin, this will correspond to opposite physical directions in normal space.

[2] More correctly, a *complete* description of a pair of entangled particles cannot be divided into a description of particle 1 and a description of particle 2.

with each other, but only if nearby. In a quantum world, however, entangled particles exhibit *nonlocality*, and need to be described as a joint system for that description to be complete, regardless of the distance between the particles. Entanglement has strange consequences: measurements made *simultaneously*, on two particles separated by *any distance*, can have totally random, yet totally correlated results.

As an example of entanglement, consider the "singlet state" of two electron spins. If one particle in a singlet state is measured, regardless of the direction in which the measurement is made, the result will be truly random. However, if both partners are subjected to the same measurement, regardless of how far they are from each other, they will *always* be measured to be in opposite directions. This type of behavior, totally random yet totally anticorrelated results, is only possible with entanglement.

1.1.5 Decoherence

True randomness is one of the strangest features of a quantum world, and violates the fundamental classical assumption of *realism*[3]. It occurs even in separable states, like a single, unentangled electron spin. Normal spins are described by a direction, and the results of measurements made along that direction's axis will not be random. This is really only true of what are called *pure states*. A spin which has been completely *decohered* will give random results in *every* direction. This type of spin is called *mixed*, the opposite of pure.

This omnidirectional randomness is a result of this spin being *entangled* with something else, such as another spin. A state which is decohered looks random because it cannot be described completely without its partner (or partners, for multi-party entanglement). If one could describe both partners, they would appear to be in a pure (non-decohered), entangled state which is both completely random, yet completely correlated. If half of an entangled state is measured by itself, it looks completely random, or decohered.

It is important to note that this is a somewhat controversial topic. It is possible that instead of decoherence always resulting from entanglement, there is some fundamental transition from the quantum to the classical world which causes decoherence. One manifestation could be a maximum possible size for *quantum* objects, with any larger objects devolving into classical states, and with any observed randomness just the result of a complicated and noisy system. It is the opinion of the author, however, that in the absence of any experimental evidence to support these alternate theories, the former view is more useful.

[3]Realism requires that a complete description of a system be sufficient to predict the result of any measurement that is made upon that system. Realism is a slightly more general version of *determinism*; the latter additionally requires that the complete description of a system be, at least in principle, accessible to an experimenter.

1.1.6 Tomography

In some ways, quantum states are very difficult to measure. In the simple example of an electron spin, one can make a measurement in any direction and get one of two answers (up or down, relative to that axis). However, the act of making this measurement *changes* the electron's state so that its spin is now parallel with the measurement axis in one of two positions. This means that any additional measurements will not give any information about the *original* state. Unless one's first measurement happened to be carried out using the axis of the electron's spin, one's guess about what the electron's spin was *before the measurement* will be wrong.

The solution to accurately determining an unknown quantum state requires having access to not just one copy of that state, but to many. If it is possible to prepare many identical copies of an unknown state (continuing the example from above, many electrons all with the same unknown spin), then many measurements from many directions will eventually allow reconstruction of the unknown quantum state. This process of using many measurements in many bases to reconstruct the quantum state of an ensemble of particles is called *quantum state tomography*.

1.2 Synopses of thesis chapters

The investigations of nonlocality presented in this thesis can be conceptually divided into three topics: the creation, manipulation, and measurement of quantum states. However, research into each of these three areas has proceeded simultaneously, and advances in each have been very useful for continuing research into the other two. As a result, after Chapter 2's introduction to state representation, the last three chapters can be consulted in any order.

1.2.1 Chapter two synopsis: state representation

This chapter establishes a common notation for and introduces the basic theory of quantum states. It begins with the theory of single-qubit states, and makes use of the graphical representation of the Bloch/Poincaré sphere. This sphere (Bloch for electron spin, Poincaré for photon polarization) shows the space of all possible single-qubit states; pure states are on the sphere's surface, and mixed states on the inside. Chapter 2 also introduces the notation of pure and mixed states.

A pure state is the most basic form of quantum state, and is a complete description of a system which is unentangled to any other particles or degrees of freedom. Even this isolated quantum system is different than its classical analog, however, as it can exist in a superposition, in some sense simultaneously coexisting in multiple states. For example, a photon can be polarized in the horizontal or vertical direction ($|H\rangle$ or $|V\rangle$ respectively), but can also exist in

a superposition of the two such as $\frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$ (otherwise known as 45° or diagonally polarized light). Either a horizontally or vertically polarized photon has only a 50% chance of passing through a 45° polarizer, but their diagonal superposition will be transmitted every time. Two states of the same photon, such as the horizontal or vertical components of diagonal light, can only exist in a superposition with each other if they are *indistinguishable*, in other words, it is not possible to tell the two superposed states apart through any method other than their direct measurement.

If different possible states of a particle are in principle distinguishable, they contribute not to a pure state in a superposition, but rather to a *mixed* state. (A mixture of H and V is *not* diagonal light, and will not completely pass through a diagonal polarizer; some of the H and some of the V would pass through, but only a pure superposition of the two make 45° light.) When multiple terms in a pure state are added in superposition, it is sometimes said they add *coherently*. Similarly, when multiple states are distinguishable, it is said that they add *incoherently* (if the multiple states are different from each other, this will result in a mixed state). Pure states are represented using one dimensional complex vectors (e.g., $|\psi\rangle$). Mixed states are represented using two dimensional complex Hermitian matrices called *density matrices* (e.g., $\hat{\rho}$), a more general notation which can also be used to represent pure states.

In the next sections, single-particle states with more than two levels ("qudits") and multiple-particle states are covered. For two-qubit systems, the abbreviation $|AB\rangle$ is used to represent a two-particle state with the first particle in state $|A\rangle$ and the second particle in state $|B\rangle$. For example, $|HH\rangle$ refers to a two-photon state with both photons horizontally polarized.

This chapter also introduces state diagonalization and the Schmidt decomposition, which allow a quantum state to be rewritten in a form that gives insight into that state's physical meaning. For example, diagonalizing a partially mixed state reveals a combination of pure states that can be used to create that state. Alternately, the Schmidt decomposition allows one to easily see if a state is entangled. Finally, methods for representing entanglement and decoherence are discussed.

1.2.2 Chapter three synopsis: state creation

Entangled photon pairs (our two-qubit system) are created using spontaneous parametric down-conversion, whereby a high-energy photon is split into two low-energy daughter photons inside a specially designed nonlinear crystal. In our case, 351-nm pump photons are down-converted into pairs of 702-nm daughter photons. In order to fulfill energy and momentum conservation inside our nonlinear crystals, the daughter photons produced always have the same polarization, a polarization which is necessarily *orthogonal* (at right angles) to the optic axis of the crystal and the polarization of the pump photons. This allows

production of the correlated yet separable states

$$|\psi_1\rangle = |HH\rangle \quad \text{and} \quad |\psi_2\rangle = |VV\rangle, \tag{1.1}$$

respectively produced by nonlinear crystals whose optic axes are oriented in either the vertical ($|HH\rangle$) or horizontal ($|VV\rangle$) planes. By using a diagonally polarized pump beam (a beam in a superposition of horizontal and vertical light) incident on a "sandwich" of two such orthogonally oriented nonlinear crystals, both $|\psi_1\rangle$ and $|\psi_2\rangle$ can be simultaneously produced. If the crystals are thin enough[4], and the coherence length of the pump laser is long enough, then these two processes will be technically *indistinguishable*, leading to the maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\phi}|VV\rangle\right). \tag{1.2}$$

This chapter also describes efforts to make a very bright (many entangled pairs produced per second), very high quality source (very pure and very entangled states). These efforts have resulted in a source which has a 98% fidelity (defined in Appendix A) with a maximally entangled state and over 2×10^6 measured entangled photons per second, the world's brightest entanglement source reported as of this writing. In addition, a state with $99.4\% \pm 0.1\%$ fidelity and 10^4 measured pairs per second has also been produced, the highest fidelity with a maximally entangled state measured to date.

These remarkably bright and pure states are difficult to generate, due to an angle-dependent phase which is present when using the "sandwich" method described above. When a large solid angle of the produced states is collected, many different maximally entangled states are simultaneously measured. Because their angle distinguishes them, they are added incoherently, and result in a mixed state. To create bright, pure states, extra birefringent elements are added to the system; these compensate for (i.e., reverse) this angle-dependent phase shift, making the output polarization states identical, regardless of the angle at which they leave the crystals. Even though they are distinguishable, states at many different angles combine to form a single, very bright and very pure entangled state.

Finally, this chapter describes techniques for creating nonmaximally entangled states. Techniques for creating other types of two-qubit states, which run the gamut from pure to mixed and separable to entangled, are discussed in Chapter 4.

[4]While thin crystals are sufficient to obtain indistinguishability, the thickness of the crystals is not to be directly compared to the coherence length of the pump laser. Rather, the walk-off between the pump laser and the down conversion in the first crystals needs to be small compared to the coherence length of the laser, so that two processes cannot be distinguished using timing information. This condition is more easily satisfied for thinner crystals, as is the requirement that the down-converted cones spatially overlap.

1.2.3 Chapter four synopsis: state manipulation

This chapter begins by discussing the basic tools for manipulating a state: unitary operators, which transform pure states into other pure states and mixed states into other mixed states; projectors, which allow state measurement; decoherers, which introduce noise into a system (often reversibly in our case); and entangling operations, types of unitary transforms which create nonlocal links. Most important are unitary transformations, which—in the single-qubit case—correspond to rotations about an arbitrary axis through the Poincaré sphere. In fact, unitary transformations on any system, regardless of the size, are mathematically identical to rotations, though in higher dimensions they lack a simple graphical representation. These higher dimensional rotations are remarkably powerful, and can often be used to represent seemingly nonunitary operators like projectors and decoherers. When this is done, the projectors and decoherers can be seen as unitary operations that work on a larger, unseen system. For example, a CNOT gate, which can change two separable qubits into two maximally entangled qubits, appears to be a decoherer when only one of the two qubits is measured, but when seen in the two-qubit space, it is simply a unitary rotation.

These methods for state manipulation, when used together, can transform the entangled states created in Chapter 3 into any two-qubit state. The chapter describes several classes of states which have been created experimentally, and in addition, includes a theoretical scheme for creating truly general two-qubit states.

In contrast to the study of how states change, the next section of this chapter describes how states *fail* to change. Decoherence, the unwanted interaction of a system with its environment, is one of the largest hurdles to the implementation of a quantum computer. One possible solution to this problem is the use of "decoherence-free subspaces", which allow qubits to be encoded into a larger system, a subspace of which is immune to certain types of decoherence. This section describes not only a simple theory of these subspaces, but the first experimental implementation of a decoherence-free subspace (using two-qubit polarization states).

This chapter continues with the theory of the representation of quantum processes, which share a number of similarities with quantum states. Like ensembles of quantum states which are represented by density matrices, any combination of quantum processes acting together can be represented by a single matrix χ_{ij}.

The most straightforward method to measure these quantum processes is called standard quantum process tomography (SQPT); SQPT requires that a complete set of linearly independent quantum states be subjected to the unknown process and then measured. Together, these results provide a mapping which completely defines the quantum process. Somewhat more strangely, a single entangled state can be used to measure a quantum process using

entanglement-assisted process tomography (EAPT). This method takes advantage of the correlations between entangled particles to measure a quantum process acting on only one half of this entangled state, by measuring the correlations remaining in the final state. Surprisingly, this type of process tomography can be accomplished using an *unentangled* two-qubit state, as long as both the qubit undergoing the process and an additional measured ancilla are in a highly correlated state. (This more general method of using an ancillary system to measure a process is called ancilla-assisted process tomography, or AAPT).

1.2.4 Chapter five synopsis: state measurement

This chapter begins by studying how a measurement is taken, specifically, a separable two-qubit measurement. In our experimental system, this is accomplished by measuring the polarization of each photon in a two-photon state. Experimentally, it is crucial to characterize and account for inaccuracies in detectors, polarizers, waveplates (birefringent crystals that implement state rotations), and any other optics in the system.

Sequences of these projections can be used to measure a *Bell inequality*. Bell inequalities place a bound on what is a possible result for classical states when summing several measurements. Exceeding this bound creates a paradox for classical logic, a paradox which is only resolved through entanglement. Bell measurements were historically the first type of entanglement detector, and though they are not as useful as other methods for detecting entanglement, they have the crucial advantage of functioning as a fundamental test of the phenomenon of nonlocality. Two types of Bell inequalities are experimentally implemented on a class of states for which they are inequivalent. Furthermore, separate experimental data for both the largest *and* the most statistically significant violations of Bell's inequality are included. The violations are, respectively, 2417-σ in two minutes of data collection and $S = 2.826 \pm 0.005$ in four hours of data collection—see Section 5.2.

The entanglement witness is another method that is sometimes able to identify a state as entangled while requiring few measurements. Its theory and experimental implementation is discussed.

The next few sections of this chapter provide a detailed look at quantum tomography, a method of *complete* state characterization. They begin with the theory of analytic tomography, whereby ideal measurements return the exact state being measured. Because no measurements are ideal, and statistical and systematic errors are inevitable, the next section details the use of the "maximum likelihood technique", which is used to find the state *most likely* to have given rise to the measured data. Over the course of the last five years, the improvements to tomography described here have transformed the standard of state measurement from a 45-minute error-prone tomography with 1% error into a completely automated 5-minute tomography with 0.1% accuracy.

Finally, a comparison of the advantages and disadvantages of each entanglement detection method is presented, showing the measurement settings and total number of state copies that are necessary for each technique to achieve a precision of 1%, in addition to a brief discussion of the error analysis techniques that are used for this analysis. (The difficulties of studying such large Hilbert spaces analytically have made the use of Monte Carlo simulations necessary for error analysis.)

1.3 A note about source material

This thesis is intended to provide a single guide to entangled, polarization-encoded photon pairs. As such, it details a number of our experiments from the last five years, expanded and integrated into a comprehensive guide to the creation, manipulation, and measurement of two-photon states. Because this is based on our recent research, portions of this thesis have been adapted from those published works; where the original publication was already written in an extended explanatory form, sections have largely been taken verbatim. Specifically, Chapters 2 and 5 have incorporated the tomography tutorial of a chapter from the 2006 "Advances in AMO Physics" [8], which was in turn a refinement of an earlier book chapter [5]. Chapter 3 has included the results of our recent paper on an ultra-bright source [7], and Chapter 4 includes a number of results from process tomography [3] to decoherence-free subspaces [38, 4] to state synthesis [57, 68]. Chapter 5 also includes our recent work on the quantitative comparison of entanglement detection methods [6]. Almost all of the previously published data has been expanded upon and more completely explained. The most important new data included in this thesis are recent extremely precise tests of Bell's inequalities, which are reported here for the first time (see Section 5.2).

2 Representation of Quantum States

This thesis, and this chapter in particular, is written not only as a reference but also as a tutorial for those trying to replicate these results. To that end, a great deal of the material in this chapter has appeared in other sources (mainly texts on quantum mechanics or quantum information) and is repeated here both to standardize notation for the rest of this thesis and to provide a single reference for the important concepts and theorems necessary to analyzing our experiments.

2.1 Quantum states

2.1.1 Representation of single-qubit states

Pure, mixed, and diagonal representations

In general, any single qubit in a pure state can be represented by

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.1}$$

where α and β are complex and $|\alpha|^2 + |\beta|^2 = 1$ [52]. If the normalization is written implicitly and the global phase is ignored, this can be rewritten as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|1\rangle. \tag{2.2}$$

Another common state representation is the density matrix $\hat{\rho}$, which for pure states is defined as

$$\hat{\rho} \equiv |\psi\rangle\langle\psi|. \tag{2.3}$$

The density matrix for a single qubit is a 2×2 Hermitian matrix whose diagonal elements are respectively equal to to probabilities that a state is measured as either $|0\rangle$ or $|1\rangle$. (Assuming the density matrix is written in the $|0\rangle$–$|1\rangle$ basis: see below).

Example 2.1 (Pure states) *Throughout this chapter, examples will be provided using qubits encoded into the electric field polarization of photons. (The experiments described throughout this thesis also use polarization-encoded qubits, and will use the same notation.) For a single photon, this system has two levels, e.g., horizontal ($|H\rangle \equiv |0\rangle$) and vertical ($|V\rangle \equiv |1\rangle$), with all possible pure polarization states constructed from superpositions of these two states. For example,*

diagonal, antidiagonal, right-circular and left-circular light are respectively represented by

$$\begin{aligned} |D\rangle &\equiv (|H\rangle + |V\rangle)/\sqrt{2}, \\ |A\rangle &\equiv (|H\rangle - |V\rangle)/\sqrt{2}, \\ |R\rangle &\equiv (|H\rangle + i|V\rangle)/\sqrt{2}, \\ \textit{and } |L\rangle &\equiv (|H\rangle - i|V\rangle)/\sqrt{2}. \end{aligned} \tag{2.4}$$

In practice, it is quite common to describe *ensembles* of quantum states, rather than just individual qubits. Equations 2.1 and 2.2 can be used to describe an ensemble of *identical* pure states, but are insufficient to describe either an ensemble of *different* pure states or an ensemble whose members are not pure (perhaps because they are entangled to unobserved degrees of freedom). In this case the overall state is *mixed*.

In general, these mixed states may be described by a probabilistically weighted incoherent[1] sum of pure states, i.e., they behave as if any particle in the ensemble has a specific probability of being in a given pure state, and this state is distinguishably labeled in some way. If it were not distinguishable, the total state's constituent pure states would add coherently (with a definite relative phase), yielding a single pure state.

A mixed state can be represented by a density matrix $\hat{\rho}$, where

$$\hat{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i| = \begin{matrix} \\ |0\rangle \\ |1\rangle \end{matrix} \begin{matrix} \begin{matrix} \langle 0| & \langle 1| \end{matrix} \\ \begin{pmatrix} A & Ce^{i\phi} \\ Ce^{-i\phi} & B \end{pmatrix} \end{matrix}. \tag{2.5}$$

P_i is the probabilistic weighting ($\sum_i P_i = 1$), A, B and C are all real and non-negative, $A + B = 1$, and $C \leq \sqrt{AB}$ [52].

While an ensemble of pure states can be represented by infinitely many terms in this way, it is also true that *any* ensemble of single-qubit states can also be represented by an ensemble of only *two* orthogonal pure states. (Two pure states $|\psi_i\rangle$ and $|\psi_j\rangle$ are orthogonal if $|\langle\psi_i|\psi_j\rangle| = 0$). For example, if the matrix from Equation 2.5 were diagonal, then it would clearly be a probabilistic

[1]The concepts of coherence and incoherence are crucial to understanding and using the representation of quantum states. When multiple possible values (e.g., $|\psi_1\rangle$ and $|\psi_2\rangle$) of a system are combined into a single representation, there are two ways to combine them. If they are combined *coherently*, their amplitudes sum and the total state's density matrix is proportional to $(|\psi_1\rangle + |\psi_2\rangle) \otimes (\langle\psi_1| + \langle\psi_2|)$. If they are added *incoherently*, their density matrices sum, and the total state's density matrix is proportional to $|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|$. States are added coherently if the superposed states are indistinguishable and incoherently if they are distinguishable. The outcome of a measurement on indistinguishable possibilities cannot be predicted, even in principle, from outside information. Similarly the result of a measurement of distinguishable outcomes is correlated with some additional degree of freedom.

combination of two orthogonal states, as

$$\begin{array}{c} \\ |0\rangle \\ |1\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle 0| & \langle 1| \end{array} \\ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \end{array} \equiv A|0\rangle\langle 0| + B|1\rangle\langle 1|. \tag{2.6}$$

However, *any* physical density matrix can be diagonalized, such that

$$\hat{\rho} = \begin{array}{c} \\ |\psi\rangle \\ |\psi^{\perp}\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle \psi| & \langle \psi^{\perp}| \end{array} \\ \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \end{array} = E_1|\psi\rangle\langle\psi| + E_2|\psi^{\perp}\rangle\langle\psi^{\perp}|, \tag{2.7}$$

where $\{E_1, E_2\}$ are the eigenvalues of $\hat{\rho}$, and $\{|\psi\rangle, |\psi^{\perp}\rangle\}$ are the eigenvectors (recall that these eigenvectors are always mutually orthogonal, denoted here by the $\perp$ symbol). Thus the mathematical representation of any single-qubit quantum state, no matter how it is constructed, is identical to that of an ensemble of two orthogonal pure states.[2]

Example 2.2 (A mixed state) *Now consider measuring a source of photons which emits a one-photon wave packet each second, but alternates—perhaps randomly—between horizontal, vertical, and diagonal polarizations. Their emission time labels these states (in principle) as distinguishable, and so if we ignore that timing information when they are measured, we must represent their state as a density matrix $\hat{\rho}$:*

$$\begin{aligned} \hat{\rho} &= \frac{1}{3}(|H\rangle\langle H| + |V\rangle\langle V| + |D\rangle\langle D|) \\ &= \frac{1}{3}\left(\begin{array}{c} \\ |H\rangle \\ |V\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle H| & \langle V| \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{array} + \begin{array}{c} \\ |H\rangle \\ |V\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle H| & \langle V| \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{array} + \begin{array}{c} \\ |H\rangle \\ |V\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle H| & \langle V| \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array} \right) \\ &= \frac{1}{6}\left(\begin{array}{c} \\ |H\rangle \\ |V\rangle \end{array} \begin{array}{c} \begin{array}{cc} \langle H| & \langle V| \end{array} \\ \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \end{array} \right). \end{aligned} \tag{2.8}$$

[2] It is an interesting question whether all physical states described by a mixed state—e.g., Equation 2.7—are indeed completely equivalent. For example, Lehner, Leonhardt, and Kwiat discussed the notion that two types of unpolarized light could be considered, depending on whether the incoherence between polarization components arose purely due to an averaging over rapidly varying phases, or from an entanglement with another quantum system altogether [43].

When diagonalized,

$$\hat{\rho} = \frac{1}{3}\begin{pmatrix} & \langle D| & \langle A| \\ |D\rangle & 2 & 0 \\ |A\rangle & 0 & 1 \end{pmatrix} = \frac{2}{3}|D\rangle\langle D| + \frac{1}{3}|A\rangle\langle A|, \tag{2.9}$$

which, as predicted in Equation 2.7, is a sum of only two *orthogonal states.*

Henceforth, the 'bra' and 'ket' labels will be suppressed from written density matrices where the basis is $\{|0\rangle, |1\rangle\}$ or $\{|H\rangle, |V\rangle\}$.

The Stokes parameters and the Poincaré sphere

Any single-qubit density matrix $\hat{\rho}$ can be uniquely represented by three parameters $\{S_1, S_2, S_3\}$:

$$\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i. \tag{2.10}$$

The $\hat{\sigma}_i$ matrices are

$$\begin{aligned} \hat{\sigma}_0 &\equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \hat{\sigma}_1 &\equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \hat{\sigma}_2 &\equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \hat{\sigma}_3 &\equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{2.11}$$

and the S_i values are given by

$$S_i \equiv \mathrm{Tr}\left\{\hat{\sigma}_i \hat{\rho}\right\}. \tag{2.12}$$

For all pure states, $\sum_{i=1}^{3} S_i^2 = 1$; for mixed states, $\sum_{i=1}^{3} S_i^2 < 1$; for the completely mixed state, $\sum_{i=1}^{3} S_i^2 = 0$. Due to normalization, S_0 will always equal one.

Physically, each of these parameters directly corresponds to the outcome of a specific pair of projective measurements (for more information on projective measurements, see Chapters 4 and 5):

$$\begin{aligned} S_0 &= P_{|0\rangle} + P_{|1\rangle} \\ S_1 &= P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)} \\ S_2 &= P_{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)} \\ S_3 &= P_{|0\rangle} - P_{|1\rangle}, \end{aligned} \tag{2.13}$$

where $P_{|\psi\rangle}$ is the probability to measure the state $|\psi\rangle$. As we shall see below, these relationships between probabilities and S parameters are extremely useful in understanding more general operators. Because $P_{|\psi\rangle} + P_{|\psi^\perp\rangle} = 1$, these can be simplified in the single-qubit case, and

$$P_{|\psi\rangle} - P_{|\psi^\perp\rangle} = 2P_{|\psi\rangle} - 1. \tag{2.14}$$

The probability of projecting a given state $\hat{\rho}$ into the state $|\psi\rangle$ (the probability of measuring $|\psi\rangle$) is given by [27]:

$$\begin{aligned} P_{|\psi\rangle} &= \langle\psi|\hat{\rho}|\psi\rangle \\ &= \mathrm{Tr}\left\{|\psi\rangle\langle\psi|\hat{\rho}\right\}. \end{aligned} \tag{2.15}$$

In Equation 2.13 above, the S_i are defined with respect to three states, $|\phi\rangle_i$:

$$\begin{aligned} |\phi\rangle_1 &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \\ |\phi\rangle_2 &= \frac{1}{\sqrt{2}}\left(|0\rangle + i|1\rangle\right) \\ |\phi\rangle_3 &= |0\rangle, \end{aligned} \tag{2.16}$$

and their orthogonal complements, $|\phi^\perp\rangle$. Parameters similar to these and serving the same function can be defined with respect to any three arbitrary states, $|\psi_i\rangle$, as long as the matrices $|\psi_i\rangle\langle\psi_i|$ along with the identity are linearly independent. Operators analogous to the $\hat{\sigma}$ operators can be defined relative to these states:

$$\hat{\tau}_i \equiv |\psi_i\rangle\langle\psi_i| - |\psi_i^\perp\rangle\langle\psi_i^\perp|. \tag{2.17}$$

We can further define an 'S-like' parameter T, given by:

$$T_i \equiv \mathrm{Tr}\left\{\hat{\tau}_i\hat{\rho}\right\}. \tag{2.18}$$

Continuing the previous convention and to complete the set, we define $\hat{\tau}_0 \equiv \hat{\sigma}_0$, which then requires that $T_0 = 1$. Note that the S_i parameters are simply a special case of the T_i, for the case when $\hat{\tau}_i = \hat{\sigma}_i$.

Unlike the specific case of the S parameters, which describe *mutually unbiased*[3] (MUB) measurement bases, for biased measurements

$$\hat{\rho} \neq \frac{1}{2}\sum_{i=0}^{3} T_i\hat{\tau}_i. \tag{2.19}$$

In order to reconstruct the density matrix, the T parameters must first be

[3]Two measurement bases, $\{\langle\psi_i|\}$ and $\{\langle\psi_j|\}$, are mutually unbiased if $\forall_{i,j}\,|\langle\psi_i|\psi_j\rangle|^2 = \frac{1}{d}$, where d is the dimension of the system (for a system of n qubits, $d = 2^n$). A set of measurement bases are mutually unbiased if each basis in the set is mutually unbiased with respect to every other basis in the set. In single-qubit Poincaré space, the axes indicating mutually unbiased measurement bases are at right angles [42].

transformed into the S parameters (see Equation 2.22 below).

Example 2.3 (The Stokes parameters) *For photon polarization, the S_i are the famous Stokes parameters (though normalized), and correspond to measurements in the D/A, R/L, and H/V bases [63]. In terms of the $\hat{\tau}$ matrices just introduced, we would define a set of basis states $|\psi_1\rangle \equiv |D\rangle, |\psi_2\rangle \equiv |R\rangle$, and $|\psi_3\rangle \equiv |H\rangle$. For these analysis bases, $\hat{\tau}_1 = \hat{\sigma}_1, \hat{\tau}_2 = \hat{\sigma}_2$, and $\hat{\tau}_3 = \hat{\sigma}_3$ (and therefore $T_i = S_i$ for this specific choice of analysis bases).*

As the simplest example, consider the input state $|H\rangle$. Applying Equation 2.12, we find that

$$\begin{aligned} S_0 &= Tr\{\hat{\sigma}_0\hat{\rho}_H\} = 1 \\ S_1 &= Tr\{\hat{\sigma}_1\hat{\rho}_H\} = 0 \\ S_2 &= Tr\{\hat{\sigma}_2\hat{\rho}_H\} = 0 \\ S_3 &= Tr\{\hat{\sigma}_3\hat{\rho}_H\} = 1, \end{aligned} \tag{2.20}$$

which from Equation 2.10 implies that

$$\hat{\rho}_H = \frac{1}{2}(\hat{\sigma}_0 + \hat{\sigma}_3) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.21}$$

When the Stokes parameters (S_i) are used as coordinates in 3-space, all physically possible states fall within a sphere of radius one (the Poincaré sphere for polarization, the Bloch sphere for electron spin or other two-level systems; see [16]). The pure states are found on the surface, states of linear polarization on the equator, circular states at the poles, mixed states within, and the totally mixed state – corresponding to completely unpolarized photons – at the center of the sphere. This provides a very convenient way to visualize one-qubit states (see Figure 2.1). The θ and ϕ values from Equation 2.2 allow any pure state to be easily mapped onto the sphere surface. These values are the polar coordinates of the pure state they represent on the Poincaré sphere.[4] In addition to mapping states, the sphere can be used to represent any unitary operation as a rotation about an arbitrary axis. For example, waveplates implement rotations about an axis that passes through the equator.

Any state $|\psi_0\rangle$ and its orthogonal partner, $|\psi_0^\perp\rangle$, are found on opposite points of the Poincaré sphere. The line connecting these two points forms an axis of the sphere, useful for visualizing the outcome of a measurement in the $|\psi_0\rangle$–$|\psi_0^\perp\rangle$ basis. Consider a line perpendicular to this axis which intercepts the state $\hat{\rho}$. This line will intercept the $|\psi_0\rangle$–$|\psi_0^\perp\rangle$ axis at some point P. The distance from this point P to the center of the sphere corresponds to the Stokes-like parameter ($T = \langle\psi_0|\hat{\rho}|\psi_0\rangle - \langle\psi_0^\perp|\hat{\rho}|\psi_0^\perp\rangle$).

[4] These polar coordinates are by convention rotated by 90°, so that $\theta = 0$ is on the equator corresponding to the state $|H\rangle$ and $\theta = 90°, \phi = 90°$ is at the North Pole corresponding to the state $|R\rangle$. This 90° rotation is particular to the Poincaré representation of photon polarization [59]; representations of two-level systems on the Bloch sphere do not introduce it.

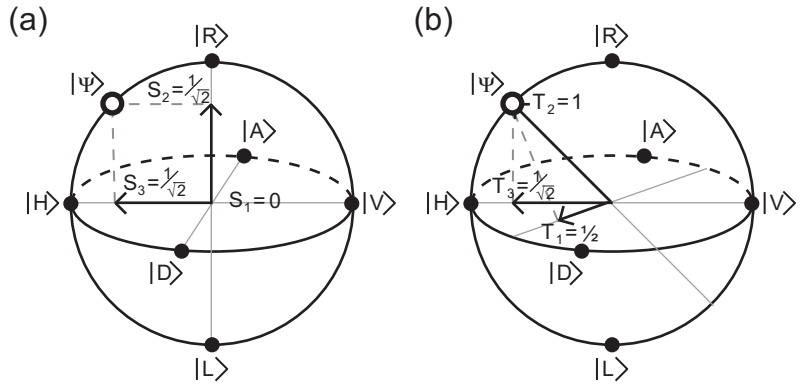

Figure 2.1: The Poincaré (or Bloch) sphere. Any single-qubit quantum state $\hat{\rho}$ can be represented by three parameters $T_i = \mathrm{Tr}\{\hat{\tau}_i\hat{\rho}\}$, as long as the operators $\hat{\tau}_i$ in addition to the identity are linearly independent. Physically, the T_i parameters directly correspond to the outcome of a specific projective measurement: $T_i = 2P_i - 1$, where P_i is the probability of success for the measurement. The T_i may be used as coordinates in 3-space. Then all 1-qubit quantum states fall on or within a sphere of radius one. The surface of the sphere corresponds to pure states, the interior to mixed states, and the origin to the totally mixed state. Shown is a particular pure state $|\psi\rangle$, which is completely specified by its projection onto a set of non-parallel basis vectors. (a) When $\hat{\tau}_i = \hat{\sigma}_i$ (the Pauli matrices), the basis vectors are orthogonal, and in this particular case the T_i are equal to the S_i, the well known Stokes parameters, corresponding to measurements of diagonal (S_1), right-circular (S_2), and horizontal (S_3) polarizations. (b) A non-orthogonal coordinate system in Poincaré space. It is possible to represent a state using its projection onto non-orthogonal axes in Poincaré space. This is of particular use when attempting to reconstruct a quantum state from mutually biased measurements. Shown here are the axes corresponding to measurements of 22.5° linear (T_1), elliptical light rotated 22.5° from H towards R (T_2), and horizontal (T_3). Taken from [5].

Thus, just as any point in three-dimensional space can be specified by its projection onto three linearly independent axes, any quantum state can be specified by the three parameters $T_i = \mathrm{Tr}\{\hat{\tau}_i\hat{\rho}\}$, where $\hat{\tau}_{i=1,2,3}$ are linearly independent matrices equal to $|\psi_i\rangle\langle\psi_i| - |\psi_i^\perp\rangle\langle\psi_i^\perp|$. The $\hat{\tau}_i$ correspond to general Stokes-like parameters for any three linearly independent axes on the Poincaré sphere. However, they can differ from the canonical Stokes axes and need not even be orthogonal. See Figure 2.1b for an example of state representation using non-orthogonal axes.

In order to use these mutually biased Stokes-like parameters, it is necessary to be able to transform a state from the mutually biased representation to the Stokes representation and vice-versa. In general, for any two representations $S_i = \mathrm{Tr}\{\hat{\sigma}_i\hat{\rho}\}$ and $T_i = \mathrm{Tr}\{\hat{\tau}_i\hat{\rho}\}$ it is possible to transform between them by using

$$\begin{pmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_0\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_1\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_2\hat{\sigma}_3\} \\ \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_0\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_1\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_2\} & \mathrm{Tr}\{\hat{\tau}_3\hat{\sigma}_3\} \end{pmatrix}\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}. \tag{2.22}$$

This relation allows S parameters to be transformed into any set of T parameters. In order to transform from T to S, we can invert the 4 by 4 matrix in Equation 2.22 and multiply both sides by this new matrix. This inversion is possible because we have chosen the $\hat{\tau}_i$ operators to be linearly independent; otherwise the T_i parameters would not specify a single point in Hilbert space.

2.1.2 Multiple qubits

With the extension of these concepts to multiple qubit systems, it becomes possible to investigate non-classical features, including the quintessentially quantum mechanical phenomenon of entanglement.

Pure, mixed, and diagonal representations

As the name implies, multiple-qubit states are constructed out of individual qubits. As such, the Hilbert space of a many-qubit system is spanned by state vectors which are the tensor product of single-qubit state vectors. A general n-qubit system in a pure state can be written as

$$|\psi\rangle = \sum_{i_1,i_2,\ldots i_n=0,1} \alpha_{i_1,i_2,\ldots i_n}|i_1\rangle \otimes |i_2\rangle \otimes \ldots \otimes |i_n\rangle. \tag{2.23}$$

Here the α_i are complex, $\sum_i |\alpha_i|^2 = 1$, and $\otimes$ denotes a tensor product, used to join component Hilbert spaces. For example, a general two-qubit pure state can be written

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle, \tag{2.24}$$

where $|00\rangle$ is shorthand for $|0\rangle_1 \otimes |0\rangle_2$.

Example 2.4 (The Bell states) *Perhaps the most famous examples of pure two-qubit states are the Bell states [12]:*

$$\begin{aligned} |\phi^{\pm}\rangle &= \frac{1}{\sqrt{2}}\left(|HH\rangle \pm |VV\rangle\right) \\ |\psi^{\pm}\rangle &= \frac{1}{\sqrt{2}}\left(|HV\rangle \pm |VH\rangle\right). \end{aligned} \tag{2.25}$$

As before, we represent a general mixed state through an incoherent sum of pure states:

$$\hat{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i|. \tag{2.26}$$

And, as before, this 2^n-by-2^n density matrix representing the n-qubit state may always be diagonalized, allowing any state to be written as

$$\hat{\rho} = \sum_{i=1}^{2^n} P_i |\phi_i\rangle\langle\phi_i|. \tag{2.27}$$

2.26 differs from 2.27 in that the ϕ_i are necessarily orthogonal ($\langle\phi_i|\phi_j\rangle = \delta_{ij}$), and there are at most 2^n of them; in 2.26 there could be an arbitrary number of $|\psi_i\rangle$.

Example 2.5 (A general two-qubit polarization state) *Any two-qubit polarization state can be written as*

$$\hat{\rho} = \begin{array}{c} \\ |HH\rangle \\ |HV\rangle \\ |VH\rangle \\ |VV\rangle \end{array} \begin{array}{c} \begin{array}{cccc} \langle HH| & \langle HV| & \langle VH| & \langle VV| \end{array} \\ \begin{pmatrix} A_1 & B_1 e^{i\phi_1} & B_2 e^{i\phi_2} & B_3 e^{i\phi_3} \\ B_1 e^{-i\phi_1} & A_2 & B_4 e^{i\phi_4} & B_5 e^{i\phi_5} \\ B_2 e^{-i\phi_2} & B_4 e^{-i\phi_4} & A_3 & B_6 e^{i\phi_6} \\ B_3 e^{-i\phi_3} & B_5 e^{-i\phi_5} & B_6 e^{-i\phi_6} & A_4 \end{pmatrix} \end{array}, \tag{2.28}$$

where $\hat{\rho}$ is positive and Hermitian with unit trace. Henceforth, the 'bra' and 'ket' labels will be omitted from density matrices presented in this standard basis. Note that this state has 15 free parameters, while a general n-qubit state has $4^n - 1$ free parameters.

Mixed states of note include the Werner states [70],

$$\hat{\rho}_W = P|\gamma\rangle\langle\gamma| + (1-P)\frac{1}{4}I, \tag{2.29}$$

where $|\gamma\rangle$ is a maximally entangled state and $\frac{1}{4}I$ is the totally mixed state, and the maximally entangled mixed states (MEMS), which possess the maximal amount of entanglement for a given amount of mixture [51].

Measures of state fidelity, entanglement, and mixture may be derived from the density matrix; for reference, Appendix A defines several such measures.

Multiple-qubit Stokes parameters

Extending the single-qubit density matrix representation of Equation 2.10, any n-qubit state $\hat{\rho}$ may be represented as

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1,i_2,\ldots i_n=0}^{3} S_{i_1,i_2,\ldots i_n} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \ldots \otimes \hat{\sigma}_{i_n}. \tag{2.30}$$

Normalization requires that $S_{0,0,\ldots 0} = 1$, leaving $4^n - 1$ real parameters (the multiple-qubit analog of the single-qubit Stokes parameters) to identify any point in Hilbert space, just as three parameters determined the exact position of a one-qubit state in the Bloch/Poincaré sphere. Already for two qubits, the state space is much larger, requiring 15 independent real parameters to describe it. For this reason, there is no convenient graphical picture of this space, as there was in the single-qubit case (see, however, the interesting approaches made by [76, 77]).

For multiple qubits the link between the multiple-qubit Stokes parameters [34, 1] and measurement probabilities still exists. The formalism of $\hat{\tau}$ operators also still holds for larger qubit systems, so that

$$T = \mathrm{Tr}\left\{\hat{\tau}\hat{\rho}\right\}. \tag{2.31}$$

For 'local' measurements (a local measurement is the tensor product of a number of single-qubit measurements: the first projecting qubit one along $\hat{\tau}_{i_1}$, the second qubit two along $\hat{\tau}_{i_2}$, etc.), $\hat{\tau} = \hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \ldots \otimes \hat{\tau}_{i_n}$. Combining Equations 2.30 and 2.31,

$$\begin{aligned} T_{i_1,i_2,\ldots i_n} &= \mathrm{Tr}\left\{\left(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \ldots \otimes \hat{\tau}_{i_n}\right)\hat{\rho}\right\} \\ &= \frac{1}{2^n} \sum_{j_1,j_2,\ldots j_n=0}^{3} \mathrm{Tr}\left\{\hat{\tau}_{i_1}\hat{\sigma}_{j_1}\right\} \mathrm{Tr}\left\{\hat{\tau}_{i_2}\hat{\sigma}_{j_2}\right\} \ldots \mathrm{Tr}\left\{\hat{\tau}_{i_n}\hat{\sigma}_{j_n}\right\} S_{j_1,j_2,\ldots j_n}. \end{aligned} \tag{2.32}$$

Recall that for single qubits,

$$\begin{aligned} T_{i=1,2,3} &= P_{|\psi_i\rangle} - P_{|\psi_i^{\perp}\rangle} \\ T_0 &= P_{|\psi\rangle} + P_{|\psi^{\perp}\rangle} = 1, \forall\psi. \end{aligned} \tag{2.33}$$

Therefore, for an n-qubit system,

$$T_{i_1,i_2,\ldots i_n} = \left(P_{|\psi_{i_1}\rangle} \pm P_{|\psi_{i_1}^{\perp}\rangle}\right) \otimes \left(P_{|\psi_{i_2}\rangle} \pm P_{|\psi_{i_2}^{\perp}\rangle}\right) \otimes \ldots \otimes \left(P_{|\psi_{i_n}\rangle} \pm P_{|\psi_{i_n}^{\perp}\rangle}\right), \tag{2.34}$$

where the plus sign is used for a zero index and the minus sign is used for a nonzero index. For a two-qubit system where $i_1 \neq 0$ and $i_2 \neq 0$, T_{i_1,i_2} simplifies

dramatically, giving

$$\begin{aligned} T_{i_1,i_2} &= (P_{|\psi_{i_1}\rangle} - P_{|\psi_{i_1}^{\perp}\rangle}) \otimes (P_{|\psi_{i_2}\rangle} - P_{|\psi_{i_2}^{\perp}\rangle}) \\ &= P_{|\psi_{i_1}\rangle|\psi_{i_2}\rangle} - P_{|\psi_{i_1}\rangle|\psi_{i_2}^{\perp}\rangle} - P_{|\psi_{i_1}^{\perp}\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}^{\perp}\rangle|\psi_{i_2}^{\perp}\rangle}. \end{aligned} \tag{2.35}$$

This relation will be crucial for reconstructing a two-qubit state from local measurements (see Chapter 5).

As before, we are not restricted to multiple-qubit Stokes parameters based only on mutually unbiased operators. Extending Equation 2.22 to multiple qubits, and again assuming two representations

$$S_{i_1,i_2,\ldots i_n} = \mathrm{Tr}\left\{\left(\hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \ldots \otimes \hat{\sigma}_{i_n}\right)\hat{\rho}\right\}, \tag{2.36}$$

and

$$T_{i_1,i_2,\ldots i_n} = \mathrm{Tr}\left\{\left(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \ldots \otimes \hat{\tau}_{i_n}\right)\hat{\rho}\right\}, \tag{2.37}$$

it therefore follows that

$$T_{i_1,i_2,\ldots i_n} = \tag{2.38}$$

$$\frac{1}{2^n} \sum_{j_1,j_2,\ldots j_n=0}^{3} \mathrm{Tr}\left\{\left(\hat{\tau}_{i_1} \otimes \hat{\tau}_{i_2} \otimes \ldots \otimes \hat{\tau}_{i_n}\right)\left(\hat{\sigma}_{j_1} \otimes \hat{\sigma}_{j_2} \otimes \ldots \otimes \hat{\sigma}_{j_n}\right)\right\} S_{j_1,j_2,\ldots j_n}.$$

In general, a given $\hat{\tau}$ operator is not uniquely mapped to a single pair of analysis states. For example, consider measurements of $|H\rangle$ and $|V\rangle$ corresponding to $\hat{\tau}_H = |H\rangle\langle H| - |V\rangle\langle V| = \hat{\sigma}_3$ and $\hat{\tau}_V = |V\rangle\langle V| - |H\rangle\langle H| = -\hat{\sigma}_3$. Therefore, $\hat{\tau}_{H,H} \equiv \hat{\sigma}_3 \otimes \hat{\sigma}_3 = -\hat{\sigma}_3 \otimes -\hat{\sigma}_3 \equiv \hat{\tau}_{V,V}$. This is a natural artifact of the mathematics (and does not, for example, affect the results of a tomography).

Example 2.6 (A separable polarization state) *Consider the state $|HH\rangle$. Following the example in Equation 2.21,*

$$\begin{aligned} \hat{\rho}_{HH} &= |HH\rangle\langle HH| \\ &= \frac{1}{2}(\hat{\sigma}_0 + \hat{\sigma}_3) \otimes \frac{1}{2}(\hat{\sigma}_0 + \hat{\sigma}_3) \\ &= \frac{1}{4}(\hat{\sigma}_0 \otimes \hat{\sigma}_0 + \hat{\sigma}_0 \otimes \hat{\sigma}_3 + \hat{\sigma}_3 \otimes \hat{\sigma}_0 + \hat{\sigma}_3 \otimes \hat{\sigma}_3). \end{aligned} \tag{2.39}$$

This implies that for this state there are exactly four non-zero two-qubit Stokes parameters: $S_{0,0}, S_{0,3}, S_{3,0}$, and $S_{3,3}$ – all of which are equal to one. (As earlier, for the special case when $\hat{\tau}_{i,j} = \hat{\sigma}_{i,j}$, we relabel the $T_{i,j}$ as $S_{i,j}$, the two-qubit Stokes parameters [34, 1].) The separable nature of this state makes it easy to calculate the two-qubit Stokes decomposition.

Example 2.7 (The singlet state) *If instead we investigate the entangled state $|\psi^-\rangle \equiv (|HV\rangle - |VH\rangle)/\sqrt{2}$, it will be necessary to calculate each two-qubit Stokes parameter from the $\hat{\sigma}$ matrices. As an example, consider $\hat{\sigma}_{3,3} \equiv \hat{\sigma}_3 \otimes \hat{\sigma}_3$, for which*

$$S_{3,3} = Tr\left\{\hat{\sigma}_{3,3}|\psi^-\rangle\langle\psi^-|\right\} = -1. \tag{2.40}$$

We could instead calculate $S_{3,3}$ directly from probability outcomes of measurements on $|\psi^-\rangle$:

$$\begin{aligned} S_{3,3} &= (P_H - P_V) \otimes (P_H - P_V) \\ &= P_{HH} - P_{HV} - P_{VH} + P_{VV} \\ &= 0 - \frac{1}{2} - \frac{1}{2} + 0 \\ &= -1. \end{aligned} \tag{2.41}$$

Continuing on, we measure $S_{0,3}$:

$$\begin{aligned} S_{0,3} &= (P_H + P_V) \otimes (P_H - P_V) \\ &= P_{HH} - P_{HV} + P_{VH} - P_{VV} \\ &= 0 - \frac{1}{2} + \frac{1}{2} - 0 \\ &= 0. \end{aligned} \tag{2.42}$$

Here the signs of the probabilities changed due to the zero index in $S_{0,3}$. These results would have been the same even if the analysis bases of the first qubit had been shifted to any other orthogonal basis, i.e., $S_{0,3} = \left(P_\psi + P_{\psi^\perp}\right) \otimes (P_H - P_V)$.

If the method above is continued for all the Stokes parameters, one concludes that

$$\begin{aligned} \hat{\rho}_{\psi^-} &= \frac{1}{2}(|HV\rangle - |VH\rangle)(\langle HV| - \langle VH|) \\ &= \frac{1}{4}(\hat{\sigma}_0 \otimes \hat{\sigma}_0 - \hat{\sigma}_1 \otimes \hat{\sigma}_1 - \hat{\sigma}_2 \otimes \hat{\sigma}_2 - \hat{\sigma}_3 \otimes \hat{\sigma}_3). \end{aligned} \tag{2.43}$$

2.1.3 Representation of non-qubit systems

Although most interest within the field of quantum information and computation has focused on two-level systems (qubits) due to their simplicity, availability, and similarity to classical bits, nature contains a multitude of many-level systems, both discrete and continuous. A discussion of continuous systems is beyond the scope of this work—see [44]—but we will briefly address here the representation and tomography of discrete, d-level systems ("qudits"). For a more detailed description of qudit tomography, see [65].

Pure, mixed, and diagonal representations

Directly extending Equations 2.1 and 2.2, a d-level qudit can be represented as

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \ldots + \alpha_{d-1}|d-1\rangle, \tag{2.44}$$

where $\sum_i |\alpha_i|^2 = 1$. Mixed qudit states can likewise be represented by generalizing Equations 2.5 and 2.7:

$$\rho = \sum_k P_k |\phi_k\rangle\langle\phi_k| \tag{2.45}$$

$$= \sum_{i=0}^{d-1} P_i |\psi_i\rangle\langle\psi_i|. \tag{2.46}$$

Here $\{|\phi_k\rangle\}$ is completely unrestricted while $|\langle\psi_i|\psi_j\rangle| = \delta_{ij}$. In other words, while any mixed state is an incoherent superposition of an undetermined number of pure states, any mixed state can be represented by an incoherent superposition of only n orthogonal states (the diagonal representation).

Example 2.8 (Orbital angular momentum modes)
The Laguerre-Gaussian modes of an optical field propagating in the z direction possess z components of orbital angular momentum that serve as the quantum numbers of a multiple-level photonic system that has recently been studied for quantum information [47, 41, 9, 11]. Consider a qudit system with an infinite number of levels representing the quantization of orbital angular momentum. A superposition of the three lowest angular momentum levels would look like

$$|\psi\rangle = |+1\rangle + |0\rangle + |-1\rangle, \tag{2.47}$$

where $|+1\rangle$ ($|-1\rangle$) corresponds to a mode where each photon has $+\hbar$ ($-\hbar$) z-component of orbital angular momentum, and $|0\rangle$ corresponds to a zero angular momentum mode, e.g. a mode having a Gaussian transverse profile. Using specially designed holograms, these states can be measured and interconverted [2].

Qudit Stokes parameters

In order to completely generalize the qubit mathematics laid out previously to the qudit case, it is necessary to find Stokes-like parameters which satisfy the following conditions:

$$\hat{\rho} = \sum_{i=0}^{n} S_i \hat{\sigma}_i, \tag{2.48}$$

$$S_i \equiv \mathrm{Tr}\left\{\hat{\sigma}_i \hat{\rho}\right\}. \tag{2.49}$$

In addition, in order to reconstruct these parameters experimentally, it will be necessary to find S_i as a function of measurable probabilities:

$$S_i = \mathcal{F}\left(\left\{P_{|\psi\rangle}\right\}\right). \tag{2.50}$$

Obviously, it would be ideal to find a simple form similar to the qubit $\hat{\sigma}$ matrices. Conveniently, the general qudit sigma matrices and corresponding S_i parameters

can be divided into three groups ($\left\{\hat{\sigma}_i^X, \hat{\sigma}_i^Y, \hat{\sigma}_i^Z\right\}$ and $\left\{S_i^X, S_i^Y, S_i^Z\right\}$), according to their similarity to $\hat{\sigma}_x = \hat{\sigma}_1$, $\hat{\sigma}_y = \hat{\sigma}_2$, and $\hat{\sigma}_z = \hat{\sigma}_3$, respectively [65]. Using these divisions, we can expand Equation 2.48:

$$\hat{\rho} = S_0\hat{\sigma}_0 + \sum_{\substack{j,k\in\{0,1,\ldots n-1\}\\ j\neq k}} \left(S_{j,k}^X\hat{\sigma}_{j,k}^X + S_{j,k}^Y\hat{\sigma}_{j,k}^Y\right) + \sum_{r=1}^{n-1} S_r^Z\hat{\sigma}_r^Z. \tag{2.51}$$

Investigating the simplest group first, it is unsurprising that

$$\hat{\sigma}_0 = I, S_0 = 1, \tag{2.52}$$

continuing the previous qubit convention. The X- and Y-related variables are defined almost identically to their predecessors:

$$\begin{aligned}
\hat{\sigma}_{j,k}^X &= |j\rangle\langle k| + |k\rangle\langle j|, && (2.53)\\
S_{j,k}^X &= P_{\frac{1}{\sqrt{2}}(|j\rangle+|k\rangle)} - P_{\frac{1}{\sqrt{2}}(|j\rangle-|k\rangle)}, && (2.54)\\
\hat{\sigma}_{j,k}^Y &= -i\left(|j\rangle\langle k| - |k\rangle\langle j|\right), && (2.55)\\
S_{j,k}^Y &= P_{\frac{1}{\sqrt{2}}(|j\rangle+i|k\rangle)} - P_{\frac{1}{\sqrt{2}}(|j\rangle-i|k\rangle)}. && (2.56)
\end{aligned}$$

The definitions for $\hat{\sigma}_i^Z$ and S_i^Z are slightly more complicated:

$$\begin{aligned}
\hat{\sigma}_r^Z &= \sqrt{\frac{2}{r\left(r+1\right)}}\left[\left(\sum_{j=0}^{r-1}|j\rangle\langle j|\right) - r|r\rangle\langle r|\right], && (2.57)\\
S_r^Z &= \sqrt{\frac{2}{r\left(r+1\right)}}\left[\left(\sum_{j=0}^{r-1}P_{|j\rangle}\right) - rP_{|r\rangle}\right]. && (2.58)
\end{aligned}$$

These definitions complete the set of n^2 sigma matrices, and have a slightly more complex form in order to satisfy $\mathrm{Tr}\left[\hat{\sigma}_i\right] = 0$ and $\mathrm{Tr}\left[\hat{\sigma}_i\hat{\sigma}_j\right] = 2\delta_{ij}$ (these definitions apply to all $\hat{\sigma}_i$ except $\hat{\sigma}_0$).

Example 2.9 (The qutrit) *For a 3-level system ($|0\rangle, |1\rangle$, and $|2\rangle$), the $\hat{\sigma}$ ma-*

trices can be defined as:

$$\hat{\sigma}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \hat{\sigma}_1^Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \hat{\sigma}_2^Z = \sqrt{\tfrac{1}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\hat{\sigma}_{1,2}^X = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \hat{\sigma}_{1,3}^X = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \hat{\sigma}_{2,3}^X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad .$$

$$\hat{\sigma}_{1,2}^Y = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \hat{\sigma}_{1,3}^Y = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \hat{\sigma}_{2,3}^Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

Expanding the S_i parameters in terms of probabilities, we find that:

$$\begin{aligned}
S_0 &= 1, \\
S_{1,2}^X &= P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)}, \\
S_{1,3}^X &= P_{\frac{1}{\sqrt{2}}(|0\rangle+|2\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-|2\rangle)}, \\
S_{2,3}^X &= P_{\frac{1}{\sqrt{2}}(|1\rangle+|2\rangle)} - P_{\frac{1}{\sqrt{2}}(|1\rangle-|2\rangle)}, \\
S_{1,2}^Y &= P_{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)}, \\
S_{1,3}^Y &= P_{\frac{1}{\sqrt{2}}(|0\rangle+i|2\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-i|2\rangle)}, \\
S_{2,3}^Y &= P_{\frac{1}{\sqrt{2}}(|1\rangle+i|2\rangle)} - P_{\frac{1}{\sqrt{2}}(|1\rangle-i|2\rangle)}, \\
S_1^Z &= P_{|0\rangle} - P_{|1\rangle}, \\
S_2^Z &= \frac{1}{\sqrt{3}} \left(P_{|0\rangle} + P_{|1\rangle} - 2P_{|2\rangle} \right).
\end{aligned}$$

2.2 Representing entanglement and decoherence

The phenomenon of entanglement deserves special attention during any discussion of quantum state representation, particularly because there appears to be no consensus among the community as to how to quantify it (most notably for systems larger than two qubits). By definition, systems A and B are not entangled if and only if they are not separable, i.e., not able to be written in the form

$$\rho_{\text{sep}} = \sum_i |\psi_{A,i}\rangle|\psi_{B,i}\rangle \otimes \langle\psi_{A,i}|\langle\psi_{B,i}|. \tag{2.59}$$

More intuitively, separable states are those which can be constructed when their constituent particles are at a distance, using only local operations. Entangled states, by contrast, only arise from joint operations. These entangling operations can only be performed when the two systems to be coupled (or parts of

the two systems) are in the same physical location[5][6]. Were this not the case, the quantum correlations produced by entangling operations could be used to communicate faster than light.

Entangled states are not the only states to possess correlations. Many classical states possess correlations, ranging from perfect correlations in one basis to imperfect correlations in many bases, but maximally entangled states are the only states which exhibit perfect correlation[7] in every basis (see two-qubit examples below). Put another way, entangled states exhibit correlated quantum fluctuations, a property which can be written mathematically as

$$\langle (A - \langle A \rangle)(B - \langle B \rangle) \rangle \geq 0, \tag{2.60}$$

where A and B are observables on particles 1 and 2, respectively.

2.2.1 Two-qubit entanglement

Because this thesis focuses on the special case of two-qubit systems, this subsection will describe entanglement in those systems. While all entangled two-qubit systems can be identified using the *tangle* (see Appendix A), there are several representations which are useful when discussing entangled states. First is the Schmidt decomposition [52], through which two-qubit pure states can be written in the following form:

$$|\psi\rangle = \cos(\theta)|\alpha\rangle|\beta\rangle + \sin(\theta)e^{i\phi}|\alpha^{\perp}\rangle|\beta^{\perp}\rangle. \tag{2.61}$$

Here $|\alpha\rangle$ and $|\beta\rangle$ represent single qubit states while $|\alpha^{\perp}\rangle$ and $|\beta^{\perp}\rangle$ represent their orthogonal complements. Pure, two-qubit entangled states, when written in this form, have exactly two terms; separable states have only one term.

This is a very convenient parametrization, but only applies to pure states. For mixed states, an entanglement rule can be constructed based on the state fidelity. The two-qubit entangled states $|\psi_{\text{ent}}\rangle$ are exactly those for which

$$\exists |\gamma\rangle \text{ such that } |\langle \psi_{\text{ent}}|\gamma\rangle|^2 > \frac{1}{2}, \tag{2.62}$$

where $|\gamma\rangle$ is a maximally entangled state. Close examination of this rule reveals that not all states which exhibit multi-basis correlation (the hallmark of entanglement) are in fact entangled (see Example 2.11). As we will see in Chapter 4,

[5]The phrase "physical location" is used somewhat loosely in this context to refer to objects which are near enough to each other that any interaction between them does not violate special relativity's prohibition on information traveling faster than the speed of light. In a laboratory setting, this meaning becomes even clearer, e.g., two photons will never become entangled with each other unless they are present in the same optical element at the same time.

[6]The exception to this rule is entanglement swapping [56]. The entanglement between two particles, A and B, can be transferred to another pair of particles, C and D, if A and C are near each other *and* B and D are near each other.

[7]Correlation here is used to mean either correlation or anticorrelation. More generally, a measurement of one particle of a maximally entangled pair will completely determine the state of its partner.

it is precisely this type of unentangled correlation in multiple bases which will facilitate ancilla-assisted process tomography (AAPT). In Chapter 5, this rule will be used to construct a simple entanglement witness, a method for detecting entangled states using few measurements.

Example 2.10 (A maximally entangled state) *Maximally entangled states possess a unique quality: more than one separable basis can be chosen such that measurements in that basis result in completely random, yet perfectly correlated results. In effect, as soon as a measurement is made on one qubit, its partner qubit is immediately known to be in a specific state corresponding to the result of the first measurement. For example, the state* $|\phi^+\rangle \equiv \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$ *exhibits perfect correlation in any linear basis and perfect anticorrelation in the circular basis. That is, if the first qubit of this state is measured to be in the state* $|H\rangle$*, the second qubit will also be measured in* $|H\rangle$*, and if the first qubit is measured in the state* $|R\rangle$*, then the second qubit will be measured in the state* $|L\rangle$*. In contrast, the singlet state* $|\psi^-\rangle \equiv \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ *exhibits anticorrelation in* every *basis. If the first qubit is measured to be in state* $|\alpha\rangle$*, then the second qubit will be measured to be in state* $|\alpha^\perp\rangle$*.*

Example 2.11 (A Werner state) *The Werner state*

$$|\rho\rangle = \frac{1}{3}|\phi^-\rangle\langle\phi^-| + \frac{2}{3}I, \tag{2.63}$$

has a fidelity with the maximally entangled state $|\phi^-\rangle$ *of exactly 0.5, which is just insufficient for entanglement. However, this state exhibits correlation in multiple bases, just as* $|\phi^+\rangle$ *does. Because it is unentangled, however, we should be able to represent it using only separable terms:*

$$\begin{aligned}\rho \quad = \quad & \frac{1}{6}(|HH\rangle\langle HH| + |VV\rangle\langle VV| + |DD\rangle\langle DD| + \\ & |AA\rangle\langle AA| + |RL\rangle\langle RL| + |LR\rangle\langle LR|)\end{aligned} \tag{2.64}$$

The state ρ *is written here as the incoherent sum of six terms, each of which provides correlation to the total state. This state exhibits the maximum two-qubit correlation that is possible while still being separable.*

2.2.2 Entanglement in higher degrees of freedom

The parametrization and quantification of the entanglement of more than two systems has generated great interest, but remains an open question. Whether this difficulty arises from the inherent complexity of the problem or because "entanglement" is not a well-defined quantity for large systems remains to be seen. Regardless, there have been proposed solutions to this problem (see, for example, [67]), but their discussion is beyond the scope of this work.

2.2.3 Subsystems and the partial trace

Often, it is important to predict the dynamics of a subsystem which is part of a larger space. Mathematically, this subsystem is described via the *partial trace*. Consider a state $\hat{\rho}^{AB}$ composed of two subsystems, A and B, each of which can be respectively described using orthonormal bases $|a_i\rangle$ and $|b_i\rangle$. Making the partial trace of $\hat{\rho}^{AB}$ over subsystem B will give a density matrix $\hat{\rho}^{A}$ which describes the subsystem A when subsystem B is ignored. Written mathematically,

$$\begin{aligned} \hat{\rho}^{A} &\equiv \mathrm{Tr_B}\left\{\hat{\rho}^{\mathrm{AB}}\right\}, & (2.65)\\ \mathrm{Tr_B}\left\{|\mathrm{a_i}\rangle\langle\mathrm{a_j}|\otimes|\mathrm{b_k}\rangle\langle\mathrm{b_l}|\right\} &\equiv |a_i\rangle\langle a_j|\mathrm{Tr}\left\{|\mathrm{b_k}\rangle\langle\mathrm{b_l}|\right\}. & (2.66)\end{aligned}$$

Written another way,

$$\hat{\rho}^{A} \equiv \mathrm{Tr_B}\left\{\hat{\rho}^{\mathrm{AB}}\right\} \equiv \sum_{\mathrm{i}} \langle \mathrm{b_i}|\hat{\rho}^{\mathrm{AB}}|\mathrm{b_i}\rangle. \quad (2.67)$$

This formalism allows one to not only predict how smaller pieces of known large systems will behave, but also to conjecture on the nature of the larger systems of which smaller observed pieces are just a small part.

Example 2.12 (The partial trace of a separable state) *The simplest example of the partial trace is for a separable state. Separable states, by definition, should be easily divided into their component subsystems. Mathematically, this means that*

$$\mathrm{Tr_B}\left\{\hat{\rho}^{\mathrm{AB}}\right\} = \mathrm{Tr_B}\left\{\hat{\rho}^{\mathrm{A}}\otimes\hat{\rho}^{\mathrm{B}}\right\} = \hat{\rho}^{\mathrm{A}}\mathrm{Tr}\left\{\hat{\rho}^{\mathrm{B}}\right\} = \hat{\rho}^{\mathrm{A}}. \quad (2.68)$$

Example 2.13 (The partial trace of a Bell state) *A more interesting example is the maximally entangled Bell state. Here we will have to calculate the result more carefully, remembering that $|HH\rangle$ is shorthand for $|H^A\rangle\otimes|H^B\rangle$:*

$$\begin{aligned} & \mathrm{Tr_B}\left\{\hat{\rho}_{\phi^+}\right\} = \mathrm{Tr_B}\left\{\frac{1}{2}\left(|\mathrm{HH}\rangle+|\mathrm{VV}\rangle\right)\left(\langle\mathrm{HH}|+\langle\mathrm{VV}|\right)\right\}\\
= \;& \mathrm{Tr_B}\left\{\frac{1}{2}\left(|\mathrm{HH}\rangle\langle\mathrm{HH}|+|\mathrm{HH}\rangle\langle\mathrm{VV}|+|\mathrm{VV}\rangle\langle\mathrm{HH}|+|\mathrm{VV}\rangle\langle\mathrm{VV}|\right)\right\}\\
= \;& \frac{1}{2}\sum_i \langle b_i|\Big(|HH\rangle\langle HH|+|HH\rangle\langle VV|+|VV\rangle\langle HH|+|VV\rangle\langle VV|\Big)|b_i\rangle\\
= \;& \frac{1}{2}\Big[\langle H^B|\Big(|HH\rangle\langle HH|+|HH\rangle\langle VV|+|VV\rangle\langle HH|+|VV\rangle\langle VV|\Big)|H^B\rangle +\\
& \langle V^B|\Big(|HH\rangle\langle HH|+|HH\rangle\langle VV|+|VV\rangle\langle HH|+|VV\rangle\langle VV|\Big)|V^B\rangle\Big]\\
= \;& \frac{1}{2}\left[|H^A\rangle\langle H^A|+|V^A\rangle\langle V^A|\right]. \end{aligned} \quad (2.69)$$

The result is the single-qubit totally mixed state. This should have been no surprise, in fact, as we have already discussed how one particle in a maximally entangled state will always give totally random measurement results in every

basis, which is only true of the totally mixed state. In fact, for a two-qubit pure state, there is a direct, monotonic relationship between the entanglement of the state and the mixture of the single-qubit subsystems which constitute it.

2.2.4 Decoherence

In Example 2.13, we saw that when a component of a maximally entangled state is investigated without attention to its partner, it appears to be maximally mixed, in other words, totally decohered. In fact, any observed decoherence in a density matrix can be mathematically modeled as the result of that system being entangled to other degrees of freedom, i.e., other systems in a larger, unobserved Hilbert space. We can in fact easily construct such a model by using the state diagonalization developed earlier in this chapter. From Equation 2.27, we know that

$$\hat{\rho} = \sum_{i=1}^{2^n} P_i |\phi_i\rangle\langle\phi_i|. \tag{2.70}$$

If the state above has more than one term, then it is at least partially decohered. Under the assumption that this decoherence is the result of entanglement in a larger Hilbert space, we can attempt to write down the larger entangled state. The larger pure state

$$|\psi\rangle = \sum_{i=1}^{2^n} e^{i\zeta_i} \sqrt{P_i} |\phi_i\rangle \otimes |X_i\rangle, \tag{2.71}$$

when subjected to a partial trace over the ancillary space X, exactly reproduces the state $\hat{\rho}$ above. There is of course some ambiguity to this process; for example, the exact phases ζ_i are totally unimportant and unspecified. Likewise, there are an infinite number of states in larger spaces which could be simplified through partial traces to the measured state $\hat{\rho}$. This is simply an example that easily applies to any decohered state, requiring an ancillary space equal in dimension to the space in which the original state $\hat{\rho}$ exists.

3 State Creation

The creation and control of quantum states is at the heart of any study of the rich, strange physics of quantum information. Light, or more precisely, photons, has been central to the study of quantum mechanics since its earliest stages, and provides a clean, controllable, easily manipulated medium for investigations of quantum information. While the behavior of single photons has been studied for centuries (classical optics is the study of single photon dynamics, albeit for many photons all undergoing the same processes), the true beauty of quantum mechanics begins with correlation and nonlocality, and for that, two-photon systems are required. The experiments described throughout this thesis utilize sources of correlated photon pairs collected from spontaneous parametric down-conversion in birefringent nonlinear crystals. Moreover, this down-conversion process can be made to yield *entangled* photon pairs, making possible the investigation of this quintessential characteristic of quantum mechanics.

This chapter will describe the physics of down-conversion, the design of a two-crystal entanglement source and the evolution of that design into a source of unequaled quality and brightness [7].

3.1 Spontaneous parametric down-conversion

Though impossible under classical electromagnetism, quantum theory predicts that birefringent crystals with a nonlinear electric susceptibility have the potential to convert high energy pump photons to pairs of low energy daughter photons, through a process known as *spontaneous parametric down-conversion*. The complete derivation of this process requires the quantization of the electric field, and will not be rederived here (although there are many references on the subject, see for example, [18]).

In brief, spontaneous parametric down-conversion is one of a broad class of nonlinear optical effects that are possible in a crystal whose dielectric polarization (dipole moment per unit volume) responds nonlinearly to an electric field. The response can be expanded as a power series of the electric field:

$$P(t) \equiv \chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t) + \dots . \tag{3.1}$$

Because down-conversion involves the transformation of a single photon into two photons, only materials with a large $\chi^{(2)}$ nonlinearity will exhibit this behavior.

This is because, in a $\chi^{(2)}$ material, two oscillating electric fields of the same frequency can produce an oscillating field of twice that frequency. In a linear material, the combination of two fields would be akin to adding two sinusoidal waves of frequency ω; the result is still a wave of frequency ω. In a $\chi^{(2)}$ material, however, the *product* of two waves of frequency ω is a wave of frequency 2ω (consider the trigonometric identity $2\cos\omega\sin\omega = \sin 2\omega$).

The phenomenology of down-conversion can be understood by observing that it must conserve both energy and momentum:

$$\omega_1 + \omega_2 = \omega_3 \quad \text{(energy conservation)}, \tag{3.2}$$

$$\vec{\kappa}_1 + \vec{\kappa}_2 = \vec{\kappa}_3 \quad \text{(momentum conservation inside the crystal)}. \tag{3.3}$$

In general, there are two ways to satisfy both of these conditions for down-conversion in a nonlinear crystal. Type-I down-conversion, which is the subject of this work, requires that an extraordinarily[1] polarized pump photon down-convert into two ordinarily polarized daughter photons. Type-II down-conversion requires instead that an extraordinarily polarized pump photon down-convert into one ordinarily polarized daughter photon and one extraordinarily polarized daughter photon.

Momentum and energy conservation in Type-I down-conversion require that the photons produced exit the crystal within the geometry of a pair of cones, shown graphically in Figure 3.1, and always on opposite sides of the propagation axis of the pump beam. The opening angles of these cones depend on the wavelength of the down-converted light, and for the special case of degenerate down-conversion ($\omega_1 = \omega_2$), these cones overlap; i.e., the daughter photons are produced on opposite sides of the same cone. While the polar angle of this single cone is defined by the geometry of the crystal, the azimuthal angle is random.

3.2 Two-crystal source of entangled photons

Our entanglement source uses degenerate Type-I down-conversion from 0.6-mm thick nonlinear β-Barium Borate (BBO) crystals, cut such that the optic axis is $\sim 33.9^\circ$ from the pump axis. In this geometry, 351.1-nm pump photons down-convert into 702.2-nm daughter photons which exit the crystal in a cone with a $\sim 3^\circ$ half opening angle. One such crystal, with its optic axis oriented in the vertical plane, will produce pairs of horizontally polarized daughter photons ($|HH\rangle$) when pumped with vertical light ($|V\rangle$). If this geometry were to be rotated by 90°, then horizontal pump photons ($|H\rangle$) would down-convert into

[1] Birefringent materials have different indices of refraction depending on both the direction that light is traveling within them and the polarization of that light. Uniaxial crystals have a single axis whose index differs from the other two. The polarization component of light which is perpendicular to this axis is said to be ordinarily polarized, and the component of the light which is orthogonal to this ordinary component is said to be extraordinarily polarized. The exception to this is light whose propagation direction is the optic axis. This light is entirely ordinary.

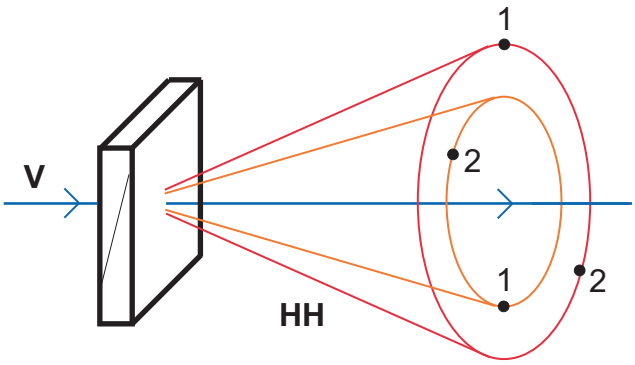

Figure 3.1: Type-I spontaneous parametric down-conversion from a nonlinear crystal. High energy pump photons occasionally transform into two lower energy daughter photons. The pump photon is extraordinarily polarized within the down-conversion crystal; the daughter photons are ordinarily polarized. The diagram shows a nonlinear crystal with vertical optic axis, vertically polarized pump beam, and horizontally polarized down-conversion. Energy conservation requires that the frequencies of these daughter photons sum to the pump frequency. Momentum conservation requires that the daughter photons exit the crystal on opposite sides of the pump axis. The down-converted mode is always in the shape of a cone, but the opening angle (with respect to the pump axis) depends on the frequency of the down-converted photons. Down-converted pairs are always emitted on opposite sides of the pump axis (e.g., the two points marked 1 and the two points marked 2 in the diagram correspond to the locations where pairs could be detected in coincidence). For the special case of degenerate down-conversion, the cones overlap and the daughter photons always exit the crystal on opposite sides of a single cone (the figure shows nondegenerate down-conversion).

pairs of vertically polarized daughter photons ($|VV\rangle$).

In order to create entangled states, we place crystals of each orientation back to back, such that the two down-conversion processes are coherent with each other [39]. (This is possible because the pump coherence length is very long compared to the difference in optical path length accumulated in the DC crystals by the pump and the down-conversion, making the two processes technically indistinguishable.) In this configuration, the 351-nm pump state

$$|\psi_{\text{pump}}\rangle = \cos(\epsilon)\,|V\rangle + e^{i\phi}\sin(\epsilon)\,|H\rangle \tag{3.4}$$

down-converts into the two-photon 702-nm state

$$|\psi_{\text{dc}}\rangle = \cos(\epsilon)\,|HH\rangle + e^{i\phi}\sin(\epsilon)\,|VV\rangle. \tag{3.5}$$

For a 45°-polarized pump photons this state is maximally entangled. Figure 3.2a shows a top-down view of the experimental setup.

Example 3.1 (Producing nonmaximally-entangled states) *Our source of entanglement, when pumped with* 45° *light, produces the maximally entangled*

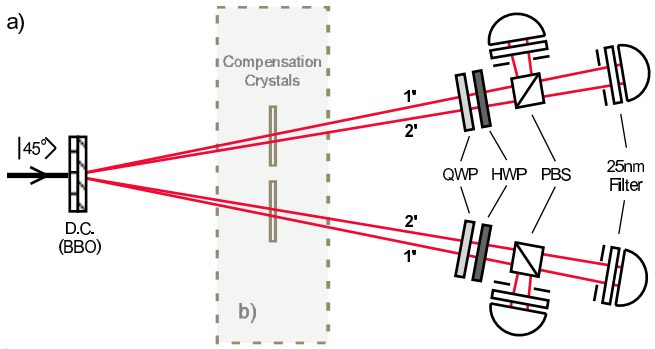

Figure 3.2: Entangled photon source. (a) A 45°-polarized 351-nm pump beam down-converts in two adjacent nonlinear crystals (BBO) into the two-photon state $|\psi(\phi)\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\phi}|VV\rangle\right)$. The phase factor ϕ depends on the angle of the down-conversion photons within the crystals, such that the states corresponding to paths 1' and 2' are $|\psi(1')\rangle$ and $|\psi(2')\rangle$, respectively. The addition of these states results in effective decoherence. The use of larger collection apertures increases the effect. (b) By placing additional specially designed birefringent crystals into the down-conversion path, this phase variation can be compensated for, largely eliminating the decohering effect of large collection apertures. The quarter waveplate (QWP), half waveplate (HWP), polarizing beam splitter (PBS) combinations in each arm allow projection into any separable polarization basis. A sequence of these projections allows complete tomographic reconstruction of any two-qubit density matrix. Quoted high count rates were collected using 9-mm irises (120 cm from the source) and 25-nm (FWHM) frequency filters.

state:

$$|\psi_{\max}\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\phi}|VV\rangle\right). \tag{3.6}$$

However, because the HH and VV processes are separately controlled by the V and H polarizations of the pump, using a different pump polarization will produce down-converted states with any amount of entanglement:

$$|\psi_{\mathrm{dc}}\rangle = \cos(\epsilon)\,|HH\rangle + e^{i\phi}\sin(\epsilon)\,|VV\rangle. \tag{3.7}$$

Here the factor ϵ controls the entanglement of the state produced. Experimentally, any value of ϵ can be achieved by sending a vertically polarized pump through a half waveplate at an angle of $\frac{\epsilon}{2}$ (followed by the down-conversion crystals) [72]. Experimentally produced nonmaximally entangled states are shown in Figure 3.3.

This geometry has proved to be very robust (it is in essence an extremely stable interferometer), and can be used to produce states which have greater than 99% fidelity with a maximally entangled state [4]. However, it is desirable to have both a very high quality as well as a very bright source; as the bit rate of any quantum protocol—as well as the speed of most experiments—is limited by state brightness. The state's brightness can be easily increased by

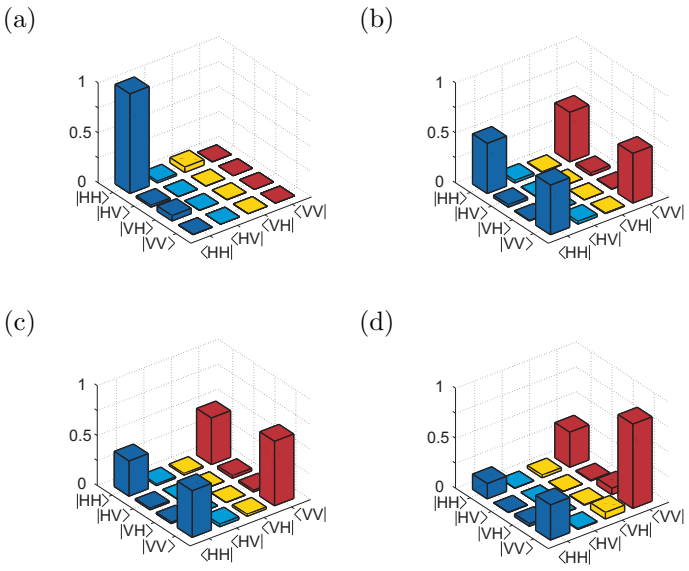

Figure 3.3: Four experimentally measured density matrices of two-qubit photon states, ranging in entanglement from separable to nonmaximally entangled to maximally entangled states. These states are of the form of Equation 3.7, with ϵ approximately equal to: (a) 0.0° (a separable state), (b) 45° (a maximally entangled state), (c) 54° (a nonmaximally entangled state), and (d) 67° (a nonmaximally entangled state). The absolute value of the density matrices are shown; the phases on the off-diagonal elements can be easily tuned using a ϕ-plate (see Example 4.1).

increasing the pump power, which is typically around 50 mW continuous wave (from a SpectraPhysics argon-ion laser running at 351.1 nm). Unfortunately, the nonlinear crystals are easily damaged by too much pump power[2], effectively limiting everyday use to less than 100 mW of power.

The next step for increasing state brightness is the collection of a larger solid angle of down-converted photons. In practice, this is accomplished by opening the collection apertures that define the $\vec{k}$-vector content of the entangled state. This quadratically increases the number of measured photon pairs with aperture diameter, but also typically induces a steady drop in the quality of the entangled states. More specifically, the coherence between the $|HH\rangle$ and $|VV\rangle$ terms degrades, leading to progressively more mixed states. This effect was measured, and the results are shown in Figure 3.4.

3.2.1 The "Migdall effect"

Equation 3.5 states that the output of our entanglement source produces only HH and VV terms, with no contribution from HV or VH counts. This is

[2]Though little data has been systematically collected concerning this effect, it is believed that it is in fact the antireflection coatings on the BBO crystals which are damaged by excessive pump power. In practice this effect requires that the experimenter periodically move the pump to fresh "spots" on the down-conversion crystals, and eventually replace them entirely with new, undamaged crystals.

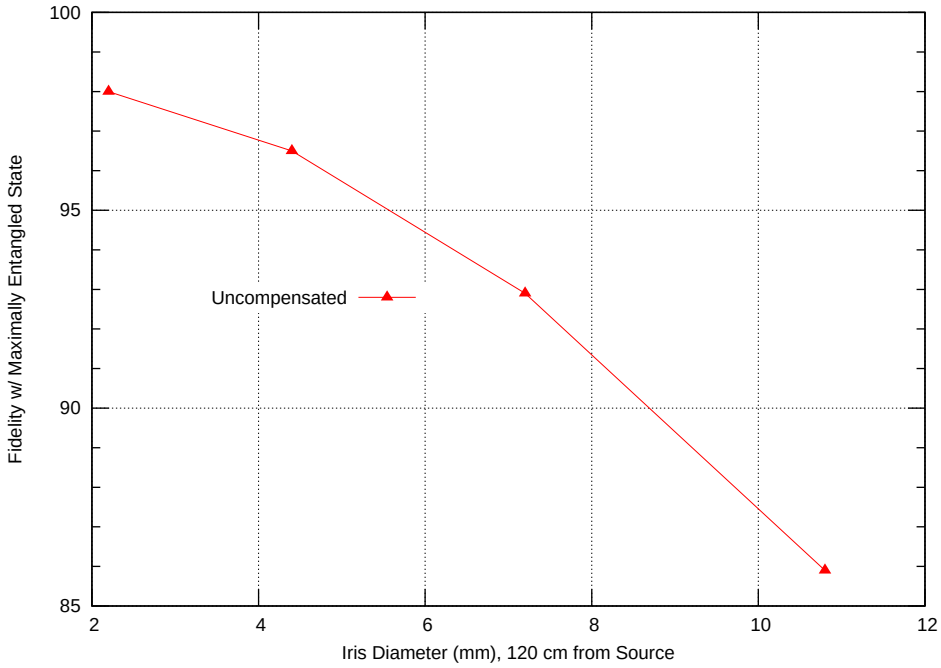

Figure 3.4: The effects of iris size on the quality of the measured entangled state. The data was taken with collection irises 120 cm from the down-conversion crystals; the irises were opened together. For larger irises, a larger solid angle is collected, which leads to state decoherence and degraded fidelity with the maximally entangled state.

in fact an approximation, which was first pointed out by Alan Migdall in [49]. Type-I down-conversion requires that the daughter photons be ordinarily polarized in the nonlinear crystal in which they are produced. However, for a nonlinear crystal which is not critically phasematched[3] (like our crystal, whose optic axis is at 33.9°), ordinary polarization will not exactly correspond to $|HH\rangle$ or $|VV\rangle$. The details of this are discussed in [49], but the essential effect arises because the ordinary polarization vector must be perpendicular to both the beam propagation vector and the crystal optic axis. For our source, the actual state produced by the first crystal is predicted to be $\sim |1.3°\rangle|-1.3°\rangle$, rather than $|HH\rangle \equiv |0°\rangle|0°\rangle$. These states have 99.9% fidelity with our HH and VV approximations, and are therefore neglected. However, this effect will account for a 0.1% error in all states that we create—an effect that becomes increasingly important as the accuracy of our state creation system improves. The actual state produced by the second crystal is $|VV\rangle \equiv |90°\rangle|90°\rangle$, as expected.

[3]Critical phasematching refers to the condition where the optic axis of the down-conversion crystal is in a plane parallel to the surface of that crystal, and hence perpendicular to the pump propagation direction.

3.3 Angle-dependent birefringent phases

The observed decoherence of the entangled state as a function of collected solid angle is due to an angle-dependent phase shift. Recall from Equation 3.5 that for experimentally produced entangled states there is a phase factor $e^{i\phi}$ between the $|HH\rangle$ and $|VV\rangle$ terms. This phase depends on the geometry of the crystals, the wavelengths of the down-converted photons and, most importantly, on the angles that the down-conversion leaves the crystals. Because the propagation direction of down-converted photons is distinguishing information, when multiple $\vec{k}$-vectors are collected simultaneously, many entangled states with different phases are added incoherently, leading to decoherence. In order to calculate this effect, it is first necessary to calculate the optical phase that a photon of a specific wavelength receives from an arbitrary birefringent crystal. It is important to note that the relevant phase is *not* derived from the total optical path length, but from the *difference* in optical path lengths between the horizontal and vertical components of light sent through the nonlinear crystal.

3.3.1 Angle-dependent phases from arbitrary crystals

In general, calculating the phase that a monochromatic photon accumulates from a birefringent crystal at arbitrary orientation requires three separate but interconnected calculations: the extraordinary phase Φ_e, the ordinary phase Φ_o, and the extra phase Φ_Δ that the extraordinary beam accumulates *outside* the crystal.

To simplify this calculation, it is assumed that the $\hat{z}$-axis is normal to the crystal surface, the incident photon is traveling in the $\hat{x}$-$\hat{z}$ plane with positive $\hat{z}$- and non-negative $\hat{x}$-momenta, and the faces of the crystal are parallel and transversely unbounded. Furthermore, it is useful to define the following angles, vectors, and variables (see Figure 3.5). Note that in this section, $\hat{}$ symbol written over a variable indicates a unit vector. In every other section, a $\hat{}$ symbol written over a variable indicates a matrix.

Using these definitions, it is possible to construct the formulae necessary to compute the desired phases. First, recall the formulae for Snell's Law and basic birefringence [24]:

$$\psi_o \equiv \operatorname{asin}\left(\frac{\sin(\alpha)}{n_o}\right), \tag{3.8}$$

$$\psi_e \equiv \operatorname{asin}\left(\frac{\sin(\alpha)}{n^e(\theta)}\right), \tag{3.9}$$

$$n^e(\theta) \equiv n_o \sqrt{\frac{1+\tan^2\theta}{1+\left(\frac{n_o}{n_e}\tan\theta\right)^2}}. \tag{3.10}$$

In birefringent materials, the extraordinary beam's Poynting vector ($\hat{S}_e$)

$\hat{k}_\alpha$	Incident photon's unit momentum outside of the crystal
$\hat{\kappa}_o$	Unit momentum for ordinarily polarized light in the crystal
$\hat{\kappa}_e$	Unit momentum for extraordinarily polarized light in the crystal
α	Angle between $\hat{z}$ and $\hat{k}_\alpha$
ψ_o	Angle between $\hat{z}$ and $\hat{\kappa}_o$
ψ_e	Angle between $\hat{z}$ and $\hat{\kappa}_e$
$\hat{O}$	Crystal optic axis
$\hat{S}_o$	Ordinary Poynting vector inside the crystal Always equal to $\hat{\kappa}_o$
$\hat{S}_e$	Extraordinary Poynting vector inside the crystal
β	Angle between $\hat{z}$ and $\hat{S}_e$
ρ	Angle between $\hat{S}_e$ and $\hat{\kappa}_e$
Δ	Free-space distance that the extraordinary beam must travel
n_o, n_e	The crystal's ordinary and extraordinary indices of refraction
$n^e(\theta)$	Index of refraction for a direction angle θ from the optic axis

Table 3.1: Variables used for calculating optical phases in birefringent crystals. Note that a ˆ symbol written over a variable indicates a unit vector.

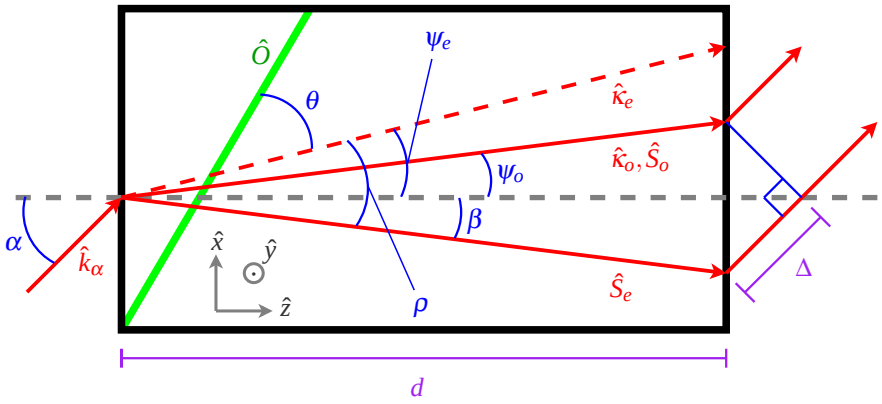

Figure 3.5: Diagram illustrating all relevant vectors, angles, and variables used for calculating angle-dependent phase differences due to birefringent crystals. Arbitrarily polarized light is incident from the left onto a negative uniaxial crystal (e.g., BBO) with its optic axis ($\hat{O}$) in the plane of the page.

deviates from its momentum vector ($\hat{\kappa}_e$) by an angle ρ [24]:

$$\rho = \left(\theta - \text{atan}\left[\frac{n_o^2}{n_e^2}\tan\theta\right]\right)\text{sgn}\left(n_o - n_e\right). \tag{3.11}$$

$\hat{S}_e$ dictates the physical path of the light as it travels through the crystal, and more importantly, when and where it *exits* the crystal. This vector is further from the optic axis than $\hat{\kappa}_e$ for negative uniaxial crystals ($n_o > n_e$, e.g., BBO), and closer for positive uniaxial crystals ($n_o < n_e$, e.g., quartz). Figure 3.5 shows the special case where $\hat{S}_e$ is in the plane of incidence.

It is now possible to compute in radians the three phases relevant to the problem, by multiplying the optical path length in waves by a factor of $\frac{2\pi}{\lambda}$ (λ

is the wavelength of the photon *outside* of the crystal):

$$\Phi_o = \frac{d}{\cos(\psi_o)} n_o \left(\hat{S}_o \cdot \hat{\kappa}_o\right) \frac{2\pi}{\lambda} = \frac{d}{\cos(\psi_o)} n_o \frac{2\pi}{\lambda}, \tag{3.12}$$

$$\Phi_e = \frac{d}{\cos(\beta)} n^e(\theta) \left(\hat{S}_e \cdot \hat{\kappa}_e\right) \frac{2\pi}{\lambda}, \tag{3.13}$$

$$\Phi_\Delta = \frac{2\pi}{\lambda}\Delta = \frac{2\pi}{\lambda}\left(\frac{d}{\cos(\psi_o)}\hat{S}_o \cdot \hat{k}_\alpha - \frac{d}{\cos(\beta)}\hat{S}_e \cdot \hat{k}_\alpha\right). \tag{3.14}$$

The dot product in Equation 3.12 and Equation 3.13 accounts for the fact that although the light travels along $\hat{S}$, its wavefronts are perpendicular to and determined by $\hat{\kappa}$, creating a smaller effective optical path length by a factor of the cosine of the angle between these vectors. In Equation 3.12, this factor is 1, but for the extraordinary case—Equation 3.13—this can be a large effect. The distance Δ depends entirely on where $\hat{S}_o$ and $\hat{S}_e$ enter and exit the crystal, which can be quite complicated for the general case where these vectors are not in the plane of incidence: to simplify this calculation, one can instead project these two vectors (using a dot product) onto $\hat{k}_\alpha$. The difference between these projections is the distance Δ.

3.3.2 The angle-dependent phase from DC crystals

Using these methods to compute ordinary and extraordinary phases in birefringent crystals, it is possible to apply them to the specific case of the down-conversion crystals used to produce entangled states. To do this, it is necessary to calculate the phase ϕ as a function of the signal direction outside of the crystals, $\hat{k}_\alpha$. To simplify this task, assume that the signal and idler photons are "born" in the center of their respective crystals, travel in directions which are completely symmetric about the pump beam, are ordinarily polarized in their own crystal, and are extraordinarily polarized in the remaining crystal (this last approximation is close to correct, but see Section 3.2.1 and reference [49]). Both the $|VV\rangle$ and $|HH\rangle$ terms receive an equal ordinary phase in their own crystals, which we can neglect as a global phase. The $|VV\rangle$ photons born in the first crystal together receive a net additional extraordinary phase in the second crystal equal to

$$\phi_{\mathrm{dc}}\left(\hat{k}_\alpha\right) = 2\left(\Phi_{\mathrm{dc},e} + \Phi_{\mathrm{dc},\Delta}\right), \tag{3.15}$$

double the phase accumulated by just one of the daughter photons. (Note that the terms in Equation 3.15 implicitly depend on $\hat{k}_\alpha$). This makes the final entangled state for a 45°-polarized pump beam:

$$|\psi_{\hat{k}_\alpha}\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i2(\Phi_{\mathrm{dc},e}+\Phi_{\mathrm{dc},\Delta})}|VV\rangle\right). \tag{3.16}$$

Because experimental setups use finite size irises which accept a range of $\hat{k}_\alpha$,

the state passed by the irises can be described by the unnormalized state:

$$|\psi\rangle_{\text{Iris}} = \int_{\text{Iris}} |\psi_{\hat{k}_\alpha}\rangle |\hat{k}_\alpha\rangle d\hat{k}_\alpha. \tag{3.17}$$

Because all pairs that pass through the irises are measured together, the actual state measurement traces over direction, producing the density matrix:

$$\hat{\rho} = \int_{\text{Iris}} |\psi_{\hat{k}_\alpha}\rangle\langle\psi_{\hat{k}_\alpha}| d\hat{k}_\alpha. \tag{3.18}$$

Incoherently summing these *individually* maximally entangled states leads to an effective decoherence, as we conjectured earlier. Direct theoretical calculation of the phase ϕ for our down-conversion crystals does in fact reveal a large, mostly one-dimensional angle-dependent phase shift, which appears to explain the experimentally observed decoherence with increased iris size. This theoretically calculated phasemap is shown in Figure 3.6a.

In order to experimentally verify this calculation, the same phasemap was measured by taking a series of 25 complete state tomographies, using very small (2-mm) irises, and recording the phase ϕ between the $|HH\rangle$ and $|VV\rangle$ components of the measured entangled state. These experimental results are shown in Figure 3.6d.

3.4 Phase compensation

The angle-dependent birefringent phase, which is the source of decoherence for the bright source of entangled photons, can be compensated for with an appropriate choice of additional birefringent optics. These compensation crystals are inserted into the path of the down-converted entangled state in each arm, parallel to the down-conversion crystals, as shown in Figure 3.2. In order to calculate the phase-shift that this type of crystal will cause, it is necessary to remember that these are different than the special case of down-conversion crystals; here all photons always travel through the *entire* crystal (e.g., no photons are born in the middle of a crystal). For compensation crystals of the same type, one in each arm, with optic axes in the horizontal plane, the compensating phase is

$$\phi_c\left(\hat{k}_\alpha\right) = 2\left(\Phi_{c,o} - \Phi_{c,e} - \Phi_{c,\Delta}\right), \tag{3.19}$$

with the final compensated state equal to

$$|\psi_{\text{comp},\hat{k}_\alpha}\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i(\phi_{\text{dc}}+\phi_c)}|VV\rangle\right). \tag{3.20}$$

In order to design crystals to compensate the measured phase slope from the down-conversion source, it is necessary to optimize compensation crystals such that the sum of the phasemaps from the down-conversion and the compensation

crystals is as flat as possible.

It was experimentally convenient to use BBO cut at the same angle as the down-conversion crystals for the compensation crystals. Note however that there is a continuum of virtually equivalent choices accessible by making a tradeoff between optic-axis angle and crystal thickness. If the optic axis is fixed at 33.9° (from crystal normal) and the thickness is optimized, we find that a thickness of 245-μm is ideal (phasemap shown in Figure 3.6a). Figure 3.6b shows the expected sum of the first two phasemaps, an extremely flat function as compared to the original down-conversion phasemap; Figure 3.6c shows the theoretical prediction of the simultaneously bright and high-fidelity compensated state (only the absolute value of the state is shown, as it is the coherence which is important rather than the particular value of the phase between $|HH\rangle$ and $|VV\rangle$).

In order to test the effectiveness of this compensation method, another experimentally measured phasemap was taken, this time with the compensation crystals in the beam path. This data, shown in Figure 3.6e, reveals a flat phasemap. This indicates that for the $\vec{k}$-vectors measured the down-converted entangled states have essentially the same phase factor ϕ. Even though these states are added incoherently, this addition does not lead to decoherence. Figure 3.6c and Figure 3.6f compare the theoretical and experimental density matrices, respectively, for a compensated entangled state collected using 1-cm irises.

Close inspection of the compensated experimental curves shows small induced phase shift in the azimuthal (y) direction, even while the radial (x) phase shift appears to have been almost completely compensated for. This residual phase shift is not predicted by theory, and could result from a number of effects, none of which have been thoroughly tested: limitations of the theory, due to its use of monochromatic light; limitations on the theory, due to the Migdall effect (see subsection 3.2.1 and reference [49]); systematic error in the tomographies, which is dependent on the $\vec{k}$-vectors of the down-conversion; misalignment of the compensation crystals relative to each other or relative to the down-conversion crystals.

3.5 The ultra-bright source

3.5.1 Results

The final test of this compensation method is its effectiveness in dealing with the original decoherence problem. Figure 3.7 shows the original data for source decoherence with iris size, superimposed with the measured data for the compensated source. For reference, the same plot also shows the quadratic increase in coincidence counts gained from using larger collection irises.

This figure's data indicates that the compensation technique has led to a simultaneously bright and high quality state. The brightest state measured using this technique and characterized with state tomography (see Chapter 5) had a

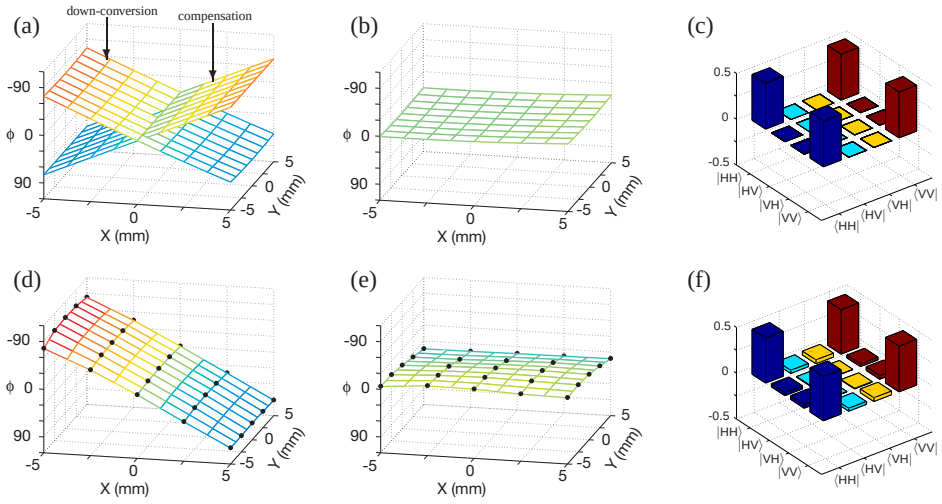

Figure 3.6: Theoretical (a-b) and experimental (d-e) plots of the phase difference ϕ as a function of signal direction $\hat{k}_\alpha$, combined with the predicted (c) and measured (f) density matrices. To match experimentally collected data, the direction $\hat{k}_\alpha$ is represented as a transverse position in space ~ 120 cm from the down-conversion crystals. Here, $x = 0$, $y = 0$ corresponds to the central 702-nm down-conversion directions, $\sim 3^\circ$ (6.35 cm) from the pump axis. Here, x refers to the radial direction and y refers to the azimuthal direction, with respect to the pump axis. (a) The phasemap due solely to the down-conversion crystals (each 600-μm thick, 33.9° optic axis, BBO), superimposed with the phasemap due to two BBO compensation crystals, one in each arm, each 245 μm thick and cut with a 33.9° optic axis. The slope of each phasemap is approximately $\pm 14^\circ$ per mm. (b) The sum of both phasemaps from (a). The flat character indicates that approximately the same entangled state will be present at each position on this plot, corresponding to a high-fidelity state measured using large irises. (c) The theoretically predicted density matrix that would result from a measurement over this flat phase surface, using 1-cm diameter irises. (d) Experimentally measured phase for the uncompensated configuration. Each black dot represents an experimental measurement of this phase, extracted from the result of a full state tomography. The mesh graphically represents these points by linearly connecting nearest neighbors. The phase difference is nearly linear, approximately 17° per millimeter, and varying in the radial direction out from the pump beam axis. (e) The compensated configuration. The surface is very flat, with a maximum slope of less than 3°/mm, and a total phase variation of approximately 25° over a centimeter, a seven-fold improvement over the uncompensated case. (f) The experimentally recorded density matrix (absolute value shown here) describing the ultra-bright entangled state: 1.02×10^6 measured pairs per second, 97.7% fidelity with a maximally entangled state.

fidelity of $97.7\% \pm 0.1\%$ with a maximally entangled state and an average measured intensity of 1.02×10^6 photon pairs per second. A series of measurements of Bell's inequalities[4] [75] were taken using a 310-mW pump beam in order

[4] A Bell inequality is a type of measurement which measures correlation and is used to characterize two particle states; it is discussed in detail in Chapter 5. For the type of inequality used here, local hidden variable theories predict values of less than or equal to 2. Quantum mechanics predicts values of up to $2\sqrt{2}$, with values of greater than 2 indicating entangled states. The largest violations of Bell's inequalities reported in Chapter 5 are 2419–σ in 2

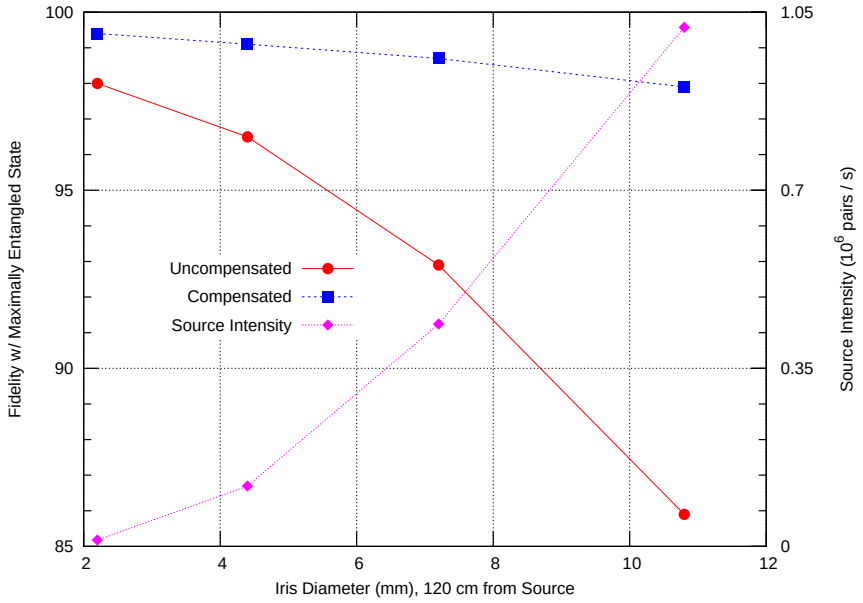

Figure 3.7: The effects of iris size on the fidelity and rate of measured entangled states. Note that for this measurement, the diameters of *both* irises are adjusted, with each positioned in the plane normal to the pump direction and 120 cm from the down-conversion crystals. Shown here are both the standard (uncompensated) and compensated configurations. The y-axis on the right describes the detected source intensity (for a 280-mW pump and 25-nm filters), quadratically increasing as a function of iris size.

to demonstrate not only the state's quality but also the extremely short time in which precise measurements can be collected. These measurements resulted in a violation of 2.7260 ± 0.00336 ($216\ \sigma$) in 0.8 s of total measurement time, and a violation of 2.7252 ± 0.000585 ($1239\ \sigma$) in 28 s with an average measured intensity of 2.01×10^6 pairs per second. (Note that because the statistical errors on these values are extremely low, it is almost certain that uncharacterized systematic errors are dominating these measurements.)

3.5.2 Experimental techniques

In order to measure this ultra-bright source, it was necessary to address the problem of detector saturation. Because the Si-avalanche photodiodes used as detectors have a maximum useful rate of approximately 4×10^6 singles counts per second, measuring high coincidence count rates requires high detection efficiencies.

One method for increasing our detection efficiencies was proposed and first implemented by A. G. White (Univ. of Queensland, Australia), and attempted to eliminate some of the loss due to high reflectivity (40%) spectral filters. By placing a QWP between the measurement PBS and the filter-detector assem-

minutes of data collection (largest statistical violation) and 2.826 ± 0.005 in four hours of data collection (largest absolute violation).

blies, and a mirror at the unused port of the PBS, all light reflected at the filters will eventually be redirected towards the filters a *second* time. This reflection technique is illustrated in Figure 3.8a, and the associated filter transmissivity improvement is plotted in Figure 3.8b. This technique effectively increases the transmission of our filters from 0.6 to $\sim (0.6 + 0.4 \times 0.6) = 0.84$. (Note that these calculations assume perfect transmission in the waveplates and polarizing beam splitter, and perfect reflection at the mirror.)

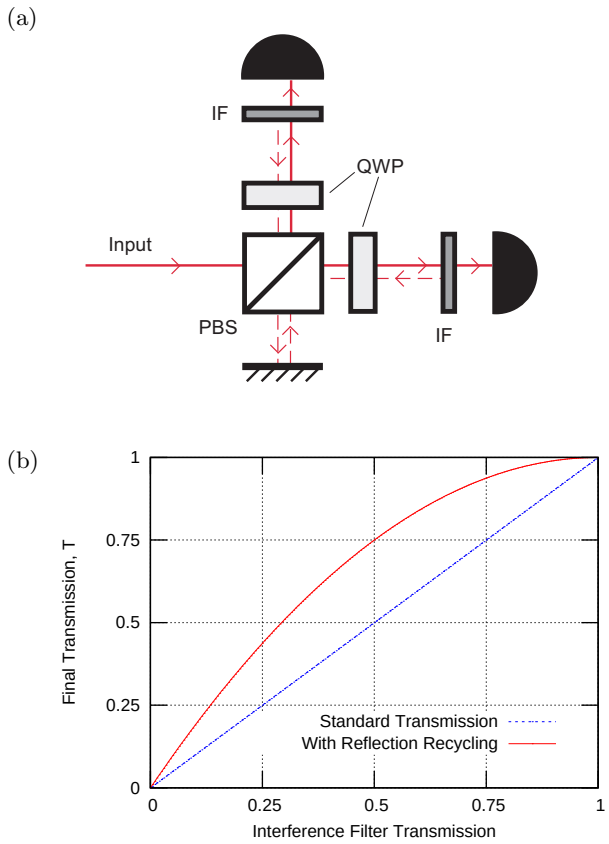

Figure 3.8: Diagram and theoretical benefits of a reflection recycling technique designed to increase the transmission of an interference filter which reflects too much light. The same technique can be applied to any partially reflective surface. (a) A diagram of the experimental setup. Consider light transmitted by the PBS, and therefore horizontally polarized. After passing through the QWP this light is transformed to right-circularly polarized light, and then either transmitted or reflected by the interference filters. If transmitted, then the device is considered to have succeeded. If reflected, the light will travel back through the quarter waveplate, be rotated to vertical polarization, reflected at the PBS, retroreflected by a mirror, and travel back through the PBS toward the filter-detector assembly for a second chance at detection. In a similar manner, light which is reflected towards the other detector will also get two attempts at transmission. (b) A theoretical plot of the effectiveness of this technique. One curve shows possible transmissions of an interference filter, T, and another shows the theoretical improved transmission T' when reflection recycling is used: $T' \equiv T + T(1 - T) = T(2 - T)$.

Taking advantage of both the high collection efficiency of large irises and this new reflection technique, detection efficiencies in excess of 30% were achieved. (This efficiency is due to detector quantum efficiency ($\sim$ 65% for the SPCM AQR-14), iris collection efficiency ($\sim$ 70% for 9-mm irises), reflection losses from the interference filters ($\sim$ 84%), saturation losses due to the 30-ns dead-time of the detectors (up to $\sim$ 5%, depending on the count rate), and other reflection and absorption losses due to miscellaneous optics.) In addition, a polarizing beam-splitter and two detectors in each arm (see Figure 3.2), simultaneously measured a complete basis of four separable projectors at a time, giving each singles count a chance to contribute to a coincidence.

It is interesting to note that because these experiments operated near the saturation limit of the detectors, the entangled pairs produced far exceeded those measured, and it is estimated that the actual rate of entangled-state *production* into 9-mm irises and 25-nm filters exceeds 20×10^6 entangled pairs per second.

4 State Manipulation

4.1 Methods for manipulating quantum states

The photon source described in Chapter 3 can produce either separable or maximally-entangled states, as well as the gamut of nonmaximally-entangled states running between those two extremes (see Example 3.1). This chapter details the experimental options for state manipulation.

To represent the operators which act on polarization-encoded qubits, it is convenient to introduce the Jones calculus [30], a representation for polarized photons corresponding to the notation of Chapter 2. In Chapter 2, pure and mixed states were represented using vectors and matrices; operators in the Jones calculus are represented by matrices which linearly act on these vectors and matrices. For an operator $\hat{O}$ acting on a pure state $|\psi\rangle$, the resulting state $|\psi'\rangle$ is given by

$$|\psi'\rangle = \hat{O}|\psi\rangle. \tag{4.1}$$

For a state represented by a density matrix $\hat{\rho}$, the resulting state after the operator is given by

$$\hat{\rho}' = \hat{O}\hat{\rho}\hat{O}^\dagger, \tag{4.2}$$

where the $\dagger$ represents a Hermitian conjugate (i.e., transposed complex conjugate). For single qubit states, the operator $\hat{O}$ will be a 2×2 matrix, for d-level qudits, a $d \times d$ matrix, and for n-qubit systems, a $2^n \times 2^n$ matrix. In all cases, each element of $\hat{O}$ will in general be complex.

4.1.1 Unitary transformations

The simplest form of operator is the unitary transformation. An operator $\hat{U}$ is unitary if and only if

$$\hat{U}\hat{U}^\dagger = I, \tag{4.3}$$

where I is the identity matrix. Unitary transformations are often referred to as *unitary rotations*, and single-qubit unitary transformations can be easily visualized as rotations of the Poincaré sphere about an axis defined by the orthogonal states defined by the eigenvectors of the transformation. In optics, unitary transforms on polarization are often implemented by *waveplates*, which most often act as either a 90° (for a quarter waveplate) or a 180° (for a half waveplate) rotation about an axis in the equator of the Poincaré sphere. More information

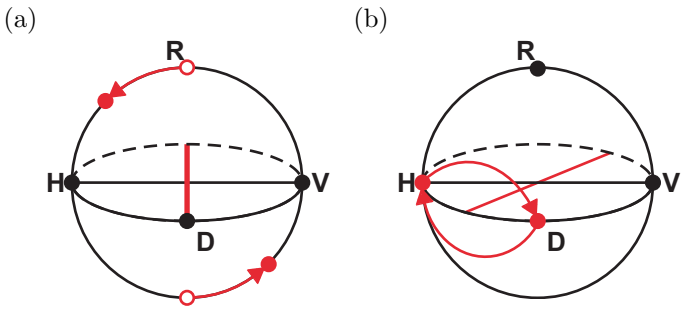

Figure 4.1: Graphical depictions of two unitary transformations on the Poincaré sphere. (a) A $\frac{\pi}{4}$ rotation about the H-V axis. This type of rotation could be implemented using a ϕ-plate (see Example 4.1). (b) A Hadamard gate. This rotation is equivalent to a half waveplate at 22.5°, and takes $|H\rangle \rightarrow |D\rangle, |V\rangle \rightarrow |A\rangle, |D\rangle \rightarrow |H\rangle, |A\rangle \rightarrow |V\rangle, |R\rangle \rightarrow |L\rangle$, and $|L\rangle \rightarrow |R\rangle$.

about waveplates can be found in Appendix C.

General unitary transformations have several qualities which are very useful to keep in mind. First, all unitary transformations have an *eigenbasis*. This is a set of orthonormal states which are unaffected (except for a phase factor $e^{i\phi}$) by the unitary transformation (for a d-level system, the eigenbasis has d elements). This is why a unitary rotation on the Poincaré sphere rotates about an axis defined by the rotation's eigenvectors. They are the only states unaffected by the rotation. Second, the unitary transformation operates by applying a phase to each of these states. Written mathematically, this looks like

$$\hat{U} \equiv e^{i\phi_0}|e_0\rangle\langle e_0| + e^{i\phi_1}|e_1\rangle\langle e_1| + \ldots + e^{i\phi_{d-1}}|e_{d-1}\rangle\langle e_{d-1}|, \tag{4.4}$$

where the states $|e_i\rangle$ are the orthonormal basis which are the eigenvectors of the operator $\hat{U}$. Thought about another way, a unitary transformation rotates the d-element canonical basis $|i\rangle$ into a new orthonormal basis $|\psi_i\rangle$ while applying a phase to each element:

$$\hat{U} \equiv e^{i\phi'_0}|\psi_1\rangle\langle 0| + e^{i\phi'_1}|\psi_2\rangle\langle 1| + \ldots + e^{i\phi'_{d-1}}|\psi_{d-1}\rangle\langle d-1|. \tag{4.5}$$

Even more generally, a unitary can always be written as a transformation from *any* orthonormal basis $|\tau_i\rangle$ to another orthonormal basis $|\psi_i\rangle$:

$$\hat{U} \equiv e^{i\phi_0}|\psi_1\rangle\langle\tau_0| + e^{i\phi_1}|\psi_2\rangle\langle\tau_1| + \ldots + e^{i\phi_{d-1}}|\psi_{d-1}\rangle\langle\tau_{d-1}|. \tag{4.6}$$

Figure 4.1 depicts several single-qubit rotations using the Poincaré sphere.

Example 4.1 (The ϕ-plate) *In general, the two-crystal entanglement source of Chapter 3 yields maximally entangled states of the form*

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\phi}|VV\rangle\right). \tag{4.7}$$

It is often experimentally useful to tune the phase ϕ in order to produce a different maximally entangled state of the same form. Experimentally, this is accomplished by placing a waveplate at $0°$, tilted about the vertical axis, in one arm of the down-conversion. Tilting the waveplate will slowly vary the retardance ϕ' of that waveplate, changing it from a quarter or a half waveplate to some other value. Written in terms of its eigenstates, this waveplate's transformation is

$$\hat{U}_\phi \equiv |H\rangle\langle H| + e^{i\phi'}|V\rangle\langle V|. \tag{4.8}$$

For a waveplate at normal incidence,

$$\phi' = 2\pi\left(n_o - n_e\right)\frac{d}{\lambda}. \tag{4.9}$$

Written as a matrix, the transformation is

$$\hat{U}_\phi \equiv \begin{pmatrix} 1 & 0 \\ 0 & e^{\phi'} \end{pmatrix}. \tag{4.10}$$

Since this is a single-qubit operator, applying it to the two-qubit state above requires writing it in a tensor product with the identity (as there is no operator being applied to the other arm of the down-conversion). The total operator is therefore

$$\hat{U} = \hat{U}_\phi \otimes I = \begin{pmatrix} 1 & 0 \\ 0 & e^{\phi'} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{\phi'} & 0 \\ 0 & 0 & 0 & e^{\phi'} \end{pmatrix}, \tag{4.11}$$

which in turn makes the final state

$$\hat{U}|\psi\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + e^{i\left(\phi+\phi'\right)}|VV\rangle\right). \tag{4.12}$$

See Figure 4.1a for a graphical depiction of this operation.

Example 4.2 (Half and quarter waveplates) *Waveplates, as some of the most useful tools in a polarization lab, are often used as single-qubit operators. The operator matrices for quarter and half-waveplates are given by*

$$\begin{aligned} U_{HWP}(\theta) &= \begin{pmatrix} \cos^2(\theta) - \sin^2(\theta) & 2\cos(\theta)\sin(\theta) \\ 2\cos(\theta)\sin(\theta) & \sin^2(\theta) - \cos^2(\theta) \end{pmatrix}, \\ U_{QWP}(\theta) &= \begin{pmatrix} \cos^2(\theta) + i\sin^2(\theta) & (1-i)\cos(\theta)\sin(\theta) \\ (1-i)\cos(\theta)\sin(\theta) & \sin^2(\theta) + i\cos^2(\theta) \end{pmatrix}, \end{aligned} \tag{4.13}$$

with θ denoting the rotation angle of the waveplate with respect to horizontal. On the Poincaré sphere, these correspond to rotations of either $\frac{\pi}{2}$ (QWP) or π (HWP) about an axis in the equatorial plane (i.e., the H-D plane), at an angle

2θ from H.

Example 4.3 (The Hadamard transformation) *The Hadamard transformation [52] is often used in the theory of quantum computation as a basic operator. It's operation on the cardinal polarization states can be summarized as* $|H\rangle \to |D\rangle, |V\rangle \to |A\rangle, |D\rangle \to |H\rangle, |A\rangle \to |V\rangle, |R\rangle \to |L\rangle$, *and* $|L\rangle \to |R\rangle$. *As an operator,*

$$\hat{H} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{4.14}$$

The Hadamard gate can be implemented for polarization-encoded qubits with a half waveplate at $22.5°$. *See Figure 4.1b for a graphical depiction of this operation.*

4.1.2 Projections

Not all useful operations are unitary. Specifically, projective measurements (first introduced in Chapter 2 as a way to interpret certain state parametrizations) allow both the preparation of pure states and the characterization of unknown systems. A projection is a lossy process, meaning that an ensemble of states subjected to projection will in general be reduced in size by the projection, and it will additionally transform all members of the ensemble which survive into the projected state. A projection[1] into the state $|\psi\rangle$ can be written as

$$P_\psi \equiv |\psi\rangle\langle\psi|. \tag{4.15}$$

Example 4.4 (A polarizing beam splitter) *A polarizing beam splitter is an optical element which splits the horizontal and vertical components of a beam of light into two different spatial modes. Most commonly, horizontal light is transmitted through the beam splitter and vertical light is reflected by* $90°$. *This effectively acts as* two *different projectors, one horizontal (*P_H*) and one vertical (*P_V*). As matrices, these are written as*

$$P_H \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_V \equiv \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.16}$$

(More correctly, the PBS entangles the polarization state with direction, and the above matrices apply when we look at only one output at a time—see Example 4.7).

[1] While we discuss projections onto a single state, real measurements almost always project onto a *subspace* of a larger Hilbert space which is composed of many states. For example, a polarizing beam splitter projects a photon into a specific polarization, but because it has limited physical extent also projects that photon into a specific range of spatial modes. This more general view of physical projectors is beyond the scope of this work.

4.1.3 Decohering elements

Decoherence is another type of non-unitary operation, and one of the chief obstacles to the realization of a quantum computer. It is generally considered something to be avoided, though as we will see later, it can also be used to engineer desired states and operations. Decoherence, like a unitary transformation, is associated with an eigenbasis, and the orthonormal states within this basis are unaffected by the decohering operation. Instead of rotating around these states, a decoherer instead destroys their coherence by making them distinguishable. This means that for a decoherer acting in the $\psi - \psi^{\perp}$ basis, all coherence between $|\psi\rangle$ and $|\psi^{\perp}\rangle$ will be destroyed, and the $|\psi\rangle\langle\psi^{\perp}|$ and $|\psi^{\perp}\rangle\langle\psi|$ terms in a state's density matrix will disappear. In a sense, decoherence is like simultaneous projection into the $|\psi\rangle$ and $|\psi^{\perp}\rangle$ bases; the decoherence described above operating on the state $\hat{\rho}$ will yield the state $\mathcal{D}_{\psi}(\hat{\rho})$, given by

$$\begin{aligned} \mathcal{D}_{\psi}(\hat{\rho}) &= P_{\psi}\hat{\rho}P_{\psi}^{\dagger} + P_{\psi^{\perp}}\hat{\rho}P_{\psi^{\perp}}^{\dagger} \\ &= |\psi\rangle\langle\psi|\hat{\rho}|\psi\rangle\langle\psi| + |\psi^{\perp}\rangle\langle\psi^{\perp}|\hat{\rho}|\psi^{\perp}\rangle\langle\psi^{\perp}|. \end{aligned} \tag{4.17}$$

For decoherence in multiple qubits, the situation becomes more complex and requires careful thought to determine the operators involved. The key element is distinguishability; when two states become (in principle) distinguishable, their coherence disappears.

Example 4.5 (Polarization-based temporal decoherers) *A thick birefringent element, with its optic axis in the horizontal or vertical plane, will induce an phase difference ($e^{i(\phi_H - \phi_V)}$) between horizontal and vertical light. Depending on the characteristics of the light passing through the element, and the size of the path length difference, the horizontal and vertical components of the light may become distinguishable. This distinguishability arises from an entanglement between frequency and polarization. The phases that a birefringent element with an optic axis at 0° apply to horizontal and vertical light (ignoring dispersion) are given by:*

$$\begin{aligned} \phi_H &= 2\pi n_e \frac{d}{\lambda} \\ \phi_V &= 2\pi n_o \frac{d}{\lambda}, \end{aligned} \tag{4.18}$$

where n_o and n_e are the ordinary and extraordinary indices of refraction, respectively. While the distinguishability can be exactly calculated by using these frequency dependent phase relationships, it is extremely helpful to think of these phase shifts as a temporal *shift, and the H-V phase difference as an induced temporal separation between the horizontal and vertical components of the light. (Recall a phase shift which is linearly proportional to frequency will act as a pulse delay). If the timing delay is larger than the coherence length of the light, then this information effectively distinguishes between the two states, $|H\rangle$ and*

$|V\rangle$, and causes decoherence, which transforms the state's density matrix $\hat{\rho}$:

$$\hat{\rho} = \begin{pmatrix} A & \alpha \\ \alpha^* & B \end{pmatrix} \quad \longrightarrow \quad \mathcal{D}_H(\hat{\rho}) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{4.19}$$

Temporal decoherence will produce a subtle phenomenon when applied to a two-qubit state. Consider applying this type of decoherence in each arm of a two-qubit, polarization-encoded state, of the type described in Chapter 3. In practice, we use thick pieces of quartz as the decoherers, which induce a $\sim 140\lambda$ path length difference between $|H\rangle$ and $|V\rangle$ photons—the typical down-conversion coherence length for 5-nm interference filters. This path length is large enough to ensure a measurable relative delay *of the light's vertical component with respect to the horizontal. However, the path length difference is much less than the coherence length of the pump photons ($\sim$ 30 cm), which determines the* absolute delay *necessary to induce decoherence. In other words, for the state $|HV\rangle$, photon A will arrive before photon B, for the state $|VH\rangle$, photon B will arrive before photon A, and for the states $|VV\rangle$ and $|HH\rangle$, the two photons will arrive at the same time. Because the coherence length of the pump is so long, the absolute arrival times of the $|VV\rangle$ and $|HH\rangle$ terms* will not be distinguishable*, and the coherence between them will not be destroyed. This makes the final operation of the twin decoherers*

$$\hat{\rho} = \begin{pmatrix} A & \alpha & \beta & \gamma \\ \alpha^* & B & \delta & \epsilon \\ \beta^* & \delta^* & C & \zeta \\ \gamma^* & \epsilon^* & \zeta^* & D \end{pmatrix} \quad \longrightarrow \quad \mathcal{D}_{\text{twin}}(\hat{\rho}) = \begin{pmatrix} A & 0 & 0 & \gamma \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ \gamma^* & 0 & 0 & D \end{pmatrix}. \tag{4.20}$$

This operation is of course an approximation. Because the pump coherence length is not infinitely larger than the induced time delay, the γ terms will be slightly decohered. Because the induced time delay is not infinitely longer than the down-conversion coherence length, the other coherences will be slightly larger than zero. An exact calculation of temporal decoherence requires using the complete frequency spectrum of the pump and the down conversion, as well as the dispersion of the quartz crystal. A full treatment is given in [15].

4.1.4 Entangling operations

All the operations discussed to this point transform single-qubit states, and are used in that capacity for most of the experiments described in this thesis. However, quantum information's real promise lies within the realm of multiple qubits. Entangling operations can couple qubits, creating entangled states from separable states and vice versa. They are, if fact, a special case of the unitary operations described in Section 4.1.1.

Example 4.6 (The CNOT gate) *The most famous example of an entangling*

operation is the CNOT gate (an abbreviation for "conditional not"). This gate implements a conditional operation. If qubit 1 *is in the state* $|V\rangle$ *then a* $\hat{\sigma}_x$ *operation is implemented on qubit* 2*. Written in operator form,*

$$\hat{U}_{\text{CNOT}} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4.21}$$

This operator transforms the separable state $|DH\rangle$ *into the maximally entangled state* $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$*.*

Interestingly, many projections and decoherers can be rewritten in terms of *unitary* entangling operations that take place in a larger dimension Hilbert space than the projection or the decoherence. For projections, an entangling operation couples the states being measured to different modes (e.g., spatial modes or $\vec{k}$-vectors), and only one of these modes is measured. This is why projections *appear* to destroy some of the members of an ensemble. Instead, these particles may have been redirected into a mode that is not measured; this type of behavior is obvious for a PBS (see Example 4.7). For a decoherer, the entangling operation can couple the states which lose coherence to another degree of freedom (see Example 4.8). Because measuring that degree of freedom would in principle distinguish between the states, they must lose coherence.

Example 4.7 (A polarizing beam splitter, revisited) *Instead of looking at the polarizing beam splitter as a projection, we can instead view it as a device which entangles polarization and momentum. If a photon is vertically polarized, it is directed to path* $\vec{k}_R$ *(reflected); if it is horizontally polarized, it is directed to path* $\vec{k}_T$ *(transmitted). Written as an operator,*

$$\hat{U}_{\text{PBS}} \equiv \begin{array}{c} \\ |H\rangle|\vec{k}_T\rangle \\ |H\rangle|\vec{k}_R\rangle \\ |V\rangle|\vec{k}_T\rangle \\ |V\rangle|\vec{k}_R\rangle \end{array} \begin{array}{c} \begin{array}{cccc} \langle H|\langle\vec{k}_T| & \langle H|\langle\vec{k}_R| & \langle V|\langle\vec{k}_T| & \langle V|\langle\vec{k}_R| \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}. \tag{4.22}$$

This is exactly the same operator matrix as the CNOT gate; both are entangling operations that can make separable states into entangled states and vice versa.

Example 4.8 (The CSIGN as a decoherer) *Consider coupling the ancillary qubit state* $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ *to a spin-*$\frac{1}{2}$ *particle initially in the pure*

state $|\psi_2\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right)$ *with the CSIGN*[2] *operation*

$$\hat{U}_{\text{CNOT}} \equiv \begin{array}{c} \\ |A\rangle|\uparrow\rangle \\ |A\rangle|\downarrow\rangle \\ |B\rangle|\uparrow\rangle \\ |B\rangle|\downarrow\rangle \end{array} \begin{array}{c} \begin{array}{cccc} \langle A|\langle\uparrow| & \langle A|\langle\downarrow| & \langle B|\langle\uparrow| & \langle B|\langle\downarrow| \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{array}. \tag{4.23}$$

This is a mathematical model for a magnetic field in the quantization direction with a random and equal chance to apply either an $e^{i2\pi}$ *(field A) or an* $e^{i\pi}$ *(field B) phase factor to* $|\downarrow\rangle$*. This produces the entangled state*

$$|\psi\rangle_{\text{total}} = \frac{1}{2}\left[\left(|\uparrow\rangle + |\downarrow\rangle\right) \otimes |\text{Field}_A\rangle + \left(|\uparrow\rangle - |\downarrow\rangle\right) \otimes |\text{Field}_B\rangle\right]. \tag{4.24}$$

Tracing over the state of the field, we find the system in the totally mixed ensemble $\rho_{\text{mixed}} = \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)$*. The presence of a degree of freedom which is in-principle measurable has made the* $|\uparrow\rangle + |\downarrow\rangle$ *and* $|\uparrow\rangle - |\downarrow\rangle$ *states distinguishable, and has destroyed their coherence. This model of decoherence can vary in either strength*[3] *or basis (with the random phases applied to a state other than* $|\downarrow\rangle$*).*

4.2 Creating arbitrary two-qubit states

Aside from implementing quantum protocols, such as decoherence-free subspaces, the tools of state manipulation that have been presented up to this point can be used to transform the Bell-state output from the double-crystal entanglement source into a wide variety of states. This section will detail a theoretical scheme for implementing truly arbitrary two-qubit state creation, as well as document several important classes of states which have been successfully constructed in the lab.

4.2.1 Creating arbitrary two-qubit pure states

A state's entanglement cannot be changed by using separable operations on its constituent qubits. The converse is also true: for a pure state, local unitary operations can transform it into any other pure state with the same entanglement. Our entanglement source can be used to create nonmaximally entangled states of the form

$$|\psi\rangle = \cos(\theta)|HH\rangle + e^{i\phi}\sin(\theta)|VV\rangle. \tag{4.25}$$

[2]The CSIGN operation is an entangling operation closely related to the CNOT; instead of conditionally applying a $\hat{\sigma}_x$ operator, the CSIGN conditionally applies a $\hat{\sigma}_z$, only if the control bit is in the state $|1\rangle$.

[3]For example, if there were only a 10% chance of the decohering magnetic fields affecting the system, it would create the partially mixed state $\rho = 0.9|\psi_i\rangle\langle\psi_i| + 0.1\rho_{mixed}$—the result of relatively weak decoherence.

Local unitary operations can transform $|H_1\rangle \rightarrow |\alpha\rangle$ and $|H_2\rangle \rightarrow |\beta\rangle$, creating the state

$$|\psi\rangle = \cos\theta|\alpha\rangle|\beta\rangle + \sin\theta e^{i\phi}|\alpha^{\perp}\rangle|\beta^{\perp}\rangle, \tag{4.26}$$

which is the Schmidt decomposition of an arbitrary two-qubit pure state (see Equation 2.61). For information on how to use waveplates to implement arbitrary local unitary operations, see Appendix C.

4.2.2 Revisiting the diagonal representation

Creating arbitrary two-qubit states requires the ability to create mixed as well as pure states. The problem is that for mixed states, an arbitrary number of types of pure states can make up an ensemble of particles. Fortunately, using Equation 2.27, it is possible to write any two-qubit state as the sum of at most four orthogonal pure states:

$$\hat{\rho}^{(2)} = \sum_{i=1}^{4} P_i|\phi_i\rangle\langle\phi_i|. \tag{4.27}$$

This sum is incoherent, which means that any mixture of these four pure states must leave them distinguishable from each other. Using this representation therefore requires the ability to simultaneously create four arbitrary and orthogonal pure states (corresponding to $|\phi_i\rangle$), incoherent with each other and with arbitrarily selected probabilities (corresponding to P_i).

4.2.3 Scheme for totally arbitrary state creation

A complete analysis of possible schemes for two-qubit schemes for state creation can be found in reference [68], and one such scheme will be presented here.

The key to this technique is the observation that a single pump photon with arbitrary polarization maps, after down-conversion, to exactly two of the four terms required for an arbitrary two-qubit pure state (see Figure 4.2a):

$$(\alpha|V\rangle + \beta|H\rangle)_{\text{pump}} \longrightarrow (\alpha|HH\rangle + \beta|VV\rangle)_{\text{dc}}\,. \tag{4.28}$$

By inserting a half waveplate at 22.5° in one arm of this down-conversion, another type of Bell state is created (see Figure 4.2b):

$$(\gamma|V\rangle + \delta|H\rangle)_{\text{pump}} \longrightarrow (\gamma|HV\rangle + \delta|VH\rangle)_{\text{dc}}\,. \tag{4.29}$$

By sending a coherent superposition of these two pump beams through the down-conversion crystals from opposite directions (so that the superposed pump photons are counter-propagating) and adding the aforementioned half waveplate to one of the (now four) down-conversion arms, the two final states from Equations 4.28 and 4.29 can simultaneously be produced, albeit traveling in opposite

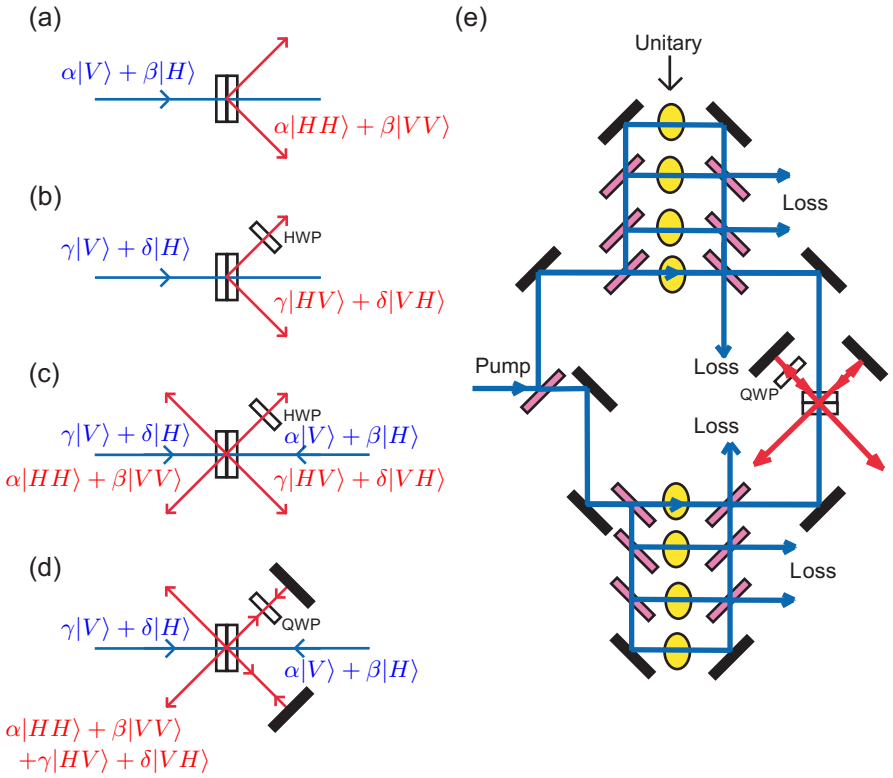

Figure 4.2: Experimental scheme for creating arbitrary two-qubit states, shown through several conceptual stages. (a) A pump beam with polarization $(\alpha|V\rangle + \beta|H\rangle)_{\text{pump}}$ down-converts into the state $(\alpha|HH\rangle + \beta|VV\rangle)_{\text{dc}}$. (b) A half waveplate at 45° is added to one arm of the down-conversion, allowing a pump beam with polarization $(\gamma|V\rangle + \delta|H\rangle)_{\text{pump}}$ to down-convert into the state $(\gamma|HV\rangle + \delta|VH\rangle)_{\text{dc}}$. (c) By counterpropagating these two pump beams through the same set of crystals, with a half waveplate at one of the four outputs the two down-converted states above can simultaneously be produced, although these two output states will be traveling in opposite directions. (d) By adding mirrors to retro-reflect one down-converted pair and replacing the half waveplate with a quarter waveplate, both down-converted states will overlap, producing the state $(\alpha|HH\rangle + \beta|VV\rangle + \gamma|HV\rangle + \delta|VH\rangle)_{\text{dc}}$, an arbitrary two-qubit pure state. (e) By arranging for four different pairs of pump pulses to enter the setup above, each timed so as to be distinguishable, an arbitrary two-qubit mixed state can be produced. (An arbitrary incoherent mixture of four arbitrary pure states is an arbitrary mixed state.) This figure, created by T. C. Wei, first appeared in [68].

directions (see Figure 4.2c).

Next, if one of these down-converted states is retroreflected using a mirror in each arm to pass back through the down-conversion crystals, a superposition of all four canonical basis terms will overlap, creating an arbitrary two-qubit pure state. This actually requires one more change. Assuming the retroreflected state is the state that was passing through the half-waveplate, it will now pass through that waveplate *twice*, negating its effect. That half waveplate must therefore be

replaced with a quarter waveplate; passing through a quarter waveplate twice has the same effect as passing through a half-waveplate once (see Figure 4.2d).

Finally, in order to complete the arbitrary source of two-qubit states, an incoherent mixture of pump light must be generated in order to cause these crystals to create four orthogonal, distinguishable pure states. Coordination of the delays between these states is essential, as the relative delay of the different pump paths completely controls the coherence (and incoherence) of all of the terms in the state. Specifically, as long as the relative pump delays (see Figure 4.2c) are greater than the coherence length of the pump, each of the four created pure states will have a random phase relationship with the other three. Figure 4.2e shows a diagram of the complete experimental setup.

4.2.4 Classes of experimentally created states

While the type of arbitrary state creation scheme detailed in Section 4.2.3 has never been experimentally implemented, several important classes of states have been experimentally reproduced with a very high fidelity. In general, all of these states are created using the methods for state manipulation that were laid out in Section 4.1. Figure 4.3 shows experimentally measured density matrices for each type of state.

Nonmaximally entangled states

The down-conversion source outlined in Chapter 3, Example 3.1, allows the single-qubit pump state

$$|\psi_{\text{pump}}\rangle = \cos(\epsilon)\,|V\rangle + e^{i\phi}\sin(\epsilon)\,|H\rangle \tag{4.30}$$

to down-convert into the two-qubit state

$$|\psi_{\text{dc}}\rangle = \cos(\epsilon)\,|HH\rangle + e^{i\phi}\sin(\epsilon)\,|VV\rangle. \tag{4.31}$$

This state can possess, with an appropriate choice of ϵ, any amount of entanglement, from separable ($\epsilon = 0°$) to nonmaximally entangled ($0° < \epsilon < 45°$) to maximally entangled ($\epsilon = 45°$). These types of states can be reliably produced with fidelities in excess of 99.5% [72, 7].

MEMS

During the early stages of quantum information, it was suspected that the Werner states possessed the maximum entanglement for a given amount of mixture (for measures of both entanglement and mixture, see Appendix A). This was a natural assumption considering that the Werner states are a linear combination of a maximally entangled state and the totally mixed state. However, it was discovered [51] that a different class of states was *more entangled* than

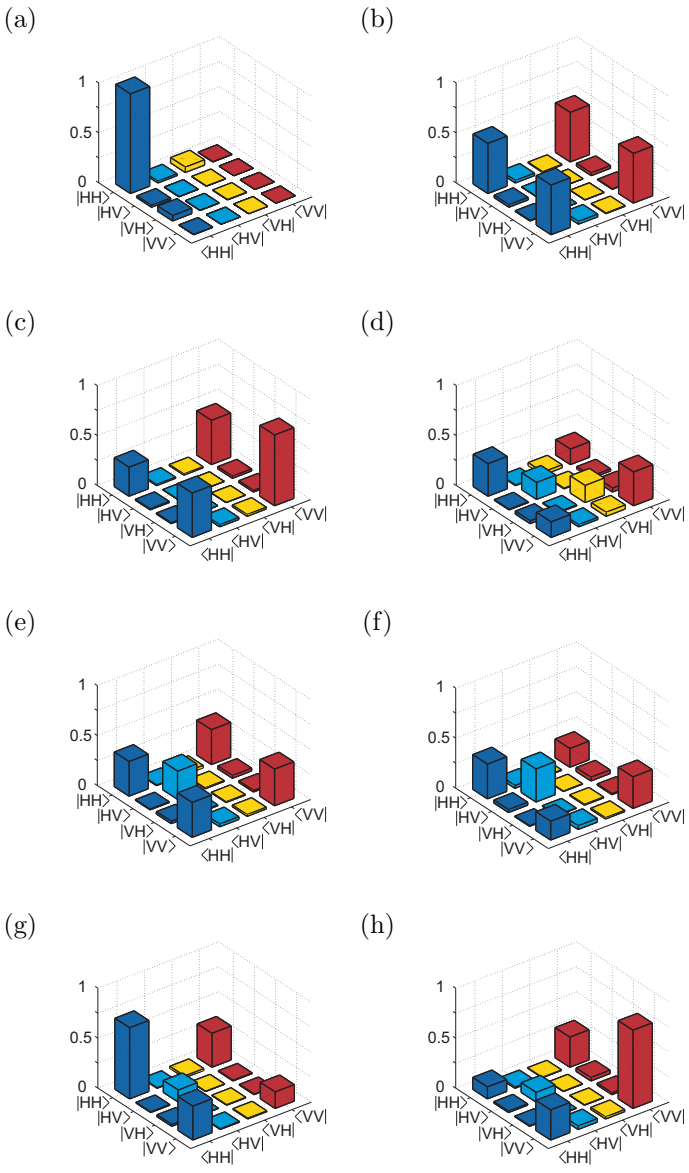

Figure 4.3: Experimentally measured density matrices for several classes of two-qubit states. The absolute values of the density matrices are shown, as the states which differ by off-diagonal phases are in practice equivalent (under local phase shifts). All states have greater than 99% fidelity with their target. For diagrams of the experimental apparatus used to create these states, see Figures 3.2 (nonmaximally entangled states), 4.10 (Werner states and MEMS), and 5.7 (CG states). (a) The state $|HH\rangle$, a separable state. (b) A maximally entangled Bell state. (c) A nonmaximally entangled state, $\cos(\epsilon)|HH\rangle + e^{i\phi}\sin(\epsilon)|VV\rangle$, with $\epsilon = 57°$. (d) A Werner state on the border between entanglement and separability. (e) A maximally entangled mixed state (MEMS) with $r = 0.71$. (f) A MEMS with $r = \frac{2}{3}$. (g) A "Collins-Gisin" state, specifically, $\hat{\rho}_{\mathrm{CG}}(\frac{\pi}{8})$. (h) $\hat{\rho}_{\mathrm{CG}}(\frac{3\pi}{8})$.

the Werner states for the same amount of mixture. These maximally entangled mixed states, or MEMS, are written as

$$\hat{\rho}_{\text{MEMS}} \equiv \begin{array}{ll} \begin{pmatrix} \frac{r}{2} & 0 & 0 & \frac{r}{2} \\ 0 & 1-r & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{r}{2} & 0 & 0 & \frac{r}{2} \end{pmatrix} & \frac{2}{3} \leq r \leq 1 \\ \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{r}{2} \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{r}{2} & 0 & 0 & \frac{1}{3} \end{pmatrix} & 0 \leq r \leq \frac{2}{3} \end{array}, \tag{4.32}$$

where r is a state parameter which can vary from 0 to 1. To see how to create these states, note that by creating the pure state

$$\hat{\rho}_{\text{pure}} = \begin{array}{ll} \begin{pmatrix} \frac{r}{2} & \sqrt{\frac{r}{2}(1-r)} & 0 & \frac{r}{2} \\ \sqrt{\frac{r}{2}(1-r)} & 1-r & 0 & \sqrt{\frac{r}{2}(1-r)} \\ 0 & 0 & 0 & 0 \\ \frac{r}{2} & \sqrt{\frac{r}{2}(1-r)} & 0 & \frac{r}{2} \end{pmatrix} & \frac{2}{3} \leq r \leq 1 \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix} & 0 \leq r \leq \frac{2}{3} \end{array}, \tag{4.33}$$

using the techniques of Section 4.2.1 and applying to it the twin temporal decoherers of Example 4.5, one creates the partially mixed state

$$\hat{\rho}_{\text{partmixed}} = \begin{array}{ll} \begin{pmatrix} \frac{r}{2} & 0 & 0 & \frac{r}{2} \\ 0 & 1-r & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{r}{2} & 0 & 0 & \frac{r}{2} \end{pmatrix} & \frac{2}{3} \leq r \leq 1 \\ \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{pmatrix} & 0 \leq r \leq \frac{2}{3} \end{array}. \tag{4.34}$$

This is almost exactly the state required by the MEMS formula, except that the remaining off-diagonal elements are too large for $r < \frac{2}{3}$. Recall that these did not decohere because the $|HH\rangle$ and $|VV\rangle$ terms were indistinguishable after the application of the twin decoherers, because even after these temporal delays, both photons in each of these states still arrived at the detectors at the same time. If *only one* of the temporal delays were to be slightly increased, these two

states would start to become distinguishable, and the last off-diagonal element would start to shrink, allowing the final element in the density matrix to be tuned to match the form of the arbitrary MEMS state. This procedure was performed in the lab, facilitating the creation of the first experimental MEMS, which had greater than 99% fidelity with their target [57].

Werner states

A Werner state is a linear mixture of a maximally entangled state and the totally mixed state, given by the formula

$$\hat{\rho}_{\text{Werner}} \equiv r|\gamma\rangle\langle\gamma| + (1-r)\hat{\rho}_I, \tag{4.35}$$

where $\hat{\rho}_I$ is the totally mixed state, $|\gamma\rangle$ is a maximally entangled state, and r is a state parameter which varies from 0 to 1.

Werner states were experimentally created [3] using almost exactly the same technique as the MEMS above. First choose an appropriate pure state to match the state diagonal for the target state, then decohere all coherences except one, and then vary the size of that element as desired by changing the temporal delay of a single decoherer. Experimentally, this technique was implemented by creating the pure state

$$\hat{\rho}_{\text{pure}} = \begin{pmatrix} \frac{1}{3} & \sqrt{\frac{1}{3}\times\frac{1}{6}} & \sqrt{\frac{1}{3}\times\frac{1}{6}} & \frac{1}{3} \\ \sqrt{\frac{1}{3}\times\frac{1}{6}} & \frac{1}{6} & \frac{1}{6} & \sqrt{\frac{1}{3}\times\frac{1}{6}} \\ \sqrt{\frac{1}{3}\times\frac{1}{6}} & \frac{1}{6} & \frac{1}{6} & \sqrt{\frac{1}{3}\times\frac{1}{6}} \\ \frac{1}{3} & \sqrt{\frac{1}{3}\times\frac{1}{6}} & \sqrt{\frac{1}{3}\times\frac{1}{6}} & \frac{1}{3} \end{pmatrix}, \tag{4.36}$$

and applying matched temporal decoherers to create the final state

$$\hat{\rho}_{\text{Werner}} = \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{pmatrix}. \tag{4.37}$$

This final state is noteworthy because it exists in Hilbert space on the border between entangled and separable states, and because of its usefulness in implementing ancilla-assisted process tomography (see Section 4.5.2).

Collins-Gisin states

A Bell inequality is one method of testing the entanglement of a state, or more specifically, if that state is forbidden by local realism. Collins-Gisin (CG) states [21] fall on the border between violating and not violating this inequality, and provide an excellent test for methods of characterizing entanglement. The CG

states are given by the equation

$$\begin{aligned} \rho_{CG}(\theta) &= \lambda|\psi(\theta)\rangle\langle\psi(\theta)| + (1-\lambda)|HV\rangle\langle HV|, \\ \text{with} \quad |\psi(\theta)\rangle &= \cos(\theta)|HH\rangle + \sin(\theta)|VV\rangle. \end{aligned} \tag{4.38}$$

For each state $\rho_{CG}(\theta)$, λ is chosen such that the CHSH violation[4] is theoretically predicted to be exactly equal to 0.

These states look very similar to MEMS (from the previous subsection). They have multiple values on the diagonal, but only a single non-zero off-diagonal element in the upper-right half of the density matrix. They can therefore be created just as MEMS were, using unitary transformations to make an arbitrary pure state, followed by twin decoherers to eliminate all the off-diagonal elements except for the corners. Of course this will not be necessary for separable ($\theta = 0, \frac{\pi}{2}$) CG states.

4.3 Decoherence-free subspaces

4.3.1 Theory of decoherence-free subspaces

Decoherence occurs when quantum bits, internal to the quantum computer (or any other quantum system), couple to (become entangled with) external degrees of freedom that are unmeasured. A pure quantum superposition of qubits is thereby transformed into a mixed state. While error-correcting codes [14] or dynamical decoupling [66] act to minimize these effects, it is possible to perform quantum operations in a space fundamentally immune to certain types of decoherence. Specifically, qubits can be embedded in a "decoherence-free subspace" [46], or DFS, in such a way as to be unaffected by collective decoherence or dissipation (energy loss), and robust against noncollective perturbations.

Even for the simplest types of decoherers (see Example 4.8), it is impossible to avoid decoherence on a single qubit without eliminating all external couplings. For two or more qubits, however, it is possible for some states to be unaffected by some types of decoherence. The singlet state, for example, is immune to a number of operations which might otherwise change a state. It has the unique quality of having the same representation in every basis ($\frac{1}{\sqrt{2}}\left(|\zeta\rangle|\zeta^{\perp}\rangle - |\zeta^{\perp}\rangle|\zeta\rangle\right)$, for any $|\zeta\rangle$), making it unaffected by single-qubit operations which are applied to both of its constituent qubits. To see how, consider that any single-qubit unitary operator can be written in the form of Equation 4.6:

$$\hat{U}^{(1)} = e^{i\phi_0}|\psi\rangle\langle\zeta| + e^{i\phi_1}|\psi^{\perp}\rangle\langle\zeta^{\perp}|. \tag{4.39}$$

When applied to each qubit of the singlet state, the result (up to an irrelevant global phase) is simply $\frac{1}{\sqrt{2}}\left(|\psi\rangle|\psi^{\perp}\rangle - |\psi^{\perp}\rangle|\psi\rangle\right)$, an identical singlet state. In

[4] More details on Bell inequalities and the CHSH violation will be given in Chapter 5; for now, we are simply concerned with how to create these states, regardless of the value of λ.

the same way, the singlet state is unchanged by any decohering operation which couples the same ancillary qubit to each of the singlet's qubits in the same way. When a decohering operation acts in the same way for both qubits in a two-qubit state, it is classified as *collective decoherence*.

Note that the state $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_1|\downarrow\rangle_2 + |\downarrow\rangle_1|\uparrow\rangle_2)$ is also immune to decoherence in the $\uparrow,\downarrow$ basis; together $|\psi^-\rangle$ and $|\psi^+\rangle$ form the simplest DFS, a 1-qubit basis immune to collective decoherence restricted to the $|0\rangle, |1\rangle$ basis ($|\uparrow\rangle, |\downarrow\rangle$ for spin, or $|H\rangle, |V\rangle$ for polarization) [37]. Note, however, that while the singlet state is decoherence-free in *every* basis, $|\psi^+\rangle$ will degrade when subjected to decoherence in any basis other than $\uparrow,\downarrow$.

Example 4.9 (A decoherence-free singlet state) *Consider the singlet state*

$$|\psi^-\rangle \equiv \frac{1}{\sqrt{2}}(|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2), \tag{4.40}$$

subject to a random magnetic field that induces arbitrary phases $e^{i\chi_A}$ or $e^{i\chi_B}$ to the $|\downarrow\rangle$ state of both *qubits 1 and 2, the resulting state is:*

$$\begin{aligned}|\psi\rangle_{\text{total}} &= \frac{1}{2}((|\uparrow\rangle_1 e^{i\chi_A}|\downarrow\rangle_2 - e^{i\chi_A}|\downarrow\rangle_1|\uparrow\rangle_2)\otimes|\text{Field}_A\rangle \\ &\quad +(|\uparrow\rangle_1 e^{i\chi_B}|\downarrow\rangle_2 - e^{i\chi_B}|\downarrow\rangle_1|\uparrow\rangle_2)\otimes|\text{Field}_B\rangle) \\ &= |\psi^-\rangle\otimes\frac{1}{\sqrt{2}}(e^{i\chi_A}|\text{Field}_A\rangle + e^{i\chi_B}|\text{Field}_B\rangle).\end{aligned} \tag{4.41}$$

Under the assumption of collective decoherence, i.e., the field acts in the same basis for both qubits, the field has no effect on the singlet state, regardless of the specific phases χ_A and χ_B.

The theory of decoherence-free subspaces for systems larger than two-qubits is beyond the scope of this work, though see reference [10]. The remainder of this section will instead detail the first experimental implementation of a decoherence-free subspace, and an experimental study of the singlet state's immunity to collective decoherence and robust response to noncollective perturbations and dissipation.

4.3.2 Implementing the first experimental DFS

The experimental realization of a two-qubit decoherence-free subspace was performed using a source of entangled photons in the state

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle|H\rangle + |V\rangle|V\rangle), \tag{4.42}$$

of the type described in Chapter 3. The other three Bell states can be generated by applying a simple unitary operation via a half wave plate in one arm. The four Bell states are then used as the input to various decohering or dissipating apparata, shown in Figure 4.4.

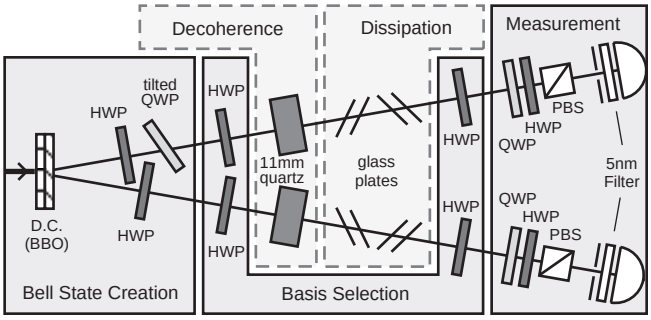

Figure 4.4: A schematic of our experimental setup. Photons in the maximally entangled state $|\psi^+\rangle = (|HH\rangle + |VV\rangle)/\sqrt{2}$ are produced when 45°-polarized pump light is directed through two adjacent nonlinear crystals [39]. Two half waveplates immediately after the crystals are used to interchange between the four Bell states within a phase factor. This phase factor is adjusted by tilting a quarter waveplate. Depending on the experiment, *either* decohering elements *or* dissipative elements are inserted into both paths. The final state of the light is determined by making a series of polarization correlation measurements in various bases, and from these deducing what the density matrix of the output light is, which may then be compared with the input density matrix.

Quantum state tomography [5] allows analysis and characterization of these Bell states both with and without decoherence. State tomography uses a series of correlation measurements (e.g., HH, HV, V45°) to reconstruct the density matrix of the incident state. Each correlation measurement is performed using a polarization analyzer in each arm, consisting of a half waveplate, quarter waveplate, and polarizing beam splitter, which together allow projection into any polarization basis.

As described in Example 4.5, decoherence in our experiment is controllably introduced using a thick (11 mm) piece of birefringent quartz in each arm; the quartz separates the ordinarily and extraordinarily polarized wave packets by approximately 140 wavelengths along the propagation axis, approximately the coherence length of the down-converted light after a 5-nm (FWHM) bandwidth interference filter. Because the H and V components of the light are separated by the coherence length, these become in principle distinguishable with respect to one another. This acts as a label (the frequency and the polarization become entangled) and therefore induces decoherence in the ordinary-extraordinary basis of the quartz crystal. Half waveplates before and after the decoherers (see Figure 4.4) induce rotations which allow each decoherer to effectively operate in any linear polarization basis. For a complete treatment of this type of photon state decoherence, see reference [15].

Collective decoherence

For collective decoherence, these waveplates rotate together[5], ensuring that the two environments always operate in the same basis. Figure 4.5 illustrates the effect of collective decoherence by plotting the fidelity (see Appendix A) between input Bell states and the same states after they pass through the decoherer; as predicted, the singlet state is decoherence-free in every collective linear basis.

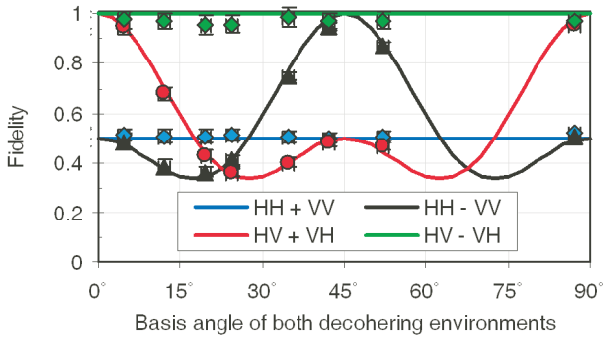

Figure 4.5: Plot showing the effect of collective decoherence on the four Bell states, when the decoherence is applied in a number of different (linear polarization) bases. Solid lines are the theoretical predictions.

Noncollective decoherence

Another predicted [10] benefit of a DFS is that it should be robust against perturbations: when the decoherence has a small noncollective component, the DFS basis states will still be largely decoherence-free. One way to investigate the dependence of a DFS's susceptibility to noncollective effects is to apply the decoherence in each arm in a different basis, using the basis-selection half waveplates (see Figure 4.4). Figure 4.6a shows the fidelity of the output state with the input Bell state when the decohering basis in arm 1 is fixed at 15° while the decohering basis in arm 2 is varied from 0° to 30°. Notice that the fidelity of the DFS state $|\psi^-\rangle$ falls off only quadratically (rather than linearly) with angle, showing that it is robust to perturbations in the normal assumption of collective decoherence.

Figure 4.6b illustrates a second type of noncollective decoherence. By varying the thickness of quartz in one arm, the *strength* of the decoherence can be changed, ranging from no decoherence (no crystal present) to total decoherence (asymptotically approached for an infinitely long crystal). For our experiment the crystal-induced separation between the o and e waves in arm 1 is fixed at 140λ, while the thickness of the crystal in arm 2 is varied. Once again, the DFS

[5]To compensate for frequency anticorrelations in the two daughter photons, the fast axes of the decohering crystals in arms 1 and 2 were oriented at 90° to each other [15].

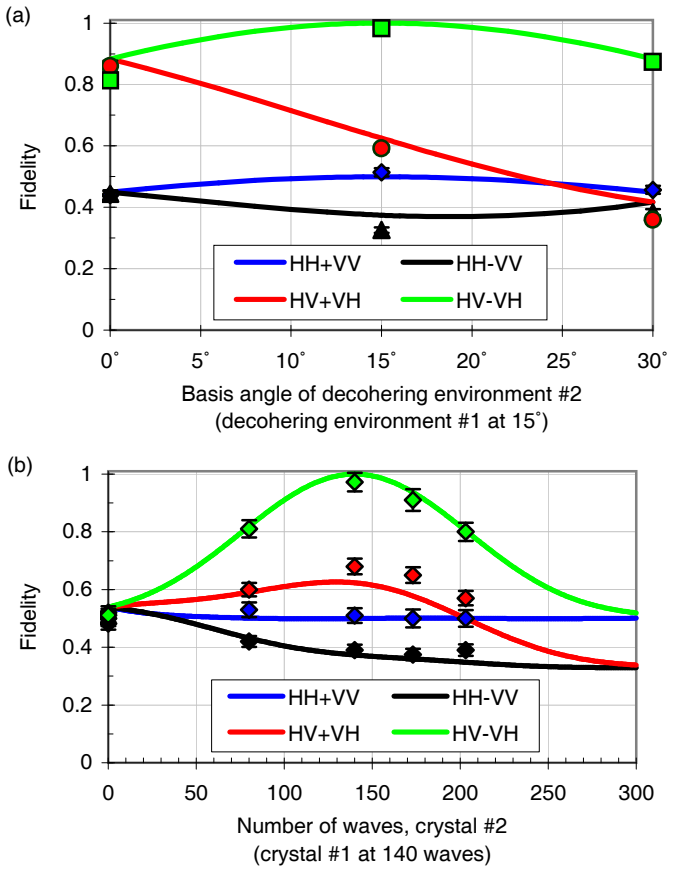

Figure 4.6: Results showing the effect of non-collective decoherence. Solid lines are theoretical predictions. (a) The strength of the decoherence affecting each qubit is the same (corresponding to the thickness of the decohering piece of quartz, measured in terms of the induced longitudinal separation between the o and e wave packets), but the relative bases in which this decoherence occurs is changed: the decoherence basis is fixed at 15° for photon number 1, while the basis is varied from 0° to 30° for photon number 2. (b) The orientation of the bases are now kept fixed (15° in both arms), while the amount of decoherence is varied for photon 2 relative to photon 1. In all cases we see that the singlet state $|\psi^-\rangle$ is robust against perturbations to the assumptions of collective decoherence, falling off quadratically rather than linearly.

state is robust (scaling quadratically) against perturbations in the assumption of collective decoherence.

Tailoring decoherence-free subspaces

The normal DFS state, the singlet state, *does* decohere somewhat under conditions of noncollective decoherence, but a different DFS state exists that can compensate for these conditions. In fact, for every pair of orientations for two equal-strength decohering environments, there exist two special DFS states which are completely decoherence-free under these conditions. For example, the

DFS for the conditions 15° basis in arm 1 and 45° basis in arm 2 is spanned by $|\psi\rangle_{\text{special}} = \frac{1}{\sqrt{2}}(|15°\rangle|45°\rangle \pm |105°\rangle|135°\rangle)$. Both the singlet state and one of these states were subjected to this environment. Figure 4.7 shows their density matrices before and after the noncollective environment and, as expected, the singlet state decoheres while $|\psi\rangle_{\text{special}}$ does not. These results, coupled with the ability to exactly characterize any source of decoherence (via quantum process tomography—see Section 4.5), allow the construction of a DFS optimized for any (static) environment.

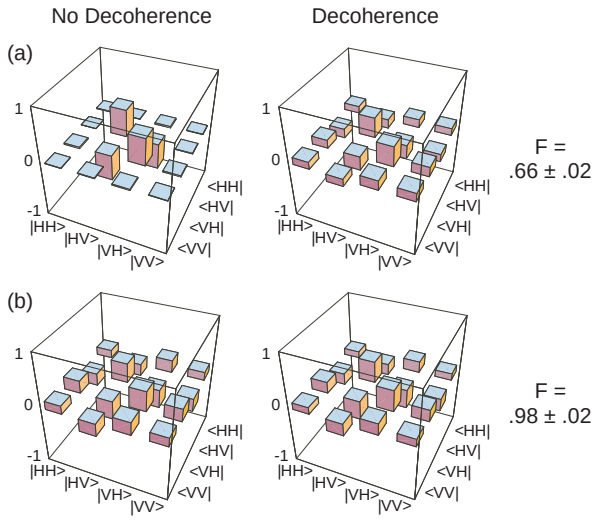

Figure 4.7: Measured density matrix elements demonstrating that DFSs exist even when the environments acting on the two qubits are very different. We show the effect of different decohering elements (15° basis in arm 1 and 45° basis in arm 2, but both the same strength) on both the singlet state and a special state specifically calculated as the DFS for these environments (see text). (a) The singlet state is heavily decohered - the fidelity between the initial and final state is $66 \pm 2\%$. (b) For these environments, a true DFS is shown. The fidelity between the initial and final state is $98 \pm 2\%$.

Dissipation

A problem separate from decoherence is dissipation, whereby entire qubits have some probability of being lost, dissipated into an unmeasured mode. Consider using a state subject to dissipation for quantum cryptography [13]. In addition to requiring additional qubits for the same size key, if the dissipation is basis dependent (e.g., dissipating the $|0\rangle$ state more frequently than the $|1\rangle$ state), Bob has a chance to incorrectly measure the $\frac{(|0\rangle \pm |1\rangle)}{\sqrt{2}}$ states sent by Alice. Dissipation of a single qubit may be characterized by a basis (e.g., $|0\rangle$ and $|1\rangle$) and a ratio (e.g., dissipates twice as much $|0\rangle$ as $|1\rangle$). As for the case of decoherence, dissipation can be collective (affecting each qubit of a multiple-qubit system identically) or noncollective (differing from qubit to qubit in either basis or

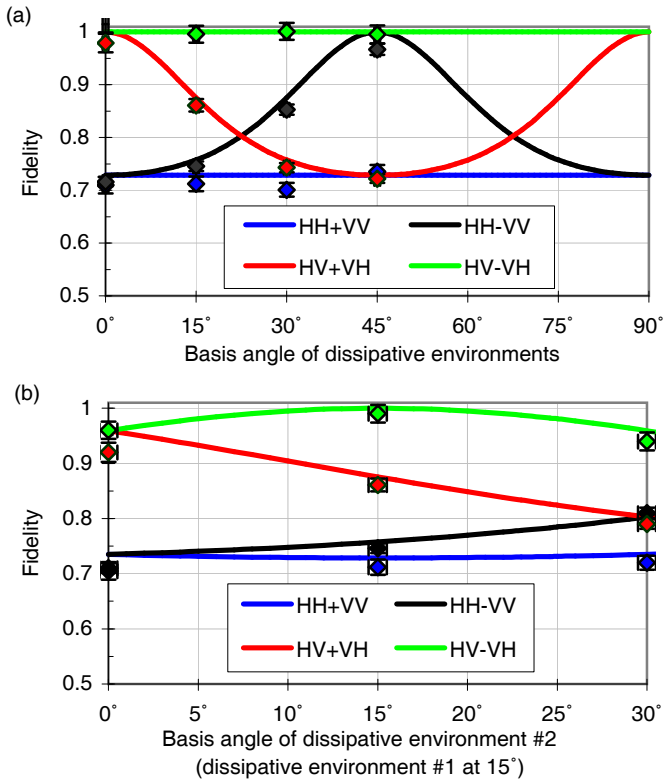

Figure 4.8: Plots showing the effect of *dissipation* on the four Bell states. (a) The effects of collective dissipation. (b) The effects of noncollective dissipation. The dissipating environment in arm 1 is applied at 15° while the environment in arm 2 is rotated between 0° and 30°. Solid lines are theoretical predictions. The singlet state $|\psi^-\rangle$ is robust against both collective and noncollective dissipation.

ratio). A DFS state subject to collective dissipative conditions will sometimes be destroyed, but never measured incorrectly.

Our dissipative environments were experimentally realized using tilted glass plates so that H polarization had different transmission than V ($T_H = 0.86, T_V = 0.21$). As before, we subjected the Bell states to both collective and noncollective environments. In Figure 4.8a the collective dissipation results show that $|\psi^+\rangle$ and $|\psi^-\rangle$ form a dissipation-free subspace (subjecting these states to unbalanced dissipation in the H-V basis causes a net loss, but never results in the states being incorrectly measured). This is to be expected, as a dissipative environment in the H-V basis causes $|H\rangle|V\rangle \pm |V\rangle|H\rangle$ to be measured as $\sqrt{T_H}|H\rangle\sqrt{T_V}|V\rangle \pm \sqrt{T_V}|V\rangle\sqrt{T_H}|H\rangle = \sqrt{T_H}\sqrt{T_V}(|H\rangle|V\rangle \pm |V\rangle|H\rangle)$, i.e., the same state after renormalization. Because the singlet state has the same representation in every basis, it is never affected by collective dissipation. Figure 4.8b shows noncollective dissipation with the environment in arm 1 fixed at 15° and the environment in arm 2 varied from 0° to 30°. As with decoherence, $|\psi^-\rangle$ is robust (scaling quadratically) against perturbations to the assumption of collective dissipation. Though not shown here, a special state can also be

constructed that is completely unperturbed by this type of stable noncollective dissipation.

4.4 Representing quantum processes

4.4.1 Limitations of the Jones calculus

Section 4.1 presented the Jones calculus, which allowed square matrices operating on both pure state vectors and mixed density matrices to represent unitary transformations and projectors. However, we have already seen in Section 4.1.3 that some operations, such as decoherers, cannot be represented using a single matrix. In general, this is because unitary operators in larger Hilbert spaces, when observed in a subspace, appear to act as several operators combined incoherently. This behavior can be modeled using the operator-sum representation [52].

4.4.2 The operator-sum representation

The operator sum representation [52] models the action of any quantum process $\mathcal{E}$ as the incoherent combination of any number of $d \times d$ matrices $\hat{E}_j$, all operating on the input state $\hat{\rho}$. Here d is the dimension of the system $\hat{\rho}$ being subjected to the process $\mathcal{E}$. In this way, the output of the process $\mathcal{E}(\hat{\rho})$ can be given by

$$\mathcal{E}(\hat{\rho}) = \sum_j \hat{E}_j \hat{\rho} \hat{E}_j^\dagger. \tag{4.43}$$

The conditions for what constitutes a legal process are somewhat ambiguous, and depend on how a "legal" process is defined. A common requirement [52] for a process to be physical is that

$$\sum_j \hat{E}_j \hat{E}_j^\dagger = I. \tag{4.44}$$

With this assumption, any process acting on the totally mixed state must yield the totally mixed state. However, for some processes, this is clearly not true. For example, a polarizer will cause a totally mixed state to be measured as a pure state because this type of projection is *lossy*, and fails to measure the entire space of the particles subjected to the process. In contrast, if both outputs of a PBS were measured, then it would indeed leave the totally mixed state unchanged.

The requirement of Equation 4.44, because of these problems, is sometimes relaxed to

$$\sum_j \hat{E}_j \hat{E}_j^\dagger \leq I, \tag{4.45}$$

where the $\leq$ symbol applies to *each element* of the matrices on either side of

the equation. For an appropriate basis choice, all projectors can satisfy this requirement, e.g., lossy processes can be characterized.

The next requirement for physicality is that the process be *completely positive*. This is a stronger version of *positivity*, which requires that a process map all physical input states to physical output states. In the Poincaré sphere picture, positivity requires that every point on the surface or inside the sphere be mapped to another point on the surface or inside the sphere, thereby disallowing the creation of illegal density matrices. Complete positivity is a stronger requirement which states that a process cannot map a state in a larger Hilbert space to an illegal state. Mathematically,

$$(\mathcal{E} \otimes I)\left(\hat{\rho}_{\text{input}}\right) = \hat{\rho}_{\text{legal}}, \tag{4.46}$$

for all legal input states.

Example 4.10 (The universal NOT gate) *The universal NOT gate is a forbidden operation which transforms any pure, single-qubit input state into its orthogonal state. This operation is positive, as it never produces an illegal single-qubit state. However, consider using this operation on one qubit of a two-qubit singlet-state, which exhibits perfect anti-correlation in every basis. After the universal NOT gate, the output state would necessarily exhibit perfect* correlation *in every basis. But there* is *no legal two-qubit state which exhibits perfect correlation in every basis; therefore, the universal NOT gate is not completely positive, and not a physical operation [19].*

4.4.3 The $\hat{\chi}$-matrix representation

The operator-sum representation, while complete, is somewhat bulky and difficult to use in practice. It is also reminiscent of the representation of mixed states as an incoherent sum of a potentially infinite list of types of pure states. Much like mixed states were concisely represented by density matrices (and those density matrices' diagonalization), there is a similar solution for quantum processes. First relate the operation elements $\hat{E}_j$ to a fixed set of operators, $\{\widetilde{E}_m\}$, where $\hat{E}_j = \sum_m e_{jm}\widetilde{E}_m$ and e_{jm} can be complex. This allows us to define a *single* matrix, $\hat{\chi}$, that fully characterizes the process: if we rewrite Equation 4.43 as

$$\mathcal{E}(\hat{\rho}) = \sum_{mn} \widetilde{E}_m \hat{\rho} \widetilde{E}_n^{\dagger} \hat{\chi}_{mn}, \tag{4.47}$$

then $\hat{\chi}$ is a positive, $d^2 \times d^2$, Hermitian matrix [52].

In general, writing down a $\hat{\chi}$ matrix requires choosing an operator basis, $\{\widetilde{E}_m\}$, just as writing down a density matrix requires choosing a basis for state vectors. We will consistently use the $\hat{\sigma}_i$ matrices as this operator basis throughout this work.

Example 4.11 ($\hat{\chi}$ matrix for a decoherer in the H,V basis) *Using Equation 4.17, we can write the operation of a decoherer in the H-V basis as*

$$\mathcal{E}_{\text{dec}}(\hat{\rho}) = |H\rangle\langle H|\hat{\rho}|H\rangle\langle H| + |V\rangle\langle V|\hat{\rho}|V\rangle\langle V|. \tag{4.48}$$

Rewriting $|H\rangle$ and $|V\rangle$ in terms of the $\hat{\sigma}$ matrices:

$$\begin{aligned}\mathcal{E}_{\text{dec}}(\hat{\rho}) &= \frac{1}{4}\left((\hat{\sigma}_0+\hat{\sigma}_3)\,\hat{\rho}\left(\hat{\sigma}_0^\dagger+\hat{\sigma}_3^\dagger\right)+(\hat{\sigma}_0-\hat{\sigma}_3)\,\hat{\rho}\left(\hat{\sigma}_0^\dagger-\hat{\sigma}_3^\dagger\right)\right)\\ &= \frac{1}{2}\left(\hat{\sigma}_0\hat{\rho}\hat{\sigma}_0^\dagger+\hat{\sigma}_3\hat{\rho}\hat{\sigma}_3^\dagger\right).\end{aligned} \tag{4.49}$$

The two terms in Equation 4.49 are the only nonzero elements in this operation's $\hat{\chi}$ matrix:

$$\hat{\chi}_{\text{dec}} = \frac{1}{2}\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}. \tag{4.50}$$

4.4.4 A geometric representation

Single-qubit processes can be visualized using the Poincaré sphere. These processes map each point on the sphere to another point, but because nearby points must be mapped to nearby points, this mapping can be modeled by showing a rotated and distended sphere inside the original state space. Unitary operations are represented as rotations of the sphere, decoherers collapse the sphere partially or completely to a spindle passing through the center of the sphere, and polarizers map all points on the sphere to a single point. This graphical approach can even be applied to lossy processes, e.g., partial polarizers, though it is important to note that it does not indicate the amount of loss, only the quantum state of the surviving qubits. Several of these sample processes are depicted in Figure 4.9.

4.5 Measuring quantum processes

The measurement of quantum processes shares a great deal in common with the measurement of quantum states, although state measurement will not be covered in detail until Chapter 5. Here three different methods for measuring quantum processes will be discussed: standard quantum process tomography (SQPT), entanglement-assisted quantum process tomography (EAPT), and ancilla-assisted quantum process tomography (AAPT). When compared to the state measurement techniques presented in the next chapter, the methods we used to reconstruct quantum processes are quite primitive, corresponding to only a linear reconstruction rather than a maximum-likelihood technique. This is partially because far more work has been spent improving state tomography,

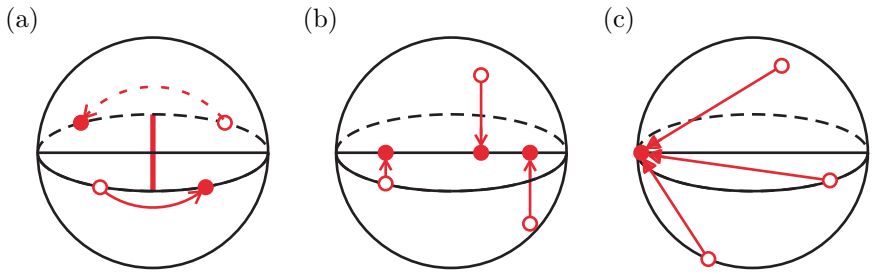

Figure 4.9: Graphical depictions of single-qubit quantum processes. (a) On the Poincaré sphere, a unitary transformation acts as a rotation about an axis defined by its eigenvectors. (b) A decoherer collapses all points on the sphere towards a spindle defined by its eigenvectors. For the partial decoherer shown, this collapse is not completed and the pure states are mapped from a sphere to an oblong ellipsoid. (c) Polarizers map all points on or inside the sphere to a single point on the sphere's surface.

but also because there are significant subtleties associated with implementing maximum-likelihood process tomography; see, however, the attempts of [55].

4.5.1 Standard quantum process tomography

In SQPT, a quantum system A experiences an unknown quantum process $\mathcal{E}$. To determine $\mathcal{E}$ we first choose some fixed set of states $\{\hat{\rho}_j\}$ which form a basis for the set of operators acting on the state space of system A, e.g., $\{\hat{\rho}_j\} = \{\hat{\rho}_H, \hat{\rho}_V, \hat{\rho}_D, \hat{\rho}_R\}$ for a polarization qubit. Each state $\hat{\rho}_j$ is then subject to the process $\mathcal{E}$, and quantum state tomography (see Chapter 5) is used to experimentally determine the output $\mathcal{E}(\hat{\rho}_j)$. Because an arbitrary state $\hat{\rho}_{\text{input}}$ is a linear combination of the set of $\{\hat{\rho}_j\}$, characterizing $\mathcal{E}(\hat{\rho})$ will completely characterize $\mathcal{E}$:

$$\begin{aligned} \hat{\rho}_{\text{input}} &= \sum_j a_j \hat{\rho}_j \\ \hat{\rho}_{\text{output}} &= \mathcal{E}(\hat{\rho}) = \sum_j a_j \mathcal{E}(\hat{\rho}_j). \end{aligned} \tag{4.51}$$

This same method will determine $\hat{\chi}$. We define the output of our measurement of $\mathcal{E}(\hat{\rho}_j)$ as:

$$\mathcal{E}(\hat{\rho}_j) = \sum_k c_{jk} \hat{\rho}_k, \tag{4.52}$$

where the set $\hat{\rho}_k$ is an arbitrary and linearly independent basis for the output states. If we define, in terms of the same output basis,

$$\widetilde{E}_m \hat{\rho}_j \widetilde{E}_n^\dagger = \sum_k \beta_{jk}^{mn} \hat{\rho}_k, \tag{4.53}$$

where β_{jk}^{mn} is another complex number matrix which we determine from our

choice of input basis states $\{\hat{\rho}_j\}$, output basis states $\{\hat{\rho}_k\}$, and operators $\{\widetilde{E}_j\}$, we can see that

$$\sum_k \sum_{mn} \hat{\chi}_{mn} \beta_{jk}^{mn} \hat{\rho}_k = \sum_k c_{jk} \hat{\rho}_k, \tag{4.54}$$

independent of $\hat{\rho}_k$. If we have linearly independent basis sets ($\{\hat{\rho}_j\}$, $\{\hat{\rho}_k\}$, and $\{\widetilde{E}_j\}$), then β is invertible and

$$\hat{\chi}_{mn} = \sum_{jk} \left(\beta^{-1}\right)_{jk}^{mn} c_{jk}. \tag{4.55}$$

4.5.2 Ancilla-assisted process tomography

In ancilla-assisted process tomography (AAPT) [3] the process $\mathcal{E}$ is characterized by preparing a *single state*, $\hat{\sigma}$, and then measuring $(\mathcal{E} \otimes \mathcal{I})(\hat{\sigma})$. Entanglement-assisted process tomography (EAPT) [48], developed first, is in fact a specific type of AAPT, wherein the input states $\hat{\sigma}$ are always maximally entangled. Standard process tomography relies on a linearly independent *set* of input states. EAPT is very similar, relying on the perfect correlations present between particles A and B in maximally entangled states: if one half of a maximally entangled state (particle B) is measured, then the state of its partner (particle A) is immediately known [58]. By making a complete set of these measurements, we are able to prepare particles of type A into a complete set of input states, which we then subject to the unknown process $\mathcal{E}$. Then, just as in SQPT, the process can be completely characterized.

Ancilla-assisted process tomography is the generalization of this process to states which have a high degree of correlation, but which are not necessarily entangled. This requires an *ancilla system*, B, with Hilbert space dimension at least as great as that of A. For an appropriate initial state, it is possible to characterize $\mathcal{E}$ by preparing the state $\hat{\sigma}$, performing the process $\mathcal{E}$ on system A—leaving system B completely isolated—and taking a tomography of the output $(\mathcal{E} \otimes \mathcal{I})(\hat{\sigma})$. The total number of measurements is the same in AAPT (16 measurements on a single 2-qubit state) as in SQPT (four measurements on each of four input states).

AAPT has potential advantages over SQPT, most notably being that preparation of only a single quantum state is necessary for its operation. Consider the possibility of using it as a diagnostic tool in a quantum computer. When an unknown effect acts on less than half of a system of qubits, knowledge of the larger state before and after the change is sufficient to exactly predict the effect this change will have on every other state. (Assuming that the larger state is usable for AAPT — see below). Alternatively, SQPT has the advantage that it is generally easier to produce and measure states with fewer qubits.

States usable for AAPT

In order to determine the class of states suitable for AAPT [6], it is necessary to introduce an operator generalization of the Schmidt decomposition for entangled states [52]. First, consider an inner product on operators, $(M, N) \equiv \mathrm{tr}(M^\dagger N)$, and define an orthonormal operator basis to be a set of operators $\{M_j\}$ such that $(M_j, M_k) = \mathrm{tr}(M_j^\dagger M_k) = \delta_{jk}$. (For example, an orthonormal basis for single-qubit operators is the set $\{I/\sqrt{2}, \hat{\sigma}_x/\sqrt{2}, \hat{\sigma}_y/\sqrt{2}, \hat{\sigma}_z/\sqrt{2}\}$). The operator-Schmidt decomposition [54] states that an operator M acting on AB can be decomposed as

$$M = \sum_l s_l A_l \otimes B_l, \tag{4.56}$$

where the s_l are non-negative real numbers, and the sets $\{A_l\}$ and $\{B_l\}$ form orthonormal operator bases for systems A and B, respectively [53]. The *Schmidt number* $\mathrm{Sch}(M)$ of an operator M is defined [54] as the number of non-zero terms in the Schmidt decomposition.

A state $\hat{\sigma}$ of AB may be used to perform AAPT *if and only if* the Schmidt number of $\hat{\sigma}$ is d_A^2, where d_A is the dimension of the state space of system A. Consider that in order to measure the mapping of the entire space, the input state must possess correlations - represented by the Schmidt number - between enough states to form a basis for the mapping. To prove this [3], expand $\hat{\sigma}$ in its Schmidt decomposition as $\hat{\sigma} = \sum_l s_l A_l \otimes B_l$. Assume $\hat{\sigma}$ has Schmidt number d_A^2, so that the A_l form an orthonormal operator basis, and $s_l > 0$ for all l. Let $\hat{\sigma}'$ be the output obtained after letting $\mathcal{E}$ act on system A, that is, $\hat{\sigma}' = (\mathcal{E} \otimes \mathcal{I})(\hat{\sigma}) = \sum_l s_l \mathcal{E}(A_l) \otimes B_l$. By the orthonormality of the B_l and the previous equation it follows that

$$\mathrm{tr}_B((I \otimes B_m^\dagger)\hat{\sigma}') = \sum_l s_l \mathcal{E}(A_l)\mathrm{tr}(B_m^\dagger B_l) = s_m \mathcal{E}(A_m), \tag{4.57}$$

and so

$$\mathcal{E}(A_m) = \mathrm{tr}_B((I \otimes B_m^\dagger)\hat{\sigma}')/s_m. \tag{4.58}$$

The fact that the Schmidt number of $\hat{\sigma}$ is d_A^2 ensures that $s_m > 0$, so there is no problem with division by zero. By performing state tomography on $\hat{\sigma}'$ and applying the above equation, it is possible to determine the action of $\mathcal{E}$.

Another proof of this result can be obtained by investigating the mapping $\mathcal{E}$. Let E_A be the space of trace-preserving quantum operations on system A, and let S_{AB} be the space of quantum states on system AB. Define a map

$$f : E_A \to S_{AB} \qquad \text{by} \qquad f(\mathcal{E}) \equiv (\mathcal{E} \otimes \mathcal{I})(\hat{\sigma}). \tag{4.59}$$

For AAPT, we require that f be a one-to-one map, i.e., there are never two

[6] The theory in this section was developed by Michael Nielsen's group at the University of Queensland, Australia, and was first presented in [3]. It is repeated here in an expanded form.

distinct operations such that $f(\mathcal{E}_1) = f(\mathcal{E}_2)$. A parameter counting argument shows that f cannot be one-to-one when $\hat{\sigma}$ has Schmidt number less than d_A^2. The dimensionality of the manifold E_A is $d_A^4 - d_A^2$. Since $f(\mathcal{E}) = \sum_l s_l \mathcal{E}(A_l) \otimes B_l$, the dimension of the image manifold $f(E_A)$ is at most $\mathrm{Sch}(M) \times (d_A^2 - 1)$, because the map $\mathcal{E} \to \mathcal{E}(A_l)$ has image of dimension at most $d_A^2 - 1$. Thus, for AAPT we require that

$$\mathrm{Sch}(M) \times (d_A^2 - 1) \geq d_A^4 - d_A^2, \tag{4.60}$$

which is only possible when $\mathrm{Sch}(M) = d_A^2$.

Note that AAPT is possible only when the dimension of system B is at least as great as the dimension of system A. When this is true, almost all states of system AB may be used for AAPT, because the set of states with Schmidt number less than d_A^2 has measure zero. That is, a maximally entangled input is not required for AAPT — indeed many of the viable input states are not entangled at all, as demonstrated below by our Werner state AAPT. However, while almost any state *can* be used for AAPT, maximally entangled states appear to be experimentally optimal in that they have perfect non-local correlations. Figure 4.11 highlights this difference, as the AAPT results have significantly greater statistical errors than the EAPT (both were from identical measurement runs).

4.5.3 Experimental results

We investigated a variety of processes, using the three methods of SQPT, EAPT, and non-entangled AAPT. Our processes operate on the polarization state of a single photon. We used spontaneous parametric down-conversion (of a 351-nm pump beam) in a nonlinear crystal (BBO) to create pairs of time-correlated photons at 702 nm. For SQPT, by triggering on one photon, the other was prepared into a single-photon state [31] with H polarization (Figure 4.10). Half and quarter waveplates converted the horizontal polarization into an arbitrary state, thus allowing preparation of the necessary input states $\hat{\rho}_H, \hat{\rho}_V, \hat{\rho}_D$, and $\hat{\rho}_R$. The tomography of the post-process states was performed by measuring (in coincidence with the trigger detector) the Stokes parameters $S_1 = P_H - P_V$, $S_2 = P_D - P_A$, and $S_3 = P_R - P_L$, and performing a maximum-likelihood estimation of the density matrix (see Chapter 5). Typical measurements yielded a maximum of 13,000 photon counts over 30 seconds.

For our EAPT results, two adjacent BBO crystals were used to prepare the maximally entangled state $|\phi^-\rangle = (|HH\rangle - |VV\rangle)/\sqrt{2})$. One of the resulting qubits was subjected to the given process, and two-qubit tomography of the pair was then performed by measuring the polarization correlations of the photons with 16 measurements, e.g., in the following bases: HH, HV, HD, HR, VH, VV, etc. (see Chapter 5). Note from Figure 4.10 that the elements used in SQPT to prepare the single-photon state are now placed (in reverse order) in the other detection arm, highlighting the symmetry of the two techniques.

We also performed AAPT using the non-entangled Werner state $\hat{\rho}_W = \frac{1}{6} I +$

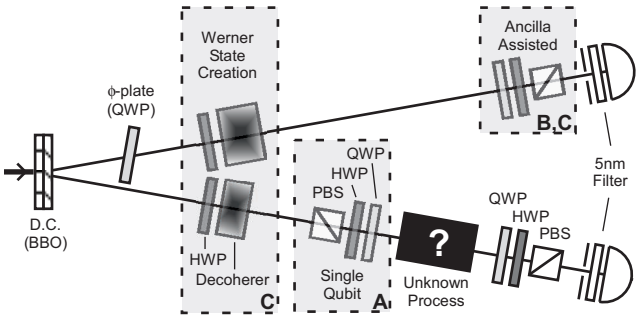

Figure 4.10: Experimental arrangements for quantum process tomography. A 351-nm pump is directed through two 0.6 mm-thick BBO crystals, giving rise to pairs of correlated 702 nm photons, which are detected using Si avalanche photodiodes and fast coincidence electronics. A, B, and C above denote which elements are present for SQPT, EAPT, and non-entangled AAPT, respectively. (a) SQPT: Polarizer (P), half waveplate (HWP) and quarter waveplate (QWP) allow preparation of required pure single-photon (conditioned on "trigger" detection) states; identical elements allow tomography of the post-process states. (b) EAPT: The source produces the maximally entangled state $(|HH\rangle - |VV\rangle)/\sqrt{2}$. A two-photon tomography of the output allows reconstruction of the process. (c) AAPT: The source produces $\hat{\rho}_W \sim \frac{1}{6}I + \frac{1}{3}|\gamma\rangle\langle\gamma|$, where $|\gamma\rangle$ is a maximally entangled state. Although there is no entanglement, the correlations in $\hat{\rho}_W$ allow AAPT.

$\frac{1}{3}|\gamma\rangle\langle\gamma|$, where $|\gamma\rangle$ is a maximally entangled state. See Section 4.2.4 for the details of Werner state creation.

We investigated several processes, including the identity, a unitary rotation, a decoherer, and both a coherent and an incoherent partial polarizer (see below). Data for these processes is shown in graphical form in Figure 4.11, and as $\hat{\chi}$ matrices in Figure 4.12. The results for the identity process measure how well the input state(s) are preserved. We used SQPT, EAPT, and AAPT to measure the same unitary rotation process (a birefringent waveplate). The results were in close agreement (Figure 4.11b); the resulting χ matrices had an average process fidelity (see Appendix A) between the three methods, of $\mathcal{F} = 100.4 \pm .8\%$. Likewise, the SQPT and EAPT measurements of a decohering process (implemented with a 6.3-mm piece of quartz) yielded $\mathcal{F} = 99.9 \pm .3\%$ (Figure 4.11c). The same process, when measured using our Werner State, appears to be a *recoherer* — a process which is not a positive map.

This Werner state was prepared using a thick piece of quartz to temporally separate the H and V components of the light, introducing decoherence. Consider adding another piece of quartz, with optic axis perpendicular to the first, after the original. This also temporally shifts the H and V components of the light, but in the opposite direction, undoing the original decoherence. Our decohering process does exactly this, effectively *recohering* the Werner State — impossible for a 1-qubit process. The resolution to this paradox lies in an unwritten assumption that that the measured process does not act on any degrees

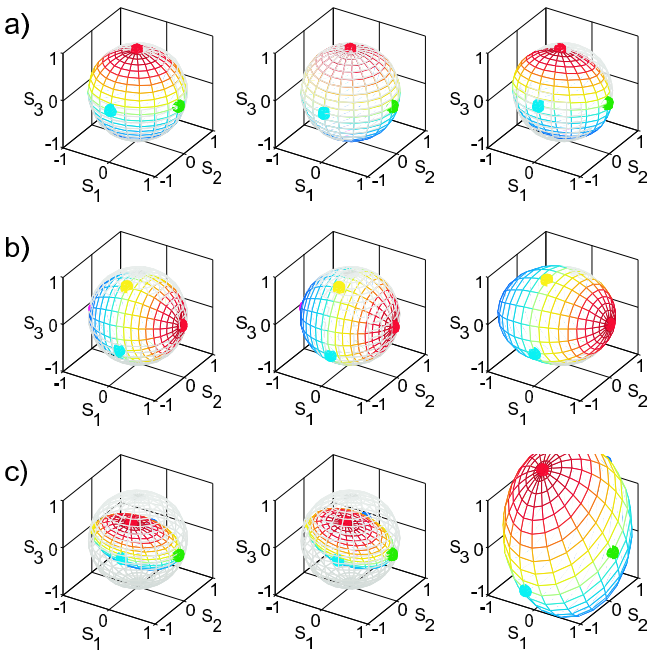

Figure 4.11: Geometric mappings for three quantum processes – (a) identity , (b) unitary transformation, and (c) decoherence – measured using SQPT (left), EAPT (center), and AAPT (right). The axes are the Stokes parameters (S_1, S_2, S_3). The colored mesh surfaces show how all pure states are transformed by the process. The initial states H, R, V, and A are shown by the green, red, yellow, and blue dots, respectively. The transformation of initial *mixed* states (inside the surface) may be interpolated from the pure state results using the linearity of quantum mechanics. The mesh coloring denotes the orientation of the transformed sphere.

of freedom used to prepare the input state other than the tested qubit. For example, if frequency is traced over to prepare a mixed input state, a process that couples to frequency cannot be measured.

Coherent and incoherent partial polarizers were measured in order to highlight the role coherence plays in lossy processes. A glass plate at Brewster's angle to an incident beam is a coherent partial polarizer, as the operation of the plate maintains the pre-existing phase relationship between the horizontal component of the light (completely transmitted) and the vertical component of the light (partially reflected). For the incoherent case, consider inserting a horizontal polarizer into the beam 50% of the time. Half the time only the horizontal component of the light will be transmitted, but more importantly, the transmitted light will have no coherence relationship with the light that does not pass through the polarizer. For the coherent partial polarizer, pure states remain pure but slide toward H along the surface of the sphere. In the incoherent case pure states travel linearly through the sphere to H, becoming mixed (Figure 4.13).

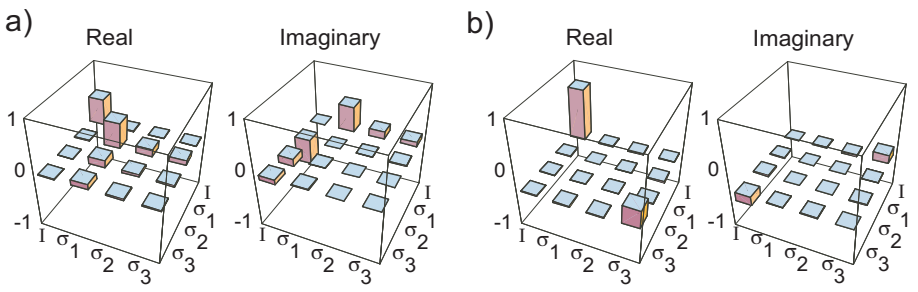

Figure 4.12: $\hat{\chi}$-matrices determined from EAPT for (a) unitary and (b) decohering processes, as shown in Figure 4.11.

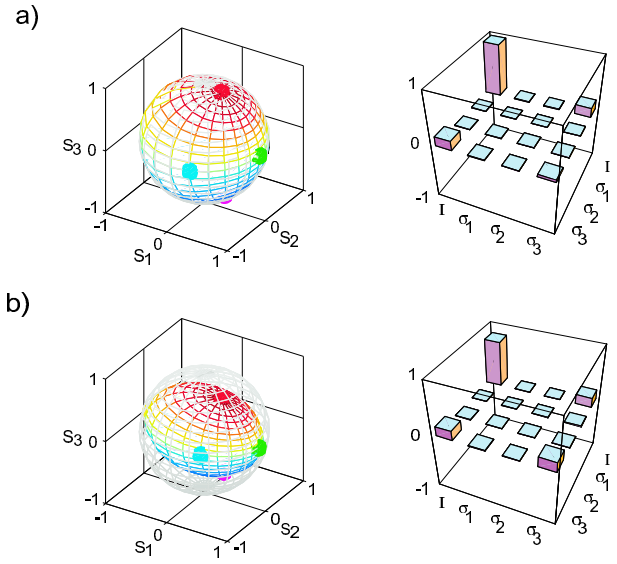

Figure 4.13: Geometric mappings and $\hat{\chi}$ matrices for (a) coherent and (b) incoherent partially polarizing processes. The former was implemented using two glass microscope slides near Brewster's angle [$T_H \sim 88\%, T_V \sim 45\%$]. The latter was simulated by inserting a horizontal polarizer 50% of the time. (Real components shown; imaginary contributions $< 1\%$.)

Notice that for all of the measured results, the data from AAPT appears to be slightly inconsistent with the data from SQPT and EAPT. Because it has imperfect correlations between its constituent particles, it requires additional measurements to receive the same accuracy. (The errors are most evident when looking at the AAPT measurement of the identity, which looks skewed compared to the other two methods, and more importantly when compared to the ideal, unchanged sphere.) If EAPT uses an ensemble of size N, then AAPT must use an ensemble of size $3N$ to achieve results of the same precision. For this reason, it appears that EAPT or SQPT are the optimal choices for process tomography.

5 State Measurement

Every experiment which utilizes quantum states requires a method to measure those states, but a plethora of techniques and protocols for state measurement exist. For tests of nonlocality, which are the focus of this thesis, there are three options which are particularly intriguing: Bell's inequalities (fundamental tests of local realism), entanglement witnesses (quick measurements which give incomplete information), and complete state tomography (a complete characterization of a quantum state which—the optimal choice for most experiments). This chapter will describe the theory and experiment of each of these measurement protocols in detail.

Each of these protocols can be realized through sets of separable, projective measurements on the ensemble of states to be tested. Section 5.1 describes the experimental details of making these measurements, and more importantly characterizing and compensating for systematic errors in those measurements. The next sections detail each of the measurement protocols in detail, both theoretically and experimentally. Each protocol has its own advantages, and this chapter will end with a quantitative comparison of each type of test of nonlocality, and a discussion of error analysis.

5.1 Collecting tomographic measurements

Before discussing the specifics of each measurement protocol, it is necessary to understand how that experimental data is collected. This chapter outlines the experimental implementation of measurements on polarization-entangled qubits generated from spontaneous parametric down-conversion [39]. We filter these photon pairs using both spatial filters (irises used to isolate a small range of $\vec{k}$-vector, necessary because our states are angle-dependent) and frequency filters (interference filters, typically 5–25 nm wide, FWHM).

After this initial filtering, measurement collection involves two central issues: projection (into an ideally arbitrary range of states) and systematic error correction (to compensate for several experimental problems ranging from imperfect optics to accidental coincidences).

5.1.1 Projection

Any measurement on a quantum system depends on state projection. While tomography could be simplified by using arbitrary projectors (e.g., joint mea-

surements on two qubits), this is experimentally difficult. Therefore, for the purposes of this chapter, these projections will be *separable*. In particular, we focus on the ability to create arbitrary single-qubit projectors which we can then easily chain together to create any separable projector.

Arbitrary single-qubit projection

The following analysis, while framed in terms of waveplates acting on photon polarization, is directly applicable to other systems, e.g., spin-$\frac{1}{2}$ particles [23, 35, 69, 40] or two-level atoms [50, 61]. In these systems, measurements in arbitrary bases are obtained using suitably phased π- and $\frac{\pi}{2}$-pulses (externally applied electromagnetic fields) to rotate the state to be measured into the desired analysis basis.

An arbitrary polarization measurement and its orthogonal complement can be realized using, in order, a quarter-waveplate, a half-waveplate, and a polarizing beam splitter. Recall from Example 4.2 that a waveplate whose optic axis is oriented at angle θ with respect to the horizontal induces a rotation on the Poincaré sphere about an axis 2θ from horizontal, in the linear plane, and that the magnitude of this rotation is equal to the waveplate's retardance (90° for quarter-waveplates and 180° for half-waveplates). Additionally, for the remainder of this chapter we adopt the convention that polarizing beam splitters transmit horizontally polarized light and reflect vertically polarized light—though for some types the roles are reversed.

To derive the settings for these waveplates as a function of the projection state desired, we use the Poincaré sphere (see Figure 5.1). For any state on the surface of the sphere, a 90° rotation about a linear axis directly below it will rotate that state into a linear polarization (see Figure 5.1b). Assume the desired projection state is

$$|\psi_P\rangle = \cos\left(\frac{\theta}{2}\right)|H\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|V\rangle. \tag{5.1}$$

Simple coordinate transforms from spherical to Cartesian coordinates reveal that a quarter-waveplate at $\theta_{QWP} = \frac{1}{2}\text{acos}\{\sin(\theta)\tan(\phi)\}$ will rotate the projection state 5.1 into a linear state

$$|\psi'_P\rangle = \cos\left(\frac{\theta'}{2}\right)|H\rangle + \sin\left(\frac{\theta'}{2}\right)|V\rangle. \tag{5.2}$$

A half-waveplate at $\frac{1}{4}\theta'$ (with respect to horizontal orientation) will then rotate this state to $|H\rangle$.[1] Finally, the PBS will transmit the projected state and reflect its orthogonal complement.

Mathematically, this process of rotation and projection can be described us-

[1] $\theta' = \text{acos}\{\sin(\theta)\tan(\phi)\} - \text{acos}\{\cot(\theta)\cot(\phi)\}$. In practice, care must be taken that consistent conventions are used (e.g., right- vs. left-circular polarization), and it may be easier to calculate this angle directly from waveplate operators and the initial state.

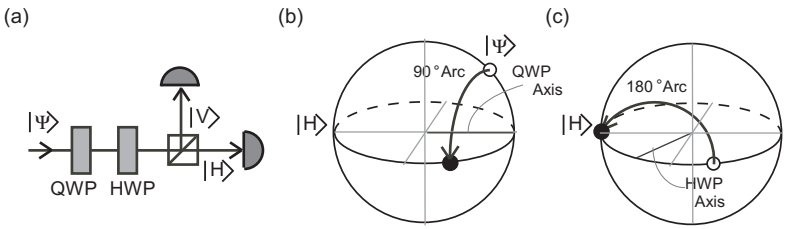

Figure 5.1: (a) A quarter-waveplate (QWP), half-waveplate (HWP), and polarizing beam splitter (PBS) are used to make an arbitrary polarization measurement. Next, the step-by-step evolution of the state on the Poincaré sphere is shown. (b) The quarter-waveplate rotates the projection state (the state we are projecting into, *not* the incoming unknown state) into the linear polarization plane (the equator). (c) The half-waveplate rotates this linear state to horizontal. The PBS transmits the projection state (now $|H\rangle$) and reflects its orthogonal complement (now $|V\rangle$), which can then both be measured.

ing unitary transformations. The unitary transformations for half- and quarter-waveplates in the H/V basis are

$$\begin{aligned} U_{HWP}(\theta) &= \begin{bmatrix} \cos^2(\theta) - \sin^2(\theta) & 2\cos(\theta)\sin(\theta) \\ 2\cos(\theta)\sin(\theta) & \sin^2(\theta) - \cos^2(\theta) \end{bmatrix}, \\ U_{QWP}(\theta) &= \begin{bmatrix} \cos^2(\theta) + i\sin^2(\theta) & (1-i)\cos(\theta)\sin(\theta) \\ (1-i)\cos(\theta)\sin(\theta) & \sin^2(\theta) + i\cos^2(\theta) \end{bmatrix}, \end{aligned} \tag{5.3}$$

with θ denoting the orientation angle of the waveplate optic axis with respect to horizontal. Assume that during the course of a tomography, the ν^{th} measurement setting requires that the QWP be set to $\theta_{\text{QWP},\nu}$ and the HWP to $\theta_{\text{HWP},\nu}$. Therefore, the total unitary[2] for the ν^{th} measurement setting will be

$$U_\nu = U_{\text{HWP}}(\theta_{\text{HWP},\nu}) U_{\text{QWP}}(\theta_{\text{QWP},\nu}). \tag{5.4}$$

For multiple qubits, we can directly combine these unitaries such that

$$U_\nu = {}^1U_\nu \otimes {}^2U_\nu \otimes \ldots \otimes {}^nU_\nu, \tag{5.5}$$

where ${}^qU_\nu$ denotes the q^{th} qubit's unitary transform due to waveplates. The total projection operator for this system is therefore $\langle 0|U_\nu$, where $|0\rangle$ is the first computational basis state (the state which passes through the beam splitters, which here we assume is $|H\rangle$). The measurement state (the state which will pass through the measurement apparatus and be measured every time) is therefore $U_\nu^\dagger|0\rangle$.

Of course, these calculations assume that we are using waveplates with retardances equal to exactly π or $\frac{\pi}{2}$ (or Rabi pulses producing perfect phase differences). Imperfect yet well characterized waveplates will lead to measurements in

[2]Note the order of the unitary matrices for the HWP and QWP. Incoming light encounters the QWP first, and therefore U_{QWP} is last when defining U_ν.

slightly different, yet known, bases. This can still yield an accurate tomography, but first these results must be transformed from a biased basis into the canonical Stokes parameters using Equation 2.22. As discussed below (see Section 5.5.2), the maximum-likelihood technique provides a different but equally effective way to accommodate for imperfect measurements. As a separate alternative, it is possible in some cases to use imperfect waveplates to make ideal measurements. For this and other experimental details concerning waveplates, see Appendix C.

Multiple-qubit projections and measurement ordering

For multiple-qubit systems, separable projectors can be implemented by using in parallel the single-qubit projectors described above. This, by construction, allows the implementation of arbitrary separable projectors. In practice, depending on the details of a specific tomography, multiple-qubit tomographies can require a large number of measurements. If the time to switch from one measurement to another varies depending on which measurements are switched between (as is the case with waveplates switching to different values for each projector), minimizing the time spent switching is a problem equivalent to the traveling salesman problem [22]. A great deal of time can be saved by implementing a simple, partial solution to this canonical problem (e.g., a genetic algorithm which is not guaranteed to find the optimal solution but likely to find a comparably good solution).

5.1.2 n vs. $2n$ Detectors

Before beginning any tomography on an n-qubit system, a fundamental experimental choice must be made: use n or $2n$ detectors[3]. The first choice requires an array of n detectors to measure a single separable projector at a time. While this is conceptually simple, we shall see below that using an array of $2n$ detectors to project every incoming n-qubit state into one of 2^n basis states can dramatically improve the efficiency and accuracy of a tomography. The $2n$ detector method is the generalization of simultaneously measuring both outputs in the single-qubit case (the two detectors used for single-qubit measurement are shown in Figure 5.1a), or all four basis states (HH, HV, VH, and VV) in the two-qubit case; in the general case $2n$ detectors will measure in n-fold coincidence with 2^n possible outcomes.

It should be emphasized that these additional detectors are not some 'trick', effectively masking a number of sequential settings of n detectors. If only n detectors are used, then over the course of a tomography most members comprising the input ensemble will never be measured. For example, consider measuring the projection of an unknown state into the $|00\rangle$ basis using two detectors. While this will give some number of counts, unmeasured coincidences will be routed

[3]If instead we are characterizing n d-level systems (qudits), then we must choose between n and $d \times n$ detectors.

into the $|01\rangle, |10\rangle$, and $|11\rangle$ modes. The information of how many coincidences are routed to which mode will be lost, unless another two detectors are in place in the '1' modes to measure it.

Returning to the notation of Section 5.1.1, recall that the state which passes through every beam splitter is $U_\nu^\dagger|0\rangle$; when $2n$ detectors are employed, the states $U_\nu^\dagger|r\rangle$ can all be measured, where r ranges from 0 to $2^n - 1$ and $|r\rangle$ denotes the r^{th} element of the canonical basis (the canonical basis is chosen/enforced by the beam splitters themselves).

Example 5.1 (The $|r\rangle$ notation for two qubits) *For two qubits each incident on separate beam splitters which transmit $|H\rangle$ and reflect $|V\rangle$, we can define the following values of $|r\rangle$, the canonical basis:*

$$|0\rangle \equiv |HH\rangle, \quad |1\rangle \equiv |HV\rangle, \quad |2\rangle \equiv |VH\rangle, \quad |3\rangle \equiv |VV\rangle. \tag{5.6}$$

The usefulness of this notation will become apparent during the discussion of the maximum-likelihood algorithm in Section 5.5.2.

The primary advantage to using $2n$ detectors is that every setting of the analysis system (every group of the projector and its orthogonal complements) generates exactly enough information to determine a single multiple-qubit Stokes vector (see Equations 2.13 and 2.34). Expanding out the probabilities that a multiple-qubit Stokes vector (which for now we will limit to those with only non-zero indices) is based on,

$$\begin{aligned} S_{i_1,i_2,\ldots i_n} &= \left(P_{\psi_1} - P_{\psi_1^\perp}\right) \otimes \left(P_{\psi_2} - P_{\psi_2^\perp}\right) \otimes \ldots \otimes \left(P_{\psi_n} - P_{\psi_n^\perp}\right) \\ &= P_{\psi_1,\psi_2,\ldots\psi_n} - P_{\psi_1,\psi_2,\ldots\psi_n^\perp} - \ldots \pm P_{\psi_1^\perp,\psi_2^\perp,\ldots\psi_n^\perp}, \end{aligned} \tag{5.7}$$

where the sign of each term on the last line is determined by the parity of the number of orthogonal ($\perp$) terms.

These probabilities are precisely those measured by a single setting of the entire analysis system followed by a $2n$ detector array. Returning to our primary decomposition of the density matrix from Equation 2.30,

$$\hat{\rho} = \frac{1}{2^n} \sum_{i_1,i_2,\ldots i_n=0}^{3} S_{i_1,i_2,\ldots i_n} \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \ldots \otimes \hat{\sigma}_{i_n},$$

we need only determine all of the multiple-qubit Stokes parameters to exactly characterize the density matrix. At first glance this might seem to imply that we need to use $4^n - 1$ settings of the analysis system, in order to find all of the multiple-qubit Stokes parameters save $S_{0,0,\ldots 0}$, which is always one.

While this is certainly sufficient to solve for $\hat{\rho}$, many of these measurements are redundant. In order to choose the smallest possible number of settings, note that the probabilities that constitute some multiple-qubit Stokes parameters overlap exactly with the probabilities for other multiple-qubit Stokes param-

eters. Specifically, any multiple-qubit Stokes parameter with at least one 0 subscript is derived from a set of probabilities that at least one other multiple-qubit Stokes vector (with no 0 subscripts) is also derived from. As an example, consider that

$$S_{0,3} = P_{|00\rangle} - P_{|01\rangle} + P_{|10\rangle} - P_{|11\rangle}, \tag{5.8}$$

while

$$S_{3,3} = P_{|00\rangle} - P_{|01\rangle} - P_{|10\rangle} + P_{|11\rangle}. \tag{5.9}$$

These four probabilities, measured simultaneously, will provide enough information to determine both values. This dependent relationship between multiple-qubit Stokes vectors is true in general, as can be seen by returning to Equation 5.7. Each S term with a non-zero subscript contributes a term to the tensor product on the right that looks like $\left(P_{\psi_i} - P_{\psi_i^\perp}\right)$. Had there been S terms with value-zero subscripts, however, they each would have contributed a $\left(P_{\psi_i} + P_{\psi_i^\perp}\right)$ term; as an aside, terms with zero subscripts are always dependent on terms will all positive subscripts. This reduces the minimum number of analysis settings to 3^n, a huge improvement in multiple qubit systems (e.g., 9 vs. 15 settings for 2-qubit tomography, 81 vs. 255 for 4-qubit tomography, etc.). Note that, as discussed earlier, this benefit is only possible if one employs $2n$ detectors, leading to a total of 6^n measurements (2^n measurements for each of 3^n analysis settings).[4]

Because Equation 2.38 can be used to transform any set of non-orthogonal multiple-qubit Stokes parameters into the canonical form, orthogonal measurement sets need not be used. One advantage of the option to use non-orthogonal measurement sets is that an orthogonal set may not be experimentally achievable, for instance, due to waveplate imperfections, as discussed in Appendix C.

5.1.3 Electronics and detectors

Single-photon detectors and their supporting electronics are crucial to any photonic tomography. Figure 5.2 shows a simple diagram of the electronics used to count in coincidence from a pair of Si-avalanche photodiodes. An electrical pulse from a single-photon-generated avalanche in the Silicon photodiode sends a signal to a discriminator, which, if the pulse has the appropriate amplitude and width, produces in a fan-out configuration several TTL signals that are fed into the coincidence circuitry. In order to avoid pulse reflections, a fan-out configuration is used in preference to repeatedly splitting one signal.

The signals from these discriminators represent physical counts, with the number of discriminator signals sent to a detector equal to the singles counts

[4]These measurements, even though they result from the minimum number of analysis settings for $2n$ detectors, are over-complete. A density matrix has only $4^n - 1$ free parameters, which implies that only $4^n - 1$ measurements are necessary to specify it (see n-detector tomography). Because the over-complete set of 6^n measurements is not linearly independent, it can be reduced to a $4^n - 1$ element subset and still completely specify an unknown state.

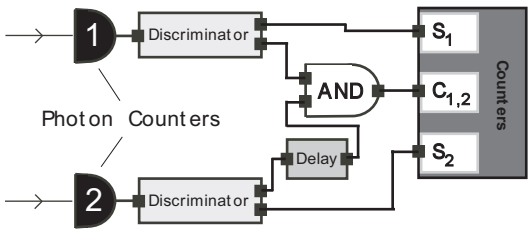

Figure 5.2: A simple diagram of the electronics necessary to operate a coincidence-based photon-counting circuit. While this diagram depicts a two-detector counting circuit, it is easily extensible to multiple detectors; by adding additional detectors each fed into a discriminator and a fan-out (shown here as the double-output of the discriminator), we gain the signals necessary for one singles counter per detector and an *AND* gate for each pair of detectors capable of recording a coincidence.

for that detector. A copy of this signal, after traveling through a variable length delay line, is input into an *AND* gate along with a similar pulse (with a static delay) from a complementary detector. The pulses sent from the discriminators are typically about 2 ns, producing a 4-ns window in which the *AND* gate can produce a signal. (The coincidence window is chosen to be as small as conveniently possible, in order to reduce the number of "accidental" coincidences, discussed below.) This signal is also sent to the counters and is recorded as a coincidence between its two parent detectors.

As with any system of this sort, the experimenter must be wary of reflected pulses generating false counts, delay lines being properly matched for correct *AND* gate operation, and system dead-time saturation effects due to, e.g., high count rates from bright sources.

5.1.4 Collecting data and systematic error correction

The projection optics and electronics described above will result in a list of coincidence counts, each tied to a single projective measurement. Incorporating the projectors defined earlier in this section, we can now make a first estimate on the number of counts we expect to receive for a given measurement of the state $\hat{\rho}$:

$$\begin{aligned} \bar{n}_{\nu,r} &= I_0 \mathrm{Tr}\left\{\tilde{M}_{\nu,r}\hat{\rho}\right\}, \\ \tilde{M}_{\nu,r} &= U_\nu^\dagger |r\rangle\langle r| U_\nu. \end{aligned} \tag{5.10}$$

Our eventually strategy (see Section 5.5.2) will be to vary our theoretical guess for $\hat{\rho}$ until our expectations optimally match our actual measured counts. Here $\bar{n}_{\nu,r}$ is the expectation value of the number of counts recorded for the ν^{th} measurement setting on the r^{th} pair of detectors (this is the pair of detectors which projects into the canonical basis state $|r\rangle$). The density matrix to be measured is denoted by $\hat{\rho}$ and I_0 is a constant scaling factor which takes into account the

duration of a measurement and the rate of state production. Note that regardless of whether n or $2n$ detectors are used, each distinct measurement setting will be indexed by ν. For n detectors, there will be a single value of r for each value of ν, as each measurement setting projects into a single state. For $2n$ detectors, there will be 2^n values of r, one for each pair of detectors capable of registering coincidences.

Throughout this Section we will modify Equation 5.10 to give a more complete estimate of the expected count rates, taking into account real errors and statistical deviations. In particular, without adjustment, the expected coincidence counts will likely be inaccurate due to experimental factors including accidental coincidences, imperfect analysis optics, mismatched detector efficiencies, and drifts in state intensity. Below we will discuss each of these in turn.

Accidental coincidences

In general, the spontaneous generation of photon pairs from down-conversion processes can result in more than one pair of photons being generated at the same time (i.e., within the detector resolution time of $\sim 1ns$). These multiple-pair generation events can lead to two uncorrelated photons being detected as a coincidence, which will tend to raise all measured counts and lead to state tomographies resulting in states closer to the maximally mixed state.[5]

We can model these accidental coincidences for the two-qubit case by considering the probability that *any* given singles count will be detected during the coincidence window of a conjugate photon. This model implies that the accidental coincidences for the ν^{th} measurement setting on the r^{th} detector pair ($n_{\nu,r}^{\text{accid}}$) will be dependent on the singles totals in each channel (${}^1S_{\nu,r}$ and ${}^2S_{\nu,r}$), the total coincidence window (Δt_r, approximately equal to twice the pulse width produced by the discriminators[6]), and the total measurement time (T_ν). When the singles channels are far from saturation (${}^{1,2}S_{\nu,r}\Delta t_{\text{dead}} \ll T_\nu$, where Δt_{dead} is the dead time of the detectors, i.e., the time it takes after a detector registers a singles count before it can register another), the percentage of time that a channel is triggered (able to produce a coincidence) is approximated by $\frac{{}^{1,2}S_{\nu,r}\Delta t_r}{T_\nu}$. The probability that the other channel will produce a coincidence within this time (again in the unsaturated regime) is proportional to the singles counts on that channel. This allows us to approximate the total accidental coincidence rate as

$$n_{\nu,r}^{\text{accid}} \simeq \frac{{}^1S_{\nu,r}\,{}^2S_{\nu,r}\Delta t_r}{T_\nu}, \tag{5.11}$$

implying that

$$\bar{n}_{\nu,r} = I_0 \text{Tr}\left\{\tilde{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}}. \tag{5.12}$$

[5]There is also a similar but generally smaller contribution from one real photon and a detector noise count, and a smaller contribution still from two detector noise counts.

[6]If the pulses are not square, or the AND logic has speed limitations, this approximation may become inaccurate.

Because the accidental rate will be necessary for analyzing the data, these expected accidental counts will need to be calculated from the singles rates for each measurement and recorded along with the actual measured coincidence counts.[7]

Beam splitter crosstalk

In most experimental implementations, particularly those involving $2n$ detectors, the polarizer used for single-qubit projection will be a beam splitter, either based on dielectric stacks, or crystal birefringence. In practice, all beam splitters function with some levels of crosstalk and absorption, i.e., some probability of reflecting or absorbing the polarization which should be transmitted and vice versa. By measuring these crosstalk probabilities and adjusting the measured counts accordingly, it is possible to recreate the approximate measurement values that would have resulted from a crosstalk-free system.

We can characterize a beam splitter using four numbers $\mathcal{C}_{r'\to r}$ which represent the probability that state r' will be measured as state r.

Example 5.2 (A faulty beamsplitter) *Assume we have measured a beam splitter which transmits* 90% *and absorbs* 10% *of incident horizontal light (state 0), while reflecting* 80% *and transmitting* 10% *of vertical light (state 1). We would therefore use*

$$\begin{aligned} \mathcal{C}_{0\to 0} &= 0.9 \\ \mathcal{C}_{0\to 1} &= 0 \\ \mathcal{C}_{1\to 0} &= 0.1 \\ \mathcal{C}_{1\to 1} &= 0.8, \end{aligned} \tag{5.13}$$

to characterize the behavior of this beam splitter.

Example 5.3 (Two-qubit crosstalk) *Consider two of the faulty beam splitters presented in Example 5.2, with crosstalk coefficients $C^A_{r'\to r}$ and $C^B_{r'\to r}$. Assuming that we label the two qubit canonical basis $|r\rangle$ as $|0\rangle \equiv |HH\rangle$, $|1\rangle \equiv |HV\rangle$, $|2\rangle \equiv |VH\rangle$, and $|3\rangle \equiv |VV\rangle$, we can derive the general two-qubit crosstalk coefficients $C_{r'\to r}$ by multiplying the single-qubit crosstalk coefficients, according to the rule:*

$$C_{(2r'_A+r'_B)\to(2r_A+r_B)} \equiv C^A_{r'_A\to r_A} C^B_{r'_B\to r_B}. \tag{5.14}$$

[7] It is advisable to initially experimentally determine Δt_r by directly measuring the accidental coincidence rate (e.g., by introducing an extra large time delay into the variable time delay before the AND gate, shown in Figure 5.2), and using Equation 5.11 to solve for Δt_r. This should be done for every pair of detectors, and ideally at several count rates, in case there are nonlinear effects in the detectors or gate electronics.

Thus, the total crosstalk matrix will be

$$C_{r'\to r} \equiv \begin{array}{c} \\ \to 0 \\ \to 1 \\ \to 2 \\ \to 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0' & 1' & 2' & 3' \end{array} \\ \begin{pmatrix} 0.81 & 0.09 & 0.09 & 0.01 \\ 0 & 0.72 & 0 & 0.08 \\ 0 & 0 & 0.72 & 0.08 \\ 0 & 0 & 0 & 0.64 \end{pmatrix} \end{array}. \tag{5.15}$$

If we use this notation to modify Equation 5.12 for predicted counts, we find that

$$\begin{aligned} \bar{n}_{\nu,r} &= I_0 \mathrm{Tr}\left\{\hat{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\mathrm{accid}} \\ \hat{M}_{\nu,r} &\equiv \sum_{r'} \left(C_{r'\to r}\right) \tilde{M}_{\nu,r'}. \end{aligned} \tag{5.16}$$

Detector-pair efficiency calibration

Because single-photon detectors will in general have different efficiencies, it may be necessary to measure the relative efficiencies of any detector pairs used in the course of a tomography. For the n-detector case, this is unnecessary, as all recorded counts will be taken with the same detectors and scaled equally. For the $2n$-detector configuration, this can be a noticeable problem, with each of the n^2 measurement bases using a different combination of detectors, with a different total coincidence efficiency. By measuring the relative efficiencies of each combination, it is possible to normalize the measured counts by dividing them by the appropriate relative efficiency.

Note that it is not necessary to know the *absolute* efficiency of each detector combination, but only the *relative* efficiencies. Knowing only the relative efficiencies leaves a single scaling factor that is applied to all counts, but as the error on a set of counts is dependent on the *measured* counts, rather than the total number of incident quantum states, this ambiguity does not affect the tomography results.

The tomography process itself may be used to conveniently determine the full set of relevant relative efficiencies. By performing enough measurements to perform an n-detector tomography while using $2n$ detectors, it is possible to perform a tomography for *each* detector combination, using only the results of that detector combination's measurements. Each of these sets will be sufficient to perform a tomography, and the tomography algorithm (see Section 5.5) will necessarily determine the total state intensity. The ratios between these state intensities (one for each detector combination) will provide the relative efficiencies of each detector combination. In the two-qubit case, this means using four detectors and 36 measurement settings, for a total of 144 measurements to calibrate the relative efficiencies.

In order to continue to update our equation for $\bar{n}_{\nu,r}$, we define an efficiency

$\hat{E}_r$ which describes the relative efficiency of the r^{th} detector combination. This allows us to correct our previous equation to

$$\bar{n}_{\nu,r} = I_0 E_0 E_r \text{Tr}\left\{\hat{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}}, \tag{5.17}$$

where E_0 is a constant scalar, which combined with the easier to measure relative efficiency E_r, gives the absolute efficiency of each detector pair.

Intensity drift

In polarization experiments based on down-conversion sources, a major cause of error can be drift in the intensity (or direction) of the pump, which causes a drift in the rate of down-conversion state production. If this intensity drift is recorded, then the prediction of the expected number of counts can be adjusted to account for this additional information. Alternatively, if $2n$ detectors are used, the sum of the counts from each of the detectors will automatically give the normalized intensity for each measurement setting, since the sum of the counts in orthonormal bases must add up to the total counts (assuming no state-dependent losses, e.g., in the polarizing beam splitters). However, when summing the counts from a complete basis like this, the measurements *should be taken at the same time*, and the summed counts must take other sources of error, like detector inefficiency, accidental counts, and beam splitter crosstalk into account.

By whatever method it is measured, assume that the relative size of the ensemble subject to the ν^{th} measurement setting is given by I_ν. Then

$$\bar{n}_{\nu,r} = I_0 E_0 I_\nu E_r \text{Tr}\left\{\hat{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}}. \tag{5.18}$$

Now it becomes clear that I_0 is the factor (not necessarily the total number of pairs produced) which, combined with the relative efficiency I_ν, gives the total number of incident states for the ν^{th} measurement setting.

Measuring probabilities

As we shall see in Section 5.5, Equation 5.18 is sufficient for the reconstruction of a density matrix; it can be used to quantify how likely it is that a given state will give an arbitrary set of data. However, there are many experimental tests (see Sections 5.2 and 5.3 and 5.4) which require probabilities, rather than count rates. The formalism of the last few subsections can be used to directly calculate probabilities from count rates.

In a crosstalk-free system,

$$P_{|\psi\rangle} = \frac{n_{\nu,r} - n_{\nu,r}^{\text{accid}}}{I_0 E_0 I_\nu E_r}, \tag{5.19}$$

where $n_{\nu,r}$ represents the actual measured data counts. The other variables will

have to be measured, though the details of this will vary from system to system. For our photonic system, for example, the product $I_0 I_\nu E_0 E_r$ is measured by using the sum of the counts from a complete orthonormal basis, weighted for detector pair efficiencies, to find the average intensity of incident photon pairs.

In a system with crosstalk, Equation 5.19 becomes incorrect because the counts $n_{\nu,r}$ are the counts *after* crosstalk has affected the system. The correct equation becomes

$$\begin{aligned} P_{|\psi\rangle} &= \frac{n'_{\nu,r'}}{I_0 E_0 I_\nu E_r}, \\ n'_{\nu,r'} &= C^{-1}_{r'\to r}\left(n_{\nu,r} - n^{\text{accid}}_{\nu,r}\right), \end{aligned} \tag{5.20}$$

where $n'_{\nu,r'}$ are the counts that would have been measured in a system with no accidental rates and no crosstalk. Notice that the crosstalk matrix $C_{r'\to r}$ has been inverted in this equation, which will only be possible when it is possible to calculate the counts before the crosstalk. Imagine for example, a beam splitter which absorbs all light. All measured counts will be zero, and the crosstalk matrix will be noninvertible. Also notice that calculating crosstalk-corrected probabilities is only easily accomplished using $2n$ detectors, as Equation 5.20 requires measurements in a full set of r bases to mathematically reverse the crosstalk effect.

5.2 Bell inequalities

Historically, one of the first uses for sequences of projective measurements was the Bell inequality [26]. First proposed in 1964 [12], the Bell inequality provides a limit on measurement correlations obtainable by any local realistic model. Here, "local realistic model" refers to any model which obeys both *locality* and *realism*. Locality is the constraint that a nearby system be completely describable without reference to distant systems. Realism requires that a complete description of a system be sufficient to predict the result of any measurement that is made upon it (even if it is not possible for an experimenter, even in principle, to have access to the complete description of a system).

5.2.1 The CHSH inequality

These assumptions eventually lead to a paradox, which is commonly written as an inequality. To see how this inequality is derived, let us define a quantity

$$E(A_i, B_j) \equiv \mathcal{M}(A_i) \times \mathcal{M}(B_j) \tag{5.21}$$

which represents the *correlation* between a measurement in basis A_i on particle A, and a measurement B_j on particle B. These measurements are assumed to give a value of either $\mathcal{M} = 1$ (the measurement succeeded) or $\mathcal{M} = -1$ (the

measurement failed). For photon polarization, this corresponds to subjecting one half of a photon pair (photon A) to a projective measurement (such as a unitary transform followed by a polarizing beam splitter) in basis $|\psi_{A_i}\rangle$, and the other half of the photon pair (photon B) to a projective measurement in basis $|\psi_{B_j}\rangle$. If photon A is transmitted through the first unitary operation and polarizing beam splitter, then $\mathcal{M}(A_i) = 1$, and if is reflected, then $\mathcal{M}(A_i) = -1$. Similarly, if photon B is transmitted through the second unitary operation and polarizing beam splitter, then $\mathcal{M}(B_i) = 1$, and if is reflected, then $\mathcal{M}(B_i) = -1$. Together, these allow a pair of separable projections on an input photon pair to give a measurement of $E(A_i, B_j)$.

For a single pair of photons, therefore, $E(A_i, B_j) = \pm 1$, regardless of the specific bases A_i or B_j that are used for the measurement. Bell's inequalities, however, are applied to *ensembles* of photon pairs (more generally, ensembles of *any* correlated measurements), rather than single pairs of photons. The value E then becomes an expectation value of this measurement applied on an ensemble, taken by averaging all results.

Local hidden variable theories, by definition, require that

$$E(A_i, B_j) = E(A_i) \times E(B_j), \tag{5.22}$$

where $E(A_i)$ and $E(B_j)$ are expectations of the results of measurements on only one particle. This assumption allows us to derive a bound on an equation involving terms of this type:

$$\begin{aligned}
& E(A_1, B_1) + E(A_1, B_2) + E(A_2, B_1) - E(A_2, B_2) \\
= \;& E(A_1)E(B_1) + E(A_1)E(B_2) + E(A_2)E(B_1) - E(A_2)E(B_2) \\
= \;& E(A_1)(E(B_1) + E(B_2)) + E(A_2)(E(B_1) - E(B_2)) \\
\leq \;& |E(A_1)\{E(B_1) + E(B_2)\} + E(A_2)\{E(B_1) - E(B_2)\}| \\
\leq \;& |E(A_1)\{E(B_1) + E(B_2)\}| + |E(A_2)\{E(B_1) - E(B_2)\}| \\
\leq \;& |E(B_1) + E(B_2)| + |E(B_1) - E(B_2)| \\
\leq \;& 2.
\end{aligned} \tag{5.23}$$

This inequality must be observed in all systems where the results of a measurement on particle A give no information about an impending measurement on particle B. Quantum mechanics predict violations of this inequality, and, for the right measurements, maximally entangled states are predicted to give a violation of up to $2 \times \sqrt{2}$.

To measure this violation using probabilities measured from separable projectors, it is possible to rewrite the CHSH inequality using the convention of [21]:

$$P_{A_1B_1} + P_{A_2B_1} + P_{A_1B_2} - P_{A_2B_2} - P_{A_1} - P_{B_1} \leq 0. \tag{5.24}$$

Here P_{A_i,B_j} is defined as the probability that photons A and B will be projected

into states A_i and B_j, respectively. It is this type of CHSH inequality that is shown in Figure 5.3 and Table 5.2, in comparison with the Collins-Gisin inequality, described below.

Experimental violations of the CHSH inequality

Since the first experimental violation of Bell's inequality [26], there has been significant improvement in the accuracy and precision of Bell measurements. Using the phase-compensated ultra-bright source described in Chapter 3, violations of Bell's inequality were measured on ensembles of larger size than ever measured before, leading to violations of:

$$\begin{aligned} 2.7260 &\pm 0.0034 \quad (216\sigma) \quad \text{in } 0.8\,\text{s} \\ 2.7252 &\pm 0.00058 \quad (1239\sigma) \quad \text{in } 28\,\text{s} \\ 2.7392 &\pm 0.00031 \quad (2417\sigma) \quad \text{in } 120\,\text{s}. \end{aligned} \tag{5.25}$$

The total experimental time to collect these violations is listed above. While the statistical violations (number of σ) for these measurements are large, they are not very accurate. By taking into account imperfections in waveplates, imperfections in measurement polarizing beam splitters, accidental counts, and detector inefficiencies (all of these are detailed in Section 5.1), we were able to make a much more accurate *absolute* measurement of Bell's inequalities:

$$2.826 \pm 0.005 \quad (165\sigma) \quad \text{in } 4\,\text{hours}. \tag{5.26}$$

Note that for all of these measurements the statistical errors are extremely low; the true errors are almost certainly dominated by uncharacterized systematic effects.

5.2.2 The Geneva inequality

If the Bell-type argument is extended to three measurement bases in each arm ($\{A,B\}_{1,2} \Rightarrow \{A,B\}_{1,2,3}$), Collins and Gisin from Geneva have shown that it is possible to construct another inequality [21]:

$$\begin{aligned} P_{A_1B_1} + P_{A_2B_1} + P_{A_3B_1} + P_{A_1B_2} + P_{A_2B_2} + \\ P_{A_1B_3} - P_{A_3B_2} - P_{A_2B_3} - P_{A_1} - 2P_{B_1} - P_{B_2} \leq 0. \end{aligned} \tag{5.27}$$

What is most interesting about this Geneva inequality is that it is *inequivalent* to the CHSH inequality; there exist states which violate the Geneva inequality but do not violate the CHSH inequality and vice versa.

In order to experimentally show this difference, we prepared a class of states which lie on the border of violating the CHSH inequality, within a very small

region of Hilbert space. These states are of the form [21]

$$\begin{aligned} \hat{\rho}_{CG}(\theta) &= \lambda|\psi(\theta)\rangle\langle\psi(\theta)| + (1-\lambda)|HV\rangle\langle HV| \\ |\psi(\theta)\rangle &= \cos(\theta)|HH\rangle + \sin(\theta)|VV\rangle. \end{aligned} \tag{5.28}$$

For each state $\hat{\rho}_{CG}(\theta)$, λ is chosen such that the CHSH violation of $\hat{\rho}_{CG}(\theta)$ is theoretically predicted to be exactly equal to 0. These states range from pure to mixed and from entangled to separable, and together exemplify the inequivalency between the Geneva and the CHSH inequalities (see Figure 5.3a for experimental results and Section 4.2.4 for information on how to create Collins-Gisin states).

The primary advantage of either the CHSH or the Geneva inequality is their function as a test of local realism. Both require previous knowledge of the state in order to choose measurement settings that maximize the value of the inequalities. For both the CHSH and the Geneva inequalities, we used a numerical search to find the optimal measurement settings (although in the case of the CHSH inequality, a simple analytic proscription for the optimal settings has been found—see [32]).

The Geneva and CHSH inequalities are experimentally compared in Figure 5.3 and Table 5.2.

5.3 Entanglement witnesses

Recently, entanglement witnesses [45] have been suggested as a use for projective measurements, and more specifically, as a test for entanglement. Their advantage over complete tomography is speed—they require few measurement settings. As we shall see in Section 5.5, however, this does not make up for the incomplete information they provide: only when successful do entanglement witnesses reveal if a state is entangled; when unsuccessful, they reveal nothing.

5.3.1 Standard entanglement witnesses

Technically, an entanglement witness W (which must be hermitian and non-positive) is an operator whose overlap with product states is non-negative, i.e., for any separable state $|\alpha\beta\rangle$,

$$\langle\alpha\beta|W|\alpha\beta\rangle \geq 0. \tag{5.29}$$

This behavior guarantees a useful quality: if $\mathrm{Tr}(W\hat{\rho}) < 0$, then $\hat{\rho}$ must be entangled. Unfortunately, however, if $\mathrm{Tr}(W\hat{\rho}) \geq 0$, no information about the state is gained, as there are both entangled and separable states in this category.

The usefulness of an entanglement witness, relative to other entanglement detection techniques is not immediately clear. Entanglement witnesses detect more states than a Bell inequality (for each entangled two-qubit state, there

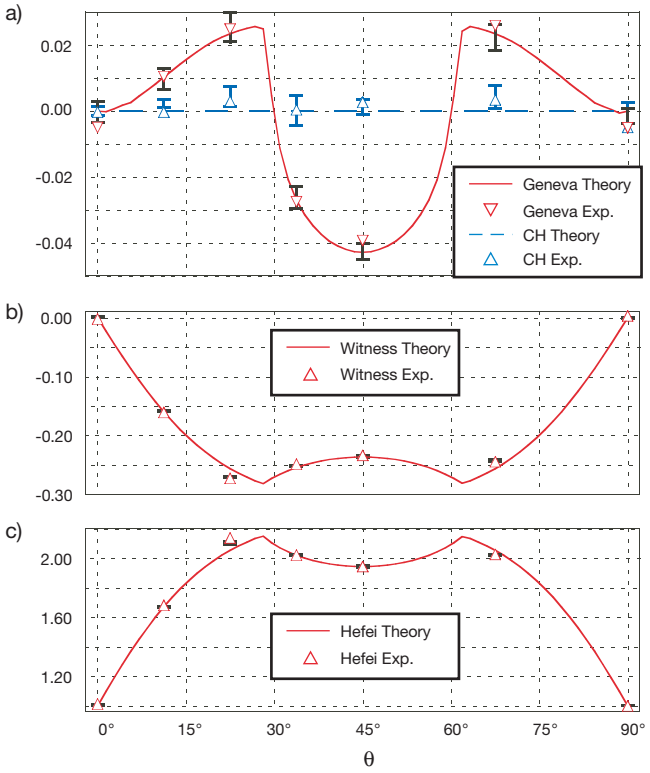

Figure 5.3: Experimentally measured values for four different entanglement detection methods, each applied to the same set of states. These states were all of the form $\rho_{CG}(\theta)$ (see equation 5.28). In all cases the solid lines represent theoretical values for ideal states, calculated numerically. In practice, the experimental states deviated slightly from the exact form of ρ_{CG} (they all maintained $> 99\%$ fidelity with the target). The error bars indicate the *state tomography's* 1-σ error region for these experimentally created states, and so for outliers may fail to bound the experimentally measured violations. (a) Measured violations of the Geneva and CHSH inequalities, which are manifestly inequivalent. ρ_{CG} is intentionally constructed so that the theoretical value of the CHSH violation is always exactly zero. Both the CHSH and Geneva inequalities have been renormalized so that 0 represents the border between a violation and nonviolation. The renormalized range of the Geneva inequality is -2 to 0.25 and the range of the CHSH inequality is -1 to $\frac{(2\sqrt{2}-2)}{4}$. (b) Experimentally measured values for an entanglement witness based on the negativity, as explained in Section 5.3.1. Negative values indicate entanglement. (c) Experimentally measured values of the Hefei inequality (see Section 5.3.2, a nonlinear Bell-type inequality based on the negativity. Because of a fundamental similarity, the value of curve 5.3c is equal to 1 minus four times curve 5.3b.

exists a witness which can detect its entanglement [45]). They use fewer measurements than a full tomography (the exact number depends upon the specific witness). Unfortunately, because a specific witness only functions for specific entangled states, it is necessary to tailor a specific witness for each entangled state if it is to be used as a reliable entanglement detection method. See section

5.6 for a more detailed comparison between these methods.

Example 5.4 (Constructing an entanglement witness) *Here we construct a witness that is capable of detecting the entanglement for all* $\rho_{CG}(\theta)$ *(see Section 4.2.4). Consider the spectrum of the partial transposition [52, 32]* ρ^{T_A} *of the state* $\rho_{CG}(\theta)$*. Denote by* $\lambda_{min}(\theta)$ *its minimum eigenvalue and by* $|e_n(\theta)\rangle$ *the corresponding eigenvector. Since* $\rho_{CG}(\theta)$ *is entangled when* $0 < \theta < \pi$*,* $\lambda_{min}(\theta)$ *is negative in the same range [32] (the value of this eigenvalue is linearly related to the negativity). Moreover,* $|e_n(\theta)\rangle$ *turns out to be independent of* θ*, i.e.,* $|e_n(\theta)\rangle = |e_n\rangle$*. It follows that* $W = |e_n\rangle\langle e_n|^{T_A}$ *is an entanglement witness, with* $\mathrm{Tr}[W\rho] \geq 0$ *for all separable states.*

5.3.2 The Hefei inequality

While entanglement witnesses are based on the linear overlap between a density matrix and an operator, it is possible to construct *nonlinear* inequalities based on the same types of measurements. Consider two sets of mutually unbiased observables, $\{A_i\}$ and $\{B_i\}$, having identical orientation ($A_1A_2A_3 = B_1B_2B_3$). In other words, $\{A_i\}$ is a set of three measurements, the axes of which are mutually perpendicular in Poincaré space (the same is true for $\{B_i\}$). Having identical orientation means that the right-hand rule applies to both sets or to neither.

The Hefei Group proved [74] that a two qubit state is separable if and only if

$$\sqrt{\langle A_1B_1 + A_2B_2\rangle_\rho^2 + \langle A_3 + B_3\rangle_\rho^2} - \langle A_3B_3\rangle_\rho \leq 1 \tag{5.30}$$

for all A_i, B_i where $\langle O\rangle_\rho = \mathrm{Tr}[\rho O]$. Moreover, the maximal value of the above inequality is equal to $1 - 4\lambda_{min}$ with λ_{min} equal to the minimal eigenvalue of the partial transpose of the density matrix.

The results of a measurement of this Hefei inequality—which once again requires one to choose the correct measurement bases to match the state—are shown in figure 5.3c. The y-axis in Figure 5.3 shows the value of the violation: a value greater than one indicates entanglement, a value of three can only be obtained by a maximally entangled state.

5.4 Tomography of ideal systems

Bell inequalities and entanglement witnesses each provide methods for state characterization, but only state tomography can *completely* characterize a state, providing the density matrix which can be used to calculate the results of *any* measurements made on the state.

The goal of tomography is to reconstruct the density matrix of an ensemble of particles through a series of measurements. In practice, this can never be performed exactly, as an infinite number of particles would be required to elim-

inate statistical error. If exact measurements *were* taken on infinite ensembles, each measurement would yield an exact probability of success, which could then be used to reconstruct a density matrix. Though unrealistic, it is highly illustrative to examine this exact tomography before considering the more general treatment. Hence, this Section will treat all measurements as yielding exact probabilities, and ignore all sources of error in those measurements.

5.4.1 Single-qubit tomography

Although reconstructive tomography of any size system follows the same general procedure, beginning with tomography of a single qubit allows the visualization of each step using the Poincaré sphere, in addition to providing a simpler mathematical introduction.

Visualization of single-qubit tomography

Exact single-qubit tomography requires a sequence of three linearly independent measurements. Each measurement exactly specifies one degree of freedom for the measured state, reducing the free parameters of the unknown state's possible Hilbert space by one.

As an example, consider measuring R, D, and H on the partially mixed state

$$\hat{\rho} = \begin{pmatrix} \frac{5}{8} & \frac{-i}{2\sqrt{2}} \\ \frac{i}{2\sqrt{2}} & \frac{3}{8} \end{pmatrix}. \tag{5.31}$$

Rewriting the state using Equation 2.10 as

$$\hat{\rho} = \frac{1}{2}\left(\hat{\sigma}_0 + \frac{1}{\sqrt{2}}\hat{\sigma}_2 + \frac{1}{4}\hat{\sigma}_3\right) \tag{5.32}$$

allows us to read off the normalized Stokes parameters corresponding to these measurements:

$$S_1 = 0, S_2 = \frac{1}{\sqrt{2}}, \text{ and } S_3 = \frac{1}{4}. \tag{5.33}$$

As always, $S_0 = 1$ due to normalization. Measuring R (which determines S_2) first, and looking to the Poincaré sphere, we determine that the unknown state must lie in the $z = \frac{1}{\sqrt{2}}$ plane (as $S_2 = \frac{1}{\sqrt{2}}$). A measurement in the D basis (with the result $P_D = P_A = \frac{1}{2}$) further constrains the state to the $y = 0$ plane, resulting in a total confinement to a line parallel to and directly above the x axis. The final measurement of H pinpoints the state. This process is illustrated in Figure 5.4a. Obviously the order of the measurements is irrelevant: it is the intersection point of three orthogonal planes that defines the location of the state.

If instead measurements are made along *non*-orthogonal axes, a very similar picture develops, as indicated in Figure 5.4b. The first measurement always

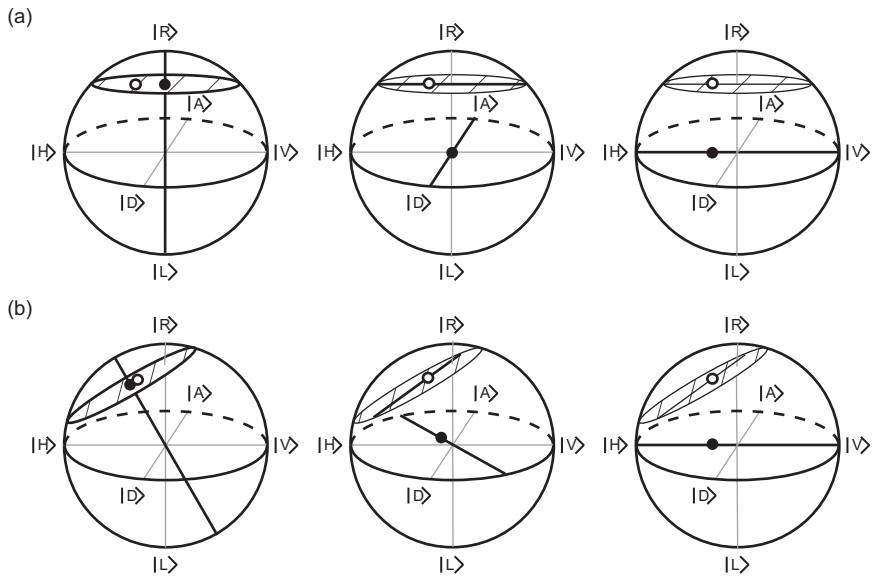

Figure 5.4: A sequence of three linearly independent measurements isolates a single quantum state in Hilbert space (shown here as an open circle in the Poincaré sphere representation). The first measurement isolates the unknown state to a plane perpendicular to the measurement basis. Further measurements isolate the state to the intersections of non-parallel planes, which for the second and third measurements correspond to a line and finally a point. The black dots shown correspond to the projection of the unknown state onto the measurement axes, which determines the position of the aforementioned planes. (a) A sequence of measurements along the right-circular, diagonal, and horizontal axes. (b) A sequence of measurements on the same state taken using non-orthogonal projections: elliptical light rotated 30° from H towards R, 22.5° linear, and horizontal. Taken from [5].

isolates the unknown state to a plane, the second to a line, and the third to a point.

Of course, in practice, the experimenter has no knowledge of the unknown state before a tomography. The set of the measured probabilities, transformed into the Stokes parameters as above, allow a state to be directly reconstructed.

A mathematical look at single-qubit tomography

Using the tools developed in the Chapter 2, single-qubit tomography is relatively straightforward. Recall Equation 2.10, $\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i$. Considering that S_1, S_2, and S_3 completely determine the state, we need only measure them to complete the tomography. From Equation 2.14, $S_{j>0} = 2P_{|\psi\rangle} - 1$; therefore, three measurements in the $|0\rangle$, $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ bases will completely specify the unknown state. If instead measurements are made in another basis, even a non-orthogonal one, they can be easily related back to the S_i parameters, and therefore the density matrix, by means of Equation 2.22.

While this procedure is straightforward, there is one subtlety which will be-

come important in the multiple-qubit case. Projective measurements generally refer to the measurement of a single basis state and return a single value between zero and one. This corresponds, for example, to an electron beam passing through a Stern-Gerlach apparatus with a detector placed at one output. While a single detector and knowledge of the input particle intensity will – in the one-qubit case – completely determine a single Stokes parameter, one could collect data from both outputs of the Stern-Gerlach device. This would measure the relative probability of projecting not only onto the state $|\psi\rangle$, but also onto $|\psi^{\perp}\rangle$, and without needing to know the input intensity. All physical measurements on single qubits, regardless of implementation, can in principle be measured this way (though in practice measurements of some qubit systems may typically detect a population in only *one* of the states, as in [37]). We will see below that although one detector functions as well as two in the single-qubit case, this situation will not persist into higher dimensions.

5.4.2 Multiple-qubit tomography

The same methods used to reconstruct an unknown single-qubit state can be applied to multiple-qubit systems. Just as each single-qubit Stokes vector can be expressed in terms of measurable probabilities—Equation 2.13, each multiple-qubit Stokes vector can be measured in terms of the probabilities of projecting the multiple-qubit state into a sequence of separable bases—Equation 2.34.

Using the most naive method, an n-qubit system, represented by 4^n Stokes parameters, would require $4^n \times 2^n$ probabilities to reconstruct (2^n probabilities for each of 4^n Stokes parameters). Of course, because an n-qubit density matrix contains $4^n - 1$ free parameters, the $4^n \times 2^n$ measured probabilities must be linearly dependent. As expected, by using the extra information that measurements of complete orthogonal bases must sum to one (e.g., $P_{HH} + P_{HV} + P_{VH} + P_{VV} = 1$, $P_{HH} + P_{HV} = P_{HD} + P_{HA}$), we find that only $4^n - 1$ probability measurements are necessary to reconstruct a density matrix.

While we can easily construct a minimum measurement set for an n-qubit system by measuring every combination of $\{H, V, D, R\}$ at each qubit, i.e.,

$$\{M\} = \{H, V, D, R\}_1 \otimes \{H, V, D, R\}_2 \otimes \ldots \{H, V, D, R\}_n , \tag{5.34}$$

this is almost never optimal (see Section 5.6). See Section 5.4.4 for a formal method for testing whether a specific set of measurements is sufficient for tomography.

Example 5.5 (An ideal 2-qubit tomography of photon pairs) *Consider measuring a state in nine complete four-element bases, for a total of 36 measurement results. These results are compiled below, with each row representing*

a single basis, and therefore a single two-qubit Stokes parameter.

$$\begin{aligned}
S_{1,1} &= \underset{\frac{1}{3}}{+P_{DD}} \underset{\frac{1}{6}}{-P_{DA}} \underset{\frac{1}{6}}{-P_{AD}} \underset{\frac{1}{3}}{+P_{AA}} = \frac{1}{3} \\
S_{1,2} &= \underset{\frac{1}{4}}{+P_{DR}} \underset{\frac{1}{4}}{-P_{DL}} \underset{\frac{1}{4}}{-P_{AR}} \underset{\frac{1}{4}}{+P_{AL}} = 0 \\
S_{1,3} &= \underset{\frac{1}{4}}{+P_{DH}} \underset{\frac{1}{4}}{-P_{DV}} \underset{\frac{1}{4}}{-P_{AH}} \underset{\frac{1}{4}}{+P_{AV}} = 0 \\
S_{2,1} &= \underset{\frac{1}{4}}{+P_{RD}} \underset{\frac{1}{4}}{-P_{RA}} \underset{\frac{1}{4}}{-P_{LD}} \underset{\frac{1}{4}}{+P_{LA}} = 0 \\
S_{2,2} &= \underset{\frac{1}{6}}{+P_{RR}} \underset{\frac{1}{3}}{-P_{RL}} \underset{\frac{1}{3}}{-P_{LR}} \underset{\frac{1}{6}}{+P_{LL}} = -\frac{1}{3} \\
S_{2,3} &= \underset{\frac{1}{4}}{+P_{RH}} \underset{\frac{1}{4}}{-P_{RV}} \underset{\frac{1}{4}}{-P_{LH}} \underset{\frac{1}{4}}{+P_{LV}} = 0 \\
S_{3,1} &= \underset{\frac{1}{4}}{+P_{HD}} \underset{\frac{1}{4}}{-P_{HA}} \underset{\frac{1}{4}}{-P_{VD}} \underset{\frac{1}{4}}{+P_{VA}} = 0 \\
S_{3,2} &= \underset{\frac{1}{4}}{+P_{HR}} \underset{\frac{1}{4}}{-P_{HL}} \underset{\frac{1}{4}}{-P_{VR}} \underset{\frac{1}{4}}{+P_{VL}} = 0 \\
S_{3,3} &= \underset{\frac{1}{3}}{+P_{HH}} \underset{\frac{1}{6}}{-P_{HV}} \underset{\frac{1}{6}}{-P_{VH}} \underset{\frac{1}{3}}{+P_{VV}} = \frac{1}{3}
\end{aligned} \tag{5.35}$$

The six remaining required parameters, listed below, are dependent upon the same measurements.

$$\begin{aligned}
S_{0,1} &= \underset{\frac{1}{3}}{+P_{DD}} \underset{\frac{1}{6}}{-P_{DA}} \underset{\frac{1}{6}}{+P_{AD}} \underset{\frac{1}{3}}{-P_{AA}} = 0 \\
S_{0,2} &= \underset{\frac{1}{6}}{+P_{RR}} \underset{\frac{1}{3}}{-P_{LR}} \underset{\frac{1}{3}}{+P_{RL}} \underset{\frac{1}{6}}{-P_{LL}} = 0 \\
S_{0,3} &= \underset{\frac{1}{3}}{+P_{HH}} \underset{\frac{1}{6}}{-P_{HV}} \underset{\frac{1}{6}}{+P_{VH}} \underset{\frac{1}{3}}{-P_{VV}} = 0 \\
S_{1,0} &= \underset{\frac{1}{3}}{+P_{DD}} \underset{\frac{1}{6}}{+P_{DA}} \underset{\frac{1}{6}}{-P_{AD}} \underset{\frac{1}{3}}{-P_{AA}} = 0 \\
S_{2,0} &= \underset{\frac{1}{6}}{+P_{RR}} \underset{\frac{1}{3}}{+P_{LR}} \underset{\frac{1}{3}}{-P_{RL}} \underset{\frac{1}{6}}{-P_{LL}} = 0 \\
S_{3,0} &= \underset{\frac{1}{3}}{+P_{HH}} \underset{\frac{1}{6}}{+P_{HV}} \underset{\frac{1}{6}}{-P_{VH}} \underset{\frac{1}{3}}{-P_{VV}} = 0
\end{aligned} \tag{5.36}$$

These terms will not in general be zero, as in Equation 2.39, when for the state $|HH\rangle$, $S_{0,3} = S_{3,0} = 1$. *Of course,* $S_{0,0} = 1$. *Taken together, these two-qubit Stokes parameters determine the density matrix:*

$$\begin{aligned}
\hat{\rho} &= \frac{1}{4}\left(\hat{\sigma}_0 \otimes \hat{\sigma}_0 + \frac{1}{3}\hat{\sigma}_1 \otimes \hat{\sigma}_1 - \frac{1}{3}\hat{\sigma}_2 \otimes \hat{\sigma}_2 + \frac{1}{3}\hat{\sigma}_3 \otimes \hat{\sigma}_3\right) \\
&= \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned} \tag{5.37}$$

This is the final density matrix, a Werner State, as defined in Equation 2.29.

5.4.3 Tomography of non-qubit systems

By making use of the qudit extensions to the Stokes parameter formalism—Equations 2.54–2.58, we can reconstruct any qudit system in exactly the same manner as qubit systems. For a single particle d-level system, a single Stokes parameter is dependent on $d-1$ independent probabilities, and $d+1$ Stokes parameters are necessary to reconstruct the density matrix. Therefore a total of $(d+1)(d-1) = d^2 - 1$ measurements are required.

Multiple-qudit systems can be reconstructed by using separable projectors [65] upon which the multiple qudit Stokes parameters are dependent (these dependencies were laid out in Section 2.50). Likewise, the following Section on general tomography, while specific to qubits, can be easily adapted to qudit systems.

5.4.4 General qubit tomography

As discussed earlier, qubit tomography will require $4^n - 1$ probabilities in order to define a complete set of T_i parameters. In practice, this will mean that 4^n measurements are necessary in order to normalize counts to probabilities. By making projective measurements on each qubit and only taking into account those results where a definite result is obtained (e.g., the photon was transmitted by the polarizer), it is possible to reconstruct a state using the results of 4^n measurements.

Our first task is to represent the density matrix in a useful form. To this end, define a set of $2^n \times 2^n$ matrices which have the following properties:

$$\begin{aligned} \mathrm{Tr}\left\{\hat{\Gamma}_\nu \cdot \hat{\Gamma}_\mu\right\} &= \delta_{\nu,\mu} \\ \hat{A} &= \sum_\nu \hat{\Gamma}_\nu \mathrm{Tr}\left\{\hat{\Gamma}_\nu \cdot \hat{A}\right\} \quad \forall \hat{A}, \end{aligned} \tag{5.38}$$

where $\hat{A}$ is an arbitrary $2^n \times 2^n$ matrix. A convenient set of $\hat{\Gamma}$ matrices to use are tensor-products of the $\hat{\sigma}$ matrices used throughout this paper:

$$\hat{\Gamma}_\nu = \hat{\sigma}_{i_1} \otimes \hat{\sigma}_{i_2} \otimes \ldots \otimes \hat{\sigma}_{i_n}, \tag{5.39}$$

where ν is simply a short-hand index by which to label the Γ matrices (there are 4^n of them) which is more concise than $i_1, i_2, \ldots i_n$. Transforming Equation 2.30 into this notation, we find that

$$\hat{\rho} = \frac{1}{2^n} \sum_{\nu=1}^{4^n} \hat{\Gamma}_\nu S_\nu. \tag{5.40}$$

Next, it is necessary to consider exactly which measurements to use. In particular, we now wish to determine the necessary and sufficient conditions on the

4^n measurements to allow reconstruction of any state.[8] Let $|\psi_\mu\rangle$ ($\mu = 1$ to 4^n) be the measurement bases, and define the probability of the μ^{th} measurement as $P_\mu \equiv \langle\psi_\mu|\hat{\rho}|\psi_\mu\rangle$.

Combining this with Equation 5.40,

$$P_\mu = \langle\psi_\mu|\frac{1}{2^n}\sum_{\nu=1}^{4^n}\hat{\Gamma}_\nu S_\nu|\psi_\mu\rangle = \frac{1}{2^n}\sum_{\nu=1}^{4^n}B_{\mu,\nu}S_\nu, \tag{5.41}$$

where the $4^n \times 4^n$ matrix $B_{\mu,\nu}$ is given by

$$B_{\mu,\nu} = \langle\psi_\mu|\hat{\Gamma}_\nu|\psi_\mu\rangle. \tag{5.42}$$

Immediately we find a necessary and sufficient condition for the completeness of the set of tomographic states $\{|\psi_\mu\rangle\}$: if the matrix $B_{\mu,\nu}$ is nonsingular, then Equation 5.41 can be inverted to give

$$S_\nu = 2^n\sum_{\mu=1}^{4^n}\left(B^{-1}\right)_{\mu,\nu}P_\mu. \tag{5.43}$$

While this provides an exact solution if exact probabilities are known, it leads to a number of difficulties in real systems. First, it is possible for statistical errors to cause a set of measurements to lead to an illegal density matrix. Second, if more than the minimum number of measurements are taken and they contain any error, they will over-define the problem, eliminating the possibility of a single analytically calculated answer. To solve these problems it is necessary to analyze the data in a fundamentally different way, in which statistically varying probabilities are assumed from the beginning and optimization algorithms find the state most likely to have resulted in the measured data (Section 5.5.2).

5.5 Experimental tomography of real data

As discussed earlier, any real experiment will contain statistical and systematic errors which preclude the use of the ideal tomography described in Section 5.4. Instead, it is necessary to use an algorithm (the maximum-likelihood technique) which assumes some uncertainty or error in measurement results, and returns a state which is the most likely to have produced the measured results.

In order to describe real tomography, we will first discuss the types of errors which are present in an experiment, the maximum-likelihood algorithm, and some details of the optimization of the entire process using numerical search techniques. We first list in Table 5.1 the information that should have been gathered during the experimental phase, followed by the formulae used to deter-

[8] If exact probabilities are known, only $4^n - 1$ measurements are necessary. However, often only numbers of counts (successful measurements) are known, with no information about the number of counts which would have been measured by detectors in orthogonal bases. In this case an extra measurement is necessary to normalize the inferred probabilities.

mine the expected number of measured counts for the ν^{th} measurement setting on the r^{th} detector combination.

U_ν	Measurement settings
$n_{\nu,r}$	Counts recorded
$n_{\nu,r}^{\text{accid}}$	Accidental counts
$\mathcal{C}_{r'\to r}$	Crosstalk coefficients
E_r	Relative efficiencies
I_ν	Relative intensities (not used with $2n$ detectors)

$$\bar{n}_{\nu,r} \equiv I_0 E_0 I_\nu E_r \mathrm{Tr}\left\{\hat{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}}$$

$$\hat{M}_{\nu,r} \equiv \textstyle\sum_{r'} \left(C_{r'\to r}\right) \tilde{M}_{\nu,r'}$$

$$\tilde{M}_{\nu,r} \equiv U_\nu^\dagger |r\rangle\langle r| U_\nu$$

Table 5.1: Summary of data and relations required for implementing the maximum-likelihood technique for analyzing experimental data.

Given this information, we are able to numerically estimate which state was most likely to return the measured results. Note that the relative intensities I_ν are optional, and can be included as data for an n-detector tomography. For a $2n$ detector tomography, the I_ν parameters will be varied as part of the optimization algorithm, and do not need to be provided as part of the experimental data.

5.5.1 Types of Errors and State Estimation

Errors in the measurement of a density matrix fall into three main categories: errors in the measurement basis, errors from counting statistics, and errors from experimental stability. The first problem can be addressed by increasing the accuracy of the measurement apparatus (e.g., obtaining higher tolerance waveplates, better controlling the Rabi pulses, etc.) while the second problem is reduced by performing each measurement on a larger ensemble (counting for a longer time). The final difficulty is drift which occurs over the course of the tomography.[9] This drift occurs either in the state produced (in the state itself,

[9]These are the main sources of error that are likely to be present to some degree in *any* qubit implementation. In addition, each implementation may have its own unique errors, such as the wedged waveplates described in Appendix C or accidental background counts from noisy detectors. Here we neglect such system-specific difficulties.

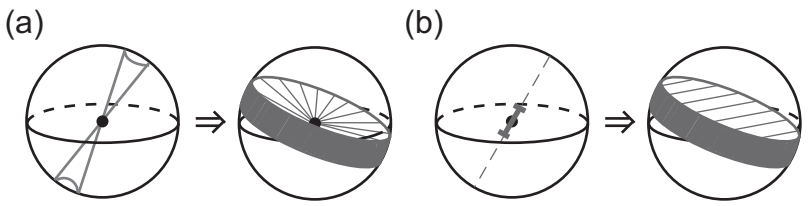

Figure 5.5: Graphical representation of errors in a single-qubit tomography. (a) Basis errors. Errors in the setting of measurement apparatus can result in an accurate measurement being taken in an unintended basis. Shown graphically is the effect that an uncertainty in the measurement basis can have on the reconstruction of a state. Instead of a single axis on the Poincaré sphere, the possible measurement axes form uncertainty cones touching at the center, since all possible measurement axes pass through the origin. This uncertainty in axis is then translated into an uncertainty in the state (shown on the right). Instead of isolating the state to a plane, all possible measurement axes trace out a volume with large uncertainty near the surface of the sphere and low uncertainty near the center. (b) Counting errors. Even if the measurement basis is exactly known, only a limited number of qubits can be measured to gain an estimate of a state's projection onto this axis (taken directly from the probability of a successful measurement). This uncertainty results in an unknown state being isolated to a one-dimensional Gaussian (approximately) in three-dimensional space, rather than to a plane.

or in the *rate* of state production) or the efficiency of the detection system, and can constrain the data-collection time.

Figure 5.5a shows what a basis error looks like on the Poincaré sphere and how that error affects the ability to isolate a state in Poincaré space. This picture indicates that a basis error is more pronounced when measuring a pure state, but actually has no effect when measuring a totally mixed state (because all bases give the same answer).

Figure 5.5b shows the same analysis of errors in counting statistics. Any real measurement can only be carried out on a limited size ensemble. Though the details of the statistics will be dealt with later, the detection events are accurately described by a Poissonian distribution, which for large numbers of counts is well approximated by a Gaussian distribution. This will cause the resultant knowledge about the unknown state to change from a plane (in the exact case) to a thick disk (uniformly thick for pure and mixed states), a one-dimensional Gaussian distribution plotted in three-dimensional space.

After all sources of error are taken into account, a single measurement results in a distribution over all possible states describing the experimenter's knowledge of the unknown state. This distribution represents the likelihood that a particular state would give the measured results, relative to another state. When independent measurements are combined, these distributions are multiplied, and ideally the knowledge of the unknown state is restricted to a small ball in Poincaré space, similar to a three-dimensional Gaussian (as a large uncertainty in any one direction will lead to a large uncertainty in the state). State isolation

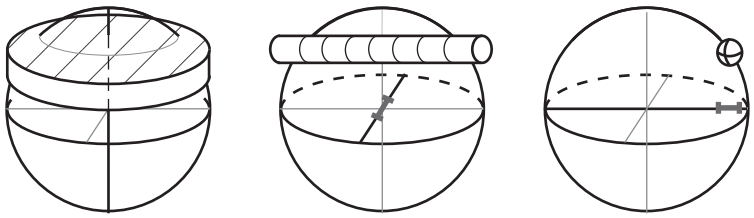

Figure 5.6: Isolation of a quantum state through inexact measurements. Although a series of real measurements (those with uncertainties) will never be able to exactly isolate an unknown quantum state, they can isolate it to a region of Hilbert space that is far more likely than any other region to contain the unknown state. Consider a series of three measurements, each containing counting errors, along orthogonal axes. From left to right, the area of Hilbert space containing the unknown state is truncated from a one-dimensional Gaussian probability distribution (the disk in the left figure) to a two-dimensional Gaussian (the cylinder in the middle figure) and finally to a three-dimensional Gaussian (the ball in the right figure). This results in an 'error ball' which approximates the position of the unknown state. The global maximum, however, can often be outside allowed Hilbert space (outside the Poincaré sphere), which is one reason a maximum-likelihood technique must be used to search over only allowed quantum states.

occurs regardless of which measurements are taken, as long as they are linearly independent, and is shown graphically in Figure 5.6 for a set of orthogonal measurements.

In contrast to the ideal case in the previous section, for which the accuracy of a reconstructed state did not depend on whether mutually unbiased measurements were made, with real measurements the advantage of mutually unbiased measurement bases becomes clear. In contrast to the measurements shown in Figure 5.6, mutually biased measurements result in a non-symmetric error ball, increasing the error in state estimation in one direction in Hilbert space.

Even after tomography returns a distribution of likelihood over Poincaré space, one final problem remains. It is very possible, especially with low counts or with the measurement of very pure states, that state estimation will return an "illegal" state. For example, in Figure 5.6, the measurements seem to place the error ball just on the edge of the sphere and slightly outside it. As all legal states have a radius of less than or equal to one in Poincaré space, it is necessary to find a way to return the most likely *legitimate* state reconstructed from a set of measurements.

5.5.2 The maximum-likelihood technique

The problem of reconstructing illegal density matrices is resolved by selecting the legitimate state most likely to have returned the measured counts [34, 33]. In practice, analytically calculating this maximally likely state is prohibitively difficult, and a numerical search is necessary. Three elements are required: a manifestly legal parametrization of a density matrix, a likelihood function which

can be maximized, and a technique for numerically finding this maximum over a search of the density matrix's parameters.

The Stokes parameters are an unacceptable parametrization for this search, as there are clearly combinations of these parameters which result in an illegal state (e.g., $S_1 = S_2 = S_3 = 1$). In this context, a legitimate state refers to a non-negative definite Hermitian density matrix of trace one. The property of non-negative definiteness for any matrix $\hat{\mathcal{G}}$ is written mathematically as

$$\langle\psi|\hat{\mathcal{G}}|\psi\rangle \geq 0 \quad \forall|\psi\rangle. \tag{5.44}$$

Any matrix that can be written in the form $\hat{\mathcal{G}} = \hat{T}^\dagger\hat{T}$ must be non-negative definite. To see that this is the case, substitute into Equation 5.44:

$$\langle\psi|\hat{T}^\dagger\hat{T}|\psi\rangle = \langle\psi'|\psi'\rangle \geq 0, \tag{5.45}$$

where we have defined $|\psi'\rangle = \hat{T}|\psi\rangle$. Furthermore, $(\hat{T}^\dagger\hat{T})^\dagger = \hat{T}^\dagger(\hat{T}^\dagger)^\dagger = \hat{T}^\dagger\hat{T}$, i.e., $\hat{\mathcal{G}} = \hat{T}^\dagger\hat{T}$ must be Hermitian. To ensure normalization, one can simply divide by the trace. Thus the matrix $\hat{g}$ given by the formula

$$\hat{g} = \hat{T}^\dagger\hat{T}/\mathrm{Tr}\{\hat{T}^\dagger\hat{T}\} \tag{5.46}$$

has all three of the mathematical properties required for density matrices.

For the one-qubit system, we have a 2×2 density matrix with 3 independent real parameters (although we will search over 4 in order to fit the intensity of the data). Since it will be useful to be able to invert relation 5.46, it is convenient to choose a tri-diagonal form for $\hat{T}$:

$$\hat{T}(\vec{t}) = \begin{pmatrix} t_1 & 0 \\ t_3 + it_4 & t_2 \end{pmatrix}, \tag{5.47}$$

where $\vec{t}$ is a vector containing each t_i. The multiple-qubit form of the same equation is given by:

$$\hat{T}(\vec{t}) = \begin{pmatrix} t_1 & 0 & \cdots & 0 \\ t_{2^n+1} + it_{2^n+2} & t_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ t_{4^n-1} + it_{4^n} & t_{4^n-3} + it_{4^n-2} & t_{4^n-5} + it_{4^n-4} & t_{2^n} \end{pmatrix}. \tag{5.48}$$

The manifestly 'physical' density matrix $\hat{\rho}_p$ is then given by the formula

$$\hat{\rho}_p(\vec{t}) = \hat{T}^\dagger(\vec{t})\hat{T}(\vec{t})/\mathrm{Tr}\{\hat{T}^\dagger(\vec{t})\hat{T}(\vec{t})\}. \tag{5.49}$$

This satisfies the first criterion for a successful maximum-likelihood search, by providing an explicitly physical parametrization for $\hat{\rho}$. The second criterion, a likelihood function, will in general depend on the specific measurement appa-

ratus used and the physical implementation of the qubit (as these will determine the statistical distributions of counts, and therefore their relative weightings). If we assume Gaussian counting statistics, then we can easily provide a suitable likelihood function.

Let $n_{\nu,r}$ be the result for the ν^{th} measurement setting on the r^{th} detector combination. Let $\bar{n}_{\nu,r}$ be the counts that would be expected from the state $\hat{\rho}$, given all information about the system:

$$\bar{n}_{\nu,r} \equiv I_0 E_0 I_\nu E_r \text{Tr}\left\{\hat{M}_{\nu,r}\hat{\rho}\right\} + n_{\nu,r}^{\text{accid}} \tag{5.50}$$

$$\hat{M}_{\nu,r} \equiv \sum_{r'} \left(C_{r'\to r}\right) \tilde{M}_{\nu,r'} \tag{5.51}$$

$$\tilde{M}_{\nu,r} \equiv U_\nu^\dagger |r\rangle\langle r| U_\nu. \tag{5.52}$$

Given that we wish to search over the parameters of $\vec{t}$, rather than $\hat{\rho}$, we will rewrite this equation as

$$\bar{n}_{\nu,r} = I_\nu E_r \text{Tr}\left\{\hat{M}_{\nu,r}\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{t}\right)\right\} + n_{\nu,r}^{\text{accid}}. \tag{5.53}$$

Notice that the unknown scalars I_0 and E_0 have been absorbed into the unnormalized $\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{t}\right)$, allowing our numerical search to discover what their combined effect is without ever knowing their individual values.

Given these definitions, the probability of obtaining the ν^{th} measurement on the r^{th} set of detectors, $n_{\nu,r}$, from the search parameters $\vec{t}$ is proportional to

$$\exp\left[-\frac{\left(\bar{n}_{\nu,r} - n_{\nu,r}\right)^2}{2\hat{\sigma}_{\nu,r}^2}\right], \tag{5.54}$$

where $\hat{\sigma}_{\nu,r}$ is the standard deviation of the ν^{th} measurement (given approximately by $\sqrt{\bar{n}_{\nu,r}}$). Therefore, the total probability of $\hat{\rho}$ yielding the counts $\{n_{\nu,r}\}$ is given by:

$$P\left(n_{\nu,r}\right) = \frac{1}{Norm}\prod_{\nu,r}\exp\left[-\frac{\left(\bar{n}_{\nu,r} - n_{\nu,r}\right)^2}{2\bar{n}_{\nu,r}}\right], \tag{5.55}$$

where *Norm* is the normalization constant. In order to find the ideal $\vec{t}$, and therefore the ideal $\hat{\rho}$, we need to maximize the probability function above. This is equivalent to maximizing the `log` of the same function, or equivalently, minimizing its negation, giving us our final likelihood function (notice that the normalization constant is ignored for this function, as it will not affect the minimum):

$$\mathcal{L}\left(\vec{t}\right) = \sum_{\nu,r}\frac{\left(\bar{n}_{\nu,r} - n_{\nu,r}\right)^2}{2\bar{n}_{\nu,r}}. \tag{5.56}$$

The final piece in the maximum-likelihood technique is an optimization routine, of which there are many available. Our implementation will be discussed

in the next subsection[10]. After a minimum is found, $\hat{\rho}$ can be reconstructed from the values of $\vec{t}$.

Example 5.6 (A single-qubit tomography) *Photon pairs generated via spontaneous parametric down-conversion from a nonlinear crystal can be used to generate single-photon states. Measuring a photon in one arm collapses the state of its partner to a single-qubit Fock state [31]. An ensemble of these photons can be characterized using the maximum-likelihood technique. The following data was taken from an experiment in "Remote State Preparation" [58]:*

$$\begin{aligned} H &= 6237 & D &= 5793 \\ V &= 8333 & R &= 6202. \end{aligned}$$

For this first example we will assume that no intensity normalization or crosstalk compensation needs to occur (see Example B.1 for a more thorough example). After minimizing the likelihood function, we obtain the following $\hat{T}$ matrix

$$\hat{T} = \begin{pmatrix} 73.4 & 0 \\ -29.0 - 1.2i & 77.1 \end{pmatrix}, \tag{5.57}$$

from which we can derive the density matrix,

$$\hat{\rho} = \frac{\hat{T}^\dagger \hat{T}}{\mathrm{Tr}\left\{\hat{T}^\dagger \hat{T}\right\}} = \begin{pmatrix} 0.5121 & 0.1837 + 0.0075i \\ 0.1837 - 0.0075i & 0.4879 \end{pmatrix}. \tag{5.58}$$

Note that the maximum-likelihood technique easily adapts to measurements in mutually biased bases (e.g., due to imperfect yet well characterized waveplates) and over-complete measurements (taking more measurements than is necessary). In the first case the set of $|\psi\rangle$ is mutually biased (i.e., not in the canonical bases), though still governed by the mathematics of tomography we have laid out; in the second case the sum in Equation 5.56 is extended beyond the minimum number of measurement settings.

5.5.3 Optimization algorithms

In order to complete a tomography, the likelihood function $\mathcal{L}\left(\vec{t}\right)$ must be minimized. A number of optimization programs exist which can search over a large number of parameters (e.g., $\vec{t}$) in order to minimize a complex function. We use the *Matlab 7.0* function `lsqnonlin`, which is optimized to minimize a sum of squares. This type of optimized algorithm is more efficient than a generic search, such as the *Matlab* function `fminunc`. In order for this minimization to work most effectively, it takes as parameters $f(\vec{t})$ and $\frac{\partial f(\vec{t})}{\partial t_i}$, where $\mathcal{L}\left(\vec{t}\right)$ is of the form

$$\mathcal{L}\left(\vec{t}\right) = \sum_x \left[f_x(\vec{t})\right]^2. \tag{5.59}$$

[10] For freely available code and further examples, see:
`http://www.physics.uiuc.edu/research/QI/Photonics/Tomography/`.

For the problem of tomography, we can write

$$f_{\nu,r} \equiv \frac{\bar{n}_{\nu,r} - n_{\nu,r}}{\sqrt{2\bar{n}_{\nu,r}}} \tag{5.60}$$

$$= \frac{I_\nu E_r \mathrm{Tr}\left\{\hat{M}_{\nu,r}\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{t}\right)\right\} + n_{\nu,r}^{\mathrm{accid}} - n_{\nu,r}}{\sqrt{2\left(I_\nu E_r \mathrm{Tr}\left\{\hat{M}_{\nu,r}\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{t}\right)\right\} + n_{\nu,r}^{\mathrm{accid}}\right)}}. \tag{5.61}$$

With some effort, we can analytically derive the partial derivatives of these terms, allowing the optimization algorithm to not only run much faster, but to converge quickly regardless of the initial search condition:

$$\begin{aligned}
\frac{\partial f_{\nu,r}}{\partial t_i} &= \frac{\left[\frac{\partial}{\partial t_i}\left(\bar{n}_{\nu,r} - n_{\nu,r}\right)\right]\sqrt{\bar{n}_{\nu,r}} - \left(\bar{n}_{\nu,r} - n_{\nu,r}\right)\left(\frac{\partial}{\partial t_i}\sqrt{\bar{n}_{\nu,r}}\right)}{\sqrt{2}\bar{n}_{\nu,r}} \\
&= \frac{\left(\frac{\partial \bar{n}_{\nu,r}}{\partial t_i}\right)\left(\bar{n}_{\nu,r}\right)^{\frac{1}{2}} - \left(\bar{n}_{\nu,r} - n_{\nu,r}\right)\left[\frac{1}{2}\left(\bar{n}_{\nu,r}\right)^{-\frac{1}{2}}\frac{\partial \bar{n}_{\nu,r}}{\partial t_i}\right]}{\sqrt{2}\bar{n}_{\nu,r}} \\
&= \frac{1}{2\sqrt{2\bar{n}_{\nu,r}}}\frac{\partial \bar{n}_{\nu,r}}{\partial t_i}\left(1 + \frac{n_{\nu,r}}{\bar{n}_{\nu,r}}\right).
\end{aligned} \tag{5.62}$$

Note that it is impossible for this function to go to zero unless $\frac{\partial \bar{n}_{\nu,r}}{\partial t_i}$ goes to zero, important when considering whether or not the maximum-likelihood function will have several local minima. (We believe that there is a single local minimum for tomographies on real sets of data, though this is unproven.) Because $\hat{T}$ is a linear function, we can easily write down

$$\begin{aligned}
\frac{\partial \bar{n}_{\nu,r}}{\partial t_i} &= I_\nu E_r \frac{\partial}{\partial t_i}\mathrm{Tr}\left\{\hat{M}_{\nu,r}\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{t}\right)\right\} \\
&= I_\nu E_r \mathrm{Tr}\left\{\hat{M}_{\nu,r}\left[\hat{T}^\dagger\left(\vec{t}\right)\hat{T}\left(\vec{\delta}_{ij}\right) + \hat{T}^\dagger\left(\vec{\delta}_{ij}\right)\hat{T}\left(\vec{t}\right)\right]\right\},
\end{aligned} \tag{5.63}$$

where $\vec{\delta}_{ij}$ is a j-element vector whose i^{th} element is equal to one. All other elements of $\vec{\delta}_{ij}$ are equal to zero. (Here j is the length of $\vec{t}$.)

Even using these derivatives (and especially if they are not used), it is important to choose an initial condition for the search which is as close as possible to the correct answer. This amounts to making the best analytic guess possible using the ideal tomographic techniques presented in Section 5.4. It is possible that those ideal techniques will result in an illegal density matrix, i.e., some of its eigenvalues will be negative. If this happens (indeed, this happening is the reason we *need* the maximum-likelihood technique), we simply set those negative eigenvalues to zero, renormalize the positive eigenvalues, and use this truncated state as the starting condition for the search.

5.6 Comparing entanglement detection methods

This chapter has presented several options for detecting entanglement using separable projective measurements, but no strong comparisons between these methods. In order to choose between them, a theoretical and experimental comparison which takes into account their response to systematic and statistical errors would be ideal.

To that end, this section describes the preparation of a set of states near the entangled-separable border in Hilbert space (the CG states described in Section 4.2.4) and their measurement using each of the aforementioned entanglement detection techniques discussed in this chapter. These experiments were the first implementations of the inequalities proposed by Collins et al. [21] and Yu et al. [74], referred to as the Geneva (Section 5.2.2) and Hefei (Section 5.3.2) inequalities, respectively (so named for the cities from which they were proposed). While analyzing each method, it is important to consider the information the test provides (e.g., quantification of entanglement, information about local realism, complete state determination), how many distinct measurement settings are required to perform the test (important if changing bases is experimentally costly), and how the test is affected by both statistical uncertainty and systematic errors (statistical factors will primarily determine the time necessary to make a measurement). This section concludes with a table quantifying the differences between these methods.

The experiments were carried out using pairs of polarization-encoded entangled photons of the type described in Chapter 3 (see Figure 5.7). The states used for experimentally testing these methods are the $\hat{\rho}_{CG}$, described in Section 4.2.4.

The first two entanglement detection methods discussed (Bell inequalities and entanglement witnesses) share two disadvantages: they require previous knowledge of the state to be effectively applied and they fail to quantify the *amount* of entanglement present. These problems can be overcome by taking a complete state tomography, which through a series of separable measurements reconstructs the full density matrix. While tomography requires no prior knowledge of the state and allows any of the above quantities to be derived from the density matrix, it does not necessarily provide a test of local realism[11] and requires a minimum of 16 separable measurements settings [71, 5].

Table 5.2 shows the number of measurements and the total number of state copies *per measurement* necessary to accurately measure, using each method, four representative two-qubit states: $I/4$, $\rho_{CG}\left(\frac{\pi}{8}\right)$, $|\psi^+\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + |VV\rangle\right)$,

[11] It is possible for tomography data which is explainable by a local hidden variable model to predict an entangled state which *could* violate local-realism. This is because the data necessary for a state tomography can be from any complete set of linearly independent data. Data for Bell inequalities must instead come from very specific bases tailored for each state to be measured.

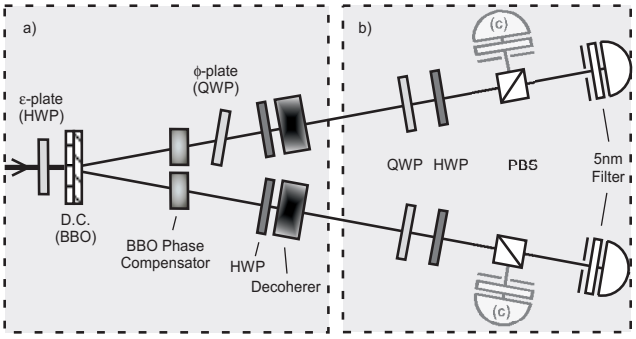

Figure 5.7: Experimental setup for the entanglement detection methods discussed. (a) State creation. A 351-nm pump beam down-converts inside two orthogonally oriented 0.6-mm BBO crystals. These crystals are designed such that the superposed down-conversion from both crystals produces the state $\psi(\epsilon, \phi) = cos(\epsilon)|HH\rangle + e^{i\phi} sin(\epsilon)|VV\rangle$, where ϵ and ϕ are respectively determined by the rotation of the ϵ-HWP (half waveplate) and the tilt of the ϕ-QWP (quarter waveplate) about its optic axis, oriented vertically. Specially designed 245-μm BBO plates compensate for any angular dependence in the phase factor $e^{i\phi}$ (as discussed in Section 3.4). Two HWP's transform this state into a state with arbitrary diagonal values in the H/V basis. Finally, as discussed in Example 4.5, two 1-cm decohering quartz crystals destroy all coherence terms in the density matrix except for $|HH\rangle\langle VV|$ and $|VV\rangle\langle HH|$ [38]. (b) Measurement. In each arm, a QWP-HWP-PBS (polarizing beam splitter) combination allows projection into any single-qubit basis. Silicon avalanche photodiodes and coincidence electronics allow the results of separable, two-qubit measurements to be recorded. (c) For some experiments, it is advantageous to add an additional detector at each of the remaining PBS ports, in order to collect not just the results of a single separable projector, but an entire four-element basis measurement.

and $|HH\rangle$. The number of measurements were minimized in each case, which for the 2-detector case leads to a substantially larger necessary ensemble size, exemplified by the factor of 2.6 ($\approx [8800 \times 16]/[1500 \times 9 \times 4]$) increase in necessary state copies between 36 and 16 measurement, 2-detector tomographies of $\rho_{CG}\left(\frac{\pi}{8}\right)$.

This point merits special attention. Quite counterintuitively, this implies that *increasing* the number of measurement settings *decreases* the total measurement time. This is possible because each measurement setting can operate on a smaller size ensemble of particles. Because the 36-measurement tomography uses nine complete bases, instead of just the single complete basis used in a 16-measurement tomography, *every measurement* taken reduces the error in the state intensity. This redundant measurement on the state intensity translates into a far more accurate tomography for the same size ensemble when using 36 as opposed to 16 measurement settings. For more information on this effect, see [8].

These results are a numerical upper bound that is highly dependent not only on the state to be measured, but the particular measurement *settings* that are

(a) Single Projector / 2 Detectors					
Method	M#	$I/4$	$\rho_{CG}\left(\frac{\pi}{8}\right)$	ψ^+	$\|HH\rangle$
CHSH	7	6800	4400	12400	200
Geneva	11	7000	5400	2600	200
Ent. Witness	8	800	400	200	500
Hefei	8	12300	2500	400	200
Tomography‡	16	23500	8800	900	900
(b) Four Projectors / 4 Detectors					
Method	M#	$I/4$	$\rho_{CG}\left(\frac{\pi}{8}\right)$	ψ^+	$\|HH\rangle$
CHSH	4	3400	1000	2100	200
Geneva	8	2300	2200	1600	200
Ent. Witness	3	800	400	200	500
Hefei	3	5500	1600	400	100
Tomography‡	9	4000	1500	400	200

Table 5.2: This table compares five different entanglement detection methods using two different experimental configurations. For each detection method, the second column (M#) indicates the number of necessary measurement settings. Each additional column shows, for each of four two-qubit states, the minimum number of distinct two-qubit systems that need to be used, *per measurement*, in order to attain a $\pm 1\%$ statistical error. Here, a $\pm 1\%$ error is measured relative to the entire range of the measured quantity. For example, the CHSH inequality ranges from -1 to ~ 0.207, making a $\pm 1\%$ error equal to a $\pm 0.01 * 1.207$ error in the violation. The minimum state copies necessary were numerically estimated using a Monte-Carlo simulation of the expected data, the results of which agreed with analytic estimates. The third through sixth columns respectively show the state copies per measurement necessary for the states $I/4$, $\rho_{CG}\left(\frac{\pi}{8}\right)$, $|HH\rangle$, and $|\psi^+\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + |VV\rangle\right)$. (a) Results for a single projector (2-detector) setup, where each two-qubit state is measured using only a single separable two-qubit projector. (b) Results for a full basis measurement (4-detector) setup, where each two-qubit state is measured simultaneously by four mutually orthogonal projectors (see figure 5.7c). (‡) Tomography returns a density matrix, from which the results of each other test can be derived. The tomography entries on this chart show the minimum state copies necessary to attain a density matrix precise enough to reduce the error on each of these derived quantities to less than $\pm 1\%$.

chosen (for any given state there may be many equivalent ways to measure a maximal violation of an inequality or a maximally efficient tomography). This is exemplified by the 2-detector CHSH results for $|\psi^+\rangle$, which appear to be quite high, and $|HH\rangle$, which are quite low. The maximally entangled state requires very specific measurements, and leaves little freedom to optimize for low errors. The violation for $|HH\rangle$, however, is theoretically zero, allowing measurement settings to be chosen which are all orthogonal to $|HH\rangle$, all resulting in probability zero, and all with low errors.

Comparing these five methods, we find that the CHSH and Geneva inequalities are useful for performing tests of local realism, the Hefei inequality and the entanglement witness can be used to quickly bound λ_{min}, and the tomography appears to be the most attractive option in general; each other method first *re-*

quires a tomography to choose its measurement settings—a tomography which can be used to derive any information about a state. In the 4-detector case, the tomography actually *outperforms* several other methods for entangled states, the states most likely to be measured using entanglement detection techniques.

5.7 Error Analysis

Error analysis of state characterization methods is in practice a non-trivial process. The traditional method of error analysis involves analytically solving for the error in each measurement due to each source of error, then propagating these errors through a calculation of any derived quantity. In the photon case, for example, errors in counting statistics and waveplate settings were analyzed in some detail in reference [34], giving errors in both density matrices and commonly derived quantities, such as the tangle and the linear entropy[12]. In practice, however, these errors appear to be too large: We have experimentally repeated some of our measurements many times, and observed a spread in the value of derived quantities which is approximately an order of magnitude smaller that the spread predicted from an analytic calculation of the uncertainty. Obviously, the correctness of the analytic calculation is questionable. Thus it is worthwhile to discuss alternate methods of error analysis.

One promising numerical method is the 'Monte Carlo' technique, whereby additional numerically simulated data is used to provide a statistical distribution over any derived quantity. Once an error distribution is understood over a single measurement (e.g., Gaussian for waveplate setting errors, or Poissonian over count statistics), a set of 'simulated' results can be generated. These results are simulated using the known error distributions in such a way as to produce a full set of numerically generated data which could feasibly have come from the same system. These data are numerically generated (at the measured counts level), and each set is used to calculate a density matrix via the maximum likelihood technique. This set of density matrices is then used to calculate the standard error on any quantity implicit in or derived from the density matrix.

As an example, consider the application of the Monte Carlo technique to the down-conversion results from Example B.1. Two polarization-encoded qubits are generated within ensembles that obey Poissonian statistics, and these ensembles are used to generate a density matrix using the maximum likelihood technique. In order to find the error on a quantity derived from this density matrix (e.g., the tangle), 36 new measurement results are numerically generated, each drawn randomly from a Poissonian distribution with mean equal to the original number of counts. These 36 numerically generated results are then fed into the maximum likelihood technique, in order to generate a new den-

[12]Beware. The error analysis and some of the gamma matrices are incorrectly calculated in this reference. See http://www.physics.utoronto.ca/ dfvj/publications.html for a list of corrections.

sity matrix, from which, e.g., the tangle may be calculated. This process is repeated many times, generating many density matrices and a distribution of tangle values, from which the error in the initial tangle may be determined. In practice, additional sets of simulated data must be generated until the error on the quantity of interest converges to a single value. For the data in Examples 5.6 and B.1, a total of 100 simulations were used.

6 Conclusions

Taken as a whole, the research presented here represents a dramatic shift in what is possible for experimental quantum information, and in many cases is the limit of the current technology (e.g., state brightness, state entanglement, Bell violations, precision of state creation). These advances would not have been possible without the simultaneous development of an intuitive sense of how the quantum mechanical world works, at least for two qubits. This intuition is far from complete, however, and there are many things still to be done.

While phase compensation has been extremely successful at generating bright, pure states, there are still unexplained second-order effects in the experimental data. Pushing state brightness and fidelity to the next level will require exploring these both theoretically and experimentally. Additionally, the theory of the construction and measurement of quantum processes is far behind what exists for states. It is, at this point, unclear if it is even possible to apply a maximum likelihood algorithm to processes without serious assumptions about the type of process being measured. State tomography of two qubits has a separate problem: if it is to be incorporated into technologically useful systems, tomography must be performed using an algorithm that can use incremental adjustments to a density matrix rather than using an optimization which can only be performed after *all* the data has been collected.

The true frontier of this research, however, is the exploration of larger Hilbert spaces. For the work presented here, state tomography provided the exquisite precision necessary to improve every other aspect of the system. When Hilbert spaces grow to only a few more qubits, tomography will simply not be feasible, and smarter characterization methods will need to be developed. These methods hold the promise of facilitating continual experimental improvements to the first stages of working quantum computers. And as for whether those computers will use photonic qubits? It is not yet clear if photons will be utilized as the basic qubits in real quantum computers. However, it is almost certain that *the connections* between distant quantum systems will use photonic quantum information, and take advantage of the strange nonlocal properties which it exemplifies. It is my hope that the work presented here has brought such a system a little closer to realization.

A Measures for Characterizing Quantum Systems

A.1 Measures on quantum states

A.1.1 Fidelity

Fidelity is a measure of state overlap:

$$F(\rho_1, \rho_2) = \left(\mathrm{Tr} \left\{ \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \right\} \right)^2, \tag{A.1}$$

which - for ρ_1 and ρ_2 pure - simplifies to $\mathrm{Tr}\{\rho_1\rho_2\} = |\langle\psi_1|\psi_2\rangle|^2$ [36][1].

A.1.2 Tangle

The concurrence and tangle are measures of the non-classical properties of a quantum state [73, 20]. For two qubits[2], concurrence is defined as follows: consider the non-Hermitian matrix $\hat{R} = \hat{\rho}\hat{\Sigma}\hat{\rho}^{\mathrm{T}}\hat{\Sigma}$ where the superscript T denotes transpose and the 'spin flip matrix' $\hat{\Sigma}$ is defined by:

$$\hat{\Sigma} \equiv \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{A.2}$$

If the eigenvalues of $\hat{R}$, arranged in decreasing order, are given by $r_1 \geq r_2 \geq r_3 \geq r_4$, then the concurrence is defined by

$$C \quad = \quad \mathrm{Max}\left\{0, \sqrt{r_1} - \sqrt{r_2} - \sqrt{r_3} - \sqrt{r_4}\right\}. \tag{A.3}$$

The tangle is calculated directly from the concurrence:

$$T \equiv C^2. \tag{A.4}$$

[1] Note that some groups use an alternate convention of fidelity, equal to the square root of the formula presented here.

[2] The analysis in this thesis applies, in large part, to the two-qubit case only. Measures of entanglement for mixed n-qubit systems are a subject of on-going research: see, for example, [64] for a recent survey. In some restricted cases it may be possible to measure entanglement directly, without quantum state tomography; this possibility was investigated in [60]. Also, one can detect the presence of non-zero entanglement, without quantifying it, using so-called "entanglement witnesses" [45]. Chapter 5 describes the trade-offs associated with these other entanglement characterization schemes.

The tangle (and the concurrence) range from 0 for product states (or, more generally, any incoherent mixture of product states) to a maximum value of 1 for Bell states.

A.1.3 Entropy and linear entropy

The Von Neumann entropy quantifies the degree of mixture in a quantum state, and is given by

$$S \equiv -\mathrm{Tr}\left\{\hat{\rho}\ln\left[\hat{\rho}\right]\right\} = -\sum_i p_i \ln\left\{p_i\right\}, \tag{A.5}$$

where the p_i are the eigenvalues of ρ. The linear entropy [71] is a more analytically convenient description of state mixture. The linear entropy for a two-qubit system is defined by:

$$\begin{aligned} S_L &= \frac{4}{3}\left(1 - \mathrm{Tr}\left\{\hat{\rho}^2\right\}\right) \\ &= \frac{4}{3}\left(1 - \sum_{a=1}^{4} p_a^2\right), \end{aligned} \tag{A.6}$$

where p_a are the eigenvalues of ρ. Note that for pure states, $\hat{\rho}^2 = \hat{\rho}$, and $\mathrm{Tr}\left[\hat{\rho}\right]$ is always 1, so that S_L ranges from 0 for pure states to 1 for the completely mixed state.

A.2 Measures on quantum processes

A.2.1 Process fidelity

The standard state fidelity is insufficient for comparing two processes. Instead, we define the process fidelity $\mathcal{F}$ [17]:

$$\mathcal{F}(\mathcal{E}_1, \mathcal{E}_2) \equiv \int d\psi F(\mathcal{E}_1(|\psi><\psi|), \mathcal{E}_2(|\psi><\psi|)). \tag{A.7}$$

For the results presented in this work, we calculated this sum numerically, as an analytic solution is prohibitively difficult.

B A Complete Example of Tomography

In order to demonstrate how the concepts presented in Chapter 5 are actually applied, we have included an example which from start to finish uses laboratory parameters and data, taken from a two-qubit entangled photon source. Throughout this example we will use our standard convention for the canonical basis: $|0\rangle \equiv |HH\rangle$, $|1\rangle \equiv |HV\rangle$, $|2\rangle \equiv |VH\rangle$, and $|3\rangle \equiv |VV\rangle$.

Example B.1 (A complete two-qubit tomography) *Before collecting tomography data, there are several measurement parameters that must be measured. After experimentally determining that each of our beam splitters has negligible absorption, a* 0.8% *chance to reflect* $|H\rangle$*, and a* 0.5% *chance to transmit* $|V\rangle$*, we can determine that*

$$C_{r'\to r} \equiv \begin{array}{c} \\ \to 0 \\ \to 1 \\ \to 2 \\ \to 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0' & 1' & 2' & 3' \end{array} \\ \begin{pmatrix} 0.9842 & 00049 & 0.0049 & 0.0000 \\ 0.0079 & 0.9871 & 0.0000 & 0.0050 \\ 0.0079 & 0.0000 & 0.9871 & 0.0050 \\ 0.0001 & 0.0079 & 0.0079 & 0.9901 \end{pmatrix} \end{array}. \tag{B.1}$$

Rather than measuring intensity fluctuations by picking off a part of the pump laser, we will choose during this tomography to fit the intensity parameters I_ν *as part of the maximum likelihood technique (we use four detectors, which will allow us to fit a relative intensity to each measurement setting by using the measured counts from each of four orthogonal projectors).*

Because this particular tomography will use a total of nine measurement settings (the minimum number required), there will not be enough information to fit for the detector-pair efficiencies. A previous tomography (using 36 measurement settings and not shown here) was used to solve for the E_r*, using a two-detector tomography applied to the 36 measurement results from* each *of the four pairs of detectors:*

$$\begin{aligned} E_1 &= 0.9998 & E_3 &= 0.9195 \\ E_2 &= 1.0146 & E_4 &= 0.9265\ . \end{aligned}$$

To simplify the example, we will make all measurements in the canonical bases (this could be accomplished using either ideal waveplates or, in some cases, imperfect waveplates—see Appendix C.

With these parameters recorded, we can now take the data. The following counts were recorded for a slightly mixed Bell state (close to $\frac{1}{\sqrt{2}}(|HH\rangle + i|VV\rangle)$*):*

$n_{1,r}$:	$HH = 3708$	$HV = 77$	$VH = 51$	$VV = 3642$
$n_{2,r}$:	$HD = 1791$	$HA = 1987$	$VD = 2096$	$VA = 3642$
$n_{3,r}$:	$HR = 2048$	$HL = 1854$	$VR = 1926$	$VL = 1892$
$n_{4,r}$:	$DH = 1766$	$DV = 1914$	$AH = 2153$	$AV = 1741$
$n_{5,r}$:	$DD = 1713$	$DA = 1945$	$AD = 2208$	$AA = 1647$
$n_{6,r}$:	$DR = 3729$	$DL = 91$	$AR = 102$	$AL = 3662$
$n_{7,r}$:	$RH = 2017$	$RV = 1709$	$LH = 1917$	$LV = 1955$
$n_{8,r}$:	$RD = 3686$	$RA = 102$	$LD = 109$	$LA = 3651$
$n_{9,r}$:	$RR = 2404$	$RL = 1474$	$LR = 1712$	$LL = 2209$,

with the corresponding accidental counts (calculated using the measured singles rates and the previously determined coincidence window Δt_r (c.f., Equation 5.11).

$n_{1,1}^{\text{accid}} = 5.4$	$n_{1,2}^{\text{accid}} = 5.6$	$n_{1,3}^{\text{accid}} = 5.9$	$n_{1,4}^{\text{accid}} = 6.0$
$n_{2,1}^{\text{accid}} = 5.2$	$n_{2,2}^{\text{accid}} = 5.5$	$n_{2,3}^{\text{accid}} = 5.6$	$n_{2,4}^{\text{accid}} = 6.0$
$n_{3,1}^{\text{accid}} = 5.3$	$n_{3,2}^{\text{accid}} = 5.5$	$n_{3,3}^{\text{accid}} = 5.6$	$n_{3,4}^{\text{accid}} = 5.9$
$n_{4,1}^{\text{accid}} = 5.2$	$n_{4,2}^{\text{accid}} = 5.3$	$n_{4,3}^{\text{accid}} = 6.0$	$n_{4,4}^{\text{accid}} = 6.1$
$n_{5,1}^{\text{accid}} = 5.2$	$n_{5,2}^{\text{accid}} = 5.4$	$n_{5,3}^{\text{accid}} = 5.9$	$n_{5,4}^{\text{accid}} = 6.2$
$n_{6,1}^{\text{accid}} = 5.2$	$n_{6,2}^{\text{accid}} = 5.3$	$n_{6,3}^{\text{accid}} = 5.9$	$n_{6,4}^{\text{accid}} = 6.1$
$n_{7,1}^{\text{accid}} = 5.3$	$n_{7,2}^{\text{accid}} = 5.4$	$n_{7,3}^{\text{accid}} = 5.9$	$n_{7,4}^{\text{accid}} = 6.1$
$n_{8,1}^{\text{accid}} = 5.4$	$n_{8,2}^{\text{accid}} = 5.9$	$n_{8,3}^{\text{accid}} = 6.1$	$n_{8,4}^{\text{accid}} = 6.6$
$n_{9,1}^{\text{accid}} = 5.3$	$n_{9,2}^{\text{accid}} = 5.4$	$n_{9,3}^{\text{accid}} = 6.0$	$n_{9,4}^{\text{accid}} = 6.2.$

After minimizing the likelihood function, we obtain the following $\hat{T}$ matrix

$$\hat{T} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2.401 + 3.167i & 2.372 & 0 & 0 \\ -6.381 - 2.649i & 3.919 - 0.897i & 2.674 & 0 \\ -8.975 + 58.630i & 1.356 - 2.106i & 1.685 - 1.514i & 60.08 \end{pmatrix}, \quad \text{(B.2)}$$

from which we can derive the density matrix,

$$\begin{aligned} \hat{\rho} &= \frac{\hat{T}^\dagger \hat{T}}{\text{Tr}\left\{\hat{T}^\dagger \hat{T}\right\}} \\ &= \begin{pmatrix} 0.50 & -0.02 - 0.01i & -0.02 - 0.01i & -0.07 - 0.49i \\ -0.02 + 0.01i & 0.00 & 0.00 + 0.00i & 0.01 + 0.02i \\ -0.02 + 0.01i & 0.00 - 0.01i & 0.00 & 0.01 + 0.01i \\ -0.07 + 0.49i & 0.01 - 0.02i & 0.01 - 0.01i & 0.50 \end{pmatrix}. \end{aligned} \quad \text{(B.3)}$$

Our search algorithm returned this density matrix because it minimized not only the main search parameters $\vec{t}$, but the intensities I_ν:

$I_1 = 7647$	$I_2 = 7745$	$I_3 = 7879$
$I_4 = 7725$	$I_5 = 7669$	$I_6 = 7754$
$I_7 = 7751$	$I_8 = 7716$	$I_9 = 7967$,

allowing us to calculate the expected counts $\bar{n}_{\nu,r}$ (for the final density matrix):

$$
\begin{array}{llll}
\bar{n}_{1,1} = 3792 & \bar{n}_{1,2} = 81 & \bar{n}_{1,3} = 67 & \bar{n}_{1,4} = 3544 \\
\bar{n}_{2,1} = 1794 & \bar{n}_{2,2} = 1956 & \bar{n}_{2,3} = 2106 & \bar{n}_{2,4} = 1735 \\
\bar{n}_{3,1} = 2046 & \bar{n}_{3,2} = 1787 & \bar{n}_{3,3} = 1933 & \bar{n}_{3,4} = 1956 \\
\bar{n}_{4,1} = 1815 & \bar{n}_{4,2} = 1895 & \bar{n}_{4,3} = 2108 & \bar{n}_{4,4} = 1758 \\
\bar{n}_{5,1} = 1618 & \bar{n}_{5,2} = 2050 & \bar{n}_{5,3} = 2247 & \bar{n}_{5,4} = 1604 \\
\bar{n}_{6,1} = 3792 & \bar{n}_{6,2} = 97 & \bar{n}_{6,3} = 103 & \bar{n}_{6,4} = 3594 \\
\bar{n}_{7,1} = 2032 & \bar{n}_{7,2} = 1699 & \bar{n}_{7,3} = 1901 & \bar{n}_{7,4} = 1966 \\
\bar{n}_{8,1} = 3758 & \bar{n}_{8,2} = 103 & \bar{n}_{8,3} = 105 & \bar{n}_{8,4} = 3583 \\
\bar{n}_{9,1} = 2271 & \bar{n}_{9,2} = 1580 & \bar{n}_{9,3} = 1751 & \bar{n}_{9,4} = 2206,
\end{array}
$$

Using the error analysis techniques presented in Section 5.7, we can estimate this state's fidelity with the Bell state $\frac{1}{\sqrt{2}}(|HH\rangle + i|VV\rangle)$ *to be* $98.4 \pm 0.2\%$.

C Waveplates

Waveplates are birefringent optics which perform unitary transformations on the polarization of input light. The optic axis of a waveplate is parallel to its surface, guaranteeing that input light which is normal to the surface, regardless of its polarization, will exit with an unchanged spatial mode. Because the waveplate is birefringent, it possesses a different index of refraction for ordinary and extraordinary polarizations, which we will denote with n_o and n_e, respectively. If the thickness of the waveplate is d, then the optical path length of the optic is $d \times n_{o,e}$. The optical phase (in radians) that a photon accumulates inside a waveplate is therefore

$$\phi_{o,e} \equiv 2\pi n_{o,e} \frac{d}{\lambda}. \tag{C.1}$$

Written in terms of a unitary transformation, the action of a waveplate is

$$\hat{U}_{\text{wp}} \equiv e^{i\phi_o}|o\rangle\langle o| + e^{i\phi_e}|e\rangle\langle e|, \tag{C.2}$$

where $|o\rangle$ and $|e\rangle$ represent the ordinary and extraordinary polarizations of photons inside the waveplate which therefore make up an orthonormal polarization basis. Because a global phase can always be ignored, the defining characteristic of a waveplate is therefore the phase difference, or retardation,

$$\Delta\phi \equiv (\phi_o - \phi_e) \bmod 2\pi. \tag{C.3}$$

The most common forms of waveplates are half waveplates ($\Delta\phi = \frac{\pi}{2}$) and quarter waveplates ($\Delta\phi = \frac{\pi}{4}$).

More information on waveplates as state manipulation tools can be found in Example 4.2, and additional information on waveplates as used in state projections can be found in Section 5.1.1. The remainder of this Appendix details strategies for dealing with imperfect waveplates.

Compensating for Imperfect Waveplates

While the previous Chapters showed that it was possible using a quarter- and half- waveplate to project into an arbitrary single qubit state, perfect quarter- and half-waveplates are experimentally impossible to obtain. More likely, the experimenter will have access to waveplates with known retardances slightly different than the ideal values of π (HWP) and $\frac{\pi}{2}$ (QWP). Even in this case, it is often possible to obtain arbitrary single-qubit projections. (Note that this is the

second solution to the problem of imperfect waveplates. Imperfect waveplates could be used at virtually any angles during a tomography—such as the same angles at which perfect waveplates would be used to measure in the canonical basis—resulting in a set of biased bases. The tomography mathematics have already been shown to function for either mutually biased or unbiased bases, as long as the set of bases is complete. In contrast, this section describes how—*even using imperfect waveplates*—one can still measure in the canonical, mutually unbiased bases.)

Analytically finding the angles where this is possible proves to be inconvenient and, for some waveplates, impossible. Rather than solve a system of equations based on the unitary waveplate matrices, we will examine the effect of these waveplates graphically using the Poincaré sphere. For the remainder of this discussion, we will assume that the experimenter has access to two waveplates, WP_1 and WP_2, which will respectively take the place of the QWP and HWP normally present in the experimental setup. We constrain the retardances of these waveplates to be $0 \leq \phi_1 \leq \phi_2 \leq \pi$.

In order to project into an arbitrary state $|\psi\rangle$, WP_1 and WP_2 must together rotate the state $|\psi\rangle$ into the state $|H\rangle$ (assuming a horizontal polarizer is used after the waveplates—any linear polarizer is equivalent). Taking a piecewise approach, first consider which states are possible after acting on the input state $|\psi\rangle$ with WP_1. Figure C.1a shows several example cases on the Poincaré sphere, each resulting in curved band of possible states that can be reached by varying the orientation of WP_1. Next consider which states could be rotated by WP_2 into the target state $|H\rangle$. Figure C.1b shows several examples of these states, which also take the form of a curved band, traversed by varying the orientation of WP_2. In order for state $|\psi\rangle$ to be rotatable into state $|H\rangle$, these two bands of potential states (shown in Figures C.1a and C.1b) must overlap.

Briefly examining the geometry of this system it appears that for most states this will be possible as long as the waveplate phases do not differ too much from the ideal HWP and QWP. Further consideration reveals that it is sufficient to be able to project into the states on the H-R-V-L great circle. There are two conditions under which this will *not* occur. First, if WP_2 is too close to a HWP, with WP_1 far from a QWP, the states at the poles (close to $|R\rangle$ and $|L\rangle$) will be unreachable from $|H\rangle$ (see Figure C.1c). Quantifying this condition:

$$2\left|\frac{\pi}{2} - \phi_1\right| \leq \pi - \phi_2. \tag{C.4}$$

Put another way, the error in the QWP must be less than half the error in the HWP. Second, the combined retardances from both waveplates can be insufficient to reach $|V\rangle$ (see Figure C.1c):

$$\phi_1 + \phi_2 \geq \pi. \tag{C.5}$$

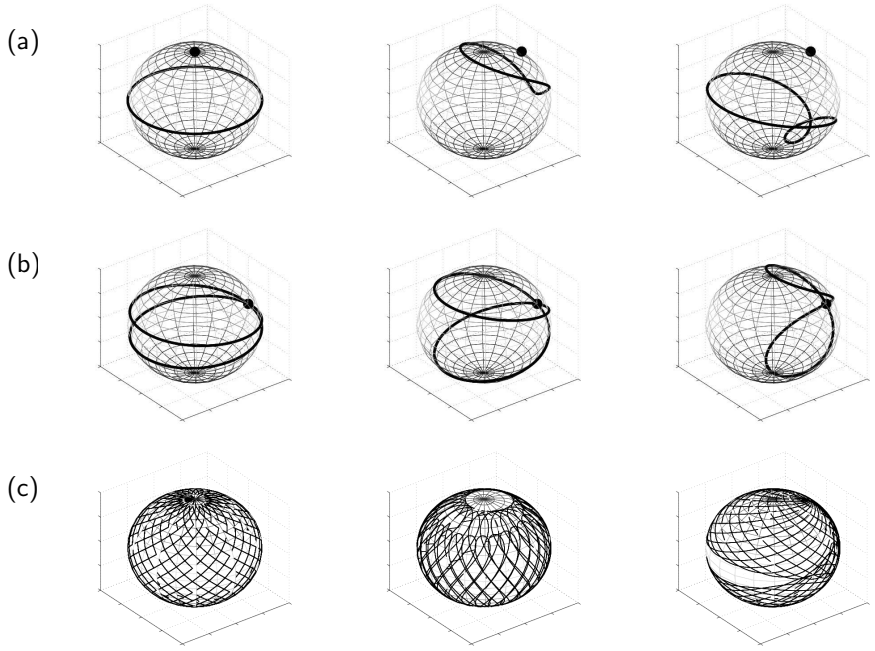

Figure C.1: Possible projectors simulated by waveplates and a stationary polarizer, graphically shown on Poincaré spheres. (a) WP_1, depending on its orientation, can rotate an incoming state into a variety of possible output states. Shown here on three Poincaré spheres are an initial incoming state (represented by a solid dot) and the set of all output states that WP_1 can rotate it into (represented by a dark band on the surface of the sphere). From left to right, the spheres depict $|R\rangle$ transformed by a $\frac{\pi}{2}$-waveplate, $|\gamma\rangle = cos\left(\frac{\pi}{8}\right)|H\rangle + isin\left(\frac{\pi}{8}\right)|V\rangle$ transformed by a $\frac{\pi}{3}$-waveplate, and $|\gamma\rangle$ transformed by a $\frac{2\pi}{3}$-waveplate. (b) WP_2, depending on its orientation, can rotate a variety of states into the target state $|H\rangle$. Shown here from left to right are the states able to be rotated into $|H\rangle$ by a $\frac{11\pi}{12}$-waveplate, a $\frac{3\pi}{4}$-waveplate, and a $\frac{\pi}{2}$-waveplate. (c) The possible projectors able to be produced by two waveplates and a horizontal polarizer. A series of arcs blanketing the Poincaré sphere show the areas of the sphere representing achievable projectors for each waveplate combination. From left to right, the spheres show the states (in this case, all of them) accessible from an ideal QWP and HWP, the states accessible using $\frac{\pi}{3}$- and $\frac{11\pi}{12}$-waveplates (groups of states near the poles are inaccessible), and the states accessible using $\frac{\pi}{3}$- and $\frac{3\pi}{5}$-waveplates (states on the equator are inaccessible). Note that the spheres shown in (c) are *not* simply combinations of the spheres above it, but include retardance values chosen to illustrate the possible failure modes of imperfect waveplates.

Given these two conditions, numerical simulations confirm that arbitrary single-qubit projectors can be constructed with two waveplates.

To clarify, as discussed previously, one does not *require* arbitrary single-qubit projectors, since an accurate tomography can be obtained with *any* set of linearly independent projectors as long as they are known. In fact, one advantage to this approach is that the exact same tomography measurement system can be used on photons with different wavelengths (on which the waveplates' birefringent phase retardances depend), simply by entering in the analysis program what the actual phase retardances are at the new wavelengths [58].

Wedged Waveplates It is an experimental reality that all commercially available waveplates have some degree of *wedge* (i.e., the surfaces of the waveplate are not parallel). This leads to a number of insidious difficulties which the experimenter must confront, grouped into two categories: (1) The thickness of the waveplate will change along its surface, providing a corresponding change in the phase retardance of the waveplate. This means that during a tomography when the waveplate is routinely rotated to different orientations, its total phase after rotation will change according to a much more complex—and often very difficult to calculate—formula. (In fact, if a large collection aperture is used, then different parts of the beam will experience different phase shifts.) (2) The direction (k-vector) of a beam will be deflected after passing through a wedged waveplate. This deflection will again depend on waveplate orientation, therefore changing throughout a tomography. This can have the effect of changing detector efficiencies (if, as in our case, a lens is used to focus to a portion of a very small detector area, different pieces of which have different efficiencies). This deflection will also affect any interferometric effects that depend on the beam direction being stable under waveplate rotation. Some of these problems can be mitigated (e.g., by taking care to pass through the center of the waveplate), but the best solution is to select waveplates with faces very close to parallel.

D Notes on Other Experimental Equipment

This Appendix contains notes on specific equipment used during the described experiments, and is intended mainly as a personal reference.

Counting electronics

We use NIM (nuclear instrumentation modules) for coincidence counting. From ORTEC, we obtained the NIM bin (powers and holds most of the electronics, part number 4001C), the inverting transformers (part number IT100), the quad counters (part number 974), and the dual counters (part number 994). From Phillips Scientific we obtained the discriminator (part number 711), the dual delay line (part number 792), and the octal logic unit which contains the actual coincidence logic (part number 758). We typically used coincidence windows between 1 and 10 ns.

Detectors

We used SPCM-CD 2801 (also known as AQR 13 or AQR 14, where the number indicates the dark count rate of the detector) Silicon avalanche photodiodes as detectors. They typically have a quantum efficiency of $\sim 65\%$. They have a typical dead-time of between 30 and 50 ns, after which there is after-pulsing. The probability of after-pulsing can vary widely between individual detectors, but should be less than 1%. The four-channel power supply for the detectors is the Agilent 6627A. The two-channel version of the same supply is the Agilent E3646A.

Down-conversion crystals

Necessary for our entanglement production, we have used a variety of down-conversion crystals. Almost all have been β-Barium Borate (BBO), purchased from Casix, NewLight Photonics, Eksma, Photox, and Cleveland Crystals. NewLight Photonics appear to be the best, with the lowest background coatings coming from MLD. BiBO is a new material we are just starting to switch to, which appears to have up to 2.5 times brighter output (13400 coincidences/s/mW as opposed to 5400 coincidences/s/mW). Unfortunately, BiBO appears to have higher background singles. BiBO has 1.5% background and BBO has 0.75% background, measured as the percentage of emitted singles counts which are

not associated with coincidences and not dependent on input pump polarization.

The typical dimensions for BBO crystals are 0.6mm × 8mm × 8mm, with an optic axis 33.9° degrees from crystal normal. The BiBO crystals (which are biaxial) are cut at $\theta = 140.97°$ and $\phi = 90°$. Both of these cuts are for a 3° half opening angle in free space for the down-converted cone.

Waveplates

Because wedged waveplates have been a consistent problem, we have ordered waveplates from several companies, including: Casix, CVI, Meadowlark Newport, OFR, Special Optics, and Tower Optics. By far the best results were obtained with Special Optics waveplates. This may be due in part to the Special Optics waveplates being air-spaced instead of glued, which we believe contributes to a low wedge angle.

References

[1] A. F. Abouraddy, A. V. Sergienko, B. E. A. Saleh, and M. C. Teich. Quantum entanglement and the two-photon stokes parameters. *Optics Communications*, 201:93, 2002.

[2] L. Allen, Stephen M. Barnett, and Miles J. Padgett. *Optical Angular Momentum*. Institute of Physics Publishing, London, 2004.

[3] J. B. Altepeter, D. Branning, E. Jeffrey, T.-C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White. Ancilla-assisted quantum process tomography. *Phys. Rev. Lett.*, 90:193601, 2003.

[4] J. B. Altepeter, P. G. Hadley, S. M. Wendelken, A. J. Berglund, and P. G. Kwiat. Experimental Investigation of a Two-Qubit Decoherence-Free Subspace. *Physical Review Letters*, 92(14):147901–+, April 2004.

[5] Joseph B. Altepeter, Daniel F. V. James, and Paul G. Kwiat. *Quantum State Estimation*, volume 649 of *Lecture Notes in Physics*. Springer, Berlin, 2004.

[6] Joseph B. Altepeter, E. R. Jeffrey, P. G. Kwiat, S. Tanzilli, N. Gisin, and A. Acin. Experimental methods for detecting entanglement. *submitted to Phys. Rev. Lett.*, 2005.

[7] Joseph B. Altepeter, Evan R. Jeffrey, and Paul G. Kwiat. Phase compensated ultra-bright source of entangled photons. *Optics Express*, 13(22), 2005.

[8] Joseph B. Altepeter, Evan R. Jeffrey, and Paul G. Kwiat. *Advances in AMO Physics, 2006, Chapter 3: Photonic State Tomography*. Elsevier, Munich, Germany, 2006.

[9] H. H. Arnaut and G. A. Barbosa. Orbital and intrinsic angular momentum of single photons and entangled pairs of photons generated by parametric down-conversion. *Phys. Rev. Lett.*, 85:286–289, 2000.

[10] D. Bacon, D. A. Lidar, and K. B. Whaley. Robustness of decoherence-free subspaces for quantum computation. *Physical Review A*, 60:1944–1955, September 1999.

[11] Julio T. Barreiro, Nathan K. Langford, Nicholas A. Peters, and Paul G. Kwiat. Generation of hyperentangled photon pairs. *Phys. Rev. Lett.*, 95:260501, 2005.

[12] J. S. Bell. On the einstein-podolsky-rosen paradox. *Physics*, 1:195–200, 1964.

[13] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. of the IEEE Int. Conf. on Computers, Systems, and Signal Processing, Bangalore, India*, Proc. IEEE, page 175. IEEE, New York, 1984.

[14] Charles H. Bennett and Peter W. Shor. Quantum information theory. *IEEE Transactions on Information Theory*, 44:2724–2742, October 1998.

[15] A. J. Berglund. Quantum coherence and control in one- and two-photon optical systems. *ArXiv Quantum Physics e-prints*, September 2000.

[16] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, Oxford, U.K., 1987.

[17] Mark D. Bowdrey, Daniel K. L. Oi, Anthony J. Short, Konrad Banaszek, and Jonathan A. Jones. Fidelity of single qubit maps. *Physics Letters A*, 294:258, 2002.

[18] Robert Boyd. *Nonlinear Optics*. Elsevier, London, 2003.

[19] V. Buzek, M. Hillery, and R. Werner. Optimal Manipulations with Qubits: Universal NOT Gate. *ArXiv Quantum Physics e-prints*, January 1999.

[20] V. Coffman, J. Kundu, and W. K. Wootters. Distributed entanglement. *Phys. Rev. A*, 61:052306, 2000.

[21] D. Collins and N. Gisin. A relevant two qubit Bell inequality inequivalent to the CHSH inequality. *J. Phys. A*, 37:1775, 2004.

[22] Thomas H Cormen, Charles E. Leiserson, Ronald L. Rivest, and Cliff Stein. *Introduction to Algorithms (Second Edition)*. MIT Press, London, U.K., 2001.

[23] D. G. Cory, A. F. Fahmy, and T. F. Havel. Ensemble quantum computing by nmr spectroscopy. *Proceedings of the National Academy of Sciences*, 94:1634, 1997.

[24] V. G. Dmitriev, G. G. Gurzadyan, and D. N. Nikogosyan. *Handbook of Nonlinear Optical Crystals*. Springer, Berlin, 1999.

[25] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47:777–780, May 1935.

[26] Stuart J. Freedman and John F. Clauser. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.*, 28:938–941, 1972.

[27] S. Gasiorowitz. *Quantum Physics*. John Wiley and Sons, Inc., 1996.

[28] Nicolas Gisin, Grgoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74:145, 2002.

[29] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79:325, 1997.

[30] E. Hecht. *Optics*. Addison-Wesley, Reading, MA, 1974.

[31] C. K. Hong and L. Mandel. Experimental realization of a localized one-photon state. *Phys. Rev. Lett.*, 56:58, 1986.

[32] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223:1–8, 1996.

[33] Zdenek Hradil and Jaroslav Rehacek. Efficiency of maximum-likelihood reconstruction of quantum states. *Fortschr. Phys.*, 49:1083, 2001.

[34] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Phys. Rev. A*, 64:052312, 2001.

[35] J. A. Jones, M. Mosca, and R. H. Hansen. Implementation of a quantum search algorithm on a quantum computer. *Nature*, 392:344, 1997.

[36] Richard Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41:2315–2323, 1994.

[37] D. Kielpinski, V. Meyer, M. A. Rowe, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. A decoherence-free quantum memory using trapped ions. *Science*, 291:1013, 2001.

[38] Paul G. Kwiat, Andrew J. Berglund, Joseph B. Altepeter, and Andrew G. White. Experimental verification of decoherence-free subspaces. *Science*, 290:498–501, 2000.

[39] Paul G. Kwiat, Edo Waks, Andrew G. White, Ian Appelbaum, and Phillipe H. Eberhard. Ultrabright source of polarization-entangled photons. *Phys. Rev. A*, 60:R773–R776, 1999.

[40] R. Laflamme, E. Knill, D.G. Cory, E.M. Fortunato, T. Havel, C. Miquel, R. Martinez, C. Negrevergne, G. Ortiz, M.A. Pravia, Y. Sharf, S. Sinha, R. Somma, and L. Viola. Introduction to nmr quantum information processing. *quant-ph/0207172*, 2002.

[41] N. K. Langford, R. B. Dalton, M. D. Harvey, J. L. O'Brien, G. J. Pryde, A. Gilchrist, S. D. Bartlett, and A. G. White. Measuring entangled qutrits and their use for quantum bit commitment. *Phys. Rev. Lett.*, 93:053601, 2004.

[42] Jay Lawrence, Caslav Brukner, and Anton Zeilinger. Mutually unbiased binary observable sets on n qubits. *Phys. Rev. A*, 65:032320, 2002.

[43] J. Lehner, U. Leonhardt, and H. Paul. Unpolarized light: Classical and quantum states. *Phys. Rev. A*, 53:2727, 1996.

[44] U. Leonhardt, editor. *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge, UK, 1997.

[45] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki. Optimization of entanglement witnesses. *Phys. Rev. A.*, 62:052310, 2000.

[46] D. A. Lidar, I. L. Chuang, and K. B. Whaley. Decoherence-Free Subspaces for Quantum Computation. *Physical Review Letters*, 81:2594–2597, September 1998.

[47] Alois Mair, Alipasha Vaziri, Gregor Weihs, and Anton Zeilinger. Entanglement of the orbital angular momentum states of photons. *Nature*, 412:313, 2001.

[48] Francesco De Martini, Andrea Mazzei, Marco Ricci, and Giacomo Mauro D'Ariano. Pauli tomography: complete characterization of a single qubit device. *quant-ph/0207143*, 2002.

[49] A. Migdall. Polarization directions of noncollinear phase-matched optical parametric downconversion output. *J. Opt. Soc. Am. B*, 14:1093–1098, May 1997.

[50] Chris Monroe. Quantum information processing with atoms and photons. *Nature*, 416:238, 2002.

[51] W. J. Munro, D. F. V. James, A. G. White, and P. G. Kwiat. Maximizing the entanglement of two mixed qubits. *Phys. Rev. A*, 64:030302, 2001.

[52] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.

[53] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines. Quantum dynamics as a physical resource. *Physical Review A*, 67(5):052301–+, May 2003.

[54] Michael Nielsen. Phd thesis: Quantum information theory. *quant-ph/0011036*, 1998.

[55] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. James, N. K. Langford, T. C. Ralph, and A. G. White. Quantum Process Tomography of a Controlled-NOT Gate. *Physical Review Letters*, 93(8):080502–+, August 2004.

[56] Jian-Wei Pan, Dik Bouwmeester, Harald Weinfurter, and Anton Zeilinger. Experimental entanglement swapping: Entangling photons that never interacted. *Phys. Rev. Lett.*, 80:38913894, 1998.

[57] Nicholas A. Peters, Joseph B. Altepeter, David Branning, Evan R. Jeffrey, Tzu-Chieh Wei, and Paul G. Kwiat. Maximally entangled mixed states: Creation and concentration. *Physical Review Letters*, April 2004.

[58] Nicholas A. Peters, Julio T. Barreiro, Michael E. Goggin, Tzu-Chieh Wei, and Paul G. Kwiat. Remote state preparation: arbitrary remote control of photon polarization. *to appear in Phys. Rev. Lett.*, 2005.

[59] Nick Peters, Joseph B. Altepeter, Evan R. Jeffrey, David Branning, and Paul G. Kwiat. Precise creation, characterization and manipulation of single optical qubits. *Quantum Information and Computation*, 3:503, 2003.

[60] J. M. G. Sancho and S. F. Huelga. Measuring the entanglement of bipartite pure states. *Phys. Rev. A*, 61:042303, 2000.

[61] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Gulde, G. P. T. Lancaster, T. Deuschle, C. Becher, C. F. Roos, J. Eschner, and R. Blatt. Realization of the ciraczoller controlled-not quantum gate. *Nature*, 422:408, 2003.

[62] P. W. Shor. Algorithms for quantum computation: Discrete logarithm and factoring. In S. Goldwasser, editor, *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, page 116. IEEE Computer Society Press, Los Alamitos, CA, 1994.

[63] G. C. Stokes. On the composition and resolution of streams of polarized light from dierent sources. *Trans. Cambr. Phil. Soc.*, 9:399, 1852.

[64] B. M. Terhal. Detecting quantum entanglement. *quant-ph/0101032*, 2001.

[65] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro. Qudit quantum-state tomography. *Phys. Rev. A*, 66:012303, 2002.

[66] L. Viola, E. Knill, and S. Lloyd. Dynamical Decoupling of Open Quantum Systems. *Physical Review Letters*, 82:2417–2421, March 1999.

[67] T.-C. Wei and P. M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Physical Review A*, 68(4):042307, October 2003.

[68] Tzu-Chei Wei, Joseph B. Altepeter, David Branning, Paul M. Goldbart, D. F. V. James, Evan R. Jeffrey, Paul G. Kwiat, Swagatam Mukhopadhyay, and Nicholas A. Peters. Synthesizing arbitrary two-photon polarization mixed states. *Physcal Review A*, 71:032329, 2005.

[69] Y. S. Weinstein, M. A. Pravia, E. M. Fortunato, S. Lloyd, and D. G. Cory. Implementation of the quantum fourier transform. *Phys. Rev. Lett.*, 86:1889, 2001.

[70] R. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277, 1989.

[71] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat. Nonmaximally entangled states: Production, characterization, and utilization. *Phys. Rev. Lett.*, 83:3103, 1999.

[72] A. G. White, D. F. V. James, W. J. Munro, and P. G. Kwiat. Exploring hilbert space: Accurate characterization of quantum information. *Phys. Rev. A*, 65:012301, 2002.

[73] W. K. Wooters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245, 1998.

[74] Sixia Yu, Jian-Wei Pan, Zeng-Bing Chen, and Yong-De Zhang. Comprehensive test of entanglement for two-level systems via the indeterminacy relationship. *Phys. Rev. Lett.*, 91:217903, 2003.

[75] A. Zeilinger. Experiment and Foundations of Quantum Physics. In B. Bederson, editor, *More Things in Heaven and Earth : A Celebration of Physics at the Millennium*, pages 482–+, 1999.

[76] K. Zyczkowski. Geometry of entangled states. *quant-ph/0006068*, 2000.

[77] K. Zyczkowski. cp^n, or, entanglement illustrated. *quant-ph/0108064*, 2001.

Author's Biography

Joe Altepeter was born and raised in St. Louis, Missouri, where his parents Shirley and Mike still live. He has a younger brother, Greg, and a younger sister, Stephanie.

D.1 Awards and honors

Third Place Team: Charlie Townes *Amazing Light* Competition	*2005*
Drickamer Research Fellowship	*2005*
Mavis Memorial Fund Scholarship for Engineering Education	*2004*
National Science Foundation Graduate Research Fellow	*2001 - 2005*
Excellence in Teaching Award, every eligible semester	*1997 - 2003*
Fulbright Scholar, Australia	*2001*
Myers Scholar (Washington Univ. Full Scholarship)	*1996 - 2000*
NNUN Stanford Summer Fellow	*1998*
Washington University Dean's List, every semester	*1996 - 2000*
National Merit Scholar	*1996*

D.2 Publications and patents

Publications Summary: 1 patent pending, 2 book chapters, 1 *Science*, 5 *Physical Review Letters*, 2 *Physical Review A*, 1 *Optics Express*, 4 conference proceedings. In addition, he has acted as a referee for *Physical Review Letters*.

E. R. Jeffrey, J. B. Altepeter, M. Colci, and P. G. Kwiat, "Optical implementation of quantum orienteering", *Physical Review Letters* **96**, 150503 (2006).

J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Phase-compensated ultrabright source of entangled photons", *Optics Express* **13**, 22 (2005).

J. B. Altepeter, E. R. Jeffrey, P. G. Kwiat, S. Tanzilli, N. Gisin, and A. Acin, "Experimental Methods for Detecting Entanglement", *Physical Review Letters* **95**, 033601 (2005).

J. B. Altepeter, D. F. V. James, and P. G. Kwiat, "Photonic Qubit Tomography", Chapter, *Advances in Atomic, Molecular, and Optical Physics, Vol.*

52 , P. Berman, ed. (Elsevier 2005).

T. C. Wei, J. B. Altepeter, D. Branning, P. M. Goldbart, D. F. V. James, E. Jeffrey, P. G. Kwiat, S. Mukhopadhyay, and N. A. Peters, "Synthesizing arbitrary two-photon polarization mixed states", *Physical Review A* **71**, 032329 (2005).

J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Quantum Random Number Generator", *Patent Pending*, Application No. 10/885503, (2004).

J. B. Altepeter, D. F. V. James, and P. G. Kwiat, "Qubit Quantum State Tomography", Chapter 4, *Lecture Notes in Physics: Quantum Detection and Estimation – Theory and Experiment*, M. Paris and J. Rehacek, eds. (Springer 2004).

T.-C. Wei, J. B. Altepeter, P. M. Goldbart, and W. J. Munro, "Measures of entanglement in multipartite bound entangled states", *Physical Review A* **70**, 022322 (2004).

J. B. Altepeter, P. G. Hadley, S. M. Wendelken, A. J. Berglund, and P. G. Kwiat, "Experimental Investigation of a Two-Qubit Decoherence-Free Subspace", *Physical Review Letters* **92**, 147901 (2004).

N. A. Peters, J. B. Altepeter, D. Branning, E. R. Jeffrey, T. C. Wei, and P. G. Kwiat, "Maximally Entangled Mixed States: Creation and Concentration", *Physical Review Letters* **92**, 133601 (2004).

P. G. Kwiat, J. B. Altepeter, J. Barreiro, D. Branning, E. Jeffrey, N. A. Peters, and A. VanDevender, "Optical Technologies for Quantum Information Science", *Proc. SPIE* 5161 (2004).

J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, Paul Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, "Ancilla-Assisted Quantum Process Tomography", *Physical Review Letters* **90**, 193601 (2003).

N. A. Peters, J. B. Altepeter, E. Jeffrey, D. Branning, and P. G. Kwiat, "Precise Creation, Characterization, and Manipulation of Single Optical Qubits", *Quantum Information and Computation* **3**, 503 (2003).

P. G. Kwiat, J. B. Altepeter, D. Branning, E. Jeffrey, N. Peters, and T. C. Wei, "Taming Entanglement", *Proceedings of the 6th International Conference on Quantum Communication, Measurement, and Computing* (QCMC '02), J. H. Shapiro and O. Hirota, eds. Rinton Press, 117 (2003).

P. G. Kwiat, A. J. Berglund, J. B. Altepeter, and A. J. White, "Experimental Verification of Decoherence-Free Subspaces", *Science* **290**, 498 (2000).

R. P. Loui, J. Norman, J. Altepeter, D. Pinkard, D. Craven, J. Linsday, M. A. Foltz, "Progress on Room 5: A Testbed for Public Interactive Semi-Formal Legal Argumentation", *Proceedings of the Sixth International Conference on Artificial Intelligence and Law* (ICAIL '97), 207-214 (1997).