IN SEARCH OF PHOTONIC BOUND ENTANGLEMENT: USING
HYPERENTANGLEMENT TO STUDY MIXED ENTANGLED STATES

BY

ADITYA NILAKANTAN SHARMA

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, 2016

Urbana, Illinois

Doctoral Committee:

 Professor John D. Stack, Chair
 Professor Paul G. Kwiat, Director of Research
 Assistant Professor Virginia Lorenz
 Professor Richard L. Weaver

# Abstract

Quantum entanglement exhibits various interesting features that emerge only in high-dimensional systems. One of the most fascinating is bound entanglement, entanglement that cannot be extracted using local operations and classical communication. This thesis describes our work towards an experimental realization of the four-qubit bound-entangled Smolin state, using the polarization and spatial mode of photon pairs. We describe a number of interesting experimental challenges that this work faced. We present our results on entangled two-qubit spatial mode states and hyperentangled four-qubit polarization-spatial-mode states generated by spontaneous parametric down-conversion from a second-order TEM mode pump. Some of the subtleties involved in preparing genuine mixed states in the lab are discussed.

*To Mom and Dad.*

# Acknowledgments

*Gaily bedight,*
*A gallant knight,*
*In sunshine and in shadow,*
*Had journeyed long,*
*Singing a song,*
*In search of Eldorado.*

*But he grew old —*
*This knight so bold —*
*And o'er his heart a shadow*
*Fell as he found*
*No spot of ground*
*That looked like Eldorado.*

*And, as his strength*
*Failed him at length,*
*He met a pilgrim shadow —*
*"Shadow," said he,*
*"Where can it be —*
*This land of Eldorado?"*

*"Over the Mountains*
*Of the Moon,*
*Down the Valley of the Shadow,*
*Ride, boldly ride,"*
*The shade replied —*
*"If you seek for Eldorado!"*

I first encountered this poem by Edgar Allan Poe when I was 12 years old, and at the time I read it in a sing-song way, interpreting it literally as some sort of fantasy. When it resurfaced in my consciousness a few years ago, the symbolism had greater relevance to me, and it acquired a much more grim tone. Grad school has been the toughest period of my life, and on several occasions I've shared the desperation of Poe's hapless horseman. Luckily for me, however, I haven't had to turn to ghosts for advice and encouragement.

Most of all, I thank my parents, who prioritized my (and my sister's) education above all else. They did their best to instill a good work ethic and spirit of inquiry in me, and I think they did pretty well (if I do say so myself!).

I thank Paul Kwiat for giving me my start in this field and for having faith in my work. I have learned so much from him about physics, problem solving, communication, and a number of other things, and looking back I'm glad I joined his research group.

I thank Soyoung Baek and Mike Goggin for each stepping in to help with the experiment, and committing a great deal of time, effort, and thought to it. Their hard work and attention to detail were inspiring at times when my own enthusiasm was flagging, and it was a pleasure to work with them.

I thank all of my labmates over the past several years for all they taught me, for staying out of the lab when I was taking measurements, and for the fun times we shared in and out of the lab.

I thank Anika Jain for constantly being there for me and for adding new dimensions to my life. I'm glad she indulged my idea of walking all the way to Meadowbrook Park to see fireflies.

I thank Arundathi Sharma for always being up for an adventure and for telling me good stories. I'm sure they will only continue to get more interesting in the future.

I thank all of my friends, especially Shah-Parvez Jamal, who is somehow always able to lift my spirits.

Finally, thanks to John Stack and Lance Cooper for their advice at the beginning and end, respectively, of my time as a grad student. John helped me find the lab that I eventually joined, and Lance helped me through an especially difficult period of grad school with thoughtful and practical advice.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| AL | Alignment laser |
| BBO | Barium borate |
| BE | Bound entanglement/entangled |
| BiBO | Bismuth triborate ($BiB_3O_6$) |
| BS | Beam-splitter |
| CCW | Counter-clockwise |
| CNOT | Controlled-NOT |
| CW | Clockwise |
| EPR | Einstein, Podolsky, and Rosen |
| F | Fidelity |
| FI | Faraday isolator |
| HWP | Half-wave plate |
| IF | Interference filter |
| LHS | Left-hand side |
| LOCC | Local operations and classical communication |
| MZI | Mach-Zehnder interferometer |
| NPT | Negative partial transpose |
| OAM | Orbital angular momentum |
| PBS | Polarizing beam-splitter |
| PM | Piezoelectric-actuated mirror |
| PPT | Positive partial transpose |
| PT | Partial transpose |
| RHS | Right-hand side |
| RNG | Random-number-generated |
| QC | Quadrant photodiode |

| | |
|---|---|
| QI | Quantum information |
| QIS | Quantum information science |
| QWP | Quarter-wave plate |
| QST | Quantum state tomography |
| SLM | Spatial light modulator |
| SMF | Single-mode fiber |
| SPDC | Spontaneous parametric downconversion |
| T | Tangle |

# Chapter 1

# Introduction

In their 1935 paper, Albert Einstein, Boris Podolsky, and Nathan Rosen (EPR) exposed a peculiarity of quantum theory [1]: when certain quantum states of compound systems are subjected to space-like separated measurements on their constituent parts, the wavefunction collapse described by the Copenhagen interpretation implies that each subsystem must "know" the basis and the outcome of the measurement on the other *instantaneously*, in apparent contradiction with special relativity. EPR regarded this conclusion as an artifact of mistaken assumptions and intended their work as a proof by contradiction that the theory is incomplete. Instead, they had discovered one of the most striking features of quantum mechanics, which Erwin Schrödinger, writing later that year, dubbed "*entanglement*" [2]. In the eighty years since EPR's work, Bell's theorem and numerous experiments have led to a consensus that entanglement is, in fact, a very real physical phenomenon (the measurement correlations that EPR pointed out are indeed instantaneous [3], but the contradiction with special relativity can be resolved —there is no possibility for superluminal communication).

Entanglement, this strongest of quantum mechanical correlations, has grown into a subject of active research that continues to fascinate us with its puzzling properties. Furthermore, we have come to recognize it as a technological resource, the sleight of hand behind many of the magic tricks of quantum information science (QIS). Entanglement allows us to seemingly transmit two bits for the price of one [4], or transmit a complex number at the cost of only two bits [5]. We can use it to harness the very randomness that so disturbed EPR for cryptography that is in principle unbreakable [6].

In 1927, even before EPR had articulated why (they believed) this quantum randomness is unphysical, John von Neumann introduced the idea of additional *classical* uncertainty in quantum mechanics [7]. As the randomness we encounter in quantum systems is so counterintuitive precisely because of its divergence from classical probability theory, it is ironic that von Neumann's "mixture" so complicates the study of entanglement. Even the simplest entangled systems can exhibit complex behavior in the presence of mixture [8]. In larger systems, there is still no clear way of determining whether or not a given mixed state is entangled at all [9]. Additionally, certain features of entanglement only emerge in mixed states of sufficiently

high dimension: one of the strangest such phenomena, is bound entanglement (BE). BE is entanglement that has been weakened almost to the point of becoming useless, but yet retains vestiges of its original capabilities, at times performing as well as maximal entanglement.

This thesis discusses our attempts to demonstrate BE. Our experiment uses the polarization and first-order spatial mode of hyperentangled photon pairs generated by spontaneous parametric downconversion (SPDC) to encode a four-qubit bound-entangled quantum state. **Chapter 2** introduces the relevant experimental tools and should make the previous sentence comprehensible. **Chapter 3** presents the theory motivating these experiments; before diving into the details of BE, we introduce the simplest case of pure-state entanglement (**Section 3.1**) and explain the physical significance and mathematical representation of mixed states (**Section 3.2**). The reader not already familiar with this material might prefer to read these sections before **Section 2.2**. **Chapter 4** presents the details of our experiment. One of the highlights of our work is a careful consideration of mixture, a detail that is crucially lacking in previous photonic BE experiments. We therefore hope that the discussion in **Section 4.5** will be interesting and illuminating.

# Chapter 2

# Qubits

The feature that distinguishes QIS from its classical counterpart is the substitution of quantum two-level systems —qubits —for classical bits. While classical bits can take on only the logic states "0" and "1", qubits can additionally assume quantum superposition states of the two logic values $|0\rangle$ and $|1\rangle$, e.g., $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$; furthermore, multiple qubits can exist in joint superposition, e.g., $\frac{|0,0\rangle+|1,1\rangle}{\sqrt{2}}$, a phenomenon called *entanglement*. In principle, we could use *any* quantum two-level system as a qubit, but in practice each physical system has certain advantages and drawbacks. For the specific case of quantum computing, David DiVincenzo proposed five requirements any candidate architecture must satisfy [10]:

1. The qubits must be well-defined, and it must be practical to scale up the computer.

2. It must be possible to initialize the qubit register to some fiducial state like $|000\ldots\rangle$.

3. The qubits must have coherence time much longer than the time required to execute gate operations.

4. It must be possible to execute a universal set of quantum gates, including two-qubit operations.

5. It must be possible to accurately measure the qubits to read out the state of the computer.

The various competing physical systems each have important weaknesses with respect to these criteria, and it is still unclear which would work best as the basis of a quantum computer. For quantum *communication*, on the other hand, photonic qubits are the most natural choice: thanks to photons' long coherence time, they can be transmitted over existing optical fiber networks at speeds on the order of $c$, over distances of over 100 kilometers. Loss becomes significant for longer distances, but addition of a matter-based quantum repeater technology could solve this problem, enabling intercontinental transmission of quantum information [11]. Admittedly, we are being somewhat simplistic here; as we see in this chapter, a single photon can carry quantum information in different physical degrees of freedom, and not all of these carriers are created equal. Nevertheless, *at their best*, photonic qubits offer this ease of transmission, and no other physical system currently being studied can compete in this regard. Section 2.1 introduces the three independent photonic

degrees of freedom and explains how we manipulate and measure the corresponding qubits. Section 2.2 shows how we prepare certain quantum states of these degrees of freedom.

## 2.1 Photonic degrees of freedom

To identify the photon's degrees of freedom, let's first consider light from a classical point of view: in free space, the Maxwell equations require that the electric field $\mathbf{E}(\mathbf{r}, t)$ satisfies [12]

$$(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2})\mathbf{E}(\mathbf{r}, t) = 0. \tag{2.1}$$

A plane wave solution to 2.1 is

$$\mathbf{E}(\mathbf{r}, t) = \hat{\mathbf{p}} \cdot E e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \hat{\mathbf{p}}^* \cdot E^* e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.2}$$

where $\hat{\mathbf{p}} \cdot \mathbf{k} = 0$, $\omega = c|\mathbf{k}|$, and the second term is included simply to ensure that $\mathbf{E}$ is real. $\hat{\mathbf{p}}$, $\mathbf{k}$, and $\omega$ represent the *polarization*, *wave-vector*, and *frequency*, respectively: these are the three[1] independent degrees of freedom of a single photon (two other parameters of the photon, its position in space and its time of emission, can also play a role in completely specifying the state of the particle —however, these observables are not additional *independent* degrees of freedom because they are complementary to $\mathbf{k}$ and $\omega$, respectively. We often refer to the *wavelength* as well, but this is equivalent to $\omega$, $\lambda = 2\pi c/\omega$). In this section we will see how we can interpret these physical attributes of the photon as qubits.

### 2.1.1 Polarization

In general, polarization is the easiest degree of freedom to work with in the lab. As we can see from Eq. 2.2, $\hat{\mathbf{p}}$ is the direction in which the classical field oscillates. The requirement that $\hat{\mathbf{p}} \cdot \mathbf{k} = 0$ means that there are two independent values of $\hat{\mathbf{p}}$ for any choice of $\mathbf{k}$ and $\omega$: in the lab, we label light polarized parallel to the surface of the optical table horizontally polarized ($\hat{\mathbf{p}}_H$), and light polarized perpendicular to the table vertically polarized ($\hat{\mathbf{p}}_V$). Linear combinations of these states include $\pm 45°$-oscillating fields ($\hat{\mathbf{p}}_D, \hat{\mathbf{p}}_A = \frac{\hat{\mathbf{p}}_H \pm \hat{\mathbf{p}}_V}{\sqrt{2}}$), and left- and right-circular polarization ($\hat{\mathbf{p}}_L, \hat{\mathbf{p}}_R = \frac{\hat{\mathbf{p}}_H \pm i\hat{\mathbf{p}}_V}{\sqrt{2}}$). Quantum mechanically, polarization is simply the photon spin, and these polarization vectors correspond to single-photon spin states $|H\rangle, |V\rangle$ and $|D\rangle, |A\rangle \equiv \frac{|H\rangle \pm |V\rangle}{\sqrt{2}}$ and $|L\rangle, |R\rangle \equiv \frac{|H\rangle \pm i|V\rangle}{\sqrt{2}}$. As massless spin-1 particles, photons have two spin eigenstates: along the direction $\mathbf{k}$, these states are $|L\rangle$ and $|R\rangle$, with angular momentum $\pm\hbar$. Thus,

---

[1]Only two components of $\mathbf{k}$ are independent, since the magnitude $k$ is determined by $\omega$.
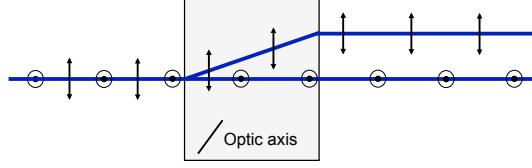
Figure 2.1: A beam displacer. Due to transverse walkoff, light polarized oblique to the principal axes of the refractive index propagates at a different angle inside birefringent media.

photon polarization perfectly exemplifies a qubit [13].

Arbitrary unitary operations[2] on the polarization can be executed easily using slabs of material with polarization-dependent refractive index, a property called *birefringence* [14]. In a *quarter-wave plate* (QWP) the thickness of the slab is such that it delays one polarization by one quarter of the wavelength $\frac{\lambda}{4}$ relative to the orthogonal polarization; in a half-wave plate (HWP), the delay is $\frac{\lambda}{2}$. We can accomplish various polarization transformations using QWPs and HWPs oriented at different angles. For example:

- $|H\rangle, |V\rangle \overset{\text{QWP at } 45°}{\leftrightarrow} |L\rangle, |R\rangle,$

- $|H\rangle, |V\rangle \overset{\text{HWP at } 45°}{\leftrightarrow} |V\rangle, |H\rangle,$

- $|H\rangle, |V\rangle \overset{\text{HWP at } 22.5°}{\leftrightarrow} |D\rangle, |A\rangle.$

Furthermore, an appropriately oriented QWP-HWP combination can convert any specific input state $|\psi\rangle$ to any desired output state $|\psi'\rangle$, and a QWP-QWP-HWP combination can execute arbitrary unitary operations.

It is also easy to make a projective measurement on the photon polarization and any device that does so is called a polarizer [14]. Polarizers come in many forms and can exploit any of several mechanisms that lead to high transmission of one polarization state and absorption (e.g., dichroic polarizer), reflection (e.g., Brewster-angle polarizer), deflection (e.g., Glan-Taylor polarizer), or spatial displacement (beam displacer) of the orthogonal state. The phenomenon behind the last of these has important implications for our experiments; a beam displacer takes advantage of *transverse walkoff* [15], in which two orthogonal polarizations can enter a birefringent medium together, propagate at different angles, and emerge spatially separated, as shown in Figure 2.1. We explain this effect in Appendix A.

### 2.1.2 Spatial mode

In practice, the plane-wave solution Eq. 2.2 can be inconvenient (the space of plane-wave states is continuous, so that infinitesimally different values of $\hat{\mathbf{k}}$ correspond to orthogonal states) and unphysical (the energy transport rate is infinite), so rather than work directly with position or momentum we use an orthonormal

---

[2]A unitary operation rotates the space of polarization states, taking any input state $|\psi\rangle \to |\psi'\rangle$. The defining feature of such operations is that they preserve the *purity* of all input states (see Section 3.2).

set of *spatial modes*. The defining feature of a mode is that the electric field amplitude retains the same transverse profile as it propagates; thus, free-space modes are eigenstates of the 2-D Fourier transform over their transverse spatial coordinates. One example is the *Hermite-Gauss* basis, which follows from separation of variables in Cartesian coordinates —it is the ideal choice when describing radiation in a rectangular cross-section laser cavity [16]. The electric field amplitudes of these modes, indexed by non-negative integers $m$ and $n$, have the following form for a beam propagating along the $z$ axis:

$$HG_{mn}(x,y,z) = \sqrt{\frac{2}{2^{m+n}\pi m!n!}}$$
$$\times H_m\left(\frac{\sqrt{2}x}{w(z)}\right) H_n\left(\frac{\sqrt{2}y}{w(z)}\right) \exp\left(-\frac{x^2+y^2}{w(z)^2}\right)$$
$$\times \exp\left(-i\left(kz + \frac{k(x^2+y^2)}{2R(z)} - \zeta(z)\right)\right). \quad (2.3)$$

where

$$H_m(x) = \left(2x - \frac{d}{dx}\right)^m \cdot 1, \quad (2.4a)$$

$$w(z) = w_0\sqrt{1 + (z/z_R)^2}, \quad (2.4b)$$

$$R(z) = z\left(1 + (z/z_R)^2\right), \text{ and} \quad (2.4c)$$

$$\zeta(z) = (m+n+1)\arctan\left(\frac{z}{z_R}\right). \quad (2.4d)$$

Let's consider Eq. 2.3 one piece at a time. The first factor is simply a normalization constant. The next three factors give (the square root of) the beam's intensity distribution in a plane transverse to its propagation: Eq. 2.4a is the Hermite polynomial of order $m$. Since $H_m$ has $m$ distinct real roots, the intensity of $HG_{mn}$ is zero for $m$ values of $x$ and $n$ values of $y$. Eq. 2.4b gives the "width" of the Gaussian part of the intensity distribution at location $z$ (i.e., the distance from the center of the cross-section to a point at which the beam intensity drops by a factor of $1/e^2$); $w_0 \equiv w(z=0)$ is the width of the beam at its *waist*, where it is narrowest, and $z_R \equiv \frac{\pi w_0^2}{\lambda}$, the *Rayleigh range*, is the propagation distance from the waist after which the beam width increases by a factor of $\sqrt{2}$. Only the *width* of the intensity distribution depends on $z$ —the shape is constant, as it must be for the $HG_{mn}$ to be legitimate modes. The intensity distributions of the first few $HG_{mn}$ are shown in Figure 2.2.

The last factor in Eq. 2.3 represents the spatial dependence of the phase. Here, the first term is the most significant contribution, the plane wave phase from Eq. 2.2: increasing $z$ by one wavelength advances this phase by $2\pi$. The second term arises because there are plane-wave components of the beam propagating
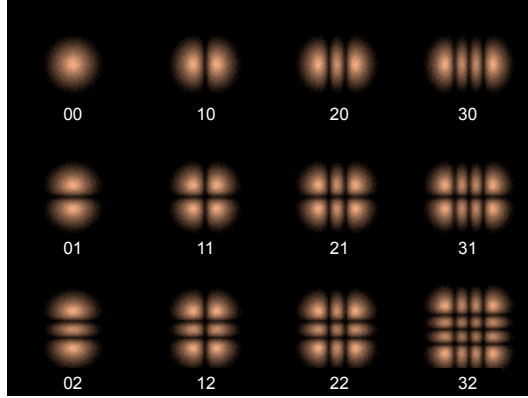
Figure 2.2: The first few Hermite-Gauss modes. The first and second indices indicate the Hermite polynomial associated with the $x$ and $y$ transverse directions, respectively. Therefore, $HG_{mn}$ has $m$ $(n)$ nulls along the horizontal (vertical) axis.

in the $x$ and $y$ directions as well, so the phase advances faster at the edges of the beam than it does along the central axis; Eq. 2.4c gives the resulting radius of curvature of the wavefront. The last term, Eq. 2.4d is called the Gouy phase. This topological phase plays an important role: since the phase factor of $HG_{mn}$ depends on $m, n$ only through the Gouy phase, arbitrary superpositions of modes of the same *order* $N = m+n$ are also modes of order $N$, while linear combinations of different-order modes change shape upon propagation. Evidently, equal-$N$ modes are degenerate propagation eigenstates with eigenvalue related to the Gouy phase. The total Gouy phase accumulated after propagation from $z = -\infty$ to $+\infty$ is $(N+1)\pi$.

Now let's return to our original problem: how do we store quantum information in the spatial degree of freedom? Although in principle we could encode an arbitrarily large number of logic levels in the spatial mode, in this work we want a qubit. It makes sense to use equal-order modes to encode $|0\rangle$ and $|1\rangle$ so that the phase between the two logic levels remains the same after propagation. Therefore, we define our qubit as the first-order spatial mode of the photon, with $|0\rangle \equiv HG_{10}$ and $|1\rangle \equiv HG_{01}$. By analogy with polarization, we make the definitions $|h\rangle \equiv HG_{10}$, $|v\rangle \equiv HG_{01}$, $|d\rangle \equiv \frac{1}{\sqrt{2}}(HG_{10} + HG_{01})$, $|a\rangle \equiv \frac{1}{\sqrt{2}}(HG_{10} - HG_{01})$, $|l\rangle \equiv \frac{1}{\sqrt{2}}(HG_{10} + iHG_{01})$, $|r\rangle \equiv \frac{1}{\sqrt{2}}(HG_{10} - iHG_{01})$, as shown in Figure 2.3). Throughout the text, we use capital letters to represent polarization states, and lower-case letters for the corresponding first-order spatial mode states.

Although it is experimentally more difficult in this case than for polarization, we can perform arbitrary rotations of the spatial-mode qubit state using the Gouy phase. We can achieve the full $(N+1)\pi$-phase shift using two identical spherical lenses to focus a collimated beam and recollimate it, but how do we apply a *relative* phase between two modes of the same order? In fact, there is more to the Gouy phase than Eq. 2.4d suggests, and this becomes relevant when we consider *astigmatic* beams (beams that are more tightly
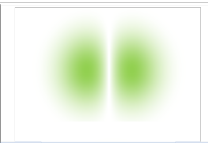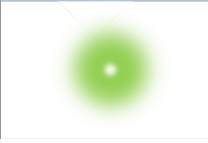
Figure 2.3: Some commonly encountered spatial-mode qubit states and the polarization qubit states they correspond to. Based on this analogy, we define similar names for the spatial mode states.

focused, for example, along the $x$ axis than along the $y$ axis). $\zeta$ actually consists of two components [17]:

$$\zeta(z) = \zeta_x(z) + \zeta_y(z), \text{ where}$$
$$\zeta_x(z) = \left(m + \frac{1}{2}\right)\arctan\left(\frac{z}{z_{R_x}}\right)$$
$$\zeta_y(z) = \left(n + \frac{1}{2}\right)\arctan\left(\frac{z}{z_{R_y}}\right), \tag{2.5}$$

where $z_{R_{x,y}}$ are the Rayleigh ranges for $x$ and $y$, respectively. If we replace the spherical lenses mentioned above with *cylindrical* lenses, we can achieve the desired effect. A cylindrical lens is curved only along one axis, flat along the other, and ideally if the curvature is along the $x$ axis, the $x$ dependence of the intensity distribution changes the way it would for a spherical lens of the same curvature, while the $y$ dependence is unchanged. Thus, $\zeta_x$ advances by $\left(m + \frac{1}{2}\right)\pi$ and $\zeta_y$ is unchanged. For the specific case of our qubit logic levels, we find that $HG_{10}$ and $HG_{01}$ acquire a relative phase of $\frac{3\pi}{2} - \frac{\pi}{2} = \pi$, so such a device functions as a spatial-mode HWP.[3] This setup is an extreme case of a scheme that can produce arbitrary relative phase shifts less than $\pi$.

To apply smaller phase shifts we can modify the system, starting with a beam that is already converging toward its waist and placing two identical cylindrical lenses symmetrically about the waist [17]. Once again the $y$ beam waist is unchanged. For a certain focal length and lens placement, we can arrange for the first lens to focus the beam to a smaller $x$ beam waist at the same location as the $y$ beam waist, and for the second lens to undo the astigmatism introduced by the first. In general, $\zeta_x$ winds further forward than $\zeta_y$ leading to a relative phase less than $\pi$. To realize a QWP, we start with a (non-astigmatic) beam with Rayleigh range $z_R$ and place cylindrical lenses of focal length $f = \frac{z_R}{1 + 1/\sqrt{2}}$ a distance $\frac{f}{\sqrt{2}}$ on either side of the original beam waist.

Finally, it is also possible to make a projective measurement of the spatial-mode state, though again experimentally more challenging than for polarization [18]. In this work, we take advantage of the fact that $HG_{mn}$ couples into single-mode optical fiber (SMF) with high efficiency if and only if $m = n = 0$; we routinely observe first-order mode coupling on the order of 0.5%, compared with 50-65% Gaussian ($HG_{00}$) coupling. We use holographic diffraction gratings with six different patterns. Each pattern converts a certain first-order spatial mode to a Gaussian in the +1-diffraction order, and converts the orthogonal first-order mode into a second-order mode (Figure 2.4). A SMF collects the +1-diffraction order, transmitting the Gaussian and rejecting the higher-order modes, thus measuring the first-order spatial mode state. This measurement apparatus is mentioned again briefly in Section 4.2.3.

---

[3]Actually, although it is possible to get arbitrarily close, a $\pi$-phase shift cannot be achieved because perfectly collimated beams don't exist, or equivalently, because it is impossible to actually propagate to $z = \infty$
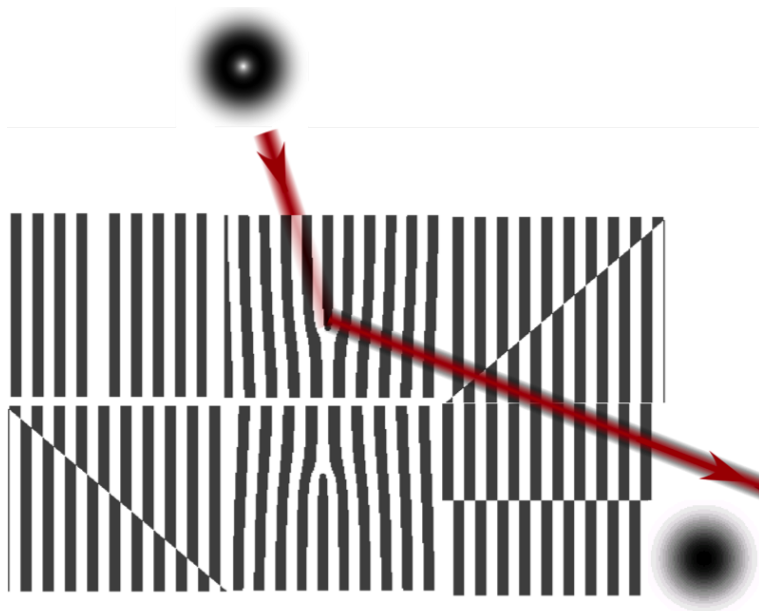
Figure 2.4: Schematic of a holographic diffraction grating used to measure the first-order spatial mode. The different regions of the hologram correspond to different measurement settings for the spatial mode; for each setting, one first-order mode is converted to a Gaussian, while the orthogonal mode is converted to a second-order mode. In the diagram the state $|l\rangle$ is incident on the region that converts $|l\rangle$ to the Gaussian mode. If instead $|r\rangle$ were incident on the same region, the output would be a second-order spatial mode. The output is coupled to a SMF that transmits the Gaussian mode with high efficiency and rejects higher-order modes, thus performing a projective measurement.

### 2.1.3 Frequency/emission time

The last photonic degree of freedom is frequency. We can discretize frequency as we did for momentum in Section 2.1.2, but the process is easier since the photon's frequency state does not evolve as it propagates through free space. A simple idea is to divide the continuous $\omega$ axis into discrete bins of width $\Delta\omega$, centered at frequencies $\omega_j$ . We can define orthonormal logic levels $|\omega_j\rangle$ for which the probability amplitude is

$$\psi_j(\omega) = \begin{cases} \sqrt{\frac{1}{\Delta\omega}} & \omega_j - \Delta\omega < \omega < \omega_j + \Delta\omega \\ 0 & \text{otherwise.} \end{cases}$$

The frequency Fourier transform space —emission time of the photon —can be binned in a similar fashion, and it is possible to encode information there as well, but it comes at the cost of the information capacity of the frequency space, since shorter pulses require larger spectral bandwidth. In this work we do use the quantum information encoded in the frequency-time degree of freedom, but we do not need to prepare or measure specific states, so it is unnecessary to define the qubits as carefully as for polarization and spatial mode.

## 2.2 Making the qubits

Now that we are familiar with the available degrees of freedom, we can see how to prepare the desired single-photon quantum states. The anisotropic behavior of birefringent media (Section 2.1.1) can be described by a *susceptibility tensor* $\chi$. Specifically, the components of the induced dipole moment per unit volume $\mathbf{P}$ are given by

$$P_j \approx \chi_{jk} E_k, \tag{2.6}$$

where repeated indices imply a sum. In reality, this linear relationship is only a lowest-order approximation of

$$P_j(t) = \chi_{jk}^{(1)} E_k(t) + \chi_{jkl}^{(2)} E_k(t) E_l(t) + \chi_{jklm}^{(3)} E_k(t) E_l(t) E_m(t) + \dots, \tag{2.7}$$

where $\chi^{(1)}$ is the linear susceptibility from Eq. 2.6, and $\chi^{(n>1)}$ is the $n$th-order correction [19]. While Eq. 2.6 is often adequate, certain phenomena only occur because of the nonlinearities in Eq. 2.7. One such process is *spontaneous parametric down-conversion* (SPDC), in which a "pump" photon interacts with a crystal with non-zero $\chi^{(2)}$ to produce a pair of lower energy photons, historically called the "signal" and "idler" [20]. In practice, because of the small $\chi^{(2)}$ values of available materials, this interaction happens with probability on

the order of $10^{-9}$. Consequently, if we use an intense laser beam to pump the crystal, the effect of SPDC on the *pump* is negligible. This observation motivates the "undepleted pump approximation", so we can write the Hamiltonian for the interaction as

$$
\hat{H} = \int_{\mathcal{V}} d^3 r \cdot \chi^{(2)} \mathbf{E}_p(\mathbf{r}, t) \hat{\mathbf{E}}_s^{(-)}(\mathbf{r}, t) \hat{\mathbf{E}}_i^{(-)}(\mathbf{r}, t) + \text{h.c.},
$$

$$
= \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3 r \cdot \chi^{(2)} \mathbf{E}_p e^{i(\mathbf{k}_p \cdot \mathbf{r} - \omega t)} \sum_{\mathbf{k}_s, m} \hat{\mathbf{p}}_{\mathbf{k}_s, m} \sqrt{\frac{\hbar \omega(\mathbf{k}_s)}{2\epsilon_0}} \hat{a}_{\mathbf{k}_s, m}^\dagger e^{i \mathbf{k}_s \cdot \mathbf{r}} \sum_{\mathbf{k}_i, n} \hat{\mathbf{p}}_{\mathbf{k}_i, n} \sqrt{\frac{\hbar \omega(\mathbf{k}_i)}{2\epsilon_0}} \hat{a}_{\mathbf{k}_i, n}^\dagger e^{i \mathbf{k}_i \cdot \mathbf{r}} + \text{h.c.}, \quad (2.8)
$$

where we have treated the pump as a classical plane wave with wave-vector $\mathbf{k}_p$, quantizing only the signal and idler fields over the interaction volume $\mathcal{V}$; the sums run over all wave-vector and polarization modes for the signal and idler. Since SPDC conserves energy and momentum,

$$
\omega_p = \omega_s + \omega_i, \tag{2.9a}
$$

$$
\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i \tag{2.9b}
$$

*inside* the crystal. As the pump traverses the crystal, SPDC can be generated anywhere along the interaction length; the requirement that the quantum mechanical probability amplitudes for generating SPDC all along this length add constructively is called *phase matching*, and it sets certain constraints on the wave-vectors $\mathbf{k}_p$, $\mathbf{k}_s$, $\mathbf{k}_i$ and polarizations $\hat{\mathbf{p}}_{\mathbf{k}_p}$, $\hat{\mathbf{p}}_{\mathbf{k}_s}$, $\hat{\mathbf{p}}_{\mathbf{k}_i}$. In this work we use the biaxial crystal bismuth triborate, $\text{BiB}_3\text{O}_6$ (BiBO). The crystal is cut such that $H$ and $V$ are the eigen-polarizations for normally incident light (i.e., these are the two polarizations that are unchanged by propagation through the crystal); $H$ experiences a smaller refractive index than $V$, so its phase velocity is larger, and we refer to $H$ ($V$) as the "fast" ("slow") polarization in the crystal[4]. We use *degenerate, non-collinear* phase matching, meaning that the signal and idler have equal energies $\omega_s = \omega_i = \omega_p/2$ and different momenta $\mathbf{k}_s \neq \mathbf{k}_i$. We use *Type-I* phase matching, meaning that a fast pump photon can give rise to a pair of slow SPDC photons[5]:

$$
|H\rangle_p \rightarrow |VV\rangle_{s,i}. \tag{2.10}
$$

---

[4]The crystal's three principal axes are traditionally labeled $x$, $y$, $z$ in order of increasing refractive index [21]. The crystal cut is specified by two angles $\theta$ and $\phi$ that indicate the normal vector to the crystal face in this coordinate system.

[5]In particular, a pump photon with $\hat{\mathbf{k}}_p = \sin(141.8°)\hat{\mathbf{y}} + \cos(141.8°)\hat{\mathbf{z}}$ and $\hat{\mathbf{p}}_p = \hat{\mathbf{x}}$ gives rise to degenerate signal-idler pairs that propagate at approximately $\pm 3°$ relative to the pump *outside* the crystal. The signal and idler can be emitted in any plane, so the SPDC emerges in a cone with opening half-angle $3°$ and signal-idler pairs can be found at any diametrically opposite points on the cone. The signal and idler have (approximately) the same polarization $\hat{\mathbf{p}}_s = \hat{\mathbf{p}}_i = \cos(141.8°)\hat{\mathbf{y}} - \sin(141.8°)\hat{\mathbf{z}}$. This equality is exact for signal and idler photons emerging in the plane that contains the $x$ axis, when the signal and idler have wave-vectors $\mathbf{k}_s, \mathbf{k}_i = \cos(3°)\hat{\mathbf{k}}_p \pm \sin(3°)\hat{\mathbf{x}}$. For pairs arriving at other points on the cone, the $\hat{\mathbf{p}}_s = \hat{\mathbf{p}}_i$ given above have components parallel to $\mathbf{k}_s$ and $\mathbf{k}_i$ and are therefore not allowed [22]. Thus, for pairs collected in other planes, $\hat{\mathbf{p}}_s$ and $\hat{\mathbf{p}}_i$ deviate slightly from this value. This effect, largest for pairs emitted in the $y$-$z$ plane, amounts to approximately $1°$ deviation from the polarization in the orthogonal plane, and it is not important in the experiments discussed here.
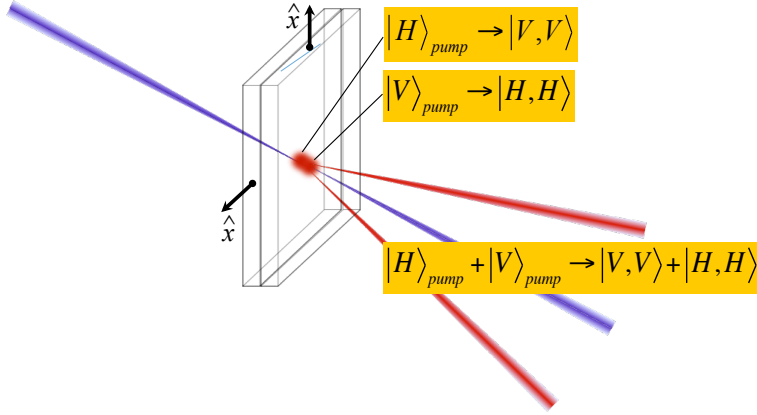
Figure 2.5: Technique used to generate polarization entangled photons using Type-I SPDC. The two crystals are identical, but rotated 90° with respect to each other, so the first (second) crystal produces horizontally (vertically) polarized photon pairs. In both cases, a fast pump photon produces a pair of slow SPDC photons. The SPDC processes combine coherently to give a maximally entangled state.

If we rotate the same crystal 90° (and ignore the small error associated with collecting SPDC photons in the "wrong" plane, discussed in Footnote 5), we can instead observe the process

$$|V\rangle_p \to |HH\rangle_{s,i}. \tag{2.11}$$

Finally, if we have *two* adjacent, orthogonally oriented crystals, we can pump them simultaneously to observe the process [23]

$$|H\rangle_p + e^{i\phi} |V\rangle_p \to |VV\rangle_{s,i} + e^{i\phi} |HH\rangle_{s,i}. \tag{2.12}$$

Here we assume that the two-crystal process Eq. 2.12 is the *coherent* sum of the two one-crystal processes Eqs. 2.10 and 2.11, which is true so long as the one-crystal processes are indistinguishable. Thus, we can prepare the SPDC photon polarization in a maximally entangled state (Figure 2.5).

The SPDC *spatial-mode* state is determined by the pump spatial mode. We can write the Hermite-Gauss components of the SPDC produced by each Hermite-Gauss mode of the pump as

$$|HG_{mn}\rangle_p \to \sum_{j,k,u,v} C_{jkuv}^{mn} |HG_{jk}\rangle_s |HG_{uv}\rangle_i, \tag{2.13}$$

where the pump Hermite-Gauss modes have beam waist radius $w_{0,p}$ at the center of the crystal, while the SPDC modes have beam waist radius $w_{0,s} = w_{0,i} = \sqrt{2}w_{0,p}$, also at the center of the crystal [24]. We choose the mode radii this way in order to match the beam divergence of the pump with that of the SPDC photons, which have twice the wavelength $\lambda_s = \lambda_i = 2\lambda_p$. The probability amplitudes $C_{jkuv}^{mn}$ can be determined by

13

first expanding $|HG_{mn}\rangle_p$ as a sum of plane waves (the *angular spectrum decomposition*) and summing all the contributions to the SPDC. Because we use thin crystals in this work ($L = 0.4$ mm), we can greatly simplify the exact expression for $C_{jkuv}^{mn}$:

$$\widetilde{C}_{jkuv}^{mn} = \begin{cases} \sqrt{\frac{2L}{\pi k_p w_{0,p}^2}}\, b(j,u,j+u-m)b(k,v,k+v-n)HG_{j+u-m,k+v-n}(\mathbf{r}=0) & \text{if } j+u \geq m \text{ and } k+v \geq n, \\ 0 & \text{otherwise,} \end{cases}$$

(2.14)

where

$$b(p,q,r) = \sqrt{\frac{(p+q-r)!r!}{2^{p+q}p!q!}}\frac{1}{r!}\frac{d^r}{dt^r}\left[(1-t)^p(1+t)^q\right]\Big|_{t=0}.$$

(2.15)

Although calculating the $\widetilde{C}_{jkuv}^{mn}$ generally requires plugging the appropriate arguments into Eq. 2.15, we can identify additional cases when the probability amplitude is zero. Since the Hermite polynomials of odd index are odd functions, the last factor in Eq. 2.14 ensures that $\widetilde{C}_{jkuv}^{mn} = 0$ when at least one of $j+u-m$ and $k+v-n$ is odd. Therefore, the pump and SPDC Hermite-Gauss modes are only non-zero when

$$j+u \geq m \text{ and } j+u \equiv m \pmod 2$$

$$k+v \geq n \text{ and } k+v \equiv n \pmod 2.$$

(2.16)

As discussed in Section 2.1.2, we wish to use the first-order spatial modes of the signal and idler as qubits. From Eqs. 2.16, we note that there are two ways to generate a pair of first-order $HG$ modes:

1. The zero-order mode pump $|HG_{00}\rangle_p \rightarrow |HG_{10}\rangle_s |HG_{10}\rangle_i + |HG_{01}\rangle_s |HG_{01}\rangle_i$.

2. A second-order mode pump:

    (a) $|HG_{20}\rangle_p \rightarrow |HG_{10}\rangle_s |HG_{10}\rangle_i$,

    (b) $|HG_{02}\rangle_p \rightarrow |HG_{01}\rangle_s |HG_{01}\rangle_i$,

    (c) $|HG_{11}\rangle_p \rightarrow |HG_{10}\rangle_s |HG_{01}\rangle_i + |HG_{01}\rangle_s |HG_{10}\rangle_i$.

Although the Gaussian mode $HG_{00}$ does produce a maximally entangled state, evidently we have no control over *which* state we get —this is inconvenient, since rotating the first-order-spatial-mode qubit state is not as easy as rotating the polarization state. In contrast, the SPDC processes from the second-order pump are reminiscent of those for polarization:

$$|HG_{20}\rangle_p + e^{i\phi}|HG_{02}\rangle_p \rightarrow |HG_{10}\rangle_s |HG_{10}\rangle_i + e^{i\phi}|HG_{01}\rangle_s |HG_{01}\rangle_i,$$

(2.17)

14

where we have control of the relative amplitudes of the two terms. Thus, in this work we use second-order pump modes to produce first-order spatial mode qubits on the signal and idler photons. As I discuss later, this transfers the difficulty of having to rotate the states from a Gaussian pump to the problem of having to prepare a second-order pump.

Finally, we come to frequency. As stated in Section 2.1.3, we do not explicitly encode quantum information in this degree of freedom, but we do make use of it as an ancilla. The condition Eq. 2.9a means that the signal and idler frequencies are *also* entangled,

$$|\omega_s, \omega_i\rangle = \int d\omega' A(\omega') \, |\omega', \omega_p - \omega'\rangle, \tag{2.18}$$

where $A$ is related to phase-matching constraints on the range of allowed frequencies. In this work, we use SPDC to generate *hyperentangled* photons, photons entangled in all three degrees of freedom.

# Chapter 3

# Bound Entanglement

Entanglement is the key ingredient in such potentially important QIS technologies as quantum computers, quantum repeaters [5] and quantum key distribution [6]. Although some applications do require the simplest possible entangled state, a maximally entangled pair of qubits, some require or at least benefit from larger numbers of qubits or a larger alphabet (i.e., $n$-level systems instead of qubits). In the one-way paradigm of quantum computing, for example, all qubits in the register are arranged on a lattice in which every pair of adjacent qubits is entangled [25]. Also common are quantum protocols that can be described *abstractly* using 2-qubit entangled states, but which benefit from or, indeed, only become feasible with the addition of ancillary qubits [26]. The ubiquitous importance of quantum error correction is an example of additional qubits making such abstract protocols potentially possible in practice. [27].

High-dimensional entanglement is also of fundamental interest. While even the simplest case of 2-qubit entanglement can display surprising complexity [8], qualitatively different features emerge in larger Hilbert spaces. In this chapter we explain the theory behind one such phenomenon, called *bound entanglement*, that only occurs in systems larger than $2 \otimes 3$ (one qubit and one three-level system) [28]. Bound entanglement was discovered theoretically 18 years ago, but it was only observed experimentally quite recently [29] because of the challenges involved in precisely preparing multi-qubit states, some of which we will see in Chapter 4.

## 3.1 Pure-state entanglement

A quantum system is said to be in a *pure* state if there exists a measurement basis for which the measurement outcome is deterministic. In many cases, it may be difficult or impossible to actually realize such a measurement, but this has no bearing on the definition. Pure states can be expressed simply by state vectors or wavefunctions. The next section discusses mixed states, which require a more detailed representation.

A *product state* of two qubits $A$ and $B$ can be described by a state vector [30]

$$|\psi_{A,B}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle, \tag{3.1}$$

so that the joint state is simply the tensor product of the single-qubit states $|\psi_A\rangle$ and $|\psi_B\rangle$. Those states that cannot be expressed as in Eq. 3.1 are *entangled*. These states can be written in the form

$$|\psi_{A,B}\rangle = \alpha\,|\psi_A,\psi_B\rangle + \beta\,\left|\psi_A^\perp,\psi_B^\perp\right\rangle, \tag{3.2}$$

where we have dropped the tensor product symbol $\otimes$ for compactness, $|\psi^\perp\rangle$ is the qubit state orthogonal to $|\psi\rangle$, and $\alpha$, $\beta$ are nonzero amplitudes. It is straightforward to show algebraically that certain states cannot be expressed as products in any basis. A physical example demonstrating the existence of at least one such state is, perhaps, more interesting. Consider the singlet state:

$$\left|\psi^-\right\rangle = \frac{|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle}{\sqrt{2}}. \tag{3.3}$$

The Pauli exclusion principle demands that two identical spin-1/2 particles with the same spatial wave-function (e.g., two electrons occupying the same atomic orbital) exist in the spin state Eq. 3.3 to ensure antisymmetry of the total state —the singlet is the *only* antisymmetric state of two qubits. Since the spin state must be antisymmetric regardless of the basis we choose to measure it in, the singlet evidently takes the same form in *every basis*:

$$\left|\psi^-\right\rangle = \frac{\left|X,X^\perp\right\rangle - \left|X^\perp,X\right\rangle}{\sqrt{2}}, \tag{3.4}$$

so that it cannot be expressed with fewer than two terms. This means that $|\psi^-\rangle$ is entangled. Note that the states of qubits $A$ and $B$ individually are *not* pure: there is probability $1/2$ of finding each particle in the state $|X\rangle$ or $\left|X^\perp\right\rangle$ regardless of measurement basis. At the same time, measurement results for the two qubits are perfectly correlated as long as the same analysis basis is used for both qubits. The singlet is, therefore, an example of *maximal entanglement* of two qubits. States of the form Eq. 3.2 with $|\alpha| \neq |\beta|$ are non-maximally entangled. Thus, by simply varying the ratio of the two coefficients, we can produce a pure state with any amount of entanglement ranging from a product state to a maximally entangled state. There is an added complication, however, that permits other types of two-qubit states, as we see in the next section.

## 3.2  Mixed states

Quantum states that give probabilistic measurement outcomes in *every* basis are called mixed states. In fact, one *never* encounters pure states in the lab, so we need to extend Eqs. 3.1, 3.2 to the mixed-state case.

To begin with, a mixed state can be represented by a *density operator* [31]

$$\rho = \sum_j p_j \left| \psi_j \middle\rangle \middle\langle \psi_j \right|, \tag{3.5}$$

where $p_j$ are classical probabilities ($\sum p_j = 1$) of finding the system in the pure state $\left| \psi_j \middle\rangle \middle\langle \psi_j \right| \to \left| \psi_j \right\rangle$, and the $\left| \psi_j \right\rangle$ are normalized and, in general, not orthogonal. We can write $\rho$ as a matrix in a basis of our choice. Density matrices have two properties that follow directly from the definition above:

1. The diagonal elements are non-negative and sum to 1.

2. The matrix is Hermitian, $\rho = \rho^\dagger$.

*Any matrix satisfying these requirements represents a valid quantum state.* The expression Eq. 3.5 is *not* unique: there are infinitely many ways to construct the same mixed state by summing pure states, and all are equally valid. Surprisingly, $\rho$ completely specifies the quantum state.

To understand the origin of the classical probabilities in Eq. 3.5, consider several different processes, each of which prepares the system in one of the pure states $\left| \psi_j \right\rangle$. If the system undergoes one of these processes, selected according to the probability distribution $p_j$, the state Eq. 3.5 results. When the different processes are coherent, we instead sum the state vectors $\left| \psi_j \right\rangle$ with amplitudes $e^{i\phi_j} \sqrt{p_j}$ (i.e., if the $\phi_j$ are well-defined we must use them in the sum) [32]. Of course, these are two extremes and in the lab we generally deal with partially coherent processes. What does it actually mean for two processes to be coherent? In general, it means that no information exists anywhere *else* as to which process occurred[1]. Conversely, if there exists classical information that labels the system as having been prepared via one process or another, there will be reduced coherence between the processes. To make these statements concrete, let's write the density matrix

---

[1]We might be able to determine which process occurred by an appropriate measurement of the system itself, but that doesn't count. For example, in Section 2.1.1 we claim that the SPDC processes generating the states $\left| HH \right\rangle$ and $\left| VV \right\rangle$ are indistinguishable, but we could determine which one happened by passing one of the photons through a horizontal polarizer. Here, however, we have identified the process by a projective measurement, changing the system's state. In contrast, we can, in principle, identify which one of two *distinguishable* processes has prepared a system without interacting with the system at all.

for the SPDC state in Eq. 2.12:

$$\begin{aligned}
\rho &= \frac{|HH\rangle + |VV\rangle}{\sqrt{2}} \cdot \frac{\langle HH| + \langle VV|}{\sqrt{2}} \\
&= \frac{|HH\rangle\langle HH| + |HH\rangle\langle VV| + |VV\rangle\langle HH| + |VV\rangle\langle VV|}{2} \\
&\to \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix},
\end{aligned} \tag{3.6}$$

where the rows and columns are indexed $HH, HV, VH, VV$ (this basis and ordering will be used throughout the text). Now let's relax the condition that the processes from the two crystals are indistinguishable. In particular, suppose there is an additional qubit $C$ that provides partial information about whether we have produced $HH$ or $VV$. This information can be parameterized by $0 \leq \alpha \leq 1$, with $\alpha = 0$ corresponding to indistinguishable processes. For example,

$$\rho = \frac{|HH\rangle \otimes |H\rangle_C + |VV\rangle \otimes (\sqrt{1-\alpha^2}\,|H\rangle_C + \alpha\,|V\rangle_C)}{\sqrt{2}}$$
$$\times \frac{\langle HH| \otimes \langle H|_C + \langle VV| \otimes (\sqrt{1-\alpha^2}\,\langle H|_C + \alpha\,\langle V|_C)}{\sqrt{2}}. \tag{3.7}$$

Note that the *three-qubit* state in Eq. 3.7 is pure —however, it's the state of the first two qubits that we're interested in: we can find this state by taking the *partial trace* with respect to the third qubit [33]:

$$\mathrm{Tr}_C\,\rho = \langle H|_C\,\rho\,|H\rangle_C + \langle V|_C\,\rho\,|V\rangle_C \tag{3.8}$$
$$= \frac{1}{2}\left(|HH\rangle\langle HH| + |VV\rangle\langle VV| + \sqrt{1-\alpha^2}\,(|HH\rangle\langle VV| + |VV\rangle\langle HH|)\right). \tag{3.9}$$

The *reduced density matrix* for the first two qubits is

$$\mathrm{Tr}_C\,\rho \to \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & \sqrt{1-\alpha^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sqrt{1-\alpha^2} & 0 & 0 & 1 \end{pmatrix}. \tag{3.10}$$

The *purity* —in general equal to $\mathrm{Tr}\,\rho^2$ for state $\rho$ —of Eq. 3.10 is $1 - \frac{\alpha^2}{2}$. Evidently, increasing the "distinguishability" $\alpha$ reduces the purity. Here, we have shown that entanglement to an external degree of freedom

leads to mixture, and in nature this appears to be the most common cause of decoherence. Nevertheless, the fact that $\rho$ contains all information about the quantum state means that any other mechanism producing the same mixture $\rho$ is an equally valid preparation.

We will also want to quantify how "close" two quantum states are. The *fidelity* of $\rho_1$ to $\rho_2$ is defined as [34]

$$F(\rho_1, \rho_2) = F(\rho_2, \rho_1) \equiv \left( \text{Tr} \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2. \tag{3.11}$$

Finally, let us define entanglement for mixed states: the state of two systems $A$ and $B$ is *separable* if the joint density operator can be written

$$\rho_{A,B} = \sum_j p_j \rho_{A,j} \otimes \rho_{B,j}, \tag{3.12}$$

where $p_j$ are probabilities, and $\rho_{A,j}$ and $\rho_{B,j}$ are valid density operators for systems $A$ and $B$, respectively. If no such expression exists, $A$ and $B$ are *entangled* [35]. We can quantify the entanglement of a $2 \otimes 2$ state $\rho$ using the *tangle* [36]:

$$T(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}^2, \tag{3.13}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq 0$ are the eigenvalues of the operator

$$\sqrt{\sqrt{\rho} \left( \sigma_y \otimes \sigma_y \right) \rho^T \left( \sigma_y \otimes \sigma_y \right) \sqrt{\rho}}, \tag{3.14}$$

where $\rho^T$ is the transpose of the density matrix, and

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.15}$$

Multipartite entanglement is harder to characterize since we have to consider correlations between all subsets of the particles, and there is no single measure that works best in all situations [9]. For example, while the maximally entangled two-qubit state Eq. 3.3 can be "converted" into any other two-qubit state (this conversion is subject to certain conditions that we discuss in Section 3.3), the situation is different for three-qubit systems: there are two distinct types of three-qubit entanglement that cannot be inter-converted [37]. Additionally, certain simplifications that are possible for qubits do not extend to subsystems with $d > 2$ levels (qudits), as in Ref. [36]. In this work we discuss *four*-qubit states, but we are only concerned with bipartite entanglement. The next sections develop the tools we need to analyze these states.

## 3.3   Entanglement distillation

Many important QIS protocols require two communicating parties Alice and Bob to possess an *ebit*, a shared pair of qubits in a maximally entangled state, such as one of the four *Bell states* (written here in the canonical $\{H, V\}^{\otimes 2}$ basis):

$$\left|\Phi^\pm\right\rangle = |HH\rangle \pm |VV\rangle, \left|\Psi^\pm\right\rangle = |HV\rangle \pm |VH\rangle. \tag{3.16}$$

An *ebit* represents potential communication between the two parties [9]. In dense coding [4], for example, shared entanglement allows Alice to encode two classical bits in one qubit, seemingly violating the Holevo bound [38]. As another example, we saw in Section 3.1 that a maximally entangled two-qubit state gives correlated, random measurement results; moreover, the strength of these correlations depends on the two systems being completely uncorrelated with the rest of the universe (entanglement is "monogamous" [39]). An ebit, therefore, represents one random bit, shared with perfect fidelity and privacy by Alice and Bob, a resource that enables one bit of secure communication [40]. We can imagine a future quantum network in which a central location generates and distributes ebits to parties who wish to communicate. In practice, we expect decoherence in the quantum channels to result in Alice and Bob sharing *non*-maximally entangled pairs. This imperfect entanglement leads to noisy communication, or equivalently to a reduction in the number of potential bits of communication. Therefore, given a non-maximally entangled state, it is natural to ask how many ebits it represents. In particular, once the entangled qubits have been delivered to Alice and Bob at their remote locations, we would like to know how many ebits they could possibly extract using only *local operations* on their respective qubits and *classical communication* to (for example) inform each other of measurement outcomes (LOCC). This extraction process is called *entanglement distillation* [41]. We say that Alice and Bob share a distillable state if using only LOCC they can extract a maximally entangled state with non-zero probability and know when they have succeeded.

A simple example of entanglement distillation is the "Procrustean method" [42, 43]. We start with a known state $\left|\psi^0\right\rangle = \alpha |HH\rangle_{A,B} + \sqrt{1-\alpha^2} |VV\rangle_{A,B}$, with $\alpha > 1/\sqrt{2}$. In order to distill a maximally entangled pair, Alice and Bob simply have to make the amplitudes for $|HH\rangle$ and $|VV\rangle$ equal. If Alice sends her half of the state through a *partial* polarizer (such a device can be realized simply using a tilted piece of glass, since the magnitude of the Fresnel reflection depends on polarization) that has probability amplitude $\epsilon\sqrt{1-\alpha^2}$ for transmitting $|H\rangle$ and amplitude $\epsilon\alpha$ for transmitting $|V\rangle$, she can balance the probabilities for the two terms. Note that Alice and Bob have to sacrifice some fraction of the copies of their initial state in order to distill maximal entanglement —it is impossible to increase the amount of shared entanglement using only LOCC [44].

## 3.4 Not all entangled states are distillable

One might guess that all entangled states permit distillation of *some* entanglement, however little. In fact, in $2 \otimes 2$ and $2 \otimes 3$ systems it is true that all entangled states are distillable; in higher dimensions, however, there exist undistillable entangled states. This section explains the observations and reasoning that lead to this conclusion.

### 3.4.1 The Peres (positive-partial-transpose) criterion

Particularly at high dimensions, it is not always apparent when a density matrix represents an entangled state. Peres developed a simple criterion for separability [45]. To understand this test, we must first define a mathematical operation, the *partial transpose* of a matrix. First of all, we can perform the (full) transpose of a density matrix as

$$\rho^T = \left( \sum_{j,k} c_{j,k} |j\rangle\langle k| \right)^T = \sum_{j,k} c_{j,k} |k\rangle\langle j|, \tag{3.17}$$

where $j$ and $k$ index the same set of orthonormal basis states. Recalling from Section 3.2 the requirements a density matrix has to satisfy, it is clear that if $\rho$ is a valid density matrix, then so is $\rho^T$.

Now we consider the joint density matrix of $A$ and $B$. We write the matrix in a separable basis (we cannot, for example, use the basis $\{\Phi^+, \Phi^-, HV, VH\}$ because the first two basis vectors are entangled). Then the partial transpose with respect to $A$ is

$$\rho^{T_A} = \left( \sum_{jklm} c_{jklm} |j,l\rangle_{A,B} \langle k,m| \right)^{T_A} = \sum_{jklm} c_{jklm} |k,l\rangle_{A,B} \langle j,m|, \tag{3.18}$$

where $j$ and $k$ ($l$ and $m$) index the basis vectors for $A$ ($B$).

**The positive-partial-transpose (PPT) criterion**: *If $A$ and $B$ are separable, then the partial transpose $\rho_{A,B}^{T_A}$ must have no negative eigenvalues.*

Proof: If $A$ and $B$ are separable, then $\rho_{A,B}$ can be written in the form Eq. 3.12. The partial transpose is

$$\rho_{A,B}^{T_A} = \sum_j p_j \rho_{A,j}^T \otimes \rho_{B,j}. \tag{3.19}$$

Since the $\rho_{A,j}$ are valid density matrices by construction, so are the $\rho_{A,j}^T$: therefore, $\rho_{A,B}^{T_A}$ is *also* a valid density matrix (in fact, it also has the form Eq. 3.12), so it cannot have any negative eigenvalues.

### 3.4.2 The negative-partial-transpose criterion

Despite its elegance, it is not immediately obvious that the PPT criterion is very helpful —after all, its utility depends on how frequently entangled states *fail* the test. We cannot yet say anything conclusive about the partial transpose when $A$ and $B$ are entangled because all we know about such states is that they do *not* take the form Eq. 3.12. Peres conjectured that the converse of the PPT criterion is also true, and less than a year after his work the Horodeckis showed that this conjecture is, in fact, correct in certain cases: *the converse of the PPT criterion is true for $2 \otimes 2$ and $2 \otimes 3$ states, but in higher dimensions it is not* [46]. That is,

$$\rho_{AB} \text{ is separable} \begin{cases} \Longleftrightarrow \rho_{AB} \text{ has PPT} & \text{when } \rho_{AB} \text{ is } 2 \otimes 2 \text{ or } 2 \otimes 3, \\ \Longrightarrow \rho_{AB} \text{ has PPT} & \text{for all larger dimensions.} \end{cases} \tag{3.20}$$

This means that larger-dimensional states can have PPT and yet not be separable. Two years later the Horodeckis built on this result, introducing [28]

**The negative-partial-transpose (NPT) criterion:** *If maximal entanglement between $A$ and $B$ can be distilled, then the partial transpose $\rho_{A,B}^{T_A}$ must have at least one negative eigenvalue.*

$$\rho_{AB} \text{ is distillable} \implies \rho_{AB} \text{ has NPT.} \tag{3.21}$$

Equivalently, all PPT states are undistillable. The Horodeckis' two results allow for the existence of undistillable entangled states: sufficiently high-dimensional (larger than $2 \otimes 3$) entangled states can be PPT, and PPT states are always undistillable. Undistillable entanglement is called *bound entanglement*, and distillable entanglement is called free entanglement.[2] The Venn diagram in Figure 3.1 summarizes the results stated in this section. In general, to establish that a state is BE, we can show that it satisfies the PPT criterion, and then we need to determine that the state is nevertheless entangled —for arbitrary mixed multipartite states, this second step is not always easy. The separability criterion given in Ref. [47] was used to construct the first examples of BE in the smallest possible dimensions $3 \otimes 3$ and $2 \otimes 4$ [48]. As shown in the next section, we only need the partial transpose criteria in this work.

---

[2]Note that while the above conditions imply the existence of BE states with PPT, they do not *constrain* BE states to be PPT. The existence of NPT BE is still an open question [9].
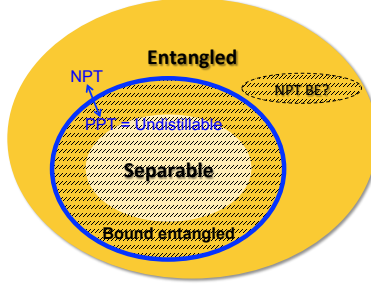
Figure 3.1: A diagram of Hilbert space. The light-colored inner region represents the separable states and the yellow outer region represents the entangled states. States outside the blue boundary have NPT and those inside have PPT; states belonging to shaded regions are undistillable. As explained in the text, for Hilbert spaces larger than $2 \otimes 3$ the entangled region extends into the PPT region and the overlap consists of bound-entangled states. These criteria say nothing about the additional shaded region, and whether or not such NPT undistillable entanglement exists is still an open question.

## 3.5   The Smolin state

Smolin identified a 4-qubit state that has become the most well-known example of bound entanglement [49]:

$$\rho_{ABCD}^{S} = \frac{1}{4}(\left|\Phi^+\right\rangle_{A,B} \left\langle\Phi^+\right| \otimes \left|\Phi^+\right\rangle_{C,D} \left\langle\Phi^+\right| +$$

$$\left|\Phi^-\right\rangle_{A,B} \left\langle\Phi^-\right| \otimes \left|\Phi^-\right\rangle_{C,D} \left\langle\Phi^-\right| +$$

$$\left|\Psi^+\right\rangle_{A,B} \left\langle\Psi^+\right| \otimes \left|\Psi^+\right\rangle_{C,D} \left\langle\Psi^+\right| +$$

$$\left|\Psi^-\right\rangle_{A,B} \left\langle\Psi^-\right| \otimes \left|\Psi^-\right\rangle_{C,D} \left\langle\Psi^-\right|), \quad (3.22)$$

where, for example, $\left|\Phi^\pm\right\rangle = \frac{\left|HH\right\rangle \pm \left|VV\right\rangle}{\sqrt{2}}, \left|\Psi^\pm\right\rangle = \frac{\left|HV\right\rangle \pm \left|VH\right\rangle}{\sqrt{2}}$ are the Bell states introduced in Section 3.3. In each of the four terms of Eq. 3.22, qubits $A$ and $B$ are in one of the Bell states, and $C$ and $D$ are in the same Bell state. To show that the Smolin state is BE, we make the following observations:

1. *The Smolin state does not change when the parties are interchanged. That is, $\rho_{ABCD}^{S} = \rho_{BCDA}^{S} = \dots$ for any permutation of the parties.*

First of all, note that the statement is obviously true for several specific cases. Since the Bell states are unchanged when the parties are swapped[3], four cases $ABCD, ABDC, BACD, BADC$ are taken care of. Furthermore, since $A$ and $B$ share the same Bell state as $C$ and $D$, the cases $CDAB, CDBA, DCAB, DCBA$ are also equivalent. 16 cases remain, but applying these simplifications again, we only need to consider one of them —for example, $ACBD$. That these remaining cases are indeed the same can be verified by simply writing out the density matrices and comparing them. In the logical basis $\{H, V\}^{\otimes 4}$ the Smolin state takes

---

[3] $\left|\Phi^\pm\right\rangle$ and $\left|\Psi^+\right\rangle$ are symmetric with respect to exchange. $\left|\Psi^-\right\rangle$ is antisymmetric, but the factor of $-1$ incurred by swapping the parties cancels out in the density matrix $\left|\Psi^-\right\rangle_{A,B} \left\langle\Psi^-\right| \rightarrow (-1)\left|\Psi^-\right\rangle_{B,A}(-1)\left\langle\Psi^-\right|$

the form

$$\rho^S_{ABCD} \rightarrow \frac{1}{8} \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{pmatrix}. \tag{3.23}$$

2. *If qubits A and B are held by a single party 'AB', the Smolin state is distillable.*

$AB$ simply needs to perform a measurement in the Bell-state basis, projecting the 4-qubit state onto one of the four terms in Eq. 3.22. $AB$ can (classically) communicate the measurement outcome to $C$ and $D$, who then know which Bell state they share. Indeed, it is straightforward to verify that the partial transpose with respect to any single party has 12 eigenvalues equal to $1/8$ and 4 eigenvalues equal to $-1/8$, satisfying the NPT criterion and indicating that each individual party shares distillable entanglement with the other three parties.[4]

3. *If qubits A and B are held by 'AB' and qubits C and D are held by 'CD', the Smolin state is separable and, therefore, undistillable.*

In fact, Eq. 3.22 is already written in the form of a separable state (Eq. 3.12) shared by parties $AB$ and $CD$. Intuitively, the reason this state is undistillable is that all the qubits that could be entangled to each other are held by single parties, so that no entanglement remains across the cut $AB : CD$. Appealing once again to observation (1) above, we can conclude that the Smolin state is separable across *every* pairwise cut. Checking the partial transpose for the case of $2 : 2$ bipartitions, we find that the transpose with respect to

---

[4]Note that the partial transpose operation does not affect any matrix elements on the diagonal, so even if the result is not a physical density matrix, the trace is still 1.

any *two* parties has 4 eigenvalues equal to 1/4 and 12 eigenvalues equal to 0, satisfying the PPT criterion for separability.

The 4-party Smolin state must also be undistillable, since the operations available to the parties in the case $A:B:C:D$ are a subset of what they can do in a $2:2$ partition. Meanwhile, (1) and (2) imply that there *is* (distillable) entanglement between each party and the other three. Therefore, the 4-party Smolin state is bound-entangled.

As mentioned above, $(\rho^S_{ABCD})^{T_{AB}}$ has 12 eigenvalues equal to zero, meaning that the state is infinitesimally close to being distillable (the zero eigenvalues are infinitesimally close to being negative). This has important experimental implications. Even arbitrarily precise measurements on a perfect Smolin state cannot distinguish it from a distillable state: instead, a convincing experimental realization must aim to produce a state in the *interior* of the BE region. To see how we can reach such a state starting from the Smolin state, note that the partial transpose (with respect to any subset of parties) is a *linear* operation: $(\rho^1_{XY} + \rho^2_{XY})^{T_X} = (\rho^1_{XY})^{T_X} + (\rho^2_{XY})^{T_X}$. Furthermore, as remarked earlier in this section, the partial transpose does not alter diagonal matrix elements. The idea, then, is to mix the Smolin state with white noise [50]:

$$\rho^S(p) = p\rho^S_{ABCD} + (1-p)\frac{\mathbb{I}}{16},\tag{3.24}$$

where the normalized identity matrix $\frac{\mathbb{I}}{16}$ is the density matrix for white noise. The desired effect of the mixture Eq. 3.24 is to raise all of the zero eigenvalues in the $2:2$ partial transposes, but it also raises the negative $1:3$ partial transpose eigenvalues: when all of these are raised beyond zero, $\rho^S(p)$ is rendered fully separable. Specifically, the least eigenvalue of the partial transpose of $\rho^S(p)$ across any bipartition is given by the weighted average of the least Smolin state partial transpose eigenvalue for that bipartition, and the white noise eigenvalue. Therefore, to maintain entanglement in the $1:3$ cuts we require

$$p\lambda^{\min}_{1:3} + (1-p)\frac{1}{16} > 0,$$

so that

$$\rho^S(p) \text{ is } \begin{cases} \text{BE} & p > \frac{1}{3}, \\ \text{separable} & p \leq \frac{1}{3}. \end{cases}\tag{3.25}$$

## 3.6 Applications of bound entanglement

As described so far, BE may be an interesting phenomenon but it's not very useful. In the context of quantum communication it seems to be exactly what we want to avoid, since we cannot extract any of

the maximally entangled pairs we would need for standard applications like dense coding, teleportation, or quantum key distribution. Framed this way, it is useful to study BE so that we can "know our enemy". However, BE does have certain capabilities, some of which are discussed in this section.

### 3.6.1 Activation of bound entanglement

Perhaps the most compelling application of BE was developed by the Horodeckis within a year of their discovery of the phenomenon [51]. Although free entanglement is (by definition) always distillable, the distillation process is not always realizable. For example, some free-entangled states can only be distilled *collectively*, i.e., by simultaneously acting on multiple copies of the state [52, 53]. Technical challenges aside, if we only have a single copy of such a free-entangled state, the distillable entanglement is fundamentally limited to zero. However, if we additionally have copies of a BE state, we can *activate* the BE to distill entanglement from the free-entangled state. BE activation was originally described for a single copy of a free-entangled $3 \otimes 3$ state and many copies of a $3 \otimes 3$ BE state, but it was later extended to activation of a noisy $2 \otimes 2$ free-entangled state by the Smolin state [54]. This version of the activation protocol requires various additional assumptions, which are discussed in more detail in Section 4.5. From a fundamental point of view, BE activation is interesting because it suggests non-additivity of distillable entanglement [51]: the combination of the undistillable single instance of free entanglement and the also undistillable multiple copies of BE is in fact distillable!

### 3.6.2 Bell inequality violation

One of the most surprising properties of the Smolin state is that it *maximally* violates the generalized CHSH-like Bell inequality derived in [55, 56]:

$$|E_{1111} + E_{1112} + E_{2221} - E_{2222}| \leq 2, \tag{3.26}$$

where

$$E_{1111} \equiv \mathrm{Tr}\left(\rho \sigma_z \otimes \sigma_z \otimes \sigma_z \otimes \frac{\sigma_z + \sigma_x}{\sqrt{2}}\right),$$

$$E_{1112} \equiv \mathrm{Tr}\left(\rho \sigma_z \otimes \sigma_z \otimes \sigma_z \otimes \frac{\sigma_z - \sigma_x}{\sqrt{2}}\right),$$

$$E_{2221} \equiv \mathrm{Tr}\left(\rho \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \frac{\sigma_z + \sigma_x}{\sqrt{2}}\right),$$

$$E_{2222} \equiv \mathrm{Tr}\left(\rho \sigma_x \otimes \sigma_x \otimes \sigma_x \otimes \frac{\sigma_z - \sigma_x}{\sqrt{2}}\right).$$

In particular for the noisy Smolin state $\rho^S_{ABCD}(p)$ from Eq. 3.24 we find that

$$|E_{1111} + E_{1112} + E_{2221} - E_{2222}| = 2\sqrt{2}p. \tag{3.27}$$

No 4-qubit state achieves a Bell parameter larger than $2\sqrt{2}$ [50]. This result is particularly unexpected, since one might intuitively think of BE as a weak form of entanglement. The maximal violation implies a maximal entanglement between each party and the other three, and yet this entanglement is not distillable.

### 3.6.3 Two-party distillation

As noted in (2) of Section 3.5, the Smolin state is distillable when any pair of parties is allowed to perform joint measurements. In case (3) of Section 3.5, the Smolin state becomes separable simply because the entanglement that could be distilled already exists at a single location. The Smolin state is unusual in this respect, as its entanglement, distillability, and bound entanglement are tied to the way the state is distributed among the parties. This is certainly not a general property of BE. For example, the $2 \otimes 4$ and $3 \otimes 3$ BE states first constructed by P. Horodecki [48] can only be distributed in one non-trivial way $A : B$. Nevertheless, every qubit of the 4-party Smolin state *is* entangled with every other qubit, and yet no entanglement can be distilled between any of the parties unless two of them are able to meet. Smolin referred to this two-party distillation capability as "unlockability" [49] This feature could conceivably be useful in some advanced quantum "secret-sharing" protocol.

### 3.6.4 Superactivation of bound entanglement

The Smolin state can also be *superactivated* (see Figure 3.2): given two copies of the Smolin state $\rho^S_{ACBD}$ and $\rho^S_{ABCE}$, $D$ and $E$ can distill a maximally entangled pair [57].[5] $A$ and $B$ perform joint measurements on their qubits, executing an entanglement swap [58]. Although $A$ and $B$ are using the unknown entangled states they share with $C$ and $D$, respectively, as the resource for this swap, those resources are the *same*, since the 4 particles are in a Smolin state. As a result, $C$ and $D$ and $E$ finally share a Smolin state, with $C$ holding two of the qubits: a joint measurement of $C$'s qubits reveals which entangled state $D$ and $E$ share.

---

[5]At first, this fact may sound alarm bells in your mind; have we made a mistake in characterizing the Smolin state as BE? Is it actually just a free-entangled state, of the type in Section 3.6.1, that can only be distilled collectively? The reason that this superactivation "doesn't count" as distillation is that $\rho^S_{ACBD}$ and $\rho^S_{ABCE}$ are two different states, *not* two instances of the same state. Recall (Sections 3.5, 3.6.3) that the properties of the Smolin state are highly dependent on how the qubits are distributed.
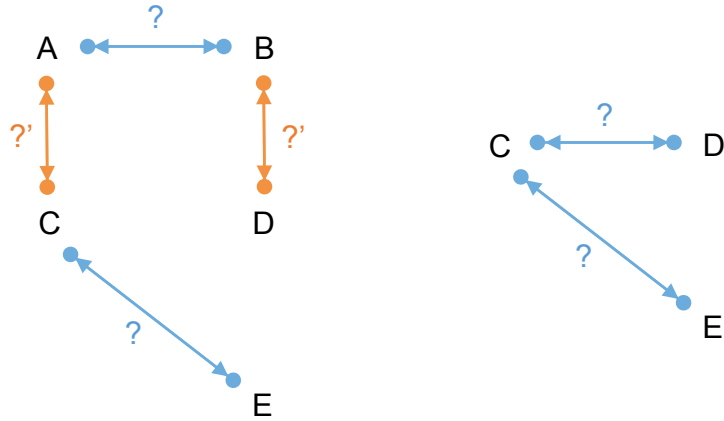
Figure 3.2: Schematic explanation of superactivation. Pairs of the same color constitute a Smolin state. As shown, the pair initially shared by $A$ and $B$ ($A$ and $C$) is in the same unknown Bell state as the pair shared by $C$ and $E$ ($B$ and $D$). $A$ and $B$ each use $\rho^S_{ACBD}$ as the resource for an entanglement swap, transferring their shares of $\rho^S_{ABCE}$ to $D$ and $E$, resulting in the final state where $C$ shares Bell states with both $D$ and $E$.

### 3.6.5 Conclusions

From a practical point of view, it may be hard to imagine BE as a technological resource. To begin with, as we see in Chapter 4, BE is not easy to generate (at least in the form of $\rho^S(p)$); if we were able to develop a sufficiently fine capability for manipulation of quantum states, it seems that we would prefer to simply deal with pure entangled states. The main interest in BE derives from the fact that it has these capabilities *despite being a "weak" form of entanglement*, challenging our understanding of the relationships between distillable entanglement, nonlocality, and cryptographic security.

# Chapter 4

# Photonic Bound Entanglement

> Quantum phenomena do not happen in
> Hilbert space. They happen in a laboratory.
>
> _____
>
> Asher Peres

## 4.1    Introduction

This chapter discusses our efforts towards the first optical realization of genuine bound entanglement. Section 4.2 gives a detailed description of our setup; Section 4.3 explains the major technical challenges we encountered along the way; Section 4.4 presents our results for each piece of the experiment: polarization entanglement, spatial-mode entanglement, the Smolin state, and BE. Finally, Section 4.5 discusses our work in the context of previous reports of BE.

## 4.2    The experimental setup

Because our experiment involves many steps requiring explanation, this section is divided into three parts for clarity. In our experiment, we need the pump photons to be in a certain polarization and spatial-mode state: Section 4.2.1 explains how we make the necessary transformations on the pump laser beam. There are 8 different SPDC processes involved in this experiment, and Section 4.2.2 shows how each of these processes arises, as well as the transformations we make to the SPDC after it is produced. Finally, we use quantum state tomography, explained in Section 4.2.3, to determine the polarization-spatial-mode state we have produced.
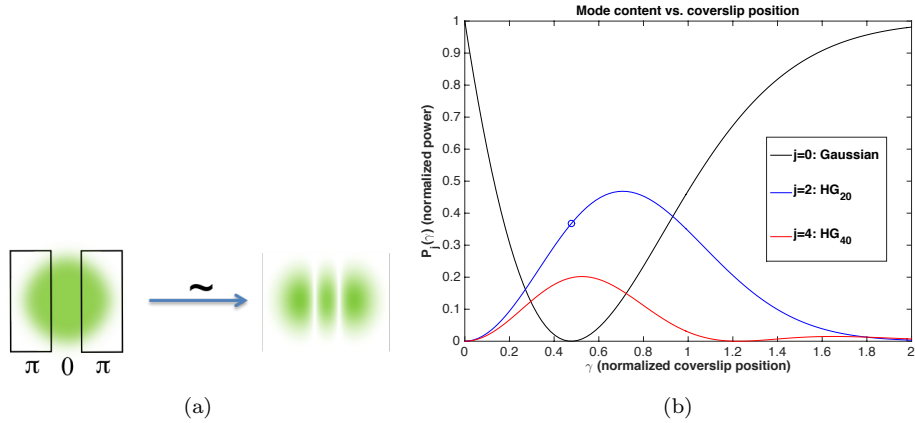
Figure 4.1: (a) Tilted coverslips apply a $\pi$-phase shift to the edges of the Gaussian output of our laser. We can achieve a high overlap with the $HG_{20}$ mode. (b) The three lowest-order mode non-zero contributions to the total power in the pump beam. We choose the coverslip edge positions $\pm\gamma$ such that the Gaussian power is zero, though that results in less power in the $HG_{20}$ mode (the point circled on the $HG_{20}$ curve). $P_4(\gamma)$ is shown for comparison.

### 4.2.1 Pump preparation

The pump polarization and spatial mode are imprinted on the SPDC state; we wish to prepare the pump photon state

$$
\rho_{\text{pump}}^{\text{ideal}} = \frac{1}{2} \left| D \right\rangle\!\left\langle D \right| \otimes \frac{\left| HG_{20} \right\rangle + \left| HG_{02} \right\rangle}{\sqrt{2}} \cdot \frac{\left\langle HG_{20} \right| + \left\langle HG_{02} \right|}{\sqrt{2}} +
$$
$$
\frac{1}{2} \left| A \right\rangle\!\left\langle A \right| \otimes \frac{\left| HG_{20} \right\rangle - \left| HG_{02} \right\rangle}{\sqrt{2}} \cdot \frac{\left\langle HG_{20} \right| - \left\langle HG_{02} \right|}{\sqrt{2}}. \quad (4.1)
$$

Here we explain how this first step of the experiment is accomplished.

**Producing the second-order spatial mode**

As shown in Section 2.2, we need a second-order pump spatial mode in order to tune the spatial mode of the SPDC. In this work we use a continuous-wave $\text{Ar}^+$ laser operating at the 351.1 nm line in the fundamental (Gaussian) mode. We can modify the laser cavity boundary conditions such that it is favorable for another spatial mode to lase instead. For example, we can position a conducting wire across the cavity such that it aligns with one of the nulls of a $HG_{20}$ mode. This enforces the condition $\mathbf{E} = 0$ along the length of the wire, which only modes $HG_{2,n}$ satisfy exactly. Unfortunately, this solution is unsuitable in our case; although the final experiment uses a second-order pump, the setup must be aligned using a Gaussian pump, so a less drastic modification of the system is desirable.

Note from Eq. 2.3 that each lobe of $HG_{mn}$ is $\pi$ out of phase from its neighbors[1]. Our approach here is to apply a spatially non-uniform phase to the Gaussian output of the laser such that the resulting phase profile matches that of a $HG_{20}$ mode. We get the required phase shifts by positioning thin pieces of glass on the left and right sides of the beam, with their edges at the nulls of the $HG_{20}$ (Figure 4.1). The coverslips can be tilted so that they apply phase shifts $(2n+1)\pi$, equivalent to an odd number of half-wavelengths.

While the optimal coverslip tilt condition is clear, the optimal edge position is *not*. The mode conversion has two objectives that are not equivalent:

1. Maximize the $HG_{20}$ mode content of the beam.

2. Minimize the Gaussian $HG_{00}$ content of the beam.

Not surprisingly, different edge positions are required to satisfy conditions (1) and (2). The second condition is important because, as shown in Section 2.2, the Gaussian is the only other mode capable of producing first-order spatial mode signal-idler pairs. A residual Gaussian component could swamp the state that we want to get from the second-order pump modes. Moreover, as we shall see, there is a simple way to determine when we have satisfied the second condition. If we constrain the coverslips to apply $\pi$-phase shifts symmetrically on either side of the beam, we only need to consider even-index modes:

$$|\psi\rangle = \sum_{m,n=0}^{\infty} |HG_{m,n}\rangle \langle HG_{m,n}|\psi\rangle = \sum_{j \text{ even}} |HG_{j,0}\rangle \langle HG_{j,0}|\psi\rangle, \tag{4.2}$$

where $|\psi\rangle$ is the spatial mode state after the coverslips. The coefficients in the sum are

$$\langle HG_{j,0}|\psi\rangle = \int_0^{\infty} dx \cdot (HG_{j,0}(x,y))^* (\phi(x)HG_{00}(x)), \tag{4.3}$$

where $\phi(x)$ is the spatially non-uniform phase applied by the coverslips, equal to $-1$ in their presence and $1$ in their absence. The coefficient Eq. 4.3 can be simplified analytically, and the optical power in the mode is given by the absolute square (Figure 4.1):

$$P_j(\gamma) = \begin{cases} (1 - 2 \cdot \text{erf}(\gamma))^2 & j = 0, \\ \frac{16}{2^j \pi j!} \left( H_{j-1}(\gamma)e^{-\gamma^2} \right)^2 & j > 0 \text{ and even,} \end{cases} \tag{4.4}$$

where the coverslip edges are located at $\pm w_0 \gamma/\sqrt{2}$ for beam width $w_0$, and $\sum P_j = 1$. Although we can achieve the maximum $HG_{20}$ content $P_2(\gamma = 1/\sqrt{2}) = 0.468$, there is still considerable power in the Gaussian

---

[1]The beam is collimated, so we can ignore the phase factor $\exp(-i\frac{k(x^2+y^2)}{2R(z)})$.
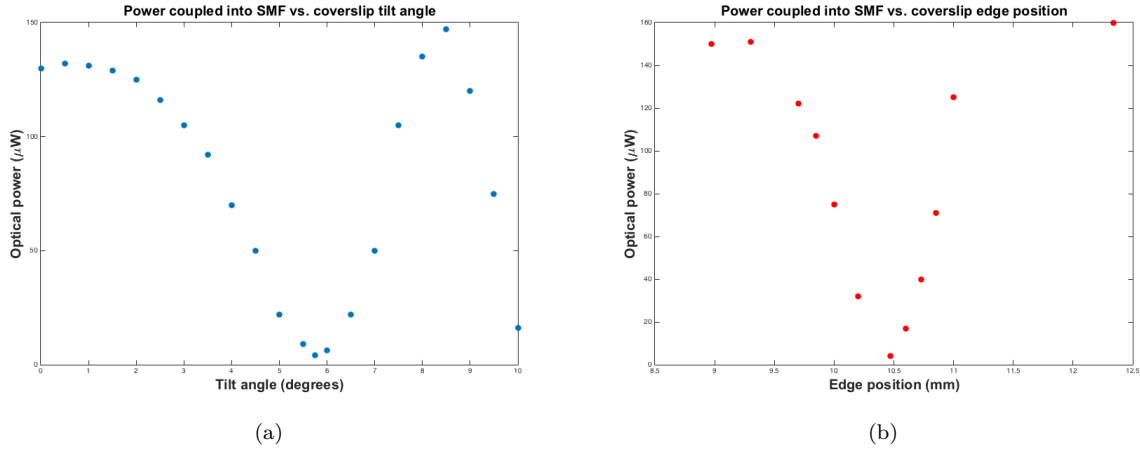
Figure 4.2: Experimental demonstration of coverslip mode conversion. As the initial Gaussian mode is converted to $HG_{10}$, SMF coupling efficiency drops. (a) With its edge near the center of the beam, the coverslip is tilted, changing the relative phase between the covered and uncovered halves of the beam. The relative phase at zero tilt depends on the thickness of the coverslip modulo $\lambda$. (b) At $6°$ tilt, the edge of the coverslip is moved across the beam. Edge position zero (3.04) corresponds to almost all of the beam being covered (uncovered).

here, $P_0(1/\sqrt{2}) = 0.133$. Instead we choose $\gamma = 0.477$, where $P_0 = 0$ and $P_2 = 0.368$.

Experimentally, we can evaluate the success of this mode conversion by coupling the unmodified laser beam into SMF, which only transmits the Gaussian mode, moving in the coverslips, and observing the drop in coupling efficiency (Figure 4.2). To indicate the higher-order mode content of the beam (i.e., higher than second order), we will refer to the state as $|HG'_{20}\rangle$; the primary effect of these higher-order modes will be to reduce the power in the desired mode and, therefore, to reduce the first-order mode SPDC rate as well.

**Entangling the polarization and spatial mode**

We first use a *Dove prism* to spatially rotate the pump beam (Figure 4.3). In a Dove prism, a beam incident on one of the slanted faces refracts and impinges on the long face at an angle larger than the total internal reflection critical angle, reflecting and following a symmetric path out of the prism. The incident beam profile is reflected in a plane that depends on the orientation of the prism, so that rotating the prism causes the reflected intensity distribution to rotate at twice the rate (e.g., the outputs of a $0°$-oriented prism and a $90°$-oriented prism differ by a $180°$ rotation). For two-fold symmetric beams like the Hermite-Gauss modes the net effect is indistinguishable from a pure rotation. In our setup, a Dove prism at $22.5°$ first rotates the $|HG'_{20}\rangle$ state by $45°$. The beam starts in polarization state $|D\rangle$.

The $|HG'_{20}\rangle$ at $45°$ then enters the polarization Sagnac interferometer diagrammed in Figure 4.3. $|H\rangle$ ($|V\rangle$) is transmitted (reflected) and takes the counter-clockwise (clockwise) path of the interferometer. The two
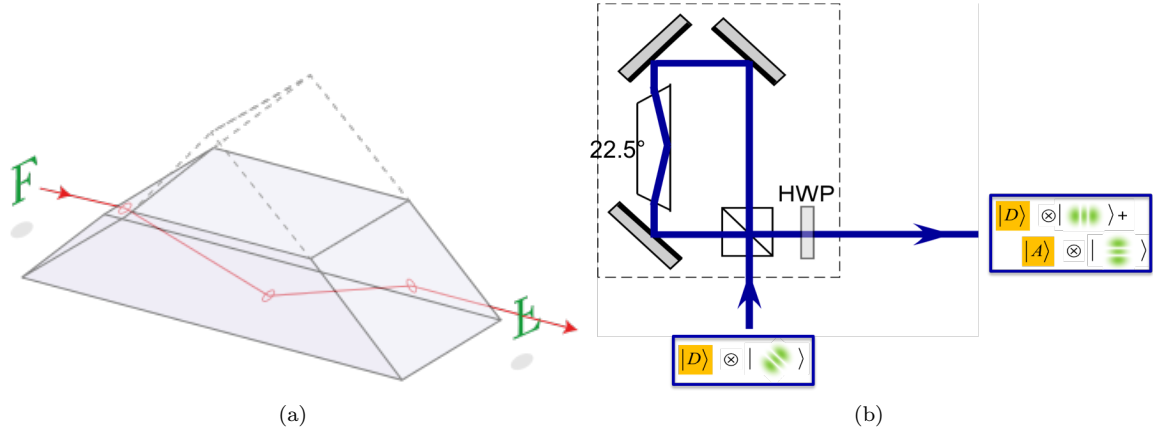
Figure 4.3: (a) We use Dove prisms to rotate our $|HG'_{20}\rangle$ beam. Note, however, that the Dove prism does *not* perform a rotation: instead it reflects the beam about a rotatable axis. This operation happens to be indistinguishable from a rotation for our two-fold symmetric beam. [*Image copied from https://en.wikipedia.org/wiki/Dove_prism on 4 June, 2016.*] (b) Our polarization Sagnac interferometer entangles the polarization and spatial mode of each pump photon. $|V\rangle$ ($|H\rangle$) takes the (counter)clockwise path and encounters a Dove prism at $\pm 22.5°$. Thus, for the two paths the effective rotation angles are $\pm 45°$.

beams undergo the same number of reflections mod 2 before reaching the Dove prism in the interferometer, so when they do, their orientations are the same. However, while $|V\rangle$ sees a Dove prism oriented at $22.5°$, rotating the clockwise (CW) beam an additional $45°$, $|H\rangle$ sees the orientation as $-22.5°$, undoing the effect of the first Dove prism for the counter-clockwise (CCW) beam. Hence, the output of the interferometer is

$$|\psi\rangle_{\text{pump}}^{\text{actual}} = \frac{1}{\sqrt{2}}(|H, HG'_{20}\rangle + |V, HG'_{02}\rangle). \tag{4.5}$$

**Dephasing the two pump components**

The next element of the pump preparation is an unbalanced polarization Mach-Zehnder interferometer (MZI). As shown in Figure 4.4, the path length difference is more than the laser *coherence length* $\sim 2\pi c/\Delta\omega$, so that there is no definite phase relationship between the beams in the two paths when they are recombined. By scanning the path length difference, we measured the laser coherence length to be approximately 1 cm. The output is similar to the ideal pump state Eq. 4.1:

$$\rho_{\text{pump}}^{\text{actual}} = \frac{1}{2}|D\rangle\langle D| \otimes \frac{|HG'_{20}\rangle + |HG'_{02}\rangle}{\sqrt{2}} \cdot \frac{\langle HG'_{20}| + \langle HG'_{02}|}{\sqrt{2}} +$$
$$\frac{1}{2}|A\rangle\langle A| \otimes \frac{|HG'_{20}\rangle - |HG'_{02}\rangle}{\sqrt{2}} \cdot \frac{\langle HG'_{20}| - \langle HG'_{02}|}{\sqrt{2}}. \tag{4.6}$$
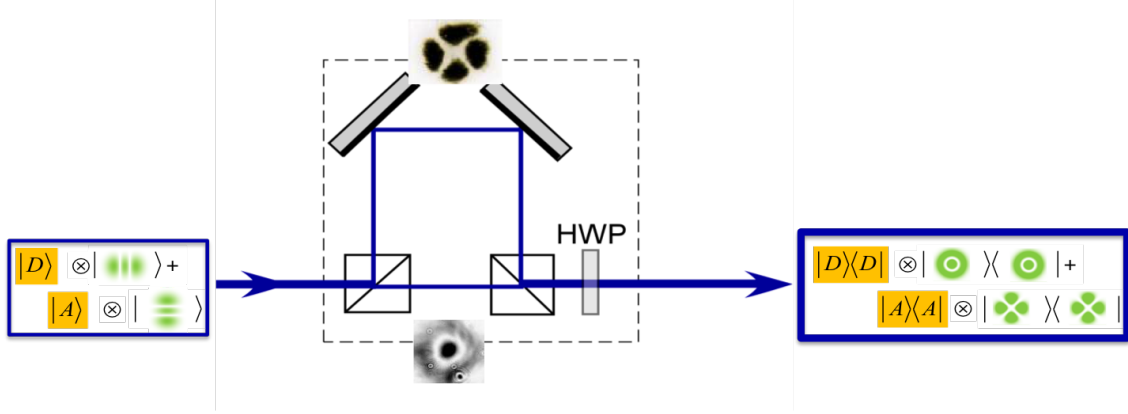
Figure 4.4: Our polarization MZI delays one polarization by longer than the coherence time of the laser. As a result, when the two polarizations are recombined their states no longer add coherently. In the two arms of the interferometer we observe the superpositions $\frac{1}{\sqrt{2}}(|HG_{20}\rangle \pm |HG_{02}\rangle)$. Shown in the diagram are intensity profiles that were observed in the lab.

### 4.2.2 SPDC

We then send the beam to our SPDC crystals (Figure 4.5). The pump beam Eq. 4.6 makes two passes through the SPDC crystals, each time producing the SPDC state

$$\rho_{\text{SPDC}}^{\text{second pass}} = \frac{1}{2} \left| \Phi^+ \right\rangle\!\!\left\langle \Phi^+ \right| \otimes \left| \phi^+ \right\rangle\!\!\left\langle \phi^+ \right| + \frac{1}{2} \left| \Phi^- \right\rangle\!\!\left\langle \Phi^- \right| \otimes \left| \phi^- \right\rangle\!\!\left\langle \phi^- \right|, \tag{4.7}$$

where, as before, $|\Phi^\pm\rangle$ ($|\phi^\pm\rangle$) represent the Bell states $\frac{|HH\rangle \pm |VV\rangle}{\sqrt{2}}$ ($\frac{|hh\rangle \pm |vv\rangle}{\sqrt{2}}$), maintaining the convention that capital letters represent polarization and lower-case letters represent first-order spatial mode states. The terms in Eq. 4.7 can be interpreted as the first two terms of the Smolin state Eq. 3.22. The SPDC produced on the first pass of the pump is reflected back through the crystals to join the second-pass SPDC. However, in the first-pass signal arm a QWP swaps $|H\rangle$ and $|V\rangle$ polarizations; similarly, in the idler arm a cylindrical lens with focal point at the back-reflecting mirror swaps $|h\rangle$ and $|v\rangle$ spatial modes. After these transformations, the SPDC produced by the first pass of the pump ends up in the state

$$\rho_{\text{SPDC}}^{\text{first pass}} = \frac{1}{2} \left| \Psi^+ \right\rangle\!\!\left\langle \Psi^+ \right| \otimes \left| \psi^+ \right\rangle\!\!\left\langle \psi^+ \right| + \frac{1}{2} \left| \Psi^- \right\rangle\!\!\left\langle \Psi^- \right| \otimes \left| \psi^- \right\rangle\!\!\left\langle \psi^- \right|, \tag{4.8}$$

so that we have produced all four terms in Eq. 3.22. But how do we know with what degree of coherence the terms in Eqs. 4.7 and 4.8 add when the two processes are combined? Two conditions are required for the processes to be coherent [59].

1. The first-pass signal and idler photons must return simultaneously to the SPDC crystals: because the second-pass signal and idler leave the crystal simultaneously, a difference in arrival time between the
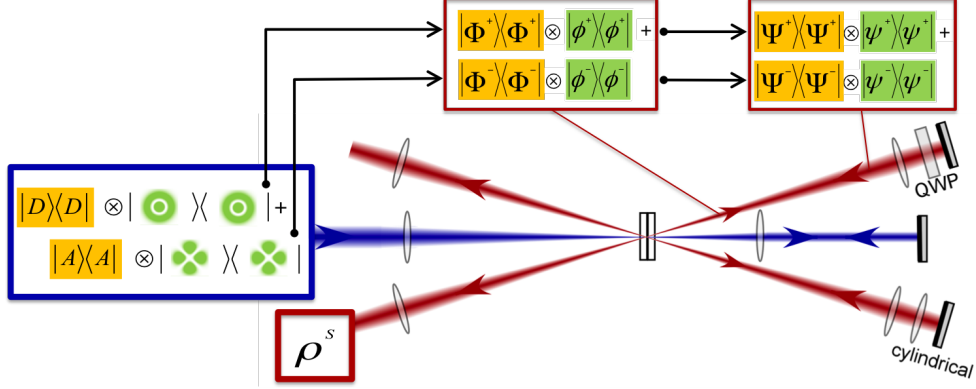
Figure 4.5: Schematic of the SPDC stage of our experiment. The pump generated in the previous stage makes two passes through a pair of nonlinear BiBO crystals. Each pass of the pump generates the SPDC polarization-spatial-mode state $\frac{1}{2}\left(|\Phi^+\rangle\langle\Phi^+| \otimes |\phi^+\rangle\langle\phi^+| + |\Phi^-\rangle\langle\Phi^-| \otimes |\phi^-\rangle\langle\phi^-|\right)$. We additionally perform a $\pi$-mode conversion on both polarization and spatial mode so that the first-pass SPDC is prepared in the state $\frac{1}{2}\left(|\Psi^+\rangle\langle\Psi^+| \otimes |\psi^+\rangle\langle\psi^+| + |\Psi^-\rangle\langle\Psi^-| \otimes |\psi^-\rangle\langle\psi^-|\right)$. The combination of the SPDC from both passes is our encoding of the Smolin state.

first-pass signal and idler would make the two processes distinguishable.[2]

2. The first- and second-pass pump photons must be coherent. That is, the second-pass pump must also return to the crystals at the same time as the first-pass signal and idler photons, but only within the (much longer) coherence time of the pump.

Therefore, to make Eqs. 4.7 and 4.8 add *incoherently*, it is sufficient to place the pump mirror more than one pump coherence length further from the crystals than the signal and idler mirrors. Although there are other ways to destroy the coherence between the two processes, we will see in Section 4.5 that this is the way we need to do it in order to demonstrate BE. Putting all the terms together, then, we get the Smolin state

$$\rho^S_{\mathrm{SPDC}} = \frac{1}{4}\left|\Phi^+\right\rangle\!\!\left\langle\Phi^+\right|\otimes\left|\phi^+\right\rangle\!\!\left\langle\phi^+\right| + \frac{1}{4}\left|\Phi^-\right\rangle\!\!\left\langle\Phi^-\right|\otimes\left|\phi^-\right\rangle\!\!\left\langle\phi^-\right| + \frac{1}{4}\left|\Psi^+\right\rangle\!\!\left\langle\Psi^+\right|\otimes\left|\psi^+\right\rangle\!\!\left\langle\psi^+\right| + \frac{1}{4}\left|\Psi^-\right\rangle\!\!\left\langle\Psi^-\right|\otimes\left|\psi^-\right\rangle\!\!\left\langle\psi^-\right|,$$

(4.9)

where the signal and idler polarizations are in one of the four entangled states $|\Phi^\pm\rangle, |\Psi^\pm\rangle$ and their spatial modes are in the "same" state $|\phi^\pm\rangle, |\psi^\pm\rangle$.[3]

---

[2]Here, "simultaneous" means "within the SPDC coherence time". In our experiment, we effectively set the bandwidth $\omega_{SPDC}$ with postselecting spectral filters. Our detection bandwidth is 5 nm, centered close to 702 nm, to detect frequency-degenerate signal-idler pairs. This corresponds to a coherence length of $100\mu m$.

[3]Of course, it doesn't really mean anything to say that the polarizations and spatial modes are in the "same" state, since they're different physical attributes of the photon. Anyway, it isn't important that the states be the same, since the BE property only depends on the states for each pair (in our case, the polarization pair and the spatial-mode pair) being orthogonal maximally entangled states.

### 4.2.3 Quantum state tomography

We use quantum state tomography (QST) to infer the density matrix of the state we have produced [60]. In QST we assume that every SPDC pair entering our measurement system is identical. We perform measurements of the mutually unbiased basis states $|H\rangle, |V\rangle, |D\rangle, |A\rangle, |L\rangle, |R\rangle$ ($|h\rangle, |v\rangle, |d\rangle, |a\rangle, |l\rangle, |r\rangle$) on each of the polarization (spatial-mode) qubits, a total of $6^4 = 1296$ measurements! For the spatial mode measurement, we send each photon to a holographic diffraction grating, which has 6 different regions. We couple the +1 diffraction order of the grating into a SMF [61]. The diffraction efficiency into the +1 order is approximately 25%, so 75% of the time the photon never makes it into the measurement system at all. When it does, each region transforms one of the 6 mutually-unbiased basis states into a Gaussian (the only mode that couples well into a SMF), and transforms the orthogonal state into a second-order mode. Thus, the hologram selects the measurement basis and the SMF executes the projective measurement. After the SMF, we perform an analogous procedure for the polarization: here, a rotatable QWP and HWP select the measurement basis and a polarizer set to transmit $|H\rangle$ executes the projective measurement. Interference filters (IFs) are then used to spectrally filter the signal and idler. Finally, the photons are detected with approximately 60% efficiency by avalanche photodetectors. At each measurement setting, a certain separable 4-qubit (2-photon) state would ideally be detected with unit probability. Therefore, we count the number of *pairs* detected at each measurement setting. In particular, when the signal and idler detectors click within 5 ns of each other, we assume that we have detected a pair from a single SPDC event (we do subtract off the average "accidental" coincidences —the average number of times we expect both detectors to click within a 5-ns time window by random chance, given the count rates on each detector —but in our experiment these constitute a negligible fraction of the signal anyway[4]).

If these coincidence measurements yielded precise *probabilities*, we could retrieve the density matrix by linear inversion[5] [62]. Since we do not have infinite copies of our state, however, the number of photon detections at each measurement setting is also affected by shot noise. Furthermore, as is common in quantum optics experiments, we are working in a regime where shot noise is a significant fraction of the average count rate. Therefore, linear inversion often returns an unphysical density matrix, especially for states with high purity, which are close to the edge of the Hilbert space. Instead, we use a "maximum-likelihood" technique, i.e., we identify the state for which the probability to produce the observed counts is highest [60]. We use a bootstrapping Monte Carlo technique to measure uncertainties on the density matrix we infer.

---

[4]Say we have coincidence window $\Delta t$ (in seconds), and signal, idler detection rates $r_s$, $r_i$ (in Hz). The average number of idler detections coinciding with a *specific* signal detection is $r_i \Delta t$ in 1 second. Therefore, the accidental coincidence rate is $r_c \approx r_s r_i \Delta t$, assuming $r_s \Delta t$, $r_i \Delta t \ll 1$.

[5]In fact, this set of measurements overdetermines the density matrix. The minimal number of measurements required is 4 per qubit (e.g., $|H\rangle, |V\rangle, |D\rangle, |L\rangle$: one per basis, plus one more for normalization), or $4^4 = 256$ for a 4-qubit state. However, the additional measurements make the results less susceptible to experimental drift.
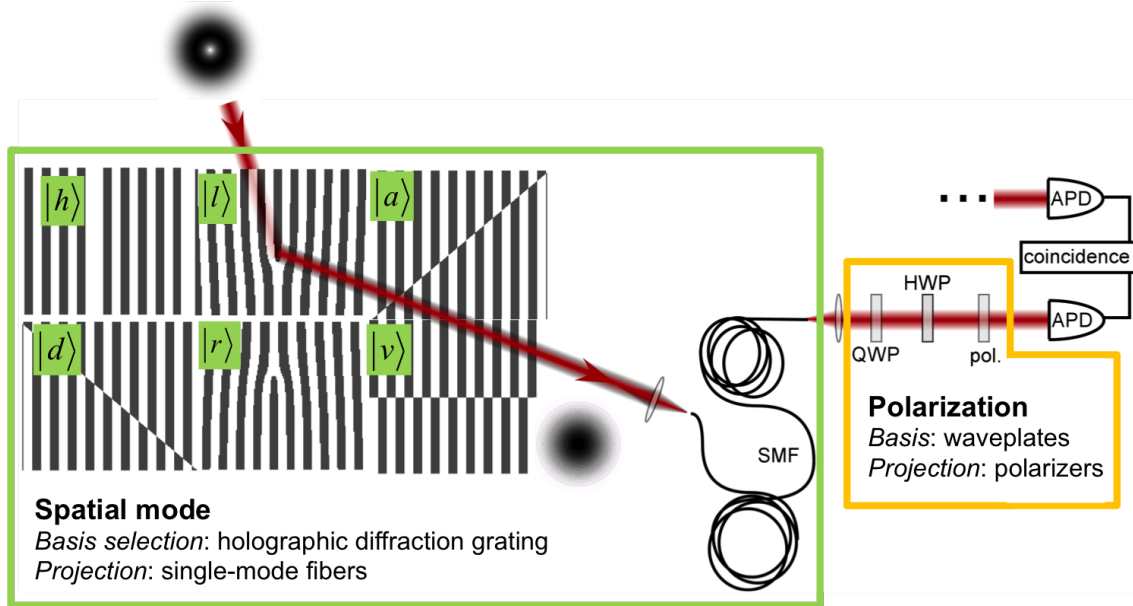
Figure 4.6: A schematic of our tomography system. We use a SMF to perform a projective measurement of the first-order spatial mode in a basis chosen by the hologram setting. After the spatial mode measurement, we make a similar polarization measurement, in which a polarizer makes a projective measurement in a basis chosen by a QWP, HWP combination. We count signal and idler detections in coincidence.

## 4.3 Technical challenges

In the previous section, we saw how the experiment *should* work —in practice, however, we encountered a number of difficulties at many points along the way. This section details the most important challenges we faced.

### 4.3.1 Sagnac interferometer visibility

The bulk of the work on this experiment was actually done without the mode-converting coverslips (Section 4.2.1), using the mode straight from the laser. For a Gaussian input, the Sagnac interferometer from Section 4.2.1 should simply split $|H\rangle$ and $|V\rangle$ and recombine them: ideally it does nothing at all. The visibility of the interferometer $V_{\text{Sagnac}}$, then, is related to the purity of the output polarization state. We developed the following procedure for aligning the interferometer:

1. Remove the Dove prism from the interferometer and align the three mirrors to maximize the visibility.

2. Move the Dove prism in, but keep it at $0°$. Again align the mirrors to maximize visibility.

3. Rotate the Dove prism to $22.5°$, and maximize the visibility again. It was found to be helpful to do the rotating and maximizing incrementally.

While we observed $V_{\text{Sagnac}} > 98\%$ after steps (1) and (2), we could barely achieve 90% after step (3). Evidently, the poor visibility is due to the rotated Dove prism: here we consider the complications it introduces.

The Dove prism affects the beam polarization: because the reflection happens out of the plane of the optical table, in general, the Dove prism's eigen-polarizations are *not* $|H\rangle$ and $|V\rangle$ [63]. The two eigen-polarizations get different phase shifts on total internal reflection, leading to a rotation of the polarization state in both CW and CCW paths. This rotation is ultimately erased, however, since the beams must pass through the polarizing beam-splitter (PBS) again in order to exit the interferometer: polarization errors only reduce the transmission, not the visibility.

Shifting our attention to the spatial mode, the CW and CCW beams are necessarily identical until step (3): when the Dove prism is at $0°$, the reflection happens in the horizontal plane and is thus the same for both paths, leading to the same spatial profile for the CW and CCW beams. After step (3), the reflection happens in planes at $\pm 22.5°$ for CW and CCW, respectively. Since different parts of the spatial profile are interfering with each other, the visibility will suffer if the beam is not radially symmetric, either because of a poor mode from the laser or subsequent aberrations in the system. However, beam profile intensity measurements suggest (this method is only sensitive to intensity, not phase) that the beam going into the interferometer is not the problem. It seems more likely that the effect is due to an aberration caused by the Dove prism itself, that is experienced differently by the CW and CCW paths. If the beam is not well enough collimated, the Dove prism introduces astigmatism. Because the axis of the Dove prism is different for the CW and CCW beams, so is the axis of the astigmatism. Therefore, the aberration does not "cancel out" when the beams are recombined.

Whatever the cause of the aberrations from the Sagnac interferometer, they effectively further reduce the second-order mode content of our pump Eq. 4.6, reducing the quality of the SPDC spatial-mode state. Although it seems that the Sagnac visibility was not the main limitation on our state quality, it had other implications that are discussed later in this section.

### 4.3.2 Laser instability

Perhaps the most significant (and certainly the most frustrating) problem in this experiment was the pump laser's instability. We could monitor this instability via $V_{\text{Sagnac}}$. A notable feature of Sagnac interferometers is their stability: since the two paths ideally only differ in the sense in which the beams traverse them[6],

---

[6]For the case of an odd number of mirrors, it is also often possible to construct a "displaced" Sagnac, with parallel, equal-length CW and CCW paths that do not coincide. Unfortunately, since the Dove prism constitutes another reflecting element in our interferometer, a displaced configuration is impossible here.
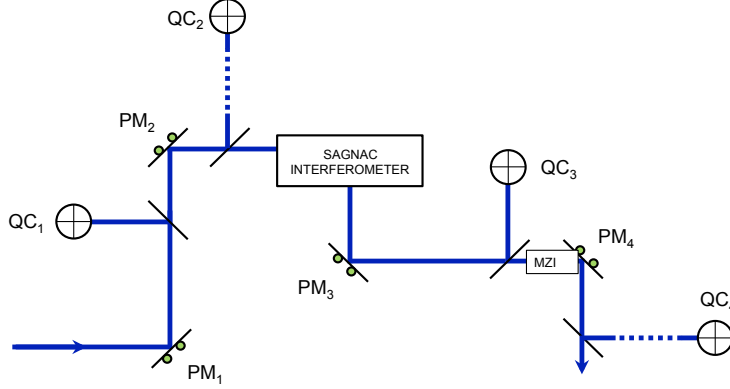
Figure 4.7: Schematic of our beam-pointing stabilizing feedback control system. Each of the $PM_j$ moves to maintain a constant signal at $QC_j$. The MZI poses a problem as the two paths are different lengths, so it is impossible for the $PM_3$ to $QC_3$ distance to match the $PM_3$ to $PM_4$ distance for both MZI paths. Nevertheless, control systems 3 and 4 did improve the system's overall stability.

Sagnacs are protected from many of the problems that Michelson and Mach-Zehnder interferometers are susceptible to, e.g., vibrations or temperature fluctuations that may have local effects on optical path lengths. As discussed in Section 4.3.1, $V_{\text{Sagnac}}$ was *low*, but in view of the geometry, we expected it to be *stable* as long as the input laser beam is stable. Unfortunately, for a number of reasons this was not the case, and the visibility would stay stable for time periods ranging from less than an hour to more than a day, presumably depending on specific conditions in the lab that we were unable to identify with certainty. Nevertheless, let us first look at the instabilities in the laser that we *do* understand.

1. *The power out of the laser was not constant:* Looking at the laser power on an oscilloscope revealed that in addition to the desired constant power there was a component at 360 Hz. Moreover, the amplitude of this oscillation was very large, with the power dropping to less than $1/2$ of its peak value —indeed, it doesn't seem quite right to describe the laser as continuous-wave at all! Since our measurement times for each setting in the experiment always spanned at least several periods, the oscillating power was not intrinsically a problem. However, it might have complicated our attempts to fix other instabilities.

2. *The laser beam direction was unstable:* This came as no surprise. As noted in Ref. [64], it had been shown earlier that the beam direction of this laser fluctuates as much as 18 $\mu$radians over a few hours; because the angular deviations of the pump are amplified $\sim 5$ times in the signal and idler wave-vectors, this fluctuation corresponds to nearly 100 $\mu$radians at our spatial-mode analyzers, where the signal and idler have beam divergence of approximately 150 $\mu$radians. Therefore, an external feedback control system is required [65].

Following Ref. [64], we set up two control systems, each consisting of a piezoelectric-actuated mirror (PM), a quadrant photodiode (quad cell, QC), and a control circuit (Figure 4.7). $QC_1$ and $PM_2$ were positioned at equal distances after $PM_1$, thereby avoiding the need for imaging optics; $QC_2$ was placed far

from $PM_2$. We found the analog circuitry —specifically, the normalization performed by the analog voltage divider AD538 —problematic, so we built a digital version of the control circuit. With the beam centroid position fixed by $PM_1$ at $QC_1$, and therefore at $PM_2$, and the beam direction fixed by $PM_2$, we expected that our beam-pointing fluctuation woes were at an end, but, alas, this was not so: the Sagnac visibility would still drop a few percent after several hours. More importantly, the alignment of our spatial mode analyzers, which will be discussed in more detail later in this section, would change by as much as $\sim 50\mu$m in both transverse directions over the course of a day. While the two beam-pointing control systems can together correct fluctuations that occur before the beam enters them, they *cannot* fix anything that happens after $PM_1$. Therefore, we suspect that in addition to the laser's beam pointing fluctuations, there were drifts somewhere in the optical system after $PM_1$, which were ultimately responsible for the inability of these two control systems to stabilize the experiment. Since we were unable to fix these instabilities, and since the fluctuation in the Sagnac visibility was not our main concern, we installed an additional pair of similar feedback systems after the Sagnac (Figure 4.7). These stabilized the beam enough that we could use the same spatial mode analyzer settings for a few days before the drift became too significant.

### 4.3.3 Laser backreflection

An important drawback of our experiment —one that almost made it entirely unfeasible —is that the pump is reflected straight back toward the laser. The returning beam, having a different polarization and spatial profile than the original laser beam, can cause other modes of the laser cavity to compete, leading to further instability in power and beam pointing. We needed to optically isolate the laser from the rest of the system. We tried two different approaches. The "brute-force" solution is to use a Faraday isolator (FI) [14]. FIs use magnetic fields to seemingly violate time-reversal symmetry, thereby blocking all light propagating in the reverse direction. FIs for many wavelengths are readily available commerically, but not for 351 nm. While we were able to have one custom made, it had approximately 40% transmission, (failing to reach the 70% transmission specification). Moreover, the low damage threshold specification meant that our incident beam could have power no higher than 300 mW. Our system suffers from high losses, so this was already an unpleasant limitation, but not an insurmountable one. A more serious limitation was discovered when we noticed that the count rates on the *right-hand side* of the crystals depend on whether or not the pump is reflected to the left, back through the system (Figure 4.8). We traced this puzzle back to the FI and ultimately determined that well below its damage threshold is a lower threshold at which thermal lensing becomes significant[7]. As shown in Figure 4.8, we were forced to operate with no more than 20 mW incident

---

[7]The reduced count rate was due to the additional light from the back-reflected beam exacerbating the lensing in the FI. The increased lensing modified the forward-propagating beam such that it was less well matched to our SMF collection.
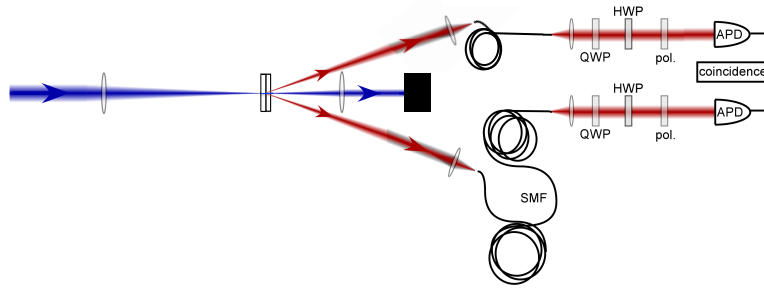
(a)
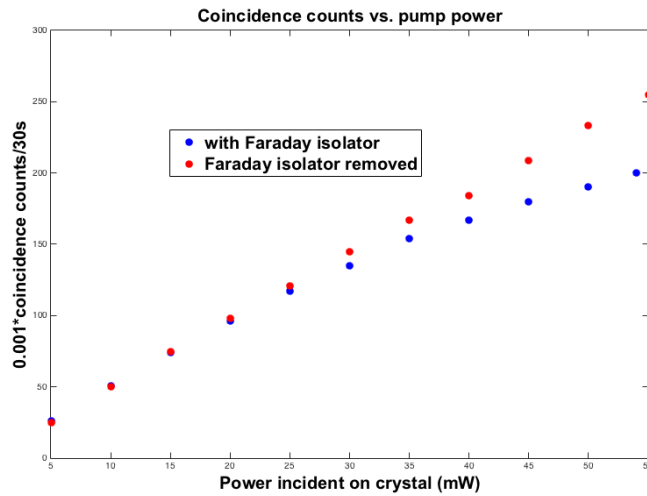


(b)

Figure 4.8: (a) We were quite shocked to find that our count rates in this setup depended on whether or not the beam block was in place before the last mirror! The FI was evidently past its threshold for thermal lensing and the back-reflected beam was therefore modifying the forward-propagating beam. (b) As the plot shows, when the FI is in the system, the SPDC collection rate deviates from its typical linear dependence on incident power.

on the crystal in order to avoid this lensing effect.

In an effort to get more power and avoid other aberrations that might have been less apparent, we eventually abandoned the FI in favor of the following approach. A linear polarizer and QWP can also work as an optical isolator, but only for one pair of orthogonal polarizations, as shown in Figure 4.9. In our case, some care is required to make this scheme work. Here, we will assume a Gaussian pump for simplicity, since a very similar analysis applies to the $|HG'_{20}\rangle$ case. Firstly, the SPDC crystals introduce transverse walkoff, so that the incident $|D\rangle$ is split into spatially separated $|H\rangle$ and $|V\rangle$ components, which are unaffected by the QWP. To remedy this, we include an identical pair of crystals oriented at $180°$ relative to the SPDC crystals. These crystals apply the same transverse walkoff in the opposite direction so that the QWP works as desired. In our experiment, the PBS shown in Figure 4.9 is the last element of the MZI discussed in Section 4.2.1, so, in fact, light enters the system through *both* ports of the PBS. Consequently, it is not quite enough to make the beam entering from the transmitted (reflected) port exit through the reflected (transmitted) port: we must follow it back through the whole system and make sure that it is finally rejected by a polarizer somewhere along the way. In our setup, the beam has to travel surprisingly far back for this to happen. The PBSs in the MZI undo what they did for the incident beam. That is, the beam that took the short path in the forward direction takes the long path in the backward direction and vice versa. Thus, the dephasing caused by the MZI on the forward beam is undone by the MZI on the backward beam, restoring the state we (ideally) get after the Sagnac in the forward direction: for a Gaussian input, this is $|D, g\rangle$. This state then traverses the Sagnac in reverse. Recall that for an ideal Gaussian beam the Sagnac *should* simply separate $|H\rangle$ and $|V\rangle$ and recombine them, ultimately doing nothing to the state. If this were so, the beam would then continue back through the system until it reached the first polarizer in our system, where we finally use only one port, and where it would finally be rejected. And this is where the poor visibility of the Sagnac (Section 4.3.1) actually becomes a problem. Since the output polarization of the Sagnac is in reality a mixed state, the polarizer *cannot* completely reject it (the very definition of mixture, Section 3.2). Therefore, this scheme is inevitably consigned to send back to the laser a fraction $(1 - V_{\text{Sagnac}})/2$ of the return beam, approximately 5% in our case.

## 4.4 Experimental results

Here we present our experimental results. As we will see, our setup was very sensitive to alignment of the SMFs in the spatial mode tomography system. Furthermore, we had difficulty finding a metric for evaluating the alignment of the fibers with adequate sensitivity. While we were eventually able to modify the system
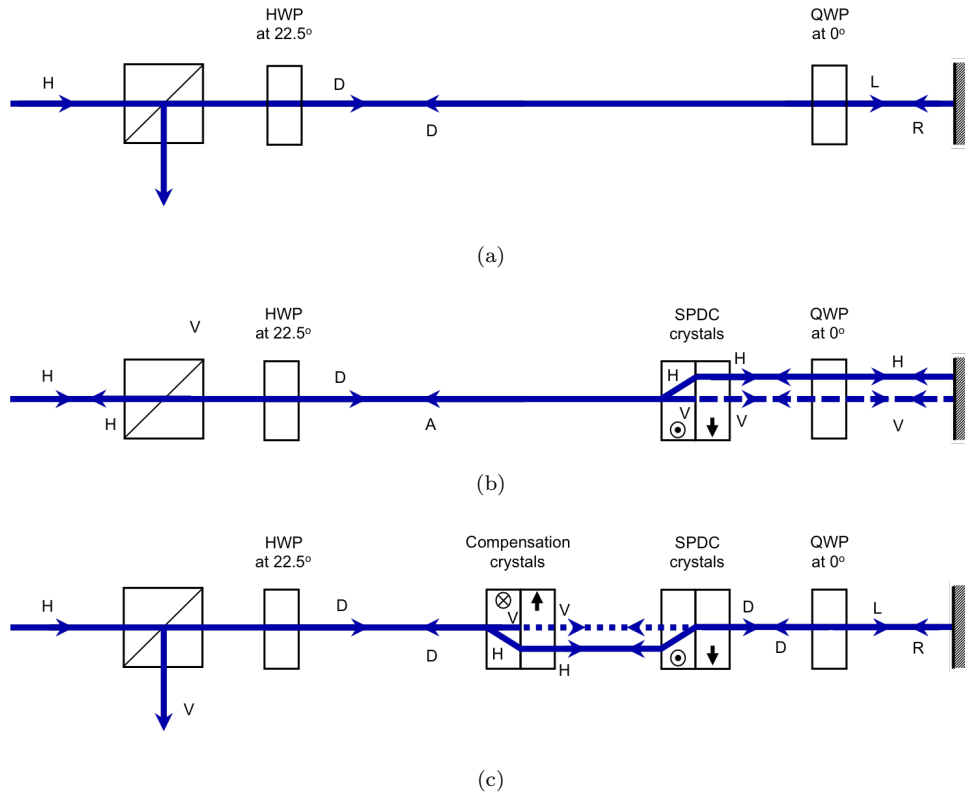
Figure 4.9: (a) A QWP can be used to prevent destabilizing reflections of the pump back into the laser. (b) The SPDC crystals cause walkoff, so that the QWP isolator no longer works. The dashed line indicates that the beam is propagating above the plane of the page. (c) We can still use this technique as long as we compensate for the walkoff. The dotted line indicates that the beam is propagating below the plane of the page.

so that we could achieve this sensitivity, the experimental results obtained in this configuration were not as good as those we obtained from the initial setup; this work is therefore still ongoing. In this section we discuss experimental details, starting with the original experimental configuration and results (Section 4.4.1) and continuing to the modified version of the setup (Section 4.4.2). The results in this section are inferred density matrices for 2-qubit polarization states, 1-qubit spatial mode states, 2-qubit spatial mode states, and 4-qubit states. In each case, the analyzers for the qubits that are not mentioned in the state description are kept fixed at a certain setting.

### 4.4.1   The original setup

In this configuration, we used the FI and restricted the pump power incident on the crystals to 20 mW (see Section 4.3.3). After the MZI (Section 4.2.1), the pump beam is focused down by a 1-m focal length biconvex lens such that it has beam waist radius 75 $\mu$m at the SPDC crystals. After the SPDC crystals, the pump propagates 1 m, where it is re-collimated by an identical biconvex lens, before reflecting back through the whole system. The signal and idler generated on the first pass propagate 500 mm before being collimated by 500-mm focal length achromatic doublets (Thorlabs AC254-500-B). After the polarization and spatial mode conversions (Section 4.2.2), they are reflected back to the SPDC crystals where they join the SPDC produced on the second pump pass. The first- and second- pass signal and idler then both propagate 500 mm to another pair of 500-mm achromats. The collimated SPDC is then sent to the tomography system. After diffracting from the hologram, the SPDC is coupled to SMF (Thorlabs SM600) through an 8.0-mm focal length aspheric lens (Thorlabs C240TME-B). In this configuration, the idler IF was a 13-nm "top-hat" filter centered at 700 nm with peak transmission over 99% (Semrock FF01-700/13-25). The signal IF had a roughly Gaussian transmission profile centered at 702 nm, with a 5 nm full width at half maximum transmission bandwidth and (unfortunately) peak transmission of 57.7%.

With a Gaussian pump mode (i.e., with the mode-converting coverslips out of the way), we collected the first-pass SPDC into SMFs on the right-hand side (RHS) of the crystals and directed these photons to our polarization tomography system. We typically observed excellent polarization-entangled states on the RHS: for example, Figure 4.10 shows the polarization density matrix inferred by QST on SPDC produced by the pump from the short arm of the MZI (the long arm is blocked; also, since we are using a Gaussian pump, the long arm is the dark port of the Sagnac interferometer). Here we had a coincidence rate of approximately 550 Hz for the bright tomography settings $HH$ and $VV$, and we integrated the counts for each setting over 1 second.

Once the RHS fibers were aligned to optimally collect the first-pass SPDC, we launched the beam
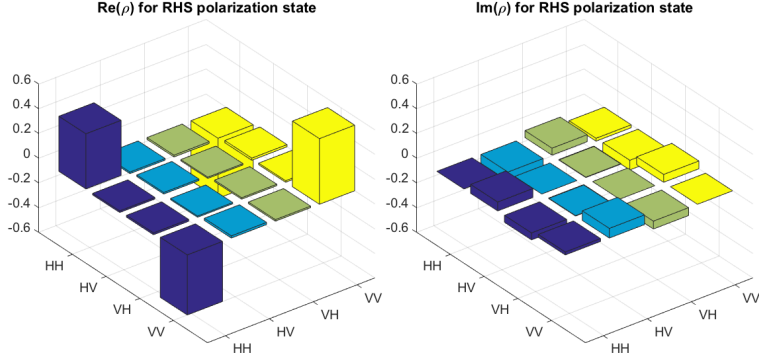
Figure 4.10: Density matrix for polarization-entangled photons collected on the RHS of the SPDC crystals. We find that $\text{Tr}(\rho^2) = 99.1(6)\%$, $T = 98(1)\%$, and $F(\rho, |\Phi^-\rangle\langle\Phi^-|) = 97.3(3)\%$, where the uncertainties are estimated using 100 Monte Carlo samples.

from a temperature-controlled diode laser at 702 nm (Hitachi HL-6738MG) from the RHS fibers toward the left-hand side (LHS) fibers. This alignment laser (AL) was used repeatedly during the experiment for various alignments. Here, we aligned the LHS fibers to maximize the coupling with the RHS fibers. We could typically achieve 65% coupling, after accounting for losses between the fibers, including the diffraction efficiency of the hologram. With the LHS now coupled to the same single mode as the RHS, we should immediately be able to collect the first-pass (second-pass) SPDC on the LHS by placing mirrors at normal incidence in the signal and idler paths (pump path). Thus, we maximized coupling of the first- and second-pass SPDC to the LHS fibers by adjusting only these three back-reflecting mirrors.

We should mention that aligning the RHS fibers to optimize SPDC collection always involves a balancing act: due to transverse walkoff, there is no positioning of the signal and idler fibers that optimally collects both the $|HH\rangle$ and $|VV\rangle$ photon pairs. Noting that the walkoff also affects the AL, we aligned the LHS using both $H$- and $V$-polarized AL beams, attempting to balance the coupling for the two cases.

With the LHS fibers aligned to collect all the SPDC processes, we came to the task of calibrating our spatial tomography system. To align the signal (idler) hologram, we kept the idler (signal) hologram fixed at a featureless region of the grating where it does not change the spatial mode at all. The signal (idler) hologram was then scanned incrementally near each of the 6 first-order mode analyzer settings, while coincidences were recorded. When the signal (idler) analyzer is optimally aligned, the tomography system should reject states orthogonal to $|x\rangle_s \otimes |g\rangle_i$ ($|g\rangle_s \otimes |x\rangle_i$), where $x$ is one of $\{h, v, d, a, l, r\}$, depending on which of the 6 hologram settings we look at. Since the most significant contribution to the SPDC spatial-mode state produced by a Gaussian pump is $|g, g\rangle$, orthogonal to both $|x, g\rangle$ and $|g, x\rangle$, we should see a minimum in the coincidence rate when we optimally align the hologram.

Due to walkoff, the optimal hologram settings were not the same for different SPDC processes. While the

| | | $\lvert h\rangle_{\mathrm{pump}}$ | | $\lvert v\rangle_{\mathrm{pump}}$ | | $\lvert d\rangle_{\mathrm{pump}}$ | | $\lvert a\rangle_{\mathrm{pump}}$ | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | signal | idler | signal | idler | signal | idler | signal | idler |
| | $h$ | 101 | 115 | 1 | 2 | 67 | 50 | 37 | 37 |
| | $v$ | 0 | 0 | 117 | 87 | 50 | 58 | 74 | 71 |
| | $d$ | 42 | 43 | 77 | 49 | 87 | 102 | 6 | 0 |
| | $a$ | 48 | 53 | 43 | 38 | 2 | 1 | 92 | 78 |
| | $l$ | 58 | 53 | 61 | 62 | 71 | 67 | 47 | 42 |
| | $r$ | 44 | 38 | 52 | 45 | 40 | 54 | 76 | 52 |
| $\mathrm{Tr}\,\rho^2$ | | 1.0 | 1.0 | 1.0 | 0.977 | 1.0 | 0.989 | 0.968 | 1.0 |
| $F(\rho,\rho_{\mathrm{pump}})$ | | 0.997 | 0.996 | 0.985 | 0.978 | 0.976 | 0.990 | 0.939 | 0.985 |

Table 4.1: Single-qubit tomography data for first-order spatial modes. The pump spatial mode state is imprinted on the SPDC pairs. For each signal (idler) tomography, the idler (signal) analyzer is fixed at $\lvert g\rangle$. The counts for each setting were obtained over 60 seconds. Because of the high purities of these states, Monte Carlo analysis gives unreliably small uncertainties.

*spacing* between the optimal settings for the different polarizations was almost invariably less than $100\,\mu$m and usually less than $50\,\mu$m, the *extinction* of the $\lvert g, g\rangle$ spatial mode state for the two polarizations was highly dependent on alignment of the SMFs. In particular, it appears that when the SMFs are optimally aligned to collect SPDC, the optimal hologram setting gives good extinction of the Gaussian mode. Therefore, if the AL couplings for $H$ and $V$ beams are not balanced well, it is impossible to get good extinction for all the SPDC modes. The problem with this original experimental configuration was that the large walkoff-induced separation between the polarizations made it difficult to balance the couplings well. Nevertheless, on one occasion we *were* successful, but once the system drifted and we lost the alignment (after approximately ten days), we were never able to recover it. Having aligned the holograms for only the second-pass $\lvert VV\rangle$ process and finding coincidence minima no larger than 1% of the full $\lvert g, g\rangle$ rate for all hologram settings, we simply assumed that these settings would work for the other SPDC processes as well. Our results validate this assumption.

We demonstrated the operation of the mode-converting coverslips one at a time, using single-qubit spatial mode tomographies. Placing one coverslip edge at the center of the beam results in a first-order pump mode. Specifically, the spatial mode in the short arm of the MZI (the long arm is blocked for these measurements) is $\cos(2\theta)\lvert h\rangle + \sin(2\theta)\lvert v\rangle$ where $\theta$ is the angle setting of the HWP before the MZI. Based on Eqs. 2.16, we expect this pump mode to produce the SPDC spatial mode state $(\cos(2\theta)\lvert h\rangle + \sin(2\theta)\lvert v\rangle)_s \otimes \lvert g\rangle_i + \lvert g\rangle_s \otimes (\cos(2\theta)\lvert h\rangle + \sin(2\theta)\lvert v\rangle)_i$. That is, a single-qubit tomography of the signal (idler) spatial mode with the idler (signal) spatial mode analyzer fixed at $\lvert g\rangle$ should give the same state as the pump. Table 4.1 shows signal and idler single-qubit tomography data for $\lvert h\rangle, \lvert v\rangle, \lvert d\rangle$ and $\lvert a\rangle$ pumps produced by "Coverslip 1".

Visibilities measured in appropriate bases for "Coverslip 2" indicate similar mode conversion efficiency.

We then move both coverslips to their respective locations for conversion to $|HG'_{20}\rangle$, fix the HWP before the MZI at its final setting, $22.5°$, and still block one arm of the MZI. Figure 4.11 shows the results of our two-qubit tomographies of the first-order spatial mode. In each case we integrated for 60 s per measurement setting, and the coincidence rate at the brightest measurement setting was approximately 0.6 Hz. The three states $|\phi^+\rangle$, $|\phi^-\rangle$, $|\psi^-\rangle$ were observed by looking at the second pass of the short arm of the MZI, the second pass of the long arm, and the first pass of the long arm, respectively. Recall from Section 2.2 that the state $|\phi^+\rangle$ can also be produced by a Gaussian pump. On the other hand, our results for $|\phi^-\rangle$ and $|\psi^-\rangle$ are noteworthy because these spatial mode states can only be produced by a second-order pump spatial mode. For one of these spatial tomographies, shown in Figure 4.11d, we calibrate the holograms specifically for the constant polarization setting $VV$ used for the measurements. This result is significantly better than what we get for the same state with the holograms calibrated to work reasonably well for all the SPDC modes (Figure 4.11c). Naturally, such an optimization is not possible when we perform QST on 4-qubit states where all the SPDC processes are involved, so we are limited by the walkoff-induced separation of the SPDC modes.

We can examine the hyperentangled state produced by just one pass of just one arm of the MZI. Figure 4.12 shows our data for the state $|\Psi^-\rangle \otimes |\psi^-\rangle$, produced by the first pass of the long arm of the MZI. While the fidelity with the desired state is relatively low because of phase errors in the polarization state, these errors can easily be corrected and are not relevant in our work. The important metrics here are the purity and tangle of the polarization and spatial mode reduced density matrices. Note that of the four hyperentangled states we produce in our setup, this one poses the greatest difficulty: it requires the initial SPDC spatial-mode state $|\phi^-\rangle$, which can only be produced by a non-Gaussian pump mode, and it also requires precise alignment of the cylindrical lens mode converter.

Finally, we performed a four-qubit tomography of SPDC produced by the combination of both MZI arms and both pump passes, inferring a Smolin-like density matrix $\rho^{\text{exp}}$ (Figure 4.13a). The fidelity of our state with the Smolin state is 44(5)%, but this is misleading. The Smolin state's matrix elements are all real, but its properties do not depend on this fact. In fact, any state

$$\rho^S_{\text{phased}} = \sum_{j,k=1}^{16} c_{jk} |j\rangle \langle k| \tag{4.10}$$

satisfying

$$\sum_{j,k=1}^{16} |c_{jk}| |j\rangle \langle k| = \rho^S \tag{4.11}$$
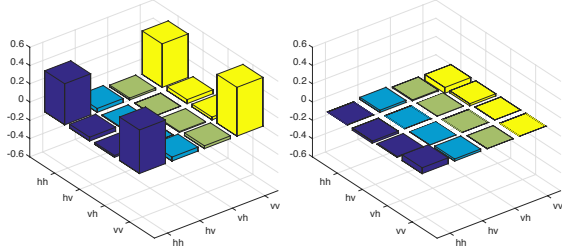
(a) $\mathrm{Tr}\,\rho^2 = 96(2)\%$, $T = 90(4)\%$, $F\left(\rho, \left|\phi^+\right\rangle\!\left\langle\phi^+\right|\right) = 97(1)\%$.
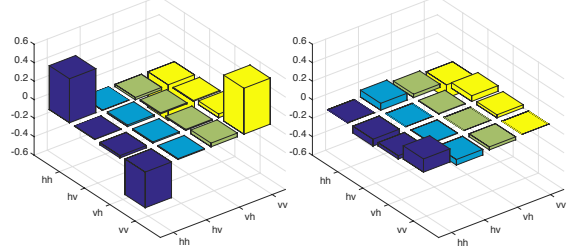
(b) $\mathrm{Tr}\,\rho^2 = 84(5)\%$, $T = 67(9)\%$, $F\left(\rho, \left|\phi^-\right\rangle\!\left\langle\phi^-\right|\right) = 87(3)\%$.
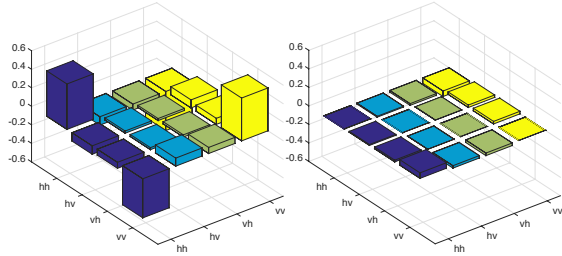
(c) $\mathrm{Tr}\,\rho^2 = 87(4)\%$, $T = 74(8)\%$, $F\left(\rho, \left|\phi^-\right\rangle\!\left\langle\phi^-\right|\right) = 90(2)\%$.
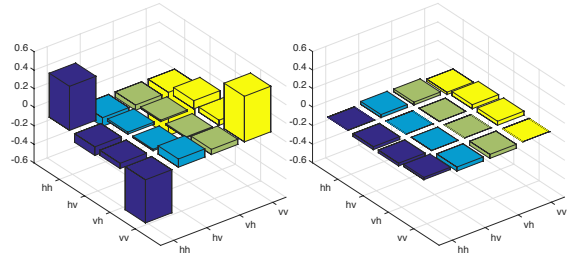
(d) $\mathrm{Tr}\,\rho^2 = 96(2)\%$, $T = 89(5)\%$, $F\left(\rho, \left|\phi^-\right\rangle\!\left\langle\phi^-\right|\right) = 95(1)\%$.
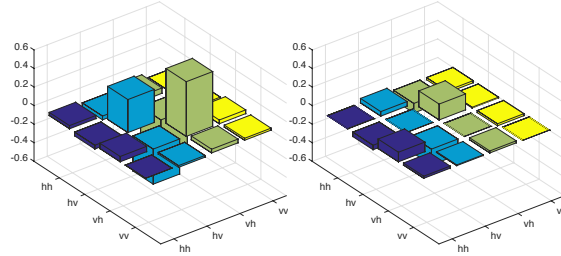
(e) $\mathrm{Tr}\,\rho^2 = 86(6)\%$, $T = 72(9)\%$, $F\left(\rho, \left|\psi^-\right\rangle\!\left\langle\psi^-\right|\right) = 85(4)\%$.

Figure 4.11: Real and imaginary parts of the inferred density matrix for SPDC produced by (a) the second pass of the short arm of the MZI, analyzed at polarization setting $VV$. This spatial mode can also be produced by a Gaussian pump, so high fidelity doesn't necessarily suggest that our system is working as desired. (b) the second pass of the long arm of the MZI, analyzed at polarization setting $HH$. In contrast to the previous state, this spatial mode can only be produced by a second-order pump mode. (c) the second pass of the long arm of the MZI, this time analyzed at polarization setting $VV$. (d) the second pass of the long arm of the MZI, again analyzed at polarization $VV$, but this time with the hologram settings optimized for the second-pass $VV$ process. We observe substantially higher state quality. (e) the first pass of the long arm of the MZI, analyzed at polarization setting $HV$. Here, in addition to relying on production of a spatial mode that can only arise from a second-order pump mode, we demonstrate the operation of the polarization and spatial-mode $\pi$-mode converters. Uncertainties were estimated using 100 Monte Carlo samples.
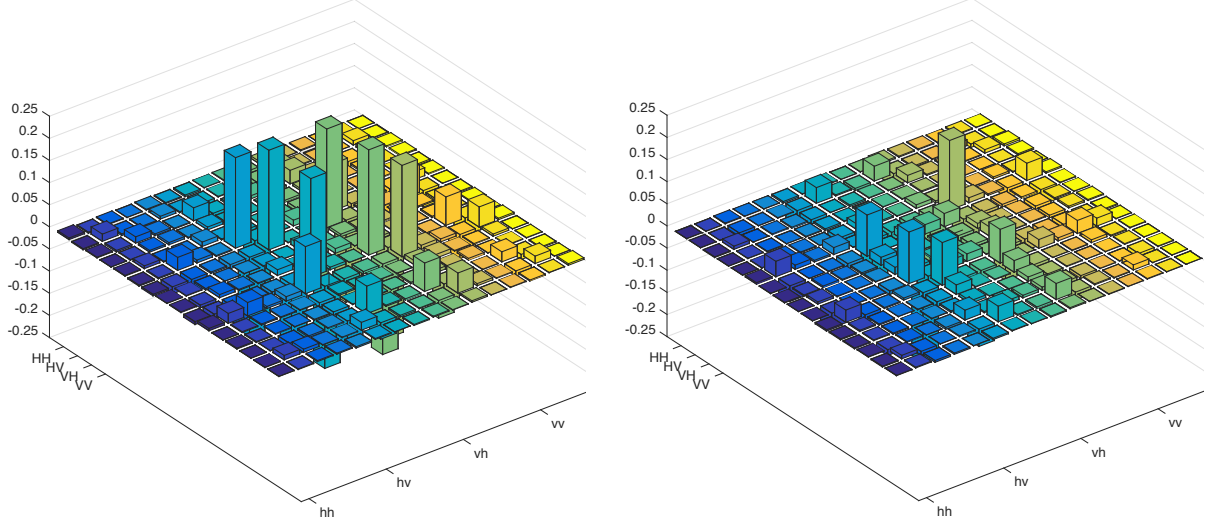
Figure 4.12: Real and imaginary parts of the inferred density matrix for hyperentangled photon pairs produced by the first pass of the long arm of the MZI. Our state deviates from the desired $|\Psi^-\rangle \otimes |\psi^-\rangle$ largely because of phase errors in the prepared polarization state. We find that $\text{Tr}(\rho^2) = 79(8)\%$ and $F(\rho, |\Psi^-\rangle\langle\Psi^-| \otimes |\psi^-\rangle\langle\psi^-|) = 75(8)\%$, where the uncertainties are estimated using 100 Monte Carlo samples. We can perform a partial trace over the polarization (spatial mode) on our inferred 4-qubit density matrix to determine the reduced density matrix for the spatial mode (polarization). This leads us to infer tangles of $T(\text{Tr}_{\text{polarization}}(\rho)) = 64\%$ and $T(\text{Tr}_{\text{spatial mode}}(\rho)) = 70\%$ for spatial mode and polarization, respectively.

has the same partial transpose eigenvalues as the Smolin state. This means that phases on the off-diagonal elements of our inferred density matrix, while reducing its fidelity with the Smolin state, have no effect on its proximity to the BE region. We therefore calculate the fidelity of our state to the nearest $\rho_{\text{phased}}^{S}$ (Figure 4.13b). As expected, since we did not add additional noise, our state is not BE, having negative eigenvalues in every $2:2$ partition. If we simply (numerically) add an identity matrix to our inferred density matrix, $p\rho^{\text{exp}} + (1-p)\mathbb{I}/16$, we *do* find a BE region for $1/(1 - 16\lambda_{1:3}^{\min}) = 0.32 < p < 0.40 = 1/(1 - 16\lambda_{2:2}^{\min})$. A more realistic test of our state's prospects, however, which also allows us to test the statistical significance with which the state satisfies the $2:2$ PPT criterion, is to add noisy counts directly to the data, i.e., to add numbers drawn from the same Poisson distribution to the counts for each measurement setting. This type of test reveals that the $1:3$ entanglement, in fact, gets destroyed before we can conclusively say that the state is separable across every $2:2$ bipartition so that we cannot reliably achieve BE. At any rate, the narrow range of $p$ that leads to BE, even for addition of the density matrices, suggests that we need a different approach.
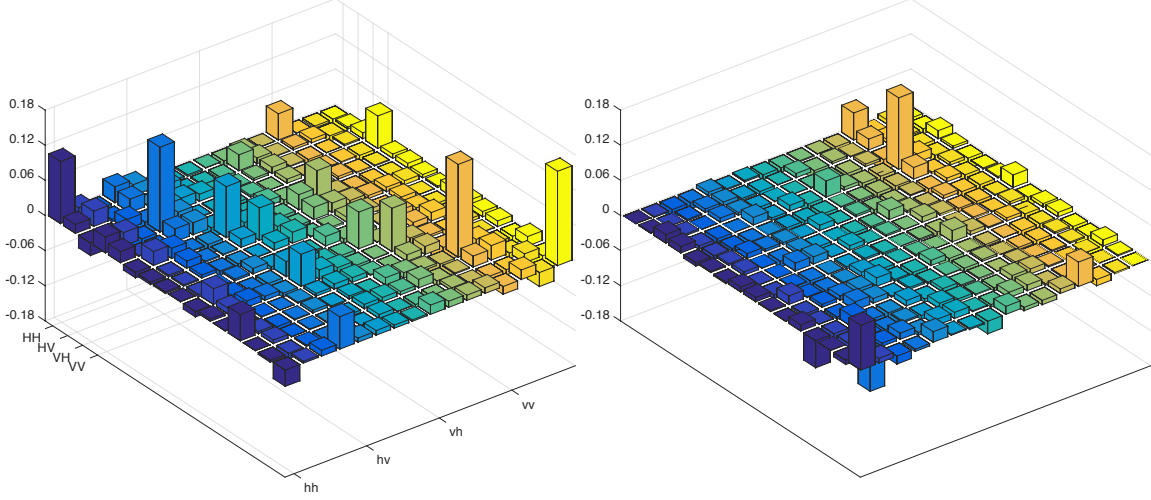
(a) Real and imaginary parts of the density matrix for $\rho^{\mathrm{exp}}$ inferred by QST
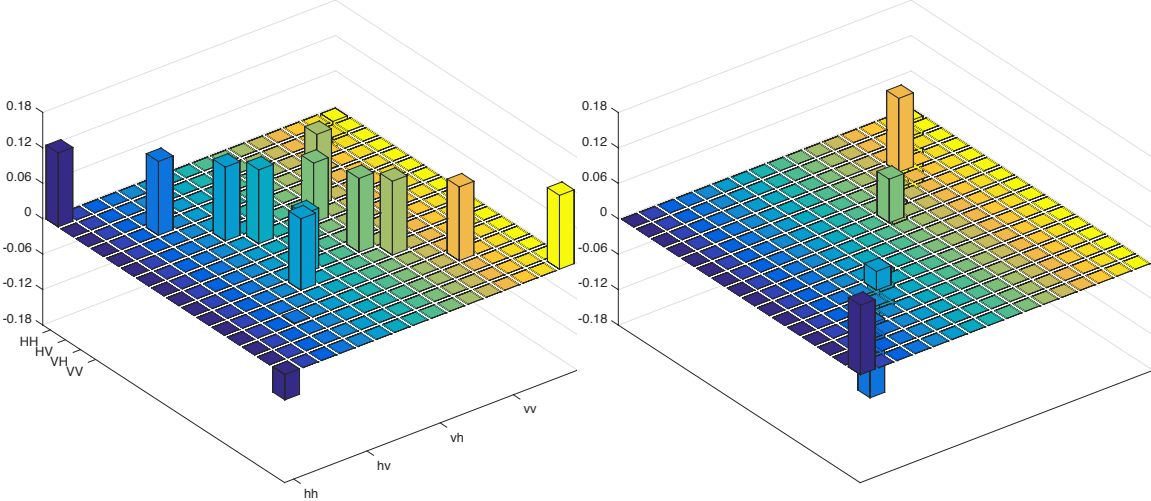


(b) Real and imaginary parts of the density matrix for the Smolin-like $\rho^{S}_{\mathrm{phased}}$ closest to $\rho^{\mathrm{exp}}$

| | 1 : 3 bipartitions | | | | 2 : 2 bipartitions | | |
| | A:BCD | B:CDA | C:DAB | D:ABC | AB:CD | AC:BD | AD:BC |
| --- | --- | --- | --- | --- | --- | --- | --- |
| $\lambda^{\min}$ | -0.15(2) | -0.15(2) | -0.14(2) | -0.13(2) | -0.05(1) | -0.09(1) | -0.08(1) |

Figure 4.13: (a)The Smolin state density matrix $\rho^{\mathrm{exp}}$ that we inferred from QST contains undesired phases, but these do not pose a problem, since they don't change the properties of the state. The Smolin-like state $\rho^{S}_{\mathrm{phased}}$, with the same phases as our state, is also BE. $F(\rho^{\mathrm{exp}}, \rho^{S}_{\mathrm{phased}}) = 61(6)\%$. The table shows the minimum partial transpose eigenvalues across all bipartitions of the state. Here $A(B)$ is the idler (signal) spatial-mode qubit, and $C(D)$ is the idler (signal) polarization qubit. The fact that the 2 : 2 bipartitions' minimum eigenvalues are all less than those of the 1 : 3 bipartitions implies that BE should be attainable by a suitable mixture with white noise. However, our low count rates make the uncertainties too large to convincingly show positive 2 : 2 minimum eigenvalues while maintaining negative 1 : 3 eigenvalues. Uncertainties were estimated using 100 Monte Carlo samples.

### 4.4.2 The modified setup

Despite repeated attempts, we were unable to reproduce the alignment that led to the results of the previous section. Specifically, we were unable to align the fibers in such a way that we could find hologram settings that give good visibilities for all the SPDC processes —rather, the optimal settings for $HH$ and $VV$ were always different, and there was no in-between position where they would both give good spatial mode visibility. Furthermore, we wanted to stabilize the pump so that we wouldn't lose alignment so quickly. Therefore, we made a number of changes to the setup. In a sense, these modifications were very successful; it became possible for us to align the system in a fairly repeatable manner. On the other hand, with the modified system we were never able to produce spatial mode states that matched the quality of those reported in the last section. In this section, I will simply explain the improvements we made and their effect on the experiment.

1. *We added two additional beam-pointing control systems after the Sagnac interferometer (see Section 4.3.2).*

   It seems that this could only have helped.

2. *We replaced the FI with a polarization basis-dependent QWP isolator (see Section 4.3.3).*

   This allowed us to increase the pump power incident on the crystals. The resulting higher count rates make accurate alignment of the SPDC and pump mirrors far easier. On the other hand, as discussed in Section 4.3.3, the QWP isolator does allow approximately 5% of the pump beam to feed back into the laser, and we did see the effect of this backreflected power. In particular, the Gaussian pump would *always* destabilize the laser when both arms of the MZI were unblocked, sending all four beam-pointing stabilization systems into large oscillations. For reasons we don't fully understand, the $|HG'_{20}\rangle$ pump had a much more benign effect; often there would be no detectable change in the control systems. When they did oscillate in response to both MZI paths being unblocked, they would often stop after a few minutes. One explanation for this is that the high-order modes contained in $|HG'_{20}\rangle$ are harder to excite in the laser. Here, we see the benefit of one of the imperfections in our experiment —if the coverslips performed an ideal mode conversion $|HG_{00}\rangle \rightarrow |HG_{20}\rangle$, then they would perform the reverse operation on the way back, and we would have a Gaussian feeding back into the cavity. Instead, the coverslips produce a state that is not a *mode*, so it does not look the same when it returns. The mode conversion in the reverse direction, thus, does *not* bring us back to a Gaussian beam (though there must still be a considerable overlap with one).

3. *We replaced the 8.0-mm focal length fiber-coupling aspheric lenses with 4.51-mm focal length aspheres.*

   The idea was to increase the area on the crystal that gets imaged onto the SMF. If we collect from a larger region on the crystal, then walkoff becomes less damaging to the state, since the collection mode

can overlap better with *both* separated modes. This modification was very successful, as we found that the hologram settings for $HH$ and $VV$ usually overlapped to within 25 $\mu$m and showed very good visibility.

4. *We implemented walkoff compensation.*

In our scheme, described in Appendix A, the LHS fibers are aligned such that they are equidistant from the modes produced by all the SPDC processes (after those modes have been appropriately compensated). Furthermore, the compensation scheme eliminates the need to balance $H$ and $V$ AL coupling between the LHS and RHS (however, we do still need to balance $HH$ and $VV$ count rates on the RHS).

5. *We added grating stabilization to the AL.* Fluctuations in the AL power, even on the order of only 5%, make it difficult to maximize the coupling between the RHS and LHS fibers, and even harder to balance the coupling for $H$ and $V$ polarizations. With the addition of Littrow-configuration grating stabilization, we suppressed mode-hopping, reducing power fluctuations below to 1%.

## 4.5   Discussion

In recent years there has been considerable experimental work on bound entanglement. In particular, Smolin-state BE has been successfully realized in Zeeman levels of four trapped $^{40}$Ca$^+$ ions [29]. There have also been reports of Smolin-state BE in photon polarization. We will see the general principle behind these experiments in the first part of this section. Photonic realizations of BE, including our own attempt, pose special challenges, and as shown in the second part of this section, none of the previous experiments represents a true example of BE. In the third part, we scrutinize our own experiment, pointing out potential flaws that could invalidate it, and how they can be addressed. Finally, we make some concluding remarks on the significance of these experimental demonstrations.

### 4.5.1   Previous photonic Smolin state experiments

To date, all experimental reports on photonic Smolin states have used 4 photon polarizations as the 4 Smolin state qubits [54, 66–68]. The general idea is the same in all of these experiments, so we will only discuss the essential experimental details here.

A laser pulse is used to pump two pairs of SPDC crystals, simultaneously producing two polarization-entangled pairs.[8] The initial state of the four photons is $|\Phi^+\rangle_{AB} \otimes |\Phi^+\rangle_{CD}$, where we have labeled the first (second) signal and idler polarizations $A$ and $B$ ($C$ and $D$), respectively. One of four unitary operations

---

[8]This is the aim, at any rate. In practice, a laser pulse generates multiple pairs, which is problematic since it is impossible to determine which signal belongs to which idler when the pairs arrive simultaneously. We can reduce the power so that multiple-pair events are sufficiently rare, but this also reduces the probability of getting one pair: the probability of simultaneously getting one pair from *both* SPDC sources —as desired —is the square of that probability.

is then chosen at random and applied to both pairs using automated rotating wave-plates or liquid-crystal retarders (LCRs), so that at any given time the state output by the system is

$$|\psi\rangle = U_j \otimes U_j \left|\Phi^+\right\rangle_{AB} \otimes \left|\Phi^+\right\rangle_{CD}, \tag{4.12}$$

where

$$U_j = \begin{cases} \mathbb{I} \otimes \mathbb{I} & j = 0, \\ \mathbb{I} \otimes \sigma_x & j = 1, \\ \sigma_z \otimes \mathbb{I} & j = 2, \\ \sigma_z \otimes \sigma_x & j = 3, \end{cases}$$

so that we end up with

$$|\psi\rangle = \begin{cases} \left|\Phi^+\right\rangle_{AB} \otimes \left|\Psi^+\right\rangle_{CD} & j = 0, \\ \left|\Psi^+\right\rangle_{AB} \otimes \left|\Psi^+\right\rangle_{CD} & j = 1, \\ \left|\Phi^-\right\rangle_{AB} \otimes \left|\Phi^-\right\rangle_{CD} & j = 2, \\ \left|\Psi^-\right\rangle_{AB} \otimes \left|\Psi^-\right\rangle_{CD} & j = 3. \end{cases} \tag{4.13}$$

The claim is that since $j$ is chosen at random throughout the experiment, the four cases in Eq. 4.13 add incoherently, giving the Smolin state. The first of these works [66] overlooked the experimental necessity for additional noise (see Section 3.5), so the density matrix they determined by QST failed to even satisfy the criteria for BE (PPT for all $2:2$ partitions, NPT for all $1:3$ partitions). This deficiency was remedied by Lavoie et al. [67] by making the two randomly generated $U_j$ imperfectly correlated, resulting instead in the state

$$|\psi\rangle = \begin{cases} U_j \otimes U_j \left|\Phi^+\right\rangle_{AB} \otimes \left|\Phi^+\right\rangle_{CD}, \text{ with probability } p \\ U_j \otimes U_{k \neq j} \left|\Phi^+\right\rangle_{AB} \otimes \left|\Phi^+\right\rangle_{CD}, \text{ with probability } 1-p, \end{cases} \tag{4.14}$$

where the "error" probability $1-p$ parameterizes the amount of white noise added to the state. This time, the claim was that the random selection of $(j,k)$ throughout the experiment gives the state $\rho^S(p)$ (Eq. 3.24), and, indeed, the authors showed (via the PPT criterion for the $2:2$ partitions, and an entanglement witness for the $1:3$ partitions) that depending on the value of $p$, the tomographically reconstructed density matrix was separable, bound-entangled, or free-entangled. Additionally, Refs. [66, 67] both report two-party distillation and violation of the Bell inequality Eq. 3.26.

In Ref. [54], Kaneda et al. report activation of a Smolin state (see Section 3.6). Their protocol differs from the original one [51], which activates a $3 \otimes 3$ BE resource to distill a single copy of a $3 \otimes 3$ free-entangled state

that cannot otherwise be distilled non-collectively. Here, three SPDC crystals are pumped to simultaneously generate three pairs. The first two pairs are prepared as in Eq. 4.13, to be shared by parties $A, B, C, D$. The third pair, shared by $B$ and $C$, remains in its original state $|\Phi^+\rangle$. The idea is that neither the Smolin state $\rho^S_{ABCD}$ nor the maximally entangled state $|\Phi^+\rangle$ permits distillation of entanglement between $A$ and $D$, and yet the combination does. Specifically, if $B$ makes a measurement of his qubits in the Bell-state basis and $C$ does the same, the remaining qubits, held by $A$ and $D$, are projected into a maximally entangled state. Kaneda et al. do not claim to have produced BE, since they produce Eq. 4.13 rather than Eq. 4.14, but they show that the amount of entanglement distilled between $A$ and $D$ in their experiment is greater than the free entanglement between $A$ and $D$ indicated by their Smolin state density matrix.

## 4.5.2  Loopholes in past photonic experiments

Now that we've seen how these random-number-generated (RNG) mixture experiments work, as well as the main results supporting their reports of BE, we will analyze them more closely. Before coming to that, however, let's look at yet another photonic BE experiment that highlights the problem more clearly.

In Ref. [69] a $3 \otimes 3$ BE state of the spatial-mode of photon pairs from SPDC is purportedly produced[9] (to our knowledge, this is the only experimental work on $3 \otimes 3$ BE). Mixture is introduced by a multi-setting optic, much as in the other experiments (this time the device is a spatial light modulator (SLM) that can programmably change the spatial mode of an incident beam), but here Hiesmayr et al. dispense with the RNG altogether, instead performing a sequence of QSTs, each with different settings of the SLMs, and summing the results. Their idea is that given this data, two parties Alice and Bob would be unable to distill entanglement. In this case, the fallacy is apparent: we should not ask what Alice and Bob can do with our *data*, but rather with the *photons* themselves: we must assume Alice and Bob have the benefit of any knowledge and technology admitted by physics, much as we do for the eavesdropper Eve when analyzing the security of quantum communication protocols. Thus, we propose the following criteria for evaluating the BE property of a state:

1. Given arbitrarily many copies of this state, can the receiving parties distinguish it from BE?

2. Can they distill it?

The answer to both of these questions must be *no* for the state to be legitimately BE. For the state in Ref. [69], however, the answer to both questions is clearly "yes." Given a finite number of copies of the state produced by each SLM setting, Alice and Bob can determine the *pure* state being produced by that setting

---

[9]To generate the qutrits, spatial mode measurements on the signal and idler included the Gaussian in addition to the two-dimensional first-order subspace used in our work. For further explanation, see [61].

and distill it. That is, the mixture is invalid because the pure state of a specific instance of the state received by Alice and Bob is *not* random from their perspective; rather, it is correlated with all the copies they have received since the last setting change.

We now return to the RNG Smolin-state experiments, where the problem is essentially the same. In Ref. [67], for example, the random setting of the LCRs happens at a rate of 10 Hz, while the coincidence detection rate for each SPDC source is 36 kHZ. Therefore, the Smolin state parties $A, B, C, D$ have the benefit of 3600 copies of the *same pure state* for each setting of the LCRs. They can sacrifice some fraction of these copies to determine which state they have, at which point the remaining copies are already in distilled form. When their maximally entangled resource ceases to function, they know that they must repeat the process of identifying which pure state they share. Refs. [54, 66–68] all suffer from this flaw. We believe the conclusions reached in these studies are invalid because the density matrices are obtained by averaging over a "mixture" that *can* be purified by $A, B, C, D$ given nothing but copies of the state.

It is natural to ask how these experiments were able to demonstrate the applications of BE despite failing to realize BE. Let us start with the two-party distillation protocol. In the RNG experiments, $A$ and $B$ share the same randomly chosen Bell state as $C$ and $D$. $A$ and $C$ perform a Bell-state measurement, which matches the result of a Bell-state measurement by $B$ and $D$. One needs neither BE nor a Smolin state to observe this effect. Each of the four pure-state terms of the Smolin state Eq. 3.22 is individually capable of this type of distillation, called entanglement swapping [58]. The same is true for violation of the Bell inequality Eq. 3.26: each of the four terms of the Smolin state maximally violates the Bell inequality. Finally, we consider the BE activation reported in Ref. [54]. Kaneda et al. deduce that BE was activated because the free entanglement in their Smolin state, though non-zero, is small. However, they come to the conclusion that their state is nearly BE because of the same faulty tomography analysis employed by Refs. [66–68]. In view of these arguments, we regard all of these experiments merely as *simulations* of BE; genuine photonic BE has yet to be demonstrated in the lab.

### 4.5.3 Possible loopholes in our experiment

Similarly, there are certain aspects of our setup that must be dealt with carefully for it to produce bona fide BE. In contrast to the RNG experiments, the mixture in our experiment is quantum mechanical in origin. This has two important implications, corresponding to the two criteria proposed in Section 4.5.2:

1. There is no correlation between which of the four processes generated the first $n - 1$ pairs and which one generates the $n^{\text{th}}$ pair. That is, given the first $n - 1$ instances of the state, parties $A, B, C, D$ can *never* predict which pure state they will get next. This is due to the random nature of SPDC, and it

is the reason we can truly call the state we produce a mixed state.

2. Information about which process occurred to generate the $n^{\text{th}}$ pair certainly *does* exist outside of the SPDC photons —that's why they are in a mixed state. However, we expect that this information is entirely contained in the pump beam[10]. Thus, as long as $A, B, C, D$ are not allowed access to the pump, there is no way for them to determine which process occurred or to distill entanglement[11].

These points mean that we *can* successfully produce BE, but we must still be careful not to inadvertently attach any extra labels to the different processes. For example, Section 4.2.2 mentions that we must place the signal and idler first-pass mirrors at equal distances (within the SPDC coherence length) from the crystals and place the pump mirror at a different distance (beyond the pump coherence length) from the crystals. While this arrangement has the desired effect of destroying the phase coherence between the first- and second-pass SPDC processes, this decoherence could also be accomplished far more easily by placing the signal and idler mirrors different distances from the crystals —indeed, given the short coherence length of the SPDC, it is hard *not* to. The problem with this simpler setup is that it introduces a label that the parties $A, B, C, D$ can use to identify which pump pass produced a given copy of the state, as long as their detectors have high enough time resolution: for example, "If the photons arrive at our locations simultaneously, they were produced by the second pump pass, but if the idler is slightly delayed, they were produced by the first pass of the pump because the first-pass idler has to travel slightly farther than the first-pass signal."

Furthermore, the noise photons we add to our bare Smolin state must also be indistinguishable from the Smolin state photons. Since we are using polarization and spatial-mode for the state encoding, we only need to worry about the distinguishability in time and frequency: the noise pairs should arrive at the signal and idler detectors with the same time delay as the SPDC pairs, and they must have the same joint spectrum as the SPDC pairs. If either of these requirements is not met, the parties can in principle filter out the noise photons by relative arrival time or frequency and retrieve the bare Smolin state, which is *not* BE. These requirements are realistic, and it should, therefore, be possible to demonstrate BE by our scheme.

### 4.5.4   Conclusions

As we've seen in this section, there has yet to be a true demonstration of photonic BE. However, we should make the following note regarding the RNG experiments: after overcoming the initial displeasure of having been scooped, and the subsequent annoyance at the inelegance of the approach, we thought long

---

[10]We can run the SPDC process in reverse, but only within the coherence length of the pump. Therefore, exactly one of the four incoherent pump modes will allow this reverse process for a given SPDC pair.

[11]Moreover, it's not clear that the parties could determine which process occurred even if they *did* have access to the pump. They wouldn't be allowed to run the SPDC in reverse, since that involves the four parties coming together, breaking the LOCC requirement, so they would have to come up with something else.

and hard about these experiments, and in the end it seems that their idea may lead to a state that is completely indistinguishable from BE under certain conditions. While the low four-fold coincidence rates in Refs. [54, 66–68] might make such a demonstration impossible in practice, with sufficiently high heralding efficiency —or alternatively, sufficiently high switching of the random unitaries —it could be arranged that less than one pair per RNG setting is produced by each SPDC source. Then it would, indeed, be the case that each instance of the RNG-Smolin state is uncorrelated with the previous instances (provided a true RNG is used, of course!).

In the abstract of Ref. [51], a year after their discovery of BE, the Horodeckis write: "Bound entanglement is the noisy entanglement which cannot be distilled to a singlet form. Thus it cannot be used alone for quantum communication purposes. Here we show that, nevertheless, the bound entanglement can be, in a sense, pumped into a single pair of free entangled particles. It allows for teleportation via the pair with the fidelity impossible to achieve without support of bound entanglement." It seems that this statement has been widely misinterpreted to mean that *only* BE can enable this sort of activation. Misconceptions regarding other applications of BE appear to be similarly widespread. For example, Lavoie et al. describe entanglement unlocking as one of the "conceptually important characteristics of Smolin states," [67]. While technically true, this statement is misleading because it suggests that a demonstration of unlocking somehow validates the Smolin state resource used. Rather, entanglement unlocking is only of interest when it has been rigorously shown that the state is not distillable. The point is that despite being a "weak" form of entanglement, BE turns out to perform just as well as a pure entangled state for this application —this fact *is* remarkable and it should motivate us to pursue a better understanding of mixed-state entanglement.

# Appendix A

# Transverse walkoff

Transverse walkoff posed an important challenge to our experiment. The first section of this appendix explains the physical origin of the effect; the second section describes our method for compensating walkoff in our setup.

## A.1   What is transverse walkoff?

In Section 2.1.1, we mention that two orthogonal polarizations incident on a birefringent crystal can propagate at different angles and emerge spatially separated. If we naïvely generalize Snell's law to birefringent media with different indices $n_1$, $n_2$ for the two incident eigen-polarizations,

$$n_{\mathrm{air}} \sin \theta_{\mathrm{air}} = n_{1,2} \sin \theta_{1,2}, \tag{A.1}$$

this behavior is hardly surprising: the index on the RHS is polarization-dependent, so the angle must be as well. However, there is a different effect that we are interested in here. Indeed, the angles $\theta_{1,2}$ indicated by Eq. A.1 correctly describe $\mathbf{k}$ inside the crystal, but the optical energy, and hence the photon itself, propagates in the direction of the Poynting vector $\mathbf{S}$. Specifically,

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \text{ while}$$

$$\mathbf{k} = \mathbf{D} \times \mathbf{H}.$$

In isotropic media this distinction is unimportant because $\mathbf{E} \parallel \mathbf{D}$. In anisotropic media, on the other hand, the linear susceptibility (and, therefore, the refractive index) is a second-order tensor with as many as three distinct eigenvalues $\chi_1$, $\chi_2$, $\chi_3$. A field $\mathbf{D} = D_1\hat{\mathbf{e_1}} + D_2\hat{\mathbf{e_2}} + D_3\hat{\mathbf{e_3}}$ with components along all three principal axes is associated with $\mathbf{E} = (1+\chi_1)D_1\hat{\mathbf{e_1}} + (1+\chi_2)D_2\hat{\mathbf{e_2}} + (1+\chi_3)D_3\hat{\mathbf{e_3}}$, and it is evident that the two are not generally parallel [15]. To make it clear that this effect is distinct from the angular deviation given
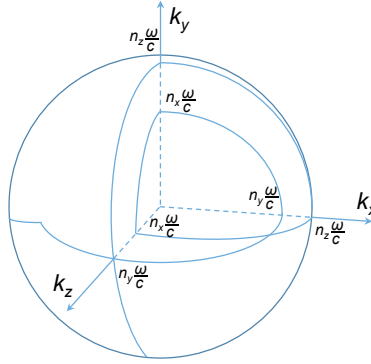
Figure A.1: The normal surface is a two-layered surface in **k**-space that specifies the relationship between **k** and **S**.

by Snell's law, note that Eq. A.1 predicts no difference between the two polarizations for normal incidence ($\theta_{\mathrm{air}} = \theta_1 = \theta_2 = 0$), but this condition does not preclude **E** $\parallel$ **D**, when transverse walkoff can still occur (this is precisely what happens in a beam displacer: see Figure 2.1).

In a uniaxial crystal, it is easy to identify the eigen-polarizations for any **k**, since regardless of **k**, one polarization state experiences the principal index $n_o$. The polarization orthogonal to this one is (confusingly) called the extraordinary polarization *for this specific value of* **k**. If the extraordinary polarization vector happens to be aligned with one of the other principal axes, $n_e(\mathbf{k})$ will take one of the values $n_o$ or $n_e$, but for arbitrary **k** the index is somewhere between these two extremes. Apparently, for uniaxial crystals only the extraordinary polarization state can experience transverse walkoff.

In this work we use the biaxial crystal BiBO. Things are not so simple in this case: because all three principal refractive indices are distinct, depending on **k** it is possible for both polarizations to walk off. For this more complicated situation, a useful construction is the *normal surface*, the **k**-space surface comprising the set of all **k** allowed in the medium. It is a *surface* because, for any direction of **k**, the frequency $\omega$ and the medium's refractive index constrain the magnitude $k$ to a single value. The normal surface generally has two layers, since there are generally two polarization eigenstates with different refractive indices for a given direction of **k**. It can be shown that the unit normal to the surface is parallel to **S**. For any point on the surface, the angle between $\mathbf{k_0}$ (the vector from the origin to the point) and $\mathbf{S_0}$ (the normal vector) is the *walkoff angle*. The product of this angle and the (geometric) path length through the crystal is the *transverse walkoff*, or simply the walkoff [15].

## A.2   Our compensation scheme

Here we describe our scheme for using additional birefringent crystals to partially undo the effects of walkoff. We use BiBO crystals with the same cut angles for all the SPDC and compensation crystals; furthermore, despite the large frequency difference and the slightly different propagation angles through the crystal[1], the dispersion in BiBO is such that the pump and SPDC experience similar walkoff. Therefore, the magnitude of the transverse walkoff in any of the crystals in our setup is the product of (the magnitude of) a common walkoff angle and the thickness of that crystal. We define $\delta \sim 35~\mu m$ as the magnitude of the walkoff in each of the SPDC crystals; the walkoff in one of the compensation crystals is then $\frac{l_{comp}}{l_{SPDC}}\delta$, where $l_{comp}$ is the thickness of that compensation crystal and $l_{SPDC} = 400~\mu m$ is the thickness of the SPDC crystals. This is the walkoff for one incident polarization. For the orthogonal polarization, the walkoff is zero: our crystals are cut such that the principal axis $\hat{\mathbf{x}}$ is parallel to the crystal face (see Section 2.2, Footnote 5 and Figure 2.5), and a normally incident beam polarized along this direction does not walk off. By analogy with walkoff in uniaxial crystals, we will refer to these two polarizations as extraordinary and ordinary, respectively.

We must also determine the direction of the walkoff in each crystal. First we define orthonormal vectors $\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}}$: here $\hat{\mathbf{w}}$ is normal to the crystal face and $\hat{\mathbf{u}}$ ($\hat{\mathbf{v}}$) points horizontally (vertically) in the transverse plane. The crystal described in Eq. 2.10 (see also Figure 2.5) causes walkoff for $V$ polarization along $\hat{\mathbf{v}}$, and similarly, the crystal in Eq. 2.11 causes walkoff for $H$ along $\hat{\mathbf{u}}$. As shown in Figure 4.9, we pre-compensate the SPDC crystals' walkoff using an identical, oppositely-oriented pair of crystals; this compensation is required for the stability of our laser, but we must also account for it when considering the SPDC modes' walkoff. Thus, these four crystals are the walkoff-inducing elements that are central to our setup. Our task is to include additional walkoff-inducing elements such that we minimize the spatial separation of the photons produced by the different SPDC processes.

Finally, let's make some notes regarding the effect of walkoff on the SPDC photons *while they are being produced.* The SPDC photons are extraordinary-polarized in the crystal that produces them, so that they experience walkoff; in contrast, the pump is ordinary-polarized here and does not walk off as it produces the SPDC. Therefore, SPDC generated just as the pump enters the crystal experiences the full walkoff amount $\delta$ and is spatially separated from SPDC produced just as the pump exits the crystal, which experiences no walkoff at all. Since the probability amplitude for SPDC adds coherently along the thickness of the crystal, a SPDC photon produced by a Gaussian pump has an elliptical transverse profile, where the major axis of the ellipse is along the walkoff direction. *Here we ignore this alteration of the spatial mode and only consider*

---

[1]The pump is normally incident on the crystal, while the SPDC emerges at approximately 3° outside the crystal. See Footnote 5 in Section 2.2 for additional detail.

*the displacement of the SPDC modes' centroids.* This displacement is $\delta/2$.

For simplicity, we will describe the walkoff on the pump and SPDC modes using their $(u, v)$ coordinates. At the SPDC crystals, where the pump and SPDC are focused to relatively small beam waists, these coordinates are simply the walkoff-induced displacements along $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$. On the other hand, at the second-pass pump mirror and the first-pass SPDC mirrors, the modes have all been collimated, so here the coordinates actually correspond to $(k_u, k_v)$, the Fourier transform **k**-space coordinates. While we can get away with this abuse of notation, we must be careful dealing with the first- and second-pass mirrors. All three mirrors are supposed to be aligned to normal incidence, but naturally this is impossible if two *different* wave-vectors are incident, as is the case when the two incident polarizations have experienced different walkoff[2]. The best we can do is to align the mirror normal to the mean of $\mathbf{k}_H$ and $\mathbf{k}_V$. For this alignment of the mirror, the two wave-vectors are *swapped* on reflection.

Now that we have described the effects of all the elements in our system, we come to the compensation scheme. Before explaining the procedure in full detail, let us outline the steps involved:

1. We maximize and balance the count rates for $HH$ and $VV$ SPDC photons collected by the RHS fibers. This ensures that our collection modes have optimal overlap with SPDC from both crystals.

2. We couple an alignment laser beam from the LHS fibers to the RHS fibers. Ultimately, our goal is to optimize LHS collection from both SPDC crystals, as we did on the RHS in Step 1. To accomplish this, we include walkoff compensation for the alignment laser and choose its polarization appropriately.

3. We maximize and balance the count rates from both crystals using the pump mirror for the second pass, and the signal and idler mirrors for the first pass. We use compensating crystals on the first pass SPDC to match the separation of the modes produced by the second pass.

Table A.1 shows the contributions of each crystal to the positions of the $HH$ and $VV$ modes at the RHS collection fibers. We assume that Step 1, above, leads to the SMFs being positioned at the centroid of the two modes $(\delta/4, -\delta/4)$. The separation of the modes $(-\delta/2, 3\delta/2)$ measures the impact of walkoff on the RHS coupling: for zero separation, the two modes coincide perfectly and it is possible to simultaneously optimize both couplings. The goals of Steps 2 and 3 become clearer now. In Step 2, assuming the pump mirror is correctly aligned to normal incidence, we wish to position the LHS fibers at the centroid of the second-pass $HH$ and $VV$ modes. In Table A.2 we find the centroid $(-\delta/4, -3\delta/4)$ and separation of the second-pass polarization modes. In Table A.3 we assume the LHS fiber has been correctly aligned to the second-pass centroid and show that the alignment laser polarization and compensation can be chosen such

---

[2]The achromatic lenses used to collimate the SPDC before the mirrors on the RHS and before the collection fibers on the LHS (see Figure 4.5) convert the two polarizations' different walkoff-induced *displacements* to different *angles*.

|  | Pump | | SPDC | |
|---|---|---|---|---|
|  | $H$ | $V$ | $HH$ | $VV$ |
| Initial | (0,0) | (0,0) | - | - |
| Pre-compensation 1 | $(-\delta,0)$ | (0,0) | - | - |
| Pre-compensation 2 | $(-\delta,0)$ | $(0,-\delta)$ | - | - |
| SPDC 1 | (0,0) | $(0,-\delta)$ | $(\delta/2,-\delta)$ | - |
| SPDC 2 | (0,0) | (0,0) | $(\delta/2,-\delta)$ | $(0,\delta/2)$ |
| **Centroid** | | | $(\delta/4,-\delta/4)$ | |
| **Separation** | | | $(\delta/2,3\delta/2)$ | |

Table A.1: Walkoff accounting for RHS collection of SPDC: each row gives $(u,v)$ coordinates after propagation through the corresponding walkoff-inducing element. The initial pump coordinates have been chosen as $(0,0)$ for simplicity. We assume that our RHS collection optimization leads to the signal and idler fibers being aligned to the centroid of the $HH$ and $VV$ mode locations. The separation between the $HH$ and $VV$ is a measure of the impact of the walkoff on our coupling.

|  | Pump | | SPDC | |
|---|---|---|---|---|
|  | $H$ | $V$ | $HH$ | $VV$ |
| Pump mirror | (0,0) | (0,0) | - | - |
| SPDC 2 | (0,0) | $(0,-\delta)$ | - | $(0,-\delta/2)$ |
| SPDC 1 | $(-\delta,0)$ | $(0,-\delta)$ | $(-\delta/2,-\delta)$ | $(0,-\delta/2)$ |
| **Centroid** | | | $(-\delta/4,-3\delta/4)$ | |
| **Separation** | | | $(\delta/2,\delta/2)$ | |

Table A.2: Walkoff accounting for LHS collection of SPDC from the second pump pass. We continue from where we left the pump after the second SPDC crystal in Table A.1. Our goal is to use the alignment laser to align the LHS fibers to the centroid of the two second-pass polarization modes.

that the beam arrives at the RHS centroid. Specifically, if we launch an $A$-polarized alignment beam from the LHS fiber, include a $-45°$-oriented compensation crystal of thickness $l_{comp} = l_{SPDC}/\sqrt{2}$, and analyze $V$ polarization at the RHS fiber output, the coupling is maximized when the RHS fiber is at the RHS SPDC centroid and the LHS fiber is at the LHS second-pass SPDC centroid.

In Step 3, assuming the signal and idler mirrors are correctly aligned to (optimal) normal incidence, we wish to add walkoff compensation such that the first-pass centroid matches the second-pass centroid, *and* such that the first-pass mode separation is as small as the second-pass mode separation. Our solution is to place a pair of crystals identical, but oppositely oriented, to the SPDC crystals in the first-pass idler arm (the arm that does *not* have a QWP). As shown in Table A.4, the resulting first-pass centroid and separation match those for the second-pass, for both signal and idler.

| | A |
|---|---|
| Initial | $(-\delta/4, -3\delta/4)$ |
| Compensation | $(\delta/4, -5\delta/4)$ |

| | H | V |
|---|---|---|
| SPDC 1 | $(5\delta/4, -5\delta/4)$ | $(\delta/4, -5\delta/4)$ |
| SPDC 2 | $(5\delta/4, -5\delta/4)$ | $(\delta/4, -\delta/4)$ |

Table A.3: Walkoff accounting for the alignment laser. The LHS fibers are positioned at the second-pass centroid. The alignment laser is launched from the LHS fibers, initially $A$ polarized so that the $-45°$-oriented compensation crystal induces walkoff in the $-45°$ direction. The SPDC crystals' walkoff, however, spatially separates the $H$ and $V$ components of the input beam. The $V$ component arrives at the RHS SPDC centroid. We can discard the $H$ component by analyzing $V$ polarization transmission through the RHS fibers.

| | Signal | | Idler | |
|---|---|---|---|---|
| | H | V | H | V |
| RHS SPDC | $(\delta/2, -\delta)$ | $(0, \delta/2)$ | $(\delta/2, -\delta)$ | $(0, \delta/2)$ |
| Compensation 1 | $(\delta/2, -\delta)$ | $(0, \delta/2)$ | $(-\delta/2, -\delta)$ | $(0, \delta/2)$ |
| Compensation 2 | $(\delta/2, -\delta)$ | $(0, \delta/2)$ | $(-\delta/2, -\delta)$ | $(0, -\delta/2)$ |
| QWP | L | R | H | V |
| Mirror | $(0, \delta/2)$ | $(\delta/2, -\delta)$ | $(0, -\delta/2)$ | $(-\delta/2, -\delta)$ |
| QWP | V | H | H | V |
| Compensation 2 | $(0, \delta/2)$ | $(\delta/2, -\delta)$ | $(0, -\delta/2)$ | $(-\delta/2, 0)$ |
| Compensation 1 | $(0, \delta/2)$ | $(\delta/2, -\delta)$ | $(\delta, -\delta/2)$ | $(-\delta/2, 0)$ |
| SPDC 2 | $(0, -\delta/2)$ | $(\delta/2, -\delta)$ | $(\delta, -\delta/2)$ | $(-\delta/2, -\delta)$ |
| SPDC 1 | $(0, -\delta/2)$ | $(-\delta/2, -\delta)$ | $(0, -\delta/2)$ | $(-\delta/2, -\delta)$ |
| **Centroid** | $(-\delta/4, -3\delta/4)$ | | $(-\delta/4, -3\delta/4)$ | |
| **Separation** | $(\delta/2, \delta/2)$ | | $(\delta/2, \delta/2)$ | |

Table A.4: Walkoff accounting for LHS collection of first-pass SPDC. We continue from where we left the signal and idler after the second SPDC crystal in Table A.1. The compensation crystal pair in the idler arm is identical to the SPDC crystals, but oppositely oriented. As shown we can compensate the first pass such that the modes from the two crystals have the same centroid and separation for both signal and idler.

# References

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical Review*, vol. 47, pp. 777–780, May 1935.

[2] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," *Naturwissenschaften*, vol. 23, Nov. 1935.

[3] J.-D. Bancal, S. Pironio, A. Acin, Y.-C. Liang, V. Scarani, and N. Gisin, "Quantum non-locality based on finite-speed causal influences leads to superluminal signalling," *Nature Physics*, vol. 8, no. 12, pp. 867–870, 2012.

[4] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," *Physical Review Letters*, vol. 69, pp. 2881–2884, Nov 1992.

[5] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, pp. 1895–1899, Mar 1993.

[6] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, vol. 67, no. 6, p. 661, 1991.

[7] J. V. Neumann, *Mathematical foundations of quantum mechanics*. No. 2, Princeton University Press, 1955.

[8] M. P. Almeida, F. de Melo, M. Hor-Meyll, A. Salles, S. Walborn, P. Souto Ribeiro, and L. Davidovich, "Environment-induced sudden death of entanglement," *Science*, vol. 316, no. 5824, pp. 579–582, 2007.

[9] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, p. 865, 2009.

[10] D. P. DiVincenzo *et al.*, "The physical implementation of quantum computation," *arXiv preprint quant-ph/0002077*, 2000.

[11] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, "Quantum repeaters based on atomic ensembles and linear optics," *Reviews of Modern Physics*, vol. 83, pp. 33–80, Mar 2011.

[12] D. J. Griffiths, *Introduction to electrodynamics*, vol. 3. Prentice Hall Upper Saddle River, NJ, 1999.

[13] G. A. Baym, *Lectures on quantum mechanics*. Benjamin, 1969.

[14] E. Hecht, *Optics*. Addison Wesley, 1998.

[15] A. Yariv and P. Yeh, *Optical waves in crystals*, vol. 10. Wiley, New York, 1984.

[16] B. Saleh and M. Teich, *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics, Wiley, 2007.

[17] M. Beijersbergen, L. Allen, H. Van der Veen, and J. Woerdman, "Astigmatic laser mode converters and transfer of orbital angular momentum," *Optics Communications*, vol. 96, no. 1, pp. 123–132, 1993.

[18] L. Allen, S. M. Barnett, and M. J. Padgett, *Optical angular momentum.* CRC Press, 2003.

[19] R. W. Boyd, *Nonlinear optics.* Academic press, 2003.

[20] C. K. Hong and L. Mandel, "Theory of parametric frequency down conversion of light," *Phys. Rev. A*, vol. 31, pp. 2409–2418, Apr 1985.

[21] V. G. Dmitriev, G. G. Gurzadyan, and D. N. Nikogosyan, *Handbook of nonlinear optical crystals*, vol. 64. Springer, 2013.

[22] A. Migdall, "Polarization directions of noncollinear phase-matched optical parametric downconversion output," *JOSA B*, vol. 14, no. 5, pp. 1093–1098, 1997.

[23] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, and P. H. Eberhard, "Ultrabright source of polarization-entangled photons," *Physical Review A*, vol. 60, pp. R773–R776, Aug 1999.

[24] S. Walborn, S. Pádua, and C. Monken, "Conservation and entanglement of hermite-gaussian modes in parametric down-conversion," *Physical Review A*, vol. 71, no. 5, p. 053812, 2005.

[25] R. Raussendorf and H. J. Briegel, "A one-way quantum computer," *Physical Review Letters*, vol. 86, pp. 5188–5191, May 2001.

[26] J. T. Barreiro, T.-C. Wei, and P. G. Kwiat, "Beating the channel capacity limit for linear photonic superdense coding," *Nature Physics*, vol. 4, no. 4, pp. 282–286, 2008.

[27] B. M. Terhal, "Quantum error correction for quantum memories," *Reviews of Modern Physics*, vol. 87, pp. 307–346, Apr 2015.

[28] M. Horodecki, P. Horodecki, and R. Horodecki, "Mixed-state entanglement and distillation: is there a "bound" entanglement in nature?," *Physical Review Letters*, vol. 80, no. 24, p. 5239, 1998.

[29] J. T. Barreiro, P. Schindler, O. Gühne, T. Monz, M. Chwalla, C. F. Roos, M. Hennrich, and R. Blatt, "Experimental multiparticle entanglement dynamics induced by decoherence," *Nature Physics*, vol. 6, no. 12, pp. 943–946, 2010.

[30] D. J. Griffiths, *Introduction to quantum mechanics.* Pearson Education India, 2005.

[31] A. Peres, *Quantum theory: concepts and methods*, vol. 57. Springer Science & Business Media, 2006.

[32] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics. 3: Quantum mechanics: Lectures on physics.* Addison-Wesley, 1965.

[33] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information.* Cambridge University Press, 2010.

[34] R. Jozsa, "Fidelity for mixed quantum states," *Journal of modern optics*, vol. 41, no. 12, pp. 2315–2323, 1994.

[35] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Physical Review A*, vol. 40, no. 8, p. 4277, 1989.

[36] V. Coffman, J. Kundu, and W. K. Wootters, "Distributed entanglement," *Physical Review A*, vol. 61, p. 052306, Apr 2000.

[37] W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways," *Physical Review A*, vol. 62, no. 6, p. 062314, 2000.

[38] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problemy Peredachi Informatsii*, vol. 9, no. 3, pp. 3–11, 1973.

[39] M. Koashi and A. Winter, "Monogamy of quantum entanglement and other correlations," *Physical Review A*, vol. 69, p. 022309, Feb 2004.

[40] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[41] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels," *Physical Review Letters*, vol. 76, pp. 722–725, Jan 1996.

[42] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, "Concentrating partial entanglement by local operations," *Physical Review A*, vol. 53, no. 4, p. 2046, 1996.

[43] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin, "Experimental entanglement distillation and "hidden" non-locality," *Nature*, vol. 409, no. 6823, pp. 1014–1017, 2001.

[44] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Physical Review A*, vol. 54, pp. 3824–3851, Nov 1996.

[45] A. Peres, "Separability criterion for density matrices," *Physical Review Letters*, vol. 77, no. 8, p. 1413, 1996.

[46] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Physics Letters A*, vol. 223, no. 1, pp. 1–8, 1996.

[47] S. Woronowicz, "Positive maps of low dimensional matrix algebras," *Reports on Mathematical Physics*, vol. 10, no. 2, pp. 165 – 183, 1976.

[48] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition," *Physics Letters A*, vol. 232, no. 5, pp. 333 – 339, 1997.

[49] J. A. Smolin, "Four-party unlockable bound entangled state," *Physical Review A*, vol. 63, no. 3, p. 032306, 2001.

[50] R. Augusiak and P. Horodecki, "Bound entanglement maximally violating bell inequalities: Quantum entanglement is not fully equivalent to cryptographic security," *Physical Review A*, vol. 74, p. 010305, Jul 2006.

[51] P. Horodecki, M. Horodecki, and R. Horodecki, "Bound entanglement can be activated," *Phys. Rev. Lett.*, vol. 82, pp. 1056–1059, Feb 1999.

[52] N. Linden, S. Massar, and S. Popescu, "Purifying noisy entanglement requires collective measurements," *Phys. Rev. Lett.*, vol. 81, pp. 3279–3282, Oct 1998.

[53] A. Kent, "Entangled mixed states and local purification," *Phys. Rev. Lett.*, vol. 81, pp. 2839–2841, Oct 1998.

[54] F. Kaneda, R. Shimizu, S. Ishizaka, Y. Mitsumori, H. Kosaka, and K. Edamatsu, "Experimental activation of bound entanglement," *Physical Review Letters*, vol. 109, p. 040501, Jul 2012.

[55] R. F. Werner and M. M. Wolf, "All-multipartite bell-correlation inequalities for two dichotomic observables per site," *Phys. Rev. A*, vol. 64, p. 032112, Aug 2001.

[56] M. Zukowski and C. Brukner, "Bell's theorem for general $N$-qubit states," *Physical Review Letters*, vol. 88, p. 210401, May 2002.

[57] P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Superactivation of bound entanglement," *Physical Review Letters*, vol. 90, p. 107901, Mar 2003.

[58] M. Zukowski, A. Zeilinger, M. Horne, and A. Ekert, ""event-ready-detectors" bell experiment via entanglement swapping," *Physical Review Letters*, vol. 71, no. 26, pp. 4287–4290, 1993.

[59] T. J. Herzog, J. G. Rarity, H. Weinfurter, and A. Zeilinger, "Frustrated two-photon creation via interference," *Physical Review Letters*, vol. 72, pp. 629–632, Jan 1994.

[60] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits," *Physical Review A*, vol. 64, p. 052312, Oct 2001.

[61] 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," *Physical Review Letters*, vol. 93, p. 053601, Jul 2004.

[62] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," *Transactions of the Cambridge Philosophical Society*, vol. 9, p. 399, 1851.

[63] M. J. Padgett and J. P. Lesso, "Dove prisms and polarized light," *Journal of Modern Optics*, vol. 46, no. 2, pp. 175–179, 1999.

[64] J. T. Barreiro, *Hyperentanglement for quantum information.* PhD thesis, University of Illinois at Urbana-Champaign, 2008.

[65] S. Grafström, U. Harbarth, J. Kowalski, R. Neumann, and S. Noehte, "Fast laser beam position control with submicroradian precision," *Optics Communications*, vol. 65, no. 2, pp. 121–126, 1988.

[66] E. Amselem and M. Bourennane, "Experimental four-qubit bound entanglement," *Nature Physics*, vol. 5, no. 10, pp. 748–752, 2009.

[67] J. Lavoie, R. Kaltenbaek, M. Piani, and K. J. Resch, "Experimental bound entanglement in a four-photon state," *Physical Review Letters*, vol. 105, p. 130501, Sep 2010.

[68] E. Amselem, M. Sadiq, and M. Bourennane, "Experimental bound entanglement through a pauli channel," *Scientific reports*, vol. 3, 2013.

[69] B. C. Hiesmayr and W. Löffler, "Complementarity reveals bound entanglement of two twisted photons," *New Journal of Physics*, vol. 15, no. 8, p. 083036, 2013.