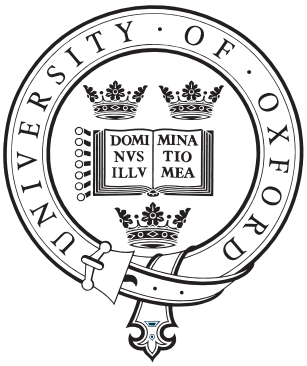
Single Photon Generation and Quantum Computing with Integrated Photonics

Justin B. Spring
Oriel College, Oxford



Submitted for the degree of Doctor of Philosophy
Trinity Term 2014

Supervised by
Prof. Ian A. Walmsley

Clarendon Laboratory
University of Oxford
United Kingdom

Single Photon Generation and Quantum Computing with Integrated Photonics

Justin Spring
Oriel College, Oxford

Submitted for the degree of Doctor of Philosophy
Trinity Term 2014

Abstract

Photonics has consistently played an important role in the investigation of quantum-enhanced technologies and the corresponding study of fundamental quantum phenomena. The majority of these experiments have relied on the free space propagation of light between bulk optical components. This relatively simple and flexible approach often provides the fastest route to small proof-of-principle demonstrations. Unfortunately, such experiments occupy significant space, are not inherently phase stable, and can exhibit significant scattering loss which severely limits their use. Integrated photonics offers a scalable route to building larger quantum states of light by surmounting these barriers.

In the first half of this thesis, we describe the operation of on-chip heralded sources of single photons. Loss plays a critical role in determining whether many quantum technologies have any hope of outperforming their classical analogues. Minimizing loss leads us to choose Spontaneous Four-Wave Mixing (SFWM) in a silica waveguide for our source design; silica exhibits extremely low scattering loss and emission can be efficiently coupled to the silica chips and fibers that are widely used in quantum optics experiments. We show there is a straightforward route to maximizing heralded photon purity by minimizing the spectral correlations between emitted photon pairs. Fabrication of identical sources on a large scale is demonstrated by a series of high-visibility interference experiments. This architecture offers a promising route to the construction of nonclassical states of higher photon number by operating many on-chip SFWM sources in parallel.

In the second half, we detail one of the first proof-of-principle demonstrations of a new intermediate model of quantum computation called boson sampling. While likely less powerful than a universal quantum computer, boson sampling machines appear significantly easier to build and may allow the first convincing demonstration of a quantum-enhanced computation in the not-distant future. Boson sampling requires a large interferometric network which are challenging to build with bulk optics, we therefore perform our experiment on-chip. We model the effect of loss on our postselected experiment and implement a circuit characterization technique that accounts for this loss. Experimental imperfections, including higher-order emission from our photon pair sources and photon distinguishability, are modeled and found to explain the sampling error observed in our experiment.

To Christina

Acknowledgements

Great teams are *really* rare–I was very lucky to be a member of one here. Thanks to Ian for putting me on an interesting and fruitful project ('Justin, we may want to look into this boson sampling thing...'). You act as an 'encouragement reservoir'–I have not yet thermalized! Nick: your hard work set the stage. I was a pretty crappy Luke to your Yoda, but I tried my best. Peter: Good in the lab AND with theory? Pick a side–we're at war! Steve: It was great to have another Boston sports fan nearby. Wait, you're from Texas?! Ben: Buddy since the beginning! We battled 3QB...and won.

To the Ultrafast team, it was truly a pleasure. One of the few downsides of having such a big group of great people is it's hard to list everybody by name here. So, I'll just paraphrase a famous Englishman...'so long, and thanks for all the tea!'

Thanks to all of the fabricators who made this work possible: Patrick and Martin in Engineering Sciences for the fs-written chips while Paolo, Dmytro, James, and Peter provided the UV-written circuits. Your jobs sound hard–I'll stick with the easy stuff.

Thanks to my family. Paul, Ryan, Dörte are the best of siblings. Savannah and Preston: your cute pictures brought a smile to my face during many a long lab session. Dad: I tend to phone during the low points but you don't complain. I'm thankful. Mom: I miss you now more than ever.

Thanks to the American taxpayer. I'd like to think you got your money's worth.

It's been real photons....time to go see about a girl!

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. government.

Author's publications

Journal publications

1. P. C. Humphreys, B. J. Metcalf, J. B. Spring, M. Moore, X-M. Jin, M. Barbieri, W. S. Kolthammer, and I. A. Walmsley. "**Linear Optical Quantum Computing in a Single Spatial Mode**" *Physical Review Letters*, **111**, 150501 (2013).

2. B. Calkins, P. L. Mennea, A. E. Lita, B. J. Metcalf, W. S. Kolthammer, A. Lamas-Linares, J. B. Spring, P. C. Humphreys, R. P. Mirin, J. C. Gates, P. G. R. Smith, I. A. Walmsley, T. Gerrits, and S. W. Nam. "**High quantum-efficiency photon-number-resolving detector for photonic on-chip information processing**" *Optics Express*, **21**, 22657 (2013).

3. J. B. Spring, P. S. Salter, B. J. Metcalf, P. C. Humphreys, M. Moore, N. Thomas-Peter, X-M. Jin, N. K. Langford, W. S. Kolthammer, M. J. Booth, and I. A. Walmsley. "**On-chip low loss heralded source of pure single photons**" *Optics Express*, **21**, 13522 (2013).

4. J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X. M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley. "**Boson Sampling on a Photonic Chip**" *Science*, **339**, 798 (2013).

5. B. J. Metcalf, N. Thomas-Peter, J. B. Spring, D. Kundys, M. A. Broome, P. C. Humphreys, X. M. Jin, M. Barbieri, W. S. Kolthammer, J. C. Gates, B. J. Smith, N. K. Langford, P. G. R. Smith, and I. A. Walmsley. "**Multiphoton quantum interference in a multiport integrated photonic device**" *Nature Communications*, **4:1356** (2013).

6. T. J. Bartley, G. Donati, J. B. Spring, X-M. Jin, M. Barbieri, A. Datta, B. J. Smith, and I. A. Walmsley. "**Multiphoton state engineering by heralded interference between single photons and coherent states**" *Physical Review A*, **86**, 043820 (2012).

7. P. S. Salter, A. Jesacher, J. B. Spring, B. J. Metcalf, N. Thomas-Peter, R. D. Simmonds, N. K. Langford, I. A. Walmsley, and M. J. Booth. "**Adaptive slit beam shaping for direct laser written waveguides**" *Optics Letters*, **37**, 470 (2012).

8. T. Gerrits, N. Thomas-Peter, J. C. Gates, A. E. Lita, B. J. Metcalf, B. Calkins, N. A. Tomlin, A. .E. Fox, A. Lamas Linares, J. B. Spring, N. K. Langford, R. P. Mirin, P. G. R. Smith, I. A. Walmsley, and S. W. Nam. "**On-chip, photon-number-resolving, telecommunication-band detectors for scalable photonic information processing**" *Physical Review A*, **84**, 060301(R) (2011).

9. N. Thomas-Peter, N. K. Langford, A. Datta, L. Zhang, B. J. Smith, J. B. Spring, B. J. Metcalf, H. B. Coldenstrodt-Ronge, M. Hu, J. Nunn, and I. A. Walmsley. "**Integrated photonic sensing**" *New Journal of Physics*, **13**, 055024 (2011).

Contents

Acknowledgements **iii**

Author's publications **v**

List of Figures **xii**

List of Tables **xvi**

1 Introduction **1**

1.1 Quantum state control 1

1.2 Interference of two photons on a beam splitter 5

1.2.1 Field quantization 6

1.2.2 Wave packets 8

1.2.3 Pure and mixed states 9

1.2.4 A beam splitter 11

1.2.5 Nonclassical interference 13

2 Theory of single-photon generation by spontaneous four-wave mixing 17

2.1 The ideal single photon source 17

2.1.1 Single emitters 18

2.1.2 Multiple emitters 19

2.2 SFWM source design 25

2.2.1 The SFWM two-photon state 25

2.2.2 Photon loss 32

2.2.3 Emission of multiple photon pairs 34

2.2.4 Interference visibility 40

2.2.5 Conversion efficiency 51

3 Demonstration of on-chip SFWM sources 55

3.1 Femtosecond-laser written SFWM source 56

3.1.1 Femtosecond-laser fabrication of waveguides 56

3.1.2 Observation of SFWM 61

3.1.3 Low-loss source of single photons 63

3.1.4 Heralded-state purity 64

3.1.5 Single-source homogeneity 67

3.1.6 Difficulty achieving multi-source homogeneity 69

3.2 Silica-on-silicon UV-written SFWM sources 73

3.2.1 Single SFWM source 74

3.2.2 Two-source interference 81

3.2.3 Source performance comparison 84

4 Boson sampling theory **87**

4.1 Models of computation 88

4.1.1 Classical Computers 88

4.1.2 A Universal Quantum Computer 89

4.1.3 Boson Sampling Machine: An Intermediate Quantum Computer . . 91

4.2 Boson Sampling Overview 93

4.2.1 A Boson Sampling Machine 93

4.2.2 Bosonic scattering amplitudes given by the matrix permanent 97

4.3 Computational Complexity of Boson Sampling 102

4.3.1 Complexity Class Zoo 102

4.3.2 Fast classical solution of exact boson sampling collapses PH 108

4.3.3 Fast classical solution of approximate boson sampling likely collapses PH 114

5 Demonstration of boson sampling **123**

5.1 Effect of loss 124

5.1.1 Lossy boson distribution still given by matrix permanent 124

5.1.2 Postselected boson sampling with loss: a case study 126

5.2 Circuit characterization 133

5.2.1 Overview 133

5.2.2 Blackbox characterization with single photons 134

5.3 Experimental results 138

5.3.1 Scope of the experiment 138

5.3.2 Experimental layout 139
5.3.3 Characterization results 140
5.3.4 Boson sampling results 141
5.4 Error analysis 144
5.4.1 Higher-order SPDC pair emission 145
5.4.2 Photon distinguishability 153
5.4.3 Final error analysis results 156

6 Conclusion and outlook **159**
6.1 SFWM sources 159
6.2 Boson sampling 162

A Circuit element characterization **169**
A.1 Characterization of UV-written elements 170
A.2 Characterization of losses 173

B Photon distinguishability characterization **179**
B.1 Overview of model 179
B.2 Determination of α_{nb} from interference visibility 181
B.3 Determination of α_{bb} from spectral overlap 183

C FPGA coincidence counting **185**
C.1 Data collection challenge 186
C.2 FPGA introduction 189
C.3 Coincidence counters 190

C.3.1 General purpose coincidence counter 191

C.3.2 High temporal resolution 197

Bibliography **203**

List of Figures

1.1 Two photons interfere on a beam splitter 5

1.2 Example of a Hong-Ou-Mandel interference dip 15

2.1 Schematic of a heralded single photon source 20

2.2 Generation of new frequencies in a $\chi^{(2)}$ dielectric 22

2.3 Modeling of loss with a beam splitter 33

2.4 Representative SFWM phasematching curve with Raman noise band 38

2.5 Experimental layout to measure $g^{(2)}(0)$ with bucket detectors 39

2.6 Experimental layout to observe HOM interference between heralded single photons from two SFWM sources 41

2.7 Representative joint spectral amplitudes for both a factorable and entangled two-photon state 44

2.8 Hard-edge dielectric boundaries lead to side lobes in the phasematching function 47

3.1 Schematic of femtosecond laser waveguide fabrication 57
3.2 Spatial light modulator beam shaping 58
3.3 THG images of the waveguide cross-section and near-field images of the spatial mode 60
3.4 Elliptical spatial mode from a femtosecond laser written SFWM source . . . 60
3.5 Experimental layout to operate and characterize a single on-chip SFWM source 61
3.6 Representative marginal spectra from a femtosecond-laser-written SFWM source 62
3.7 Number statistics from a fs-written SFWM pair source 64
3.8 Unheralded $g^{(2)}(0)$ measurements reveal spectrally factorable SFWM source 65
3.9 Spectrally factorable SFWM emission achieved from the visible to telecom wavelengths 67
3.10 Simple models of birefringence inhomogeneity indicate SFWM source uniformity 68
3.11 Three attempted methods to achieve birefringence in a femtosecond-laser-written waveguides 70
3.12 Difficulties creating identical SFWM sources with the femtosecond-laser-writing technique 71
3.13 Numerical model of the birefringence of a rectangular core waveguide 72
3.14 Schematic of the UV-written waveguide fabrication process 73
3.15 Representative SFWM marginal spectra from UV-written waveguides . . . 75
3.16 Fluorescence noise in UV-written SFWM sources 76
3.17 Spectral filters angle-tuned to minimize noise 78

3.18 Fabrication of matched UV-written SFWM sources 80
3.19 Phasematched wavelengths from UV-written sources slowly shift 81
3.20 Experimental layout for HOM interference experiments with two UV-written SFWM sources 82
3.21 High HOM interference visibilities demonstrated with a series of two-source experiments 83

4.1 Outline of a Turing Machine 89
4.2 Schematic of a photonic boson sampling machine 94
4.3 Analogy between a Galton board and boson sampling machine 95
4.4 Outline of a Nondeterministic Turing Machine 105
4.5 The Polynomial Hierarchy 109
4.6 Conjectured steps showing that an efficient classical solution to the boson sampling problem collapses PH 122

5.1 General model of a lossy linear optical network 124
5.2 Simplified model of a lossy linear optical network 129
5.3 Schematic of boson sampling experiment 140
5.4 Blackbox characterization results 141
5.5 Experimental boson sampling results 142
5.6 Loss analysis of the circuit used in the boson sampling experiment 146
5.7 Simple example of higher-order SPDC emission error analysis 149
5.8 Loss modes used in the modeling of the boson sampling experiment 150
5.9 Results of boson sampling error analysis 158

6.1 Outline of scattershot boson sampling 166

A.1 Losses in-between every linear element in the boson sampling experiment . 170

A.2 Loss independent characterization of beamsplitter reflectivity 171

A.3 Characterization of relative phase terms 172

A.4 Balanced losses commute with a beamsplitter operation 173

A.5 Consolidation of redundant loss terms-A 174

A.6 Consolidation of redundant loss terms-B 174

A.7 Consolidation of redundant loss terms-C 175

A.8 Consolidation of redundant loss terms-D 175

A.9 Schematic of experiment with minimal set of losses 176

B.1 Spectra of three SPDC modes 182

B.2 HOM interference experiment using on-chip beamsplitter 182

C.1 Two detectors input to FPGA 191

C.2 Coincidence counter electronic delay and temporal resolution signal processing steps 192

C.3 Illustration of FPGA counting coincidences 195

C.4 Coincidence counter GUI 196

C.5 Illustration of high temporal resolution FPGA counting coincidences 198

C.6 Calibration of the asynchronous FPGA coincidence counter 200

C.7 Demonstration of the asynchronous FPGA coincidence counter 201

List of Tables

3.1 η_H and $g_H^{(2)}(0)$ depends on pump power and degree of temporal filtering . . 79

C.1 Summary of Matlab commands for FPGA coincidence counter 196

Chapter 1

Introduction

We never experiment with just one electron or atom or (small) molecule. In thought experiments, we sometimes assume that we do; this invariably entails ridiculous consequences... we are not experimenting with single particles, any more than we can raise Ichthyosauria in the zoo[1].

-Erwin Schrödinger, 1952

1.1 Quantum state control

Successful physical theories predict a lot while postulating little. By this measure, quantum mechanics has been successful indeed. It accurately predicts a wide-variety of physical phenomena ranging from the sharp spectral features of atomic vapors to superconductivity. However, experimental confirmation of quantum mechanics does not imply experimental control of individual quantum states. One does not need to individually manipulate every atom in a vapor to observe sharp absorption and emission features. Nor does one need to access the individual Cooper pairs of a superconductor to demonstrate superconductivity. In these cases, the ensemble behavior is sufficient to observe quantum phenomena.

Achieving complete control of individual quantum states opens up exciting new possibilities. For example, one can construct larger quantum systems in a deliberate fashion piece-by-piece. Perhaps the most famous application of such a capability is quantum-enhanced computation. Operating a computer whose algorithmic operation is explicitly quantum mechanical, a quantum Turing machine, yields an exponential speed advantage over classical devices for certain problems. Successful construction of such a universal quantum computer would be of foundational interest in its own right, and would have added value due to its ability to solve otherwise intractable problems in a variety of disciplines [2].

Experiments on individual quantum systems must satisfy countervailing goals. The experimenter must have access to the system to prepare, manipulate, and measure its state. However, the system must generally be well-isolated from the rest of the universe. If this latter goal is not met, then undesired *decoherence* of the system occurs which frequently spoils the uniquely quantum mechanical effects that originally motivated the work.

Many physical architectures have been successfully used to demonstrate quantum state control. A few examples include include single ions, single atoms, superconductors, and nano-mechanical oscillators [3]. Minimizing decoherence in these systems generally requires cryogenics, vacuum, cavities, magneto-optical traps or some combination of the above. Developing new quantum devices that involve these demanding technologies is a major research challenge that impedes progress towards precision control of large quantum states.

Photons present an attractive alternative, as they interact only very weakly with their environment and thus decoherence can be minimized with less experimental effort. For this reason, benchtop photonic experiments have historically played a leading role in testing both fundamental quantum phenomena [4–9] and proof-of-principle demonstrations of quantum

technologies[10–14]. In addition to their coherence benefits, the rapid spatial propagation of photons makes them uniquely well-suited to experiments that require distribution of quantum systems to separated parties[15–18].

Many of these early photonic experiments were implemented in 'bulk' optical systems; the optical components of the experiment are not part of a monolithic structure. A common arrangement involves many separate mirrors and beam splitters separately bolted to an optical table. The photons traverse a beam-like path in free space between these many components. These arrangements take up substantial space, require careful alignment, frequently require sub-wavelength path stability for long time periods, and can incur significant scattering loss. While these drawbacks may not be prohibitive for small proof-of-principle experiments, they significantly hinder scaling to larger systems.

Large quantum states are of interest since in many situations the contrast between classical and quantum physics becomes striking at larger scales. Quantum computation again provides a convenient example. Quantum and classical computers can solve the same class of problems, however quantum machines offer the possibility of doing so more efficiently than their classical analogues[2]. However, if the quantum machine, and thus the size of the problem, is small, then this difference in computational efficiency becomes unimportant as the classical machine can quickly return a solution too. Only for larger systems and problem sizes does the quantum computer provide a meaningful advantage.

Integrated photonics, where optical components are consolidated into a single assembly, provides a scalable route to constructing large quantum states of light. These miniaturized devices can greatly simplify optical alignment, minimize loss, and naturally provide the optical phase stability required in many applications[19,20]. Integrated quantum optics

experiments have benefited by leveraging mature technologies developed for the telecommunications industry. For these reasons, integrated platforms have become an increasingly common route to the construction of more complicated quantum optical states in recent experiments[21–26].

Fully taking advantage of these benefits requires integration of all components in an experiment including state preparation, manipulation, and detection. Unfortunately, it is difficult to find a single substrate that is well-suited to all three of these tasks. For example, waveguides in the crystalline materials that are well-suited for preparation of single photons can be fairly lossy and thus poor materials to construct the large interferometric networks required for state manipulation. One can adopt a hybrid approach, where different substrates are chosen for each step of the experiment. However, this introduces the new problem of minimizing the losses at the interfaces between materials[20]. While researching solutions to these problems is ongoing, currently many 'integrated' experiments still require a significant number of bulk optical components.

This thesis describes the use of photonic chips to construct larger and more complex nonclassical states of light. In the first half, we detail the on-chip integration of low-loss sources of high purity single photons produced by Spontaneous Four-Wave Mixing (SFWM). Due to the use of a silica substrate, this source could feasibly be interfaced in the near-term with the silica chips that have been used in the majority of on-chip quantum optics experiments performed to date[23–25,27–29]. In the second half of this thesis, we detail the use of a large on-chip interferometric network to implement a novel quantum computational model called boson sampling. While boson sampling machines only solve a very specific type of problem, they appear far easier to build than universal quantum computers. They present

a possible route to the first definitive quantum-enhanced computation in the not-distant future. Before discussing these topics in detail, we use the second half of this introductory chapter to briefly review those aspects of quantum mechanics that will prove most pertinent to the later chapters. We ask a simple question about two-photon interference and review the quantum mechanics required to provide an adequate answer.

1.2 Interference of two photons on a beam splitter

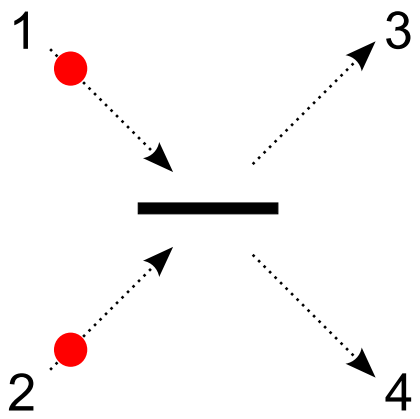


Figure 1.1: Two photons incident on a beam splitter in modes 1 and 2 will emerge on outputs 3 and 4.

What happens when two single photons interfere on a 50% reflectivity beam splitter? The beginning of this chapter will be spent sketching the answer to this deceptively simple question, which is depicted in Figure 1.1. Along the way, we will introduce many of the relevant quantum mechanical concepts that will be used regularly throughout the remainder of this thesis. Aside from introducing physical concepts, the answer to the question is important in its own right; nonclassical interference between the two photons leads to an entangled output. Generating entangled states is one of the main motivations for building

single photon sources, which is the focus of the first half of this thesis. In addition, entanglement can act as an important resource in allowing quantum computers to gain an exponential speed advantage over their classical counterparts for certain tasks. A model of quantum computation particularly well-suited to photonics is the focus of the latter half of this thesis.

1.2.1 Field quantization

To address the above question, we must first define what is meant by a *photon*. Classical physics, which allows a mode of the electromagnetic field to assume a continuous range of energies, does not have an answer. One must quantize the electromagnetic field to introduce the concept of a photon.

Consider a one-dimensional cavity of length L that has conducting walls at $z = \pm L/2$. The transverse polarization of the electric field inside the cavity must vanish at the boundaries, yielding a set of modes whose electric field is given by

$$E(z,t) = \frac{A_m}{2}\left[e^{-i(k_m z - \omega_m t)} + c.c.\right], \tag{1.1}$$

where the wavevector is given by $k_m = (2m-1)\pi/L$ with $m = 1, 2, ...$ enumerating the allowed modes. The angular frequency ω_m is determined by the dispersion relation $\omega_m = ck_m$. While there are a discrete number of modes allowed in Equation (1.1), the amplitude of a mode A_m can take any value. This implies the energy in the electromagnetic field

$$\mathcal{E}(t) = \frac{1}{2}\int_0^L dz\left[\epsilon_0 E(z,t)^2 + \mu_0^{-1} B(z,t)^2\right], \tag{1.2}$$

can also assume any value, where ϵ_0 and μ_0 are the electric permittivity and magnetic permeability of free space while $B(z,t)$ is the magnetic field.

In quantum mechanics, each of the above modes of the field is treated as a quantum harmonic oscillator whose Hamiltonian takes the form

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \tag{1.3}$$

where $\hbar$ is the reduced Planck constant, and $\hat{a}^\dagger$ and $\hat{a}$ are operators obeying the commutation relations $[\hat{a}, \hat{a}^\dagger] = 1$. A quantum state $|\psi\rangle$ is described by a vector in Hilbert space, whose temporal evolution is described by the Schrödinger equation

$$i\hbar\frac{d\,|\psi\rangle}{dt} = \hat{H}\,|\psi\rangle\,. \tag{1.4}$$

The allowed energy values of this quantized oscillator are found by writing the eigenvalue equation where $|n\rangle$ is an energy eigenstate with corresponding energy eigenvalue E_n. The commutation relation implies that the allowed energies are quantized

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \tag{1.5}$$

where $n = 0, 1, 2, ...$ [30] . Furthermore, the operators $\hat{a}^\dagger$ and $\hat{a}$ act to raise or lower the

energy eigenstate according to

$$\hat{a}^\dagger |n\rangle = (n+1)^{1/2} |n+1\rangle$$
$$\hat{a} |n\rangle = n^{1/2} |n-1\rangle , \tag{1.6}$$

where the leading coefficients ensure the bosonic commutation relation is preserved.

This quantization has surprising consequences. The electromagnetic field can only gain and lose energy in discrete packets, called photons, that correspond to excitations of a quantum harmonic oscillator. Building sources that emit single photons and subsequently using them to observe nonclassical phenomena encapsulates all the work in this thesis.

1.2.2 Wave packets

The previous quantization example was presented in terms of a monochromatic electric field. This is unphysical because a field with a single frequency ω would have infinite temporal duration. We can generalize to the more realistic situation of a single-photon wave packet that is temporally finite. This wave packet is then made up of multiple frequency components

$$\left|1_{f(\omega)}\right\rangle = \hat{A}^\dagger_{f(\omega)} |0\rangle = \int d\omega f(\omega) \hat{a}^\dagger(\omega) |0\rangle , \tag{1.7}$$

where $\hat{a}^\dagger(\omega)$ is the creation operator for a mode with frequency ω and the complex-valued $f(\omega)$ describes the weighting of these spectral components with the property that $\int d\omega |f(\omega)|^2 = 1$ to ensure $\langle\psi|\psi\rangle = 1$. We use the capitalized notation $\hat{A}^\dagger_{f(\omega)}$ to describe the creation of a photon in the wave packet defined by $f(\omega)$. To simplify notation, the

subscript will often be dropped and $\hat{A}^\dagger$ used when the form of the wave packet spectrum is implied.

If we measure the frequency of the single-photon wave packet in Equation (1.7), we can't predict the result with certainty. In quantum mechanics, the probability of obtaining a measurement result can be expressed as the expectation value of a projector. The probability of measuring the frequency to be in the range $\omega_a \leq \omega \leq \omega_b$ is given by

$$\begin{aligned} P(\omega_a \leq \omega \leq \omega_b) &= \left\langle 1_{f(\omega)} \right| \Pi_{\Delta\omega} \left| 1_{f(\omega)} \right\rangle \\ &= \left\langle 1_{f(\omega)} \right| \int_{\omega_a}^{\omega_b} d\omega' \hat{a}^\dagger(\omega') \left| 0 \right\rangle\!\left\langle 0 \right| \hat{a}(\omega') \left| 1_{f(\omega)} \right\rangle \\ &= \int_{\omega_a}^{\omega_b} d\omega \left| f(\omega) \right|^2, \end{aligned} \tag{1.8}$$

where $\Pi_{\Delta\omega}$ is the projector onto the subspace of frequencies in the range $\omega_a \leq \omega \leq \omega_b$, and we have used the commutation relation appropriate for this continuous frequency space $[\hat{a}(\omega), \hat{a}^\dagger(\omega')] = \delta(\omega - \omega')$. If $\omega_a = 0$ and $\omega_b \to \infty$ then we obtain $\int_0^\infty d\omega \left| f(\omega) \right|^2 = 1$ due to the normalization condition. This is reasonable since if you measure the frequency of the single photon, a result is guaranteed. However, for narrower frequency ranges we can only state probabilities of observing certain results, given by Equation (1.8).

1.2.3 Pure and mixed states

A *pure state* is any quantum state that can be described by a single vector $|\psi\rangle$ in Hilbert space. This includes coherent superpositions of states, for example the single-photon wave packet in Equation (1.7). While this wave packet is made up of many frequencies, they have a fixed phase relationship to one another defined by $f(\omega)$ which allows the resulting sum

to be expressed by a single vector $\left|1_{f(\omega)}\right\rangle$. Despite exact knowledge of $|\psi\rangle$, we can often only state the probabilities of obtaining certain measurement results. This uncertainty is intrinsic to quantum mechanics and cannot be resolved through improved measurements.

In practice, measurement uncertainties are often larger than this intrinsic lower limit. For example, imperfect knowledge on the part of the experimenter may leave them unsure whether they have prepared the state $|\psi_1\rangle$, $|\psi_2\rangle$, ... , or $|\psi_n\rangle$. If they had performed a more careful experiment, they would know that their state is, for example, $|\psi_1\rangle$. However, this information is not available and so they are left with a mixed state that accounts for all possibilities

$$\hat{\rho} = \sum_m p_m \left|\psi_m\right\rangle\!\left\langle\psi_m\right|, \tag{1.9}$$

where the trace must satisfy $\mathrm{Tr}(\hat{\rho}) = 1$ to maintain normalization. The density operator $\hat{\rho}$ properly accounts for the uncertainty in which pure state has been prepared. This classical uncertainty, described by the probabilities p_m, is not intrinsic to quantum mechanics but instead reflects imperfect knowledge on the part of the experimenter. This is accounted for when calculating expectation values by performing a weighted average over the possibilities

$$\begin{aligned}\langle\Pi\rangle &= \sum_m p_m \left\langle\psi_m\right| \Pi \left|\psi_m\right\rangle \\ &= \mathrm{Tr}(\hat{\rho}\Pi).\end{aligned} \tag{1.10}$$

One can quantify the purity of a state according to

$$\mathcal{P}_\rho = \mathrm{Tr}(\hat{\rho}^2) = \sum_m p_m^2. \tag{1.11}$$

A pure state can be described by the density operator $\hat{\rho} = |\psi\rangle\langle\psi|$ and has a purity $\mathcal{P} = 1$. A maximally mixed state, obtained in the limit $m \to \infty$ and $p_m \to 0$ in Equation (1.9), corresponds to $\mathcal{P} = 0$.

1.2.4 A beam splitter

A partially reflective surface can be used to split an electromagnetic wave into two outputs. Enforcing conservation of energy at this lossless beam splitter is sufficient to deduce the unitary transformation relating input and output fields shown in Figure 1.1 [30,31].

For a beam splitter with reflection and transmission coefficients $\mathcal{R}$ and $\mathcal{T}$, the creation operators transform according to

$$\begin{aligned}\hat{a}_3^\dagger &= \mathcal{R}\hat{a}_1^\dagger + \mathcal{T}\hat{a}_2^\dagger \\ \hat{a}_4^\dagger &= \mathcal{T}\hat{a}_1^\dagger - \mathcal{R}\hat{a}_2^\dagger,\end{aligned} \tag{1.12}$$

where we can assume $\mathcal{R}$ and $\mathcal{T}$ are real and modes are labelled according to Figure 1.1. For the work described in this thesis, we can assume that the beam splitter reflectivity is constant over the spectral range of a wave packet. Therefore, the monochromatic operators $\hat{a}^\dagger$ in (1.12) can be replaced with wave packet operators $\hat{A}^\dagger$.

What happens when a single photon in a pure state hits a balanced beam splitter ($\mathcal{R}=\mathcal{T}=1/\sqrt{2}$)? Our initial wave packet transforms to

$$\begin{aligned}|\psi\rangle &= \hat{A}_1^\dagger |0\rangle \\ &= \frac{1}{\sqrt{2}}\left(\hat{A}_3^\dagger + \hat{A}_4^\dagger\right)|0\rangle = \frac{1}{\sqrt{2}}\left(|10\rangle + |01\rangle\right),\end{aligned} \tag{1.13}$$

where the ket notation $|10\rangle$ represents the state with one photon in output three and zero in output four, and similarly for $|01\rangle$. Our single photon is now in a superposition of both outputs. It is not a well-posed question to ask which output the photon is *really* in. One must account for both possibilities.

It is impossible to write the state of our total system as a product of what is individually happening in output mode 3 and mode 4 alone: $|\psi\rangle \neq |\psi_3\rangle\,|\psi_4\rangle$. The subsystems can each have zero or one photons. Regardless of which choice is made, the product will never capture the superposition of a photon distributed amongst both modes in Equation (1.13). This inability to express the global state of a system as a product of its constituent subsystems is known as entanglement [32].

This superposition of a single photon over two modes is not a classical mixture. To demonstrate this, we can write out the density matrices for the pure superposition state of Equation (1.13) and the mixed state of a single photon over the two modes. For the former case, we use Equations (1.9) and (1.13) to write

$$\begin{aligned}\hat{\rho}_{pure} &= \frac{1}{2}\Big(|10\rangle + |01\rangle\Big)\Big(\langle 10| + \langle 01|\Big) \\ &= \frac{1}{2}\Big(|10\rangle\langle 10| + |10\rangle\langle 01| + |01\rangle\langle 10| + |01\rangle\langle 01|\Big),\end{aligned} \tag{1.14}$$

which has a purity of $\mathcal{P} = \mathrm{Tr}\left(\hat{\rho}^2_{pure}\right) = \mathrm{Tr}\left(\hat{\rho}_{pure}\right) = 1$. For comparison, imagine a state that is definitely either in the state $|10\rangle$ or $|01\rangle$ with equal probability, $p=1/2$. According to Equation (1.9), this state is described by

$$\hat{\rho}_{mixed} = \frac{1}{2}\Big(|10\rangle\langle 10| + |01\rangle\langle 01|\Big), \tag{1.15}$$

with a corresponding purity of $\mathcal{P} = 1/2$, which is less than unity reflecting the mixedness of this state.

1.2.5 Nonclassical interference

After this brief review of key quantum mechanics concepts, we return to our original question: what happens when two photons interfere on a beam splitter?

We will start with the simplest case of a balanced beam splitter with pure state inputs. Assuming spectrally identical wave packets the two-photon output is

$$\begin{aligned} |\psi\rangle &= \hat{A}_1^\dagger \hat{A}_2^\dagger |0\rangle \\ &= \frac{1}{2}\left(\hat{A}_3^\dagger + \hat{A}_4^\dagger\right)\left(\hat{A}_3^\dagger - \hat{A}_4^\dagger\right)|0\rangle \\ &= \frac{1}{2}\left(\hat{A}_3^\dagger \hat{A}_3^\dagger - \hat{A}_3^\dagger \hat{A}_4^\dagger + \hat{A}_3^\dagger \hat{A}_4^\dagger - \hat{A}_4^\dagger \hat{A}_4^\dagger\right)|0\rangle \\ &= \frac{|20\rangle - |02\rangle}{\sqrt{2}}. \end{aligned} \tag{1.16}$$

It's clear that the photons coalesce and always emerge from the beam splitter in the same spatial mode. This effect is known as Hong-Ou-Mandel (HOM) interference[33] and the entangled output in Equation (1.16) is a critical resource in many photonic quantum-enhanced applications.

This nonclassical HOM interference reveals itself in the correlations between measurements, such as simultaneous detection of a photon at each output. Commonly-used single photon detectors, such as avalanche photodiodes (APDs), cannot resolve frequency and so

we use an incoherent sum over frequency to find this projector for coincident detections

$$\Pi_{cc} = \iint d\omega d\omega' \hat{a}_3^\dagger(\omega)\hat{a}_4^\dagger(\omega') \left|0\right\rangle\!\left\langle 0\right| \hat{a}_3(\omega)\hat{a}_4(\omega'), \tag{1.17}$$

where subscripts again indicate the relevant spatial mode. Using this projector to find the probability of a coincident detection given the state in Equation (1.16) we find

$$P_{cc}(0) = \langle\psi| \Pi_{cc} |\psi\rangle = 0, \tag{1.18}$$

which is unsurprising since the photons always 'bunch' and never emerge on separate outputs.

The perfect HOM interference described above in Equation (1.16) relies on the two photons being indistinguishable. If the photons are instead completely distinguishable, for example if one photon is delayed with respect to the other so that they arrive at the beam splitter at different times, then there is no longer any interference. We introduce the notation $P_{cc}(\infty)$ to express the probability of coincident photon detection at the two outputs of the beam splitter when the photons are distinguishable. Since there is no interference in this case, each photon has an equal probability of emerging on either output leading to $P_{cc}(\infty) = 1/2$. Our detectors don't need to resolve these two temporal modes for the photons to be categorized as distinguishable. If there exists a possible measurement that could distinguish the photons in-principle, then they are distinguishable even if that measurement isn't implemented. In between these two extremes, there is partial interference dependent on the degree of overlap of the photon wave packets.

Adjusting the distinguishability of the photons, often by temporally delaying one of the photons, leads to a characteristic HOM interference 'dip' as shown in Figure 1.2. One can define the visibility of this interference according to

$$V = \frac{P_{cc}(\infty) - P_{cc}(0)}{P_{cc}(\infty)}, \tag{1.19}$$

such that $V = 1$ when there are indistinguishable photons ($P_{cc}(0) = 0$).

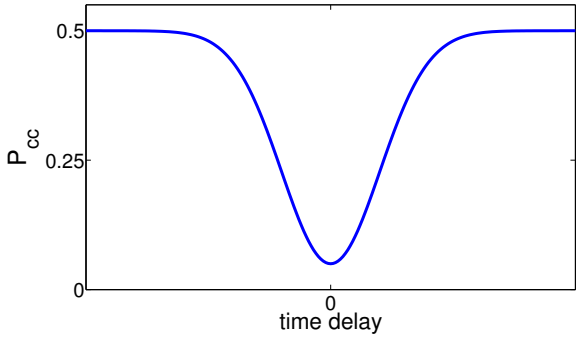


Figure 1.2: An example of a Hong-Ou-Mandel interference dip, with visibility $V = 0.9$.

If the photons incident on the beam splitter are in mixed states, the interference visibility will be reduced. The experiments described in this thesis primarily encounter mixedness in the spectral degree of freedom. For this reason, we extend Equation (1.9) to the continuous frequency basis and express these input states as $\hat{\rho}_1 = \iint d\omega_1 d\omega_1' g(\omega_1, \omega_1') \hat{a}_1^\dagger(\omega_1) \left|0\right\rangle\!\left\langle 0\right| \hat{a}_1(\omega_1')$ and $\hat{\rho}_2 = \iint d\omega_2 d\omega_2' h(\omega_2, \omega_2') \hat{a}_2^\dagger(\omega_2) \left|0\right\rangle\!\left\langle 0\right| \hat{a}_2(\omega_2')$ yielding the total input

$$\hat{\rho}_{in} = \iiiint d\omega_1 d\omega_1' d\omega_2 d\omega_2' g(\omega_1, \omega_1') h(\omega_2, \omega_2') \hat{a}_1^\dagger(\omega_1) \hat{a}_2^\dagger(\omega_2) \left|0\right\rangle\!\left\langle 0\right| \hat{a}_1(\omega_1') \hat{a}_2(\omega_2').$$

Assuming a balanced beam splitter, and transforming operators according to Equation (1.12), one finds the probability of detecting the photons at separate outputs using Equa-

tions (1.17) and (1.10)

$$P_{cc}(0) = \frac{1}{2}\left(1 - \iint d\omega d\omega' g(\omega,\omega')h(\omega',\omega)\right). \tag{1.20}$$

From Equation (1.19), this implies the HOM interference visibility will be

$$V = \iint d\omega d\omega' g(\omega,\omega')h(\omega',\omega), \tag{1.21}$$

since $P_{cc}(\infty) = 1/2$ for mixed state inputs. If the two photons are in identical, but mixed, states then

$$\begin{aligned} V &= \iint d\omega d\omega' g(\omega,\omega')g(\omega',\omega) \\ &= \mathrm{Tr}(\hat{\rho}^2) = \mathcal{P}_{\rho_{1,2}}. \end{aligned} \tag{1.22}$$

Thus, the HOM interference visibility is a direct measure of state purity for identical photons. In practice, there will always be some distinguishability between the photons which will lower the visibility below the purity of either photon individually. Thus, the interference visibility acts as a lower bound for $\mathcal{P}_{\rho_1}$ and $\mathcal{P}_{\rho_2}$.

Chapter 2

Theory of single-photon generation by spontaneous four-wave mixing

We want to construct and manipulate large quantum states of light. A basic building block in this effort is a single photon; unfortunately these are not trivial to produce. The beginning of this chapter reviews those qualities an ideal single photon source should possess. Many physical systems have been used to build single photon sources; we present a brief review of their advantages and disadvantages. This analysis leads us to choose SFWM in a silica waveguide as a promising single photon source architecture. The bulk of the chapter is then spent deriving the quantum state produced by SFWM and using this as a guide to optimize our source design.

2.1 The ideal single photon source

What characteristics would an ideal source of single photons possess?

First, an ideal source should emit exactly one photon at a time. This immediately disqualifies many candidates. For example, a pulsed laser, even when attenuated to an average of one photon per pulse, still emits zero or more than one photon some of the time. Second, an ideal source should operate on demand; every time the user requests a single photon, one should be delivered. Third, such a source would emit the pure states of light required for nonclassical interference. Fourth, one should have the ability to build many identical sources satisfying the above three criteria.

A variety of physical systems[34], each with their own advantages and challenges, have been used in attempts to meet the above requirements. It is convenient to classify these into two families: single and multiple emitters.

2.1.1 Single emitters

A single photon source can be built by isolating a single emitter and manipulating its quantum state in a way that creates one photon at a time. In the ideal case, many physical systems could deterministically emit a single photon on demand. In reality, there are a variety of challenges that limit performance.

Single atoms[35–37] and ions[38,39] are perhaps the most obvious choice for a single photon source. Lasers can be used to perturb the atom or ion such that a single photon is emitted with energy corresponding to an electronic transition of the system. This implies only a limited set of emission wavelengths are available, as dictated by the energy level structure of the emitter. Unfortunately, these wavelengths lie far from the near-infrared that would be ideal for compatibility with the lowest-loss optical fibers and photonic chips. Furthermore, such sources are very challenging to build, generally requiring cryogenic and vacuum condi-

tions with an array of cooling and trapping lasers. This complexity hinders construction of identical sources on a large scale. Furthermore, atoms are difficult to trap for long periods of time forcing such sources to be operated in bursts followed by a reset period with no photon emission.

Quantum dots[40–42] and nitrogen vacancy (NV) centers in diamond[43–45] mimic the discrete energy level structure of an atom and are an increasingly popular way to construct single photon sources. Unlike the band structure produced by an effectively spatially infinite bulk semiconductor, excitations in a nano-sized quantum dot experience particle-in-a-box boundary conditions and exhibit corresponding discrete energy levels. The level structure, and corresponding photon wavelength, can be engineered by simply altering the size and shape of the dot. NV centers, where a nitrogen atom and a vacancy substitute for two carbon atoms in a diamond lattice, similarly present an 'atom-like' level structure in a solid-state system. This convenience has a downside; both quantum dots and NV centers are sensitive to their local lattice environment, making it difficult to reliably fabricate identical sources that emit indistinguishable photons. In addition, quantum dots still require cryogenic operating conditions. These factors have impeded widespread adoption of these sources in quantum optics experiments.

2.1.2 Multiple emitters

One can employ a system of multiple emitters that each contribute to the emission of a single photon. These sources are probabilistic; the probability of emitting a single photon during any given trial is necessarily less than one. However, a successful trial is guaranteed to emit a pair of photons. If the two photons are emitted into two different modes, they can

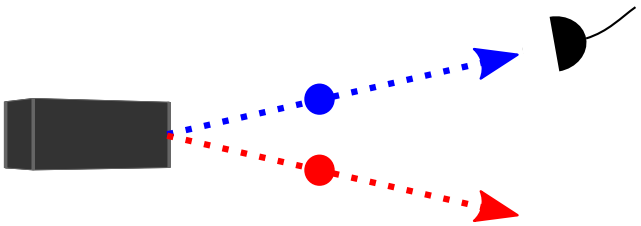


Figure 2.1: Taking advantage of the pairwise emission of photons, one can use a probabilistic emitter as a heralded single photon source. When a detector at the top measures the presence of a single photon (the herald), one knows that another photon is present in the bottom mode (the heralded photon).

be deterministically separated. When a heralding detector monitoring one of the modes clicks, then a photon is known to be in the heralded mode as shown in Figure 2.1.

An ensemble of trapped atoms is one example of a multiple emitter-based single photon source[46–49]. Whenever a pump pulse is applied to the ensemble, there is a small chance of an atomic excitation and release of a single photon. Though this excitation is probabilistic, detection of this first photon heralds when it has occured. A second strong pump pulse can then be applied at a later time to read out a second single photon in a deterministic manner. While such systems were amongst the earliest examples of single photon sources, they retain the significant experimental complexity of trapped single atom sources and have not been implemented on a large scale in single photon experiments.

Nonlinear optics provides another route to heralded single photon production. Simple models of dielectrics assume that the displacement of bound electrons varies linearly in the applied electric field; this is equivalent to the electron sitting in a harmonic potential well. In reality, the electrons experience an anharmonic potential that implies the electron

displacement is nonlinear in the applied field. One can perform the series expansion

$$\begin{aligned} P(t) &= \epsilon_0 \left[\chi^{(1)} E(t) + \chi^{(2)} E(t)^2 + \chi^{(3)} E(t)^3 + \ldots \right] \\ &= P^{(1)}(t) + P^{(2)}(t) + P^{(3)}(t) + \ldots, \end{aligned} \tag{2.1}$$

where $P(t)$ and $E(t)$ are the dipole moment density and applied electric field respectively, $\chi^{(j)}$ is the j-th order susceptibility[50], and we've assumed all fields are copolarized. The series expansion in Equation (2.1) is often truncated after the linear term, with the real and imaginary parts of $\chi^{(1)}$ yielding the index of refraction and absorption coefficient of the dielectric.

One must retain the higher order terms in Equation (2.1) when $E(t)$ has a sufficiently large amplitude. These nonlinear contributions can generate light at frequencies not present in the applied field. Assume a field with two frequencies, $E_{in}(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + \text{c.c.}$, is applied to a dielectric. If that dielectric lacks inversion symmetry, which is the case in many crystals, then its $\chi^{(2)}$ susceptibility will be nonzero. Using Equation (2.1), the second-order term in the polarization density will then be

$$\begin{aligned} P^{(2)}(t) = \epsilon_0 \chi^{(2)} \Big[& |E_1|^2 + |E_2|^2 \\ & + (E_1^2 e^{-2i\omega_1 t} + E_2^2 e^{-2i\omega_2 t} + 2E_1 E_2 e^{-i(\omega_1+\omega_2)t} + 2E_1 E_2^* e^{-i(\omega_1-\omega_2)t} + \text{c.c.}) \Big], \end{aligned} \tag{2.2}$$

where the first line gives a constant offset and the latter terms describe the bound electron's oscillation at the second harmonic frequencies $2\omega_1$ and $2\omega_2$ as well as the sum-difference frequencies $\omega_1 \pm \omega_2$[50]. Since $P(t)$ acts as a source term in the dielectric wave equation,

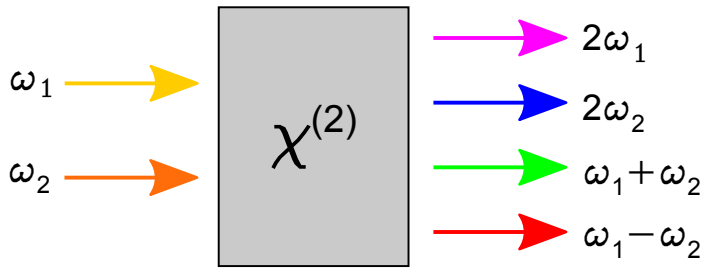


Figure 2.2: Coupling $E_{in}(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t}$+c.c. into a dielectric with a nonzero $\chi^{(2)}$ susceptibility can introduce a variety of new frequency components. Generally, only a subset of these will be efficiently generated due to phasematching constraints.

the output field can acquire these new frequency components as shown in Figure 2.2.

The processes diagrammed in Figure 2.2 are only efficiently generated at specific sets of frequencies due to phasematching constraints. We introduce the concept in terms of second-harmonic generation (SHG), though it applies to all the processes in Figure 2.2 and our SFWM sources. The output second-harmonic field is the coherent superposition of contributions from the many oscillating dipoles along the length of the dielectric. Maximum SHG efficiency occurs when these contributions constructively interfere. Dispersion in the dielectric implies that in general this will not happen; the phase of the field with frequency ω driving a dipole is usually not the same as the phase of the field at 2ω generated by previous dipoles. The field emitted by the many dipoles destructively interfere at the output leading to negligible emission at the second harmonic frequency. However, with careful choice of dielectric and operating frequencies, one can achieve a constant phase relationship between all the fields in the dielectric. Frequencies that satisfy this property, and conservation of energy, are said to be *phasematched*.

The use of $\chi^{(2)}$ materials to generate new frequency components is frequently employed in photonics. For example, second-harmonic and sum-frequency generation can

be used to convert commonly available near-infrared laser sources to ultraviolet and visible wavelengths. Difference frequency generation is the basis of optical parametric amplifiers (OPA), where a strong pump at ω_1 and strong signal at ω_2 produce a new idler at $\omega_1 - \omega_2$, along with a depleted pump and an amplified signal.

What happens if the strong signal field is removed from an OPA such that only the pump is applied to the $\chi^{(2)}$ dielectric? Since the signal mode can be modeled as a quantum harmonic oscillator, its vacuum fluctuations can still stimulate emission of photons at the idler and signal wavelengths in a process known as spontaneous parametric down-conversion (SPDC). In a downconversion event, a pump photon is annihilated while a signal and idler photon are created at frequencies that conserve energy $\omega_p = \omega_s + \omega_i$. While SPDC is very inefficient and occurs randomly, one can take advantage of the pairwise emission to employ the heralding strategy shown in Figure 2.1. When a detector monitoring the heralding mode (either signal or idler) clicks, then the presence of a photon in the other mode is heralded.

Heralded single photon sources based on SPDC have been the workhorses for quantum optics experiments for many years [34]. These SPDC sources frequently involve a bulk crystal that is pumped at approximately 400 nm and produces pairs of photons at 800 nm which are eventually coupled to optical fibers leading to silicon avalanche photodiode (APD) single photon detectors. These sources are commonly used because of their relative simplicity; the crystals are inexpensive and can be pumped at room temperature without vacuum conditions.

Energy and momentum conservation frequently lead to entanglement between multiple spatial and spectral modes of the emitted photon pairs [51,52]. If the herald photon is detected

but its mode is not resolved, then the heralded photon is left in a mixed quantum state of all possible modes. Equation (1.22) shows this is unsuitable for applications requiring high visibility nonclassical interference.

One can restore heralded state purity by placing filters before the detector so that it responds to only a single spatio-spectral mode. This filtering is most obvious in the spatial degree of freedom. Bulk crystal SPDC sources often emit signal and idler into separate cones, a portion of which are coupled into single-mode fibers for experimental use. Spectral filtering is also frequently required to eliminate correlations in the broadband signal and idler emission. These filtering strategies increase heralded state purity at the cost of higher loss. This loss reduces the rate at which single photons are emitted and may also preclude quantum enhanced performance[53–56]; alternative solutions are desirable.

A waveguide on a $\chi^{(2)}$ substrate is an attractive refinement for SPDC sources. Single photon emission can be restricted to a single spatial mode which potentially eliminates the need for lossy spatial filtering. The small mode sizes also imply that pump intensities sufficient for SPDC can be achieved with drastically reduced powers. Furthermore, it can be easier to achieve spectrally factorable emission in an appropriately designed waveguide than in a bulk crystal[57–59]. Loss remains an issue however. Despite the potential for single spatial mode operation, that mode remains poorly matched in size to optical fibers which would introduce significant coupling losses. In addition, waveguide fabrication in many $\chi^{(2)}$ crystals is a fairly new technology and it remains unclear how far scattering losses can be lowered.

Silica waveguides provide a promising route to achieving low-loss. Unfortunately, silica does not possess an appreciable $\chi^{(2)}$ nonlinearity which precludes SPDC. However, one can

construct a heralded single photon source via spontaneous four-wave mixing (SFWM) as an alternative to SPDC. In SFWM, two pump photons are annihilated to produce a signal and idler photon in a $\chi^{(3)}$ substrate such that the frequencies satisfy $\omega_{p1} + \omega_{p2} = \omega_s + \omega_i$ with subscripts indicating the four photons. Silica can have very low scattering loss, making waveguides with spatial modes well matched to optical fibers is straightforward, and there is a simple route to achieving spectrally factorable emission. All of these factors allow for extremely low loss heralding of high purity single photons.

2.2 SFWM source design

This sections begins with a derivation of the two-photon state produced by SFWM. We then use this photon pair state to investigate the design choices required to achieve desired source performance. The source must be low-loss and this leads to our choice of a silica substrate. We want to minimize heralding of more than one photon at a time and, given the noise sources in silica, this leads to a phase-matching scheme that employs a birefringent waveguide. We want our heralded single photons to be in pure states and this is easily achieved in birefringent silica SFWM sources. Finally, we review source efficiency, specifically how to compensate for silica's low $\chi^{(3)}$ to obtain sufficient brightness.

2.2.1 The SFWM two-photon state

One must derive the expected two-photon state $|\psi_2\rangle$ to optimize the design of low-loss heralded sources of high purity single photons. Heralded single-photon generation in the SFWM sources discussed in this thesis rely on the small $P^{(3)}(t)$ term in Equation (2.1). It

is thus convenient to express the Hamiltonian $\hat{H}$ describing the temporal evolution of the system as

$$\hat{H}(t) = \hat{H}_0(t) + \hat{H}_{int}(t), \tag{2.3}$$

where $\hat{H}_{int}(t)$ is a small perturbation due to $P^{(3)}(t)$. Working in the interaction picture, such that only $\hat{H}_{int}(t)$ appears in Equation (1.4), allows one to isolate the effect of this perturbation on the evolution of $|\psi_2\rangle$. Solving Equation (1.4) then gives

$$|\psi\rangle = \hat{T}\text{exp}\left(\frac{1}{i\hbar}\int_0^t dt'\hat{H}_{int}(t')\right)|\psi_0\rangle\,, \tag{2.4}$$

where $\hat{T}$ is the time-ordering operator. In SFWM, the signal and idler modes are initially in the vacuum state $|0\rangle$ and, to minimize the undesired production of two pairs of photons, the interaction strength is deliberately kept low. This allows one to truncate the Taylor series expansion of Equation (2.4), leaving

$$|\psi\rangle \approx \left(1 + \frac{1}{i\hbar}\int_0^t dt'\hat{H}_{int}(t')\right)|0\rangle\,. \tag{2.5}$$

We find $\hat{H}_{int}(t)$ by first evaluating the energy in a dielectric due to its third-order nonlinearity. We use scalar fields throughout this treatment. While the electric fields in our SFWM source will not be copolarized, this only changes which element of the $\chi^{(3)}_{ijkl}$ tensor is appropriate and so the scalar approximation remains accurate. Classically, the energy density in a dielectric is given by $U(\vec{r},t) = \frac{1}{2}E(\vec{r},t)\cdot D(\vec{r},t)$ where the electric displacement is given by $D(\vec{r},t) = \epsilon_0 E(\vec{r},t) + P(\vec{r},t)$. Expanding $P(\vec{r},t)$ as in Equation

(2.1) and finding the contribution to the Hamiltonian from the quantum analogue to the third-order polarization $\hat{P}^{(3)}(\vec{r},t)$ gives

$$\hat{H}_3(t) = \frac{1}{2}\int dV \hat{E}(\vec{r},t)\hat{P}^{(3)}(\vec{r},t) = \frac{\epsilon_0\chi^{(3)}}{2}\int dV \hat{E}(\vec{r},t)^4, \tag{2.6}$$

where $\hat{E}(\vec{r},t)$ is the quantized electric field operator and we assume all fields are linearly polarized. Anticipating our interest in four-wave mixing, we express the total electric field as a superposition of four independent waves

$$\hat{E}(\vec{r},t) = \frac{1}{2}\sum_{m=1}^{4}\left(\hat{E}_m^{(+)}(\vec{r},t) + \hat{E}_m^{(-)}(\vec{r},t)\right) \tag{2.7}$$

where $\hat{E}^{(-)} = \left(\hat{E}^{(+)}\right)^{\dagger}$. Since our sources use single spatial mode waveguides, one can perform a one dimensional quantization to find [30]

$$\hat{E}^{(+)}(\vec{r},t) = i\, g(x,y)\int d\omega \ell(\omega)\exp[-i(\omega t - kz)]\hat{a}(\omega), \tag{2.8}$$

where $g(x,y)$ describes the single transverse spatial mode satisfying the normalization $\iint |g(x,y)|^2 dxdy$=1. Several terms are consolidated into the function

$$\ell(\omega) = \sqrt{\frac{2\hbar\omega}{\epsilon_0 n^2(\omega)}}, \tag{2.9}$$

where $n(\omega)$ indicates that the index of refraction depends on the frequency.

The full expansion of $\hat{H}_3(t)$ is found by inserting Equations (2.7) and (2.8) into Equation (2.6). This results in $8^4 = 4096$ terms of the form $\hat{E}_a^{(+,-)}\hat{E}_b^{(+,-)}\hat{E}_c^{(+,-)}\hat{E}_d^{(+,-)}$, where

each of the subscripts is in the range $1-4$ and the superscripts take one of the two values indicated. Only a few of these terms are of interest here. Our goal is to find the perturbations relevant to four-wave mixing. Therefore, we're only interested in terms where all four fields appear with no repeats.[1] Henceforth, we'll assume two of these fields are from the pump along with the signal and idler single-photon modes leaving permutations of the term $\hat{E}_1^{(+,-)}\hat{E}_2^{(+,-)}\hat{E}_s^{(+,-)}\hat{E}_i^{(+,-)}$. We anticipate that two pump photons will be annihilated to create a signal and idler photon, further restricting the terms of interest to $\hat{E}_1^{(+)}\hat{E}_2^{(+)}\hat{E}_s^{(-)}\hat{E}_i^{(-)}$ (and its Hermitian conjugate). In the expansion of Equation (2.6), this term will occur with all $4! = 24$ orderings of the four field operators. Due to the bosonic commutation relations

$$\begin{aligned}\left[\hat{a}_\omega, \hat{a}_{\omega'}^\dagger\right] &= \delta(\omega-\omega') \\ \left[\hat{a}_\omega, \hat{a}_{\omega'}\right] &= \left[\hat{a}_\omega^\dagger, \hat{a}_{\omega'}^\dagger\right] = 0,\end{aligned} \tag{2.10}$$

one can combine these terms leaving us with the SFWM Hamiltonian

$$\hat{H}_{int}(t) = \frac{3}{4}\epsilon_0\chi^{(3)}\int dV \hat{E}_1^{(+)}(\vec{r},t)\hat{E}_2^{(+)}(\vec{r},t)\hat{E}_s^{(-)}(\vec{r},t)\hat{E}_i^{(-)}(\vec{r},t) + \text{h.c.} \tag{2.11}$$

One can now find the form of the two-photon state due to SFWM by inserting Equation

[1] Our sources employ a degenerate pump; two photons from the same pump laser are annihilated to create a signal and idler photon. Accordingly, we will later assume two of these fields are identical, $\hat{E}_1 = \hat{E}_2 = \hat{E}_p$, but for now we treat all four fields as distinguishable.

(2.11) into the second term of Equation (2.5) to give

$$\begin{aligned}|\psi_2\rangle &= \frac{1}{i\hbar}\int_0^t dt' \hat{H}_{int}(t')\,|0\rangle \\ &= \frac{1}{i\hbar}\frac{3\epsilon_0\chi^{(3)}}{4}\int_0^t dt' \int dV E_1^{(+)}(\vec{r},t)E_2^{(+)}(\vec{r},t)\hat{E}_s^{(-)}(\vec{r},t)\hat{E}_i^{(-)}(\vec{r},t)\,|0\rangle\,, \end{aligned} \tag{2.12}$$

where the operator notation for the pumps is dropped as they are strong fields and it is a good approximation to treat them classically. In addition, the Hermitian conjugate term goes to zero because $\hat{E}_s^{(+)}(\vec{r},t)\hat{E}_i^{(+)}(\vec{r},t)\,|0\rangle \propto \hat{a}_s\hat{a}_i\,|0\rangle = 0$. Our sources will employ a single pump laser, so $E_1(\vec{r},t) = E_2(\vec{r},t) = E_p(\vec{r},t)$, where the positive frequency component is expressed as

$$E_p^{(+)}(\vec{r},t) = A_p g_p(x,y)\int d\omega\alpha(\omega)\exp\{-i[\omega t - k_p(\omega)z]\}, \tag{2.13}$$

where A_p is the pump amplitude, and $\alpha(\omega)$ describes the shape of the spectral envelope with normalization $\int d\omega|\alpha(\omega)|^2 = 1$. Inserting Equations (2.8) and (2.13) into (2.12) then gives

$$\begin{aligned}|\psi_2\rangle = i\frac{3\epsilon_0\chi^{(3)}}{4\hbar}A_p^2\Big\{&\iiiint d\omega_{p1}d\omega_{p2}d\omega_s d\omega_i \int_0^t dt' \exp[i(\omega_s+\omega_i-\omega_{p1}-\omega_{p2})t'] \\ &\int dxdy\, g_{p1}(x,y)g_{p2}(x,y)g_s^*(x,y)g_i^*(x,y)\int_0^L dz\exp[-i(k_s+k_i-k_{p1}-k_{p2})z] \\ &\ell(\omega_s)\ell(\omega_i)\alpha(\omega_{p1})\alpha(\omega_{p2})\hat{a}^\dagger(\omega_s)\hat{a}^\dagger(\omega_i)\,|0\rangle\Big\}, \end{aligned} \tag{2.14}$$

where L is the length of the $\chi^{(3)}$ dielectric.

We use an ultrashort pulsed pump whose temporal envelope $\alpha(t') = \int d\omega\alpha(\omega)\exp[-i\omega t']$

has sub-picosecond duration. Since the transit time of a photon through our experiment is much longer than this, several nanoseconds, we can safely extend the limits of the temporal integral to infinity giving

$$\int_{-\infty}^{\infty} dt' \exp\left[i(\omega_s + \omega_i - \omega_{p1} - \omega_{p2})t'\right] = 2\pi\delta(\omega_s + \omega_i - \omega_{p1} - \omega_{p2}) = 2\pi\delta(\Delta\omega), \tag{2.15}$$

which enforces the energy conservation condition $\omega_{p1} + \omega_{p2} = \omega_s + \omega_i$. The integral along the z axis can be performed explicitly

$$\begin{aligned}\int_0^L dz \exp[-i(k_s + k_i - k_{p1} - k_{p2})z] &= \int_0^L dz \exp\left(-i\Delta k z\right) \\ &= L\exp\left(i\frac{\Delta kL}{2}\right)\mathrm{sinc}\left(\frac{\Delta kL}{2}\right) = L\phi(\omega_s, \omega_i),\end{aligned} \tag{2.16}$$

where $\mathrm{sinc}(x) = \sin(x)/x$ and we have defined the phasematching function

$$\phi(\omega_s, \omega_i) = \exp\left(i\frac{\Delta kL}{2}\right)\mathrm{sinc}\left(\frac{\Delta kL}{2}\right). \tag{2.17}$$

Substituting Equations (2.15)-(2.16) into (2.14) gives our final result for the two-photon state produced by SFWM

$$\begin{aligned}|\psi_2\rangle = &\left[i\frac{6\pi\chi^{(3)}\epsilon_0 L}{4\hbar}A_p^2\ell(\omega_{s0})\ell(\omega_{i0})\iint dxdy g_p(x,y)^2 g_s^*(x,y)g_i^*(x,y)\right] \\ &\iint d\omega_s d\omega_i \int d\omega_p \alpha(\omega_p)\alpha(\omega_s + \omega_i - \omega_p)\phi(\omega_s, \omega_i)\hat{a}^\dagger(\omega_s)\hat{a}^\dagger(\omega_i)\left|0\right\rangle \\ = &\,\zeta \iint d\omega_s d\omega_i F(\omega_s, \omega_i)\hat{a}^\dagger(\omega_s)\hat{a}^\dagger(\omega_i)\left|0\right\rangle,\end{aligned} \tag{2.18}$$

where the slowly varying terms $\ell(\omega_{s,i})$ are taken outside the integral and evaluated at the central signal and idler frequencies $\omega_{s0,i0}$. We have consolidated all the terms in square brackets on the top line of Equation (2.18) into ζ. Finally, the joint spectral amplitude

$$F(\omega_s, \omega_i) = \int d\omega_p \alpha(\omega_p)\alpha(\omega_s + \omega_i - \omega_p)\phi(\omega_s, \omega_i), \tag{2.19}$$

has been introduced which contains information about the degree of spectral entanglement between signal and idler photons. In our experiments, we detect the signal photon to herald the presence of the idler. Since the heralding detector cannot resolve frequency, the marginal state of the idler is obtained by tracing over the signal mode spectrum to obtain

$$\begin{aligned}\hat{\rho}_i &= \xi|\zeta|^2 \,\mathrm{Tr}_s \left[\iiiint d\omega_s d\omega_{s'} d\omega_i d\omega_{i'} F(\omega_s, \omega_i) F^*(\omega_{s'}, \omega_{i'}) \hat{a}^\dagger(\omega_s)\hat{a}^\dagger(\omega_i)\, |0\rangle\langle 0|\, \hat{a}(\omega_{s'})\hat{a}(\omega_{i'})\right] \\ &= \xi|\zeta|^2 \iiint d\omega_s d\omega_i d\omega_{i'} F(\omega_s, \omega_i) F^*(\omega_s, \omega_{i'}) \hat{a}^\dagger(\omega_i)\, |0\rangle\langle 0|\, \hat{a}(\omega_{i'}),\end{aligned} \tag{2.20}$$

where ξ is a normalization constant such that $\mathrm{Tr}(\hat{\rho}_i) = 1$. Equations (2.18) and (2.20) are the central results of this section. They describe the dependence of the two-photon and the marginal heralded state, respectively, on a variety of experimental parameters and are vital to our successful design of a heralded single photon source.

There are four performance criteria that we require each of our sources to individually satisfy. First, we would like our heralded source to be low loss so that, given a herald event, there is very likely an emitted photon. Second, we seek to minimize undesired heralding of more than one photon at a time. Third, these heralded single photons should be in pure states. Finally, it would be desirable for the source to emit photons at a reasonably high

rate. In the following sections, we discuss how to achieve these goals given the forms of the two-photon and heralded states in Equations (2.18) and (2.20). We also introduce the measurements used to verify this performance.

2.2.2 Photon loss

It is important that our source emits truly single photons. For a lossless heralded source, every time a herald photon is detected there will be a photon in the heralded mode. In practice, inevitable losses degrade these photon number correlations and often must be minimized to give any hope of achieving quantum-enhanced performance[8,9,53–56].

We first consider the ideal case where all losses in the heralded (idler) mode are zero. Using a perfectly efficient, but non-frequency resolving, detector on the marginal idler state in Equation (2.20) gives the probability of detecting an idler photon of

$$\begin{aligned} P_1 = \mathrm{Tr}(\hat{\rho}_i \hat{\Pi}_i) &= \xi|\zeta|^2 \iint d\omega_s d\omega_i |F(\omega_s, \omega_i)|^2 \\ &= \mathrm{Tr}(\hat{\rho}_i) = 1. \end{aligned} \tag{2.21}$$

This result is not surprising; since photons are emitted in pairs then, in a lossless system, detection of a herald must mean a detector in the heralded mode will always 'click'.

A realistic model should account for inevitable losses. These are modeled by adding a fictional beam splitter whose reflectivity is equivalent to the loss as shown in Figure 2.3. We can thus transform the operators in Equation 2.20 according to the beam splitter relations in (1.12)

$$\hat{a}^\dagger_{in} = \eta \hat{a}^\dagger_{out} + \sqrt{1-\eta^2}\hat{a}^\dagger_{loss} \tag{2.22}$$

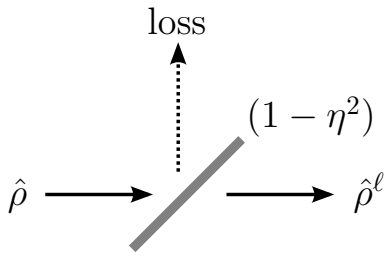


Figure 2.3: Loss can be modeled by a beam splitter whose reflectivity $(1-\eta^2)$ is equivalent to the power loss in the mode. The quantum state after the loss $\hat{\rho}^\ell$ is found by tracing over the inaccessible loss mode.

where the power loss is $(1-\eta^2)$. The resulting state of the idler is then obtained by tracing over the loss mode, yielding

$$\hat{\rho}_i^\ell = (1-\eta^2)\,|0\rangle\langle 0| + \eta^2\hat{\rho}_i, \tag{2.23}$$

which reveals that loss has the effect of mixing in vacuum. The new probability of detecting an idler photon is given by

$$P_1^\ell = \mathrm{Tr}(\hat{\rho}_i^\ell\hat{\Pi}_i) = \eta^2. \tag{2.24}$$

A successful single photon source should minimize the vacuum component in Equation (2.23), which is equivalent to minimizing the loss. While this discussion has been in terms of heralded sources, an analogous argument can be made for single emitters. While these latter sources can be deterministic emitters, high system losses frequently spoil this advantage.

We introduce a metric called the heralding efficiency, η_h, which is equivalent to this conditional probability of detecting a single photon given a herald event. In our experiments, we typically use the signal photon to herald the presence of the idler. The heralding efficiency is then

$$\eta_h = \frac{N_{si}}{N_s} \tag{2.25}$$

where N_s is the number of signal (herald) photons and N_{si} is the number of heralded single photons as measured by temporally coincident signal and idler detection. It is apparent from Equation (2.24) that the heralding efficiency is equivalent to η^2 of the heralded (idler) mode. For a lossless system, we have the relation $\eta^2 = \eta_h = 1$.

Minimizing loss, and thus maximizing η_h, is a critical performance parameter. This loss occurs either through scattering in the dielectric itself or in poor coupling to the optical fibers which are ubiquitous in quantum optics experiments. We use silica waveguides on photonic chips to minimize both contributions. Silica exhibits extremely low scattering loss and the spatial modes from silica chips are naturally well-matched to optical fibers.

2.2.3 Emission of multiple photon pairs

An ideal single photon source would emit exactly one photon at a time. We have already observed how loss degrades this performance by mixing vacuum into the emitted state. However, we also seek to minimize production of more than one photon at a time. There are two main ways in which more than one photon could be produced in our sources: higher-order SFWM emission and spontaneous Raman scattering. We discuss how to mitigate both contributions below.

While the full state emitted by our SFWM source is given by Equation (2.4), we assumed a weak interaction that allowed us to keep only a single term from the Taylor series expansion in Equation (2.5). In this approximation, our output contains at most a single pair of photons whose state is given by Equation (2.18). Inclusion of the higher order terms in Equation (2.4) leads to possible emission of more than one pair of photons at a time. This higher-order emission, combined with loss and the use of detectors that cannot resolve

photon number, can lead to incorrect heralding of more than one photon [60].

We minimize this undesired production of more than one photon by deliberately keeping the strength of the SFWM interaction small. For the sources discussed here, the joint spectral amplitude is approximately factorable such that $F(\omega_s, \omega_i) \approx S(\omega_s)I(\omega_i)$. This simplifies the expression for the higher order terms, allowing one to write the full state

$$|\psi\rangle = \sqrt{1-\lambda^2} \sum_{n=0}^{\infty} \lambda^n (\hat{A}_s^\dagger \hat{A}_i^\dagger)^n |0\rangle , \tag{2.26}$$

where $0 \leq \lambda < 1$ and the wavepacket operators $\hat{A}_{s,i}^\dagger$ include the spectral dependence from Equation (2.18). By deliberately employing a low pump power, we keep λ small such that any herald events most likely arise from emission of a single photon pair with probability $P_{pair} \approx \lambda^2$. We typically operate our sources in the range $\lambda^2 \leq 0.04$.

The production of noise photons, by processes other than SFWM, can also lead to accidental heralding of more than one photon. Spontaneous Raman scattering is frequently the dominant noise source for silica sources [61]. The strong laser used to pump SFWM can also excite phonons in the silica lattice along with a new optical frequency, traditionally called the Stokes beam. Each of these Raman scattering events conserves energy, $\omega_p = \omega_{St} + \Omega$, where ω_{St} and Ω represent the Stokes and phonon frequencies respectively. The Stokes noise can overlap with the idler that is also at a lower frequency relative to the pump. This Raman noise can seriously degrade the quality of the single photon source, as herald events will often correspond to the presence of both a Stokes and idler photon.

These negative effects can be minimized if the idler photon is produced at a frequency where there is minimal Raman scattering. The amorphous composition of silica fibers and

chips contributes to spectrally broad Stokes emission, with a peak 12 THz detuned from the pump. For the pump wavelength of $\lambda_p = 730$ nm used in our experiments, this corresponds to a wavelength shift of $\Delta\lambda = 22$ nm. If the idler photon is produced at a wavelength significantly more detuned from the pump than this, the Raman noise should be minimal.

The phasematching function $\phi(\omega_s, \omega_i)$ in Equation (2.17) determines what signal and idler frequencies will be emitted for a given pump. It is apparent from the form of $\phi(\omega_s, \omega_i)$ that the central frequencies of the signal and idler will occur when $\Delta k{=}2k_p{-}k_s{-}k_i{=}0$. Ideally, we would find a method of phasematching SFWM in a silica waveguide that could achieve significant idler detuning from the pump to minimize Raman noise.

A variety of strategies have been employed to phasematch SFWM. Standard optical fibers have used modulation instability in the anomalous dispersion region of silica[62,63], however this method generally produces significant Raman noise as the idler remains spectrally close to the pump. Furthermore, commonly-available silicon-based single-photon detectors do not work at the longer wavelengths, greater than 1300 nm, where silica exhibits the required anomalous dispersion. Photonic crystal fibers (PCFs), with engineered waveguide dispersion, have circumvented these problems by phasematching SFWM in the visible region and achieving a large idler-pump detuning[64–69]. However, PCFs are not well mode-matched to standard optical fibers and thus increase system losses. In addition, fabrication of long lengths of homogeneous PCF is not a well-developed technology which impedes construction of identical sources[67,68].

Phasematching SFWM in a birefringent waveguide presents an attractive alternative. The signal and idler photons can be phasematched in the visible region where common silicon APDs are most efficient. Noise can be minimized as the idler can lie far beyond

silica's Raman gain peak. Finally, this can be achieved in a 'standard' solid clad waveguide with a spatial mode identical to that of standard fibers and the silica chips used in many quantum optics experiments[23–25,27–29]. At visible wavelengths, where silica exhibits normal dispersion, SFWM is phasematched when the pump is linearly polarized along the slow axis (higher index) while signal and idler are along the fast axis (lower index). The phasematching condition can accordingly be expressed as

$$\Delta k = \frac{2\left[n(\omega_{p0}) + \Delta n)\right]\omega_{p0}}{c} - \frac{n(\omega_{s0})\omega_{s0}}{c} - \frac{n(2\omega_{p0} - \omega_{s0})[2\omega_{p0} - \omega_{s0}]}{c} = 0, \tag{2.27}$$

where we have used conservation of energy to eliminate ω_{i0}, Δn is the birefringence, and $n(\omega_m)$ represents the refractive index of bulk silica of mode m. These indices can be calculated using the empirically-derived Sellmeier equations. The effects of waveguide dispersion are negligible for solid-clad silica structures at visible wavelengths[61] and thus omitted here. Given Δn and ω_p, one can numerically solve Equation (2.27) for ω_s and use energy conservation to deduce ω_i. An example phasematching contour is shown in Figure 2.4. It is apparent that signal and idler detuning from the pump grows with increasing Δn. However, even for the conservative birefringence values used in Figure 2.4, higher values are routinely achieved in polarization maintaining (PM) fibers, it is obvious that the idler can be phasematched well beyond the main Raman band.

We quantify the success with which we have suppressed heralding of more than one photon via the normalized second-order correlation function of the field

$$g^{(2)}(\tau) = \frac{\langle \hat{E}^{(-)}(t)\hat{E}^{(-)}(t+\tau)\hat{E}^{(+)}(t+\tau)\hat{E}^{(+)}(t)\rangle}{\langle \hat{E}^{(-)}(t)\hat{E}^{(+)}(t)\rangle^2}. \tag{2.28}$$

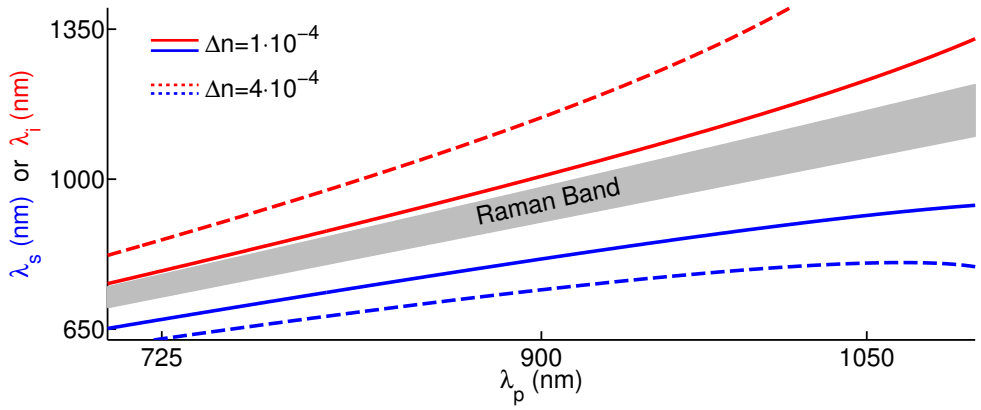


Figure 2.4: Theoretical phasematching curves show the signal (blue) and idler (red) wavelengths that satisfy the condition $\Delta k = 0$ for a range of pump wavelengths. The solid and dashed lines demonstrate that higher birefringence values lead to larger spectral detuning from the pump. The gray band shows the spectral region, with a frequency detuning $\Delta\nu \leq 30$ THz from the pump, that is significantly contaminated by spontaneous Raman noise.

Note that the spatial dependence from Equation (2.8) has been omitted. We assume a constant single transverse spatial mode so that $g(x, y)$ cancels in the ratio in Equation (2.28) and we define our axis so that our detectors are at $z = 0$. The quantity $g^{(2)}(\tau)$ measures how strongly the field intensity at times t and $t + \tau$ are correlated. We will be particularly interested in $g^{(2)}(0)$. We note that the finite bandwidth of all detectors implies that measuring exactly $g^{(2)}(0)$ is infeasible. However, we obtain a good approximation as long as the detector temporal resolution is well below the pulse period. A true single photon source is characterized by $g^{(2)}(0) = 0$ because Equation (2.8) indicates there will be a term of the form $\hat{a}^2 |1\rangle = 0$ in Equation (2.28)[30].

In our single-photon source work, measuring $g^{(2)}(0)$ requires a detector that can resolve up to two photons. While intrinsically number-resolving detectors exist, they are an active area of research[70,71]. We instead measure $g^{(2)}(0)$ by spatially multiplexing together two bucket detectors as shown in Figure 2.5. These detectors do not resolve photon number but instead fire when one or more photons are present. The mode of interest is coupled to a

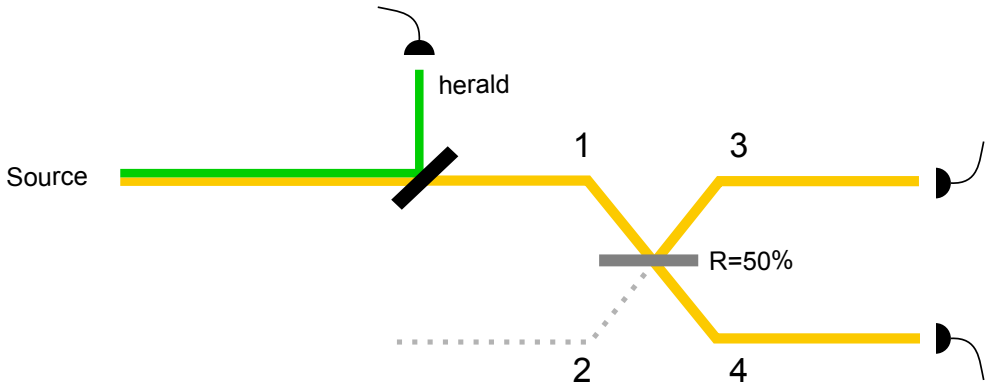


Figure 2.5: One can measure $g^{(2)}(0)$ by coupling the mode of interest (yellow) to a 50% reflectivity beam splitter with bucket detectors (APDs) monitoring both outputs. For heralded single photon sources, all measurements are conditional on also detecting a herald photon (green).

50% reflectivity beam splitter and APD detectors monitor the number statistics of the two outputs. Using Equation (1.12), one can show that the correlation function in Equation (2.28) is equivalent to

$$g^{(2)}(0) = \frac{\langle \hat{n}_3 \hat{n}_4 \rangle}{\langle \hat{n}_3 \rangle \langle \hat{n}_4 \rangle} = \frac{N_{34} N_t}{N_3 N_4}, \tag{2.29}$$

where the mean number of photons in a given mode is equivalent to the ratio of detected photons to the number of trials N_t. For example, the mean number of simultaneous photon events on both outputs is given by $\langle \hat{n}_3 \hat{n}_4 \rangle = N_{34}/N_t$ where N_{34} is the total number of times both detectors register coincident events. For our pulsed sources, N_t is equivalent to the number of pump pulses. Since our single-photon sources are heralded, then all the terms in Equation (2.29) become conditional on detecting a herald photon

$$g_H^{(2)}(0) = \frac{N_{34H} N_H}{N_{3H} N_{4H}}. \tag{2.30}$$

Classical electromagnetic fields, for example originating from thermal sources or lasers, always satisfy $g^{(2)}(0) \geq 1$ [30]. These sources exhibit bunched photon emission; if p_1 is the

probability of observing a single photon at time $t = 0$ then the probability of observing two photons $p_2 \geq p_1^2$. However, an ideal single photon source will never output more than one photon at a time. This means $N_{34} = 0$ in Equation (2.29) and thus $g^{(2)}(0) = 0$ for such a source. In practice, higher photon number terms, possibly originating from emission of multiple photon pairs via SFWM or noise processes like Raman scattering, increase $g^{(2)}(0)$ above this lower bound. However, we will use a $g^{(2)}(0) < 1$ as a clear sign of non-classical emission and will strive to keep this parameter as low as possible to approach ideal single photon emission.

2.2.4 Interference visibility

Ultimately, we want to achieve high visibility interference between the heralded photons from multiple sources. This places strict requirements on the purity of the heralded state from each individual source as well as the homogeneity between the sources.

It is illustrative to consider the archetypical example shown in Figure 2.6 where the heralded emission from two different SFWM sources are interfered on a 50% reflectivity beam splitter. In this example, the signal photons act as heralds and the idlers are interfered. Four non-frequency resolving detectors monitor all four spatial modes. What is the probability that all four detectors fire (P_4) at the same time? In the ideal case of indistinguishable and pure heralded photons from both sources, we anticipate $P_4 = 0$ due to the perfect HOM interference described in Equation (1.16).

We use the SFWM state in Equation (2.18) to investigate how to achieve this ideal performance. The four photon state produced when each SFWM source emits a single pair

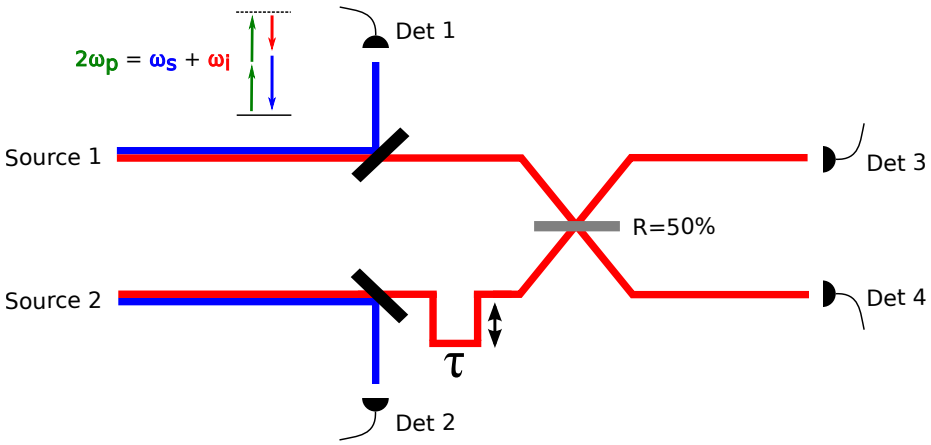


Figure 2.6: The heralded single photons from two SFWM sources are interfered on a 50% reflectivity beam splitter. Signal photons act as heralds while the idler photons are interfered. In the ideal case of indistinguishable and pure heralded photons from both sources, the probability that all four detectors fire simultaneously should be $P_4 = 0$ due to HOM interference.

of photons is given by

$$|\psi_4\rangle = \zeta_1\zeta_2 \iiiint d\omega_{s_1} d\omega_{i_1} d\omega_{s_2} d\omega_{i_2} \\ F_1(\omega_{s_1}, \omega_{i_1}) F_2(\omega_{s_2}, \omega_{i_2}) \hat{a}_1^\dagger(\omega_{s_1}) \hat{a}_1^\dagger(\omega_{i_1}) \hat{a}_2^\dagger(\omega_{s_2}) \hat{a}_2^\dagger(\omega_{i_2}) |0\rangle , \tag{2.31}$$

where $\hat{a}_i^\dagger(\omega_j)$ operates on spatial mode i and frequency ω_j and the subscripts $s_{1,2}$ and $i_{1,2}$ indicate signal and idler modes for the two sources. Detection of all four photons is described by the operator

$$\hat{\Pi}_4(t_1, t_2, t_3, t_4) = \hat{E}_1^{(-)}(t_1) \hat{E}_2^{(-)}(t_2) \hat{E}_3^{(-)}(t_3) \hat{E}_4^{(-)}(t_4) |0\rangle \\ \langle 0| \hat{E}_1^{(+)}(t_1) \hat{E}_2^{(+)}(t_2) \hat{E}_3^{(+)}(t_3) \hat{E}_4^{(+)}(t_4), \tag{2.32}$$

where the four detector modes are arranged as shown in Figure 2.6 and we have assumed all the waveguides are single spatial mode and thus only consider the temporal degree of

freedom. One can use the beam splitter relations in Equation (1.12) to express all operators in Equation (2.32) in terms of the modes before the beam splitter. This yields

$$\hat{E}^{(+)}_{1,2}(t_{1,2}) = \int d\omega_{s_1,s_2} \hat{a}_{1,2}(\omega_{s_1,s_2}) e^{-i\omega_{s_1,s_2} t_{1,2}} \tag{2.33}$$

$$\hat{E}^{(+)}_{3,4}(t_{3,4}) = \frac{1}{\sqrt{2}} \int d\omega_{i_3,i_4} \left[\hat{a}_1(\omega_{i_3,i_4}) e^{-i\omega_{i_3,i_4}\tau} \pm \hat{a}_2(\omega_{i_3,i_4})\right] e^{-i\omega_{i_3,i_4} t_{3,4}}, \tag{2.34}$$

where Equation (2.33) gives the field operators for the herald modes and (2.34) represents the idler fields backpropagated through the balanced beam splitter. Using Equations (2.31)-(2.34), we can calculate the probability of a four-fold coincidence at specified times t_1,t_2,t_3,t_4

$$P_4(t_1,t_2,t_3,t_4) = \langle\psi_4| \hat{\Pi}_4(t_1,t_2,t_3,t_4) |\psi_4\rangle . \tag{2.35}$$

However, our detectors have a slow response compared to the duration of the optical pulses and so we integrate over all possible detection times to find [72]

$$P_4 = \left(\frac{\zeta_1\zeta_2}{2}\right)^2 \iiiint d\omega_{s_1} d\omega_{i_1} d\omega_{s_2} d\omega_{i_2}$$
$$|F_1(\omega_{s_1},\omega_{i_2}) F_2(\omega_{s_2},\omega_{i_1}) - F_1(\omega_{s_1},\omega_{i_1}) F_2(\omega_{s_2},\omega_{i_2}) e^{-i(\omega_{i_2}-\omega_{i_1})\tau}|^2. \tag{2.36}$$

For temporally coincident photons ($\tau = 0$), perfect HOM interference requires $P_4 = 0$. Using Equation (2.36), this implies that

$$F_1(\omega_{s_1},\omega_{i_2}) F_2(\omega_{s_2},\omega_{i_1}) = F_1(\omega_{s_1},\omega_{i_1}) F_2(\omega_{s_2},\omega_{i_2}). \tag{2.37}$$

Equation (2.37) is only satisfied for all possible frequencies if two conditions are met. First,

the two sources must be identical so that

$$F_1(\omega_s, \omega_i) = F_2(\omega_s, \omega_i) = F(\omega_s, \omega_i). \tag{2.38}$$

Furthermore, the sources must be spectrally factorable so that

$$F(\omega_s, \omega_i) = S(\omega_s)I(\omega_i). \tag{2.39}$$

This latter condition of spectral factorability implies that the signal and idler are not spectrally entangled, so the use of a non-frequency resolving heralding detector does not leave the idler in a mixed state.

We seek to build identical, factorable sources that enable this high-visibility interference. The construction of identical sources is facilitated by mature silica waveguide fabrication methods. Therefore, we focus here on how to achieve spectral factorability given the form of the joint spectral amplitude in Equation (2.19). The integral in this equation does not, in general, have an analytic solution. We therefore simplify the analysis by expanding the phase mismatch Δk in a first-order Taylor series about the central phase-matched frequencies. This expansion can be expressed as

$$\begin{aligned}\Delta k_{lin} &= (2k_p^0 - k_s^0 - k_i^0) + \left[2k_p^{(1)}(\omega_p - \omega_p^0) - k_s^{(1)}(\omega_s - \omega_s^0) - k_i^{(1)}(\omega_i - \omega_i^0)\right] \\ &= \nu_s\kappa_s + \nu_i\kappa_i,\end{aligned} \tag{2.40}$$

where $\nu_m = \omega_m - \omega_m^0$ represents the detuning from the central frequency ω_m^0 and $\kappa_m = k_p^{(1)} - k_m^{(1)}$ with $k^{(1)} = dk/d\omega$. The pumps used in our experiments are approximated with

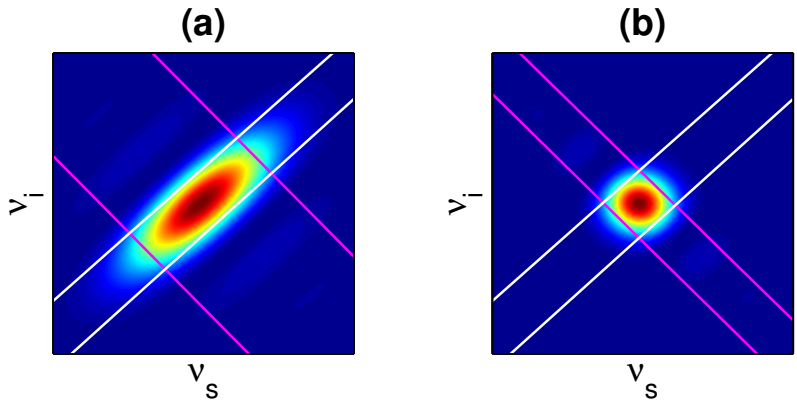


Figure 2.7: Two example joint spectral amplitude $F(\nu_s, \nu_i)$ plots show (a) spectral correlations between signal and idler, and (b) an uncorrelated (factorable) state. The half-maximum contours of the pump and phasematching contributions are indicated by the purple and white lines respectively. The phasematching function has been approximated as a Gaussian.

a Gaussian spectral envelope

$$\alpha(\omega) = \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{\sigma_p}} \exp\left[-\frac{(\omega - \omega_p^0)^2}{\sigma_p^2}\right], \tag{2.41}$$

where σ_p is the bandwidth. Note that Equation (2.41) satisfies the required normalization $\int d\omega |\alpha(\omega)|^2 = 1$. Substituting Equation (2.40) and (2.41) into (2.19) yields

$$F(\nu_s, \nu_i) \propto e^{\frac{iL}{2}(\nu_s\kappa_s + \nu_i\kappa_i)} \text{sinc}\left[\frac{L}{2}(\nu_s\kappa_s + \nu_i\kappa_i)\right] e^{-\frac{(\nu_s+\nu_i)^2}{2\sigma_p^2}}. \tag{2.42}$$

Figure 2.7 shows a plot of Equation (2.42) and illustrates how correlations between the signal and idler can arise. The convolved pump $e^{-(\nu_s+\nu_i)^2/2\sigma_p^2}$ will always lie along a $-45°$ line in the joint spectral space due to energy conservation, and will have a width determined by σ_p. The peak of the phasematching function is given by the condition $\nu_s\kappa_s + \nu_i\kappa_i = 0$ and is thus determined by the dispersive properties of the dielectric. The width of the phasematching function is proportional to $1/L$ which becomes obvious after approximating

the sinc function as a Gaussian

$$\text{sinc}\left[\frac{L}{2}\left(\nu_s\kappa_s+\nu_i\kappa_i\right)\right]\approx e^{-\gamma\left[\frac{L}{2}\left(\nu_s\kappa_s+\nu_i\kappa_i\right)\right]^2}, \tag{2.43}$$

where setting $\gamma = 0.193$ ensures the Gaussian and main sinc peak have equivalent FWHM[72]. The product of representative pump and phasematching functions then leads to the 'tilted' joint spectral amplitude example shown in Figure 2.7(a). It is clear that the major and minor axes of the ellipse describing the biphotonic state are not parallel with the $\omega_{s,i}$ axes. Therefore, the joint spectral amplitude cannot be written in the factorable form shown in Equation (2.39) and would lead to heralding of photons in mixed states.

One can produce spectrally factorable photon pairs through appropriate source engineering. Using the Gaussian approximation in Equation (2.43), entanglement is introduced by $\nu_s\nu_i$ terms in the exponent of Equation (2.42). These terms vanish if

$$\frac{1}{\sigma_p^2}+\frac{\gamma L^2}{2}\kappa_s\kappa_i=0. \tag{2.44}$$

There are two possible ways to satisfy this factorability condition. The first method is best illustrated if we re-write Equation (2.44) as $1/(L\kappa_{s,i}\sigma_p)+\gamma L\sigma_p\kappa_{i,s}/2=0$, where each term is dimensionless. If either the signal or idler photon is group velocity matched to the pump then the second term is zero. By also employing a large pump bandwidth, so that $1/(L\kappa_{s,i}\sigma_p)\ll 1$, the emission will be nearly factorable. While this strategy has been widely employed in SFWM sources, the group-velocity matching condition often requires custom PCF designs that increase system loss and are difficult to repeatably manufacture[67–69].

The second way to achieve spectral factorability is to have the two terms in Equation (2.44) cancel one another [73]. Since our source strategy relies on silica photonic chips, $\kappa_{s,i}$ are not easily tunable and so we adjust σ_p and L so that they satisfy the relation

$$\sigma L = \sqrt{-\frac{2}{\gamma \kappa_s \kappa_i}}, \tag{2.45}$$

which implies the group velocity of the pump is in between that of the signal and idler so that either κ_s or κ_i is negative. When this condition is satisfied, one obtains a factorable joint spectral amplitude resembling Figure 2.7(b).

We use Equation (2.45) to estimate what type of pump is required to achieve spectral factorability in a silica chip source. We operate at visible wavelengths to take advantage of readily-available silicon APDs, the single photon modes are spectrally detuned by > 50 nm from the pump to mitigate Raman noise, and our photonic chips are typically a few centimeters long. We use this information and Equation (2.45) to estimate that σ must be several nanometers. Thus, our spectrally factorable silica-chip SFWM sources require ultrashort pulsed pump lasers.

Even when the factorability condition in Equation (2.45) is satisfied, our sources will only emit approximately factorable states. This is because the true phasematching function does not have exactly the Gaussian dependence shown in Equation (2.43). Accurate joint spectral amplitudes, originating from the true sinc dependence of the phasematching function, are shown in Figure 2.8. The source in Figure 2.8(b) still satisfies Equation (2.44), however the secondary lobes in the phasematching function spoil perfect factorability. While there are strategies to restore factorable emission [74,75], they introduce extra

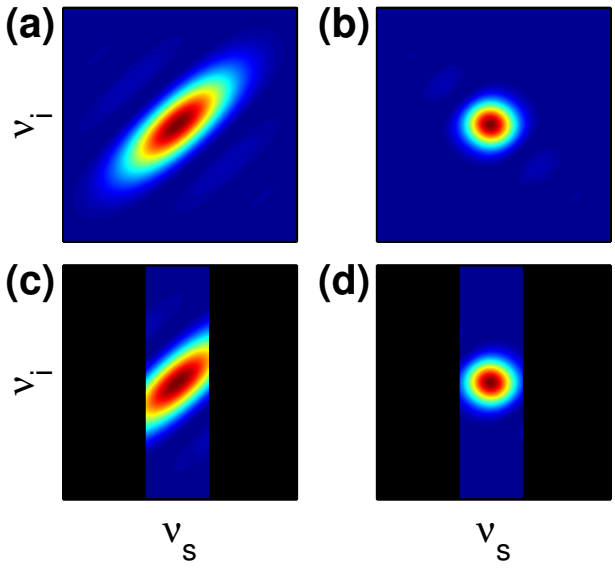


Figure 2.8: Joint spectral amplitude $F(\nu_s, \nu_i)$ plots show (a) spectral correlations between signal and idler, and (b) a nearly uncorrelated state. The sinc dependence of the phasematching function leads to secondary lobes along a $-45°$ line that makes perfect factorability difficult to achieve. However, the plot in (d) shows that filtering out the side lobes (black) on the heralding signal mode can restore near-perfect factorability while still transmitting the majority of photons. The same filter on the entangled state (c) transmits a smaller fraction of the two-photon state with poorer results; the filtered state still clearly exhibits spectral correlations.

experimental complexity and so we simply use spectral filters on the herald mode. Since the central portion of the joint spectral amplitude is factorable, Figure 2.8(d) shows that we only need to filter the minor side lobes. This filtering has no effect on the heralding efficiency and only causes a small reduction in count rate. If our source is not engineered to be spectrally factorable then a much larger fraction of the biphoton state must be discarded to achieve high heralded purity. This situation is shown in Figure 2.8(d).

We need a method to verify that a source is spectrally factorable and thus heralds high purity photons. At first glance, it isn't clear how one might achieve this. One could measure the joint spectral intensity, $|F(\omega_s, \omega_i)|^2$, with a frequency-resolving single photon detector[58,76,77]. Unfortunately, this measurement is insensitive to possible correlations in

the phase of $F(\omega_s, \omega_i)$. One could perform a two-source HOM experiment and look for a reduction in the interference visibility. However, this method can only bound the spectral factorability as any reduction in visibility could be due to either spectral entanglement or source distinguishability. In addition, it also requires a two-source experiment even if we're only interested in the purity of either source in isolation.

Surprisingly, one can use the photon number statistics from a single heralded source to quantify the degree of correlations in its joint spectral amplitude. This is explained by first re-writing the SFWM Hamiltonian in Equation (2.3) in terms of a minimal set of broadband modes via the Schmidt decomposition [78]

$$\hat{H}_{int} = \sum_{m=0}^{\infty} c_m \hat{A}_m^\dagger \hat{B}_m^\dagger + \text{h.c.}, \tag{2.46}$$

where $\hat{A}_m^\dagger = \int d\omega_s \xi_m(\omega_s)\hat{a}^\dagger(\omega_s)$, $\hat{B}_m^\dagger = \int d\omega_i \chi_m(\omega_i)\hat{a}^\dagger(\omega_i)$, and the Schmidt coefficients c_m are both real and satisfy the normalization $\sum_m c_m = 1$. The set of functions ξ_m and χ_m are called Schmidt modes, which define an orthonormal basis for the signal and idler Hilbert spaces respectively. This decomposition shows that the evolution due to SFWM is equivalent to an ensemble of independent two-mode squeezing operations $e^{i\hat{H}_{int}} = \bigotimes_{m=0}^{\infty} \hat{S}_m$ where $\hat{S}_m = e^{\eta_m \hat{A}_m^\dagger \hat{B}_m^\dagger + \eta_m^* \hat{A}_m \hat{B}_m}$ is the two-mode squeezing operator on modes ξ_m and χ_m [76]. These squeezing operators then produce the state [79]

$$|\psi\rangle = \prod_{m=0}^{\infty} \hat{S}_m |0\rangle = \bigotimes_{m=0}^{\infty} \sqrt{1-\lambda_m^2} \sum_{n_m=0}^{\infty} \frac{(\lambda_m \hat{A}_m^\dagger \hat{B}_m^\dagger)^{n_m}}{n_m!} |0\rangle . \tag{2.47}$$

If our source can be described by a single two-mode squeezing operation then Equation

(2.47) simplifies to

$$|\psi_0\rangle = \hat{S}_0 |0\rangle = \sqrt{1-\lambda^2} \sum_{n=0}^{\infty} \frac{(\lambda \hat{A}_0^\dagger \hat{B}_0^\dagger)^n}{n!} |0\rangle . \tag{2.48}$$

In this single squeezer case, the $n = 1$ term that is of interest for heralded single photon sources is clearly spectrally factorable. Thus, enginering a spectrally factorable pair source is equivalent to constructing a single two-mode squeezer. The number statistics of $|\psi_0\rangle$ indicate when this goal has been achieved. The marginal state of either mode is $\hat{\rho}_{s,i}$=$(1 - |\lambda|^2) \sum_n \lambda^{2n} |n_{s,i}\rangle\langle n_{s,i}|$. This yields a thermal number distribution $P_n = (1-|\lambda|^2)\lambda^{2n}$ which, using Equation (2.29), gives $g^{(2)}(0) = 2$.

If multiple squeezers are excited, then Equation (2.47) reveals that the single photon pair state takes the form $\left(\hat{A}_0^\dagger \hat{B}_0^\dagger + ... + \hat{A}_N^\dagger \hat{B}_N^\dagger\right) |0\rangle$, where we have ignored multiplicative factors. This state is spectrally entangled over N Schmidt mode pairs. These correlations between signal and idler alter the marginal number statistics. This multimode source outputs N independent thermal distributions, corresponding to the N two-mode squeezers. However, our detectors cannot resolve Schmidt modes and observe overall number statistics that are the convolution of these many thermal sources. In the limit $N \to \infty$, it is known that this leads to a Poissonian distribution with corresponding $g^{(2)}(0) = 1$ [80].

The marginal photon number statistics thus allow one to discriminate between spectrally factorable and maximally entangled states. These two cases correspond to heralded single photon purities of $\mathcal{P} = 1$ and 0 respectively. However, in reality all of our observed states will lie in between these two extremes. In the weak interaction limit appropriate for our

sources, where the probability of multi-pair emission is small, one finds

$$g^{(2)}(0) \approx 1 + \sum_m |c_m|^4 \approx 1 + \mathcal{P}_{s,i}, \tag{2.49}$$

where $\mathcal{P}_{s,i}$ is the heralded purity of the signal or idler mode[81,82]. Thus, the marginal number statistics give a direct measurement of the purity of the heralded state. In the ideal limit of a perfectly factorable source that involves only a single squeezer, $c_0 = 1$ and thus $g^{(2)}(0) = 2$. In the other limit of $N \to \infty$ squeezing operations, $c_m \to 0$ and $g^{(2)}(0) = 1$. Since $g^{(2)}(0)$ is straightforward to measure, this will prove to be a very convenient way to experimentally confirm the success of our spectral engineering efforts.

The singular value decomposition (SVD) is an efficient numerical method to calculate the Schmidt decomposition and thus the heralded purity for a modeled joint spectral amplitude. We can numerically express the joint spectral amplitude $F(\omega_s, \omega_i)$ as a matrix $\mathbf{F}$. The element F_{mn} gives the value $F(\omega_{i,m}, \omega_{s,n})$, where $\omega_{i,m}$ and $\omega_{s,n}$ are numerical values for the idler and signal frequencies. Each row (column) of $\mathbf{F}$ thus represents a constant idler (signal) frequency, as shown in Figures 2.7-2.8. The SVD decomposes $\mathbf{F}$ into the product

$$\mathbf{F} = \mathbf{UDV}^\dagger. \tag{2.50}$$

Both $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices that describe the Schmidt modes; the m^{th} column of $\mathbf{U}$ represents $\chi_m(\omega_i)$, while the n^{th} row of $\mathbf{V}^\dagger$ describes $\xi_n(\omega_s)$. The diagonal matrix $\mathbf{D}$ contains the Schmidt coefficients c_m. Multiplying the three together sums over the Schmidt mode pairs with accurate weightings, yielding the correct joint spectral amplitude.

Since the SVD is easily evaluated numerically, we regularly use it as a way to predict the heralded state purity. Using Equation (2.19), we first model what the joint spectral amplitude would be for a particular pump spectrum and chip length. We then take the SVD of this modeled joint spectral amplitude and use the resulting Schmidt coefficients and Equation (2.49) to predict $\mathcal{P}_{s,i}$ and bound the HOM interference visibility.

2.2.5 Conversion efficiency

It is desirable for our heralded source to emit photons at a high rate. In this section we derive an expression for the efficiency of the SFWM process which shows how this goal can be achieved. Due to the number correlations between signal and idler, it is sufficient to consider just one of these modes when calculating the conversion efficiency; here we focus on the idler mode. Furthermore, our goal here is to express how the conversion efficiency varies with certain experimental parameters, not calculate an exact numerical value[83]. This allows us to considerably simplify the analysis by ignoring all constants and slowly varying terms.

We wish to calculate the SFWM conversion efficiency

$$\eta = \frac{\langle \hat{n}_i \rangle}{N_p}, \tag{2.51}$$

where $\hat{n}_i$ is the number operator for the idler mode and N_p is the number of pump photons per pulse. Using Equation (2.20), we find

$$\langle \hat{n}_i \rangle \propto \left(\frac{\chi^{(3)} L A_p^2}{A_{\text{eff}}} \right)^2 \iint d\omega_s d\omega_i \left| F(\omega_s, \omega_i) \right|^2, \tag{2.52}$$

where we have ignored all constant factors and have introduced the effective area

$$A_{\text{eff}} = \frac{1}{\iint dxdy\, g_p(x,y)^2 g_s^*(x,y) g_i^*(x,y)}, \tag{2.53}$$

that accounts for both the size and overlap of the four transverse spatial modes. We note that our expression for the conversion efficiency will only be valid in the weak interaction limit because we have used Equation (2.20) which resulted from truncating the Taylor series expansion of the Hamiltonian in Equation (2.5).

It is convenient to re-express the pump amplitudes A_p found in Equation (2.52) in terms of the number of pump photons N_p. For a pulsed laser, N_p is proportional to the pulse energy U_p which can be expressed in terms of A_p using Equation (2.13). We accordingly find

$$A_p^2 \propto \int dV \int dt |E_p^{(+)} + E_p^{(-)}|^2 \propto U_p \propto N_p. \tag{2.54}$$

We also wish to investigate the effect of the pump's spectral-temporal characteristics, in particular the bandwidth σ_p, on SFWM efficiency. Equations (2.19) and (2.41) reveal that σ_p appears four times in Equation (2.52). To explicitly reveal the σ_p dependence, we define a new joint spectral amplitude

$$f(\omega_s, \omega_i) = \left(\frac{\pi\sigma_p^2}{2}\right)^{1/2} F(\omega_s, \omega_i), \tag{2.55}$$

that only contains exponentials and the sinc function without any prefactors.

We use Equations (2.54) and (2.55) to find the final expression for the SFWM efficiency.

We first re-express (2.52)

$$\langle \hat{n}_i \rangle \propto \left(\frac{\chi^{(3)} L N_p}{A_{\text{eff}} \sigma_p} \right)^2 \iint d\omega_s d\omega_i \left| f(\omega_s, \omega_i) \right|^2 . \tag{2.56}$$

Inserting this expectation value into Equation (2.51) gives

$$\eta = \frac{\langle \hat{n}_i \rangle}{N_p} \propto \frac{N_p}{\sigma_p^2} \left(\frac{\chi^{(3)} L}{A_{\text{eff}}} \right)^2 \iint d\omega_s d\omega_i |f(\omega_s, \omega_i)|^2. \tag{2.57}$$

Approximating the phasematching function as a Gaussian, we use Equations (2.42) and (2.43) to analytically solve the integral in Equation (2.57). This gives our final result for the SFWM conversion efficiency

$$\eta \propto \frac{N_p L \sigma_p}{|k_i^{(1)} - k_s^{(1)}|} \left(\frac{\chi^{(3)}}{A_{\text{eff}}} \right)^2 . \tag{2.58}$$

Silica has a low $\chi^{(3)}$, which impedes construction of sufficiently efficient sources. This problem is compounded by our use of a birefringent phase-matching scheme in two ways. First, to minimize Raman noise we try to maximize the pump-idler detuning which frequently increases $\left| k_i^{(1)} - k_s^{(1)} \right|$. Secondly, $\chi^{(3)}$ is actually a tensor $\chi^{(3)}_{ijkl}$ where the subscripts indicate the polarizations of the four fields in Equation (2.12). Our birefringent phasematching strategy, where the pump is orthogonally polarized to the signal and idler, uses the $\chi^{(3)}_{xxyy}$ element. A co-polarized scheme in PCF could access the larger nonlinearity $\chi^{(3)}_{xxxx} = 3\chi^{(3)}_{xxyy}$ [61]. Therefore, the birefringent phasematching approach offers significantly reduced Raman noise but at the cost of a lower conversion efficiency.

We overcome these efficiency difficulties by maximizing the pump intensity inside the

source. This is accomplished by using a high-average power pump so that N_p is large. Furthermore, this pump is focused into a single spatial-mode waveguide so that A_{eff} is small. These design choices allow us to operate bright SFWM sources in a silica substrate.

If we wish to maintain high heralded state purity, then we cannot increase source brightness by separately increasing σ_p or L. This is evident from Equation (2.45) which shows that the product $L\sigma_p$ is constrained to a constant value defined by the group velocities of the modes.

Chapter 3

Demonstration of on-chip SFWM sources

We seek to build low-loss heralded sources of high-purity single photons. According to theoretical arguments in the previous chapter, a promising approach is to use SFWM in a single-mode silica waveguide. A wide variety of SFWM sources in optical fibers have been constructed, however fabrication repeatability problems with PCF [67–69] or sensitivity to the local environment with standard solid-clad fibers [73,82,84,85] have limited the scaling of this approach. We use photonic chips, rather than fibers, to facilitate the construction of many identical sources. Furthermore, the integration on-chip facilitates source interfacing with the silica circuits that have formed the basis of many recent quantum optics experiments. [23–29]

We used two different fabrication methods to build these sources. Femtosecond laser writing of waveguides successfully demonstrated low-loss and high heralded purity from a single source. However, it was difficult to achieve repeatable fabrication results and this precluded construction of multiple identical sources. A second method, UV-writing in a silica-on-silicon substrate, demonstrated greater repeatability and was employed to fabricate matched sources. This latter work culminated in HOM interference experiments

using heralded photons originating from two sources on the same chip.

3.1 Femtosecond-laser written SFWM source

3.1.1 Femtosecond-laser fabrication of waveguides

Femtosecond laser pulses focused into a silica substrate can cause a permanent modification to the glass that corresponds to an increase in the refractive index that is localized to the focal spot[86]. By translating a chip through the focus, one can fabricate a waveguide as shown in Figure 3.1. While investigating the causes of this material modification remains an active research area[87], the popularity of this fabrication method has greatly increased in recent years. This is largely due to the convenience of rapid prototyping, enabled by short fabrication times without the need for cleanroom facilities. In addition, this method is well-suited to making 3-D circuit geometries since it does not rely on lithography. These benefits have led to the use of femtosecond laser writing to construct a wide variety of photonic circuits[26–28,77].

The characteristics of the femtosecond laser writing pulses, in particular the pulse repetition frequency (PRF), have a significant impact on fabrication results. With a high PRF, particularly >100 kHz, thermal effects at the focus generally lead to a circularly symmetric waveguide cross-section regardless of the focal shape[87,88]. This approach has proven popular since a circularly symmetric core is often desired and this can be achieved without the added complexity of beam-shaping optics. If a low PRF ($\approx$ 1 kHz) is used, then the fabricated waveguide is similar in cross-section to the focal volume. This can be an attractive feature since it allows control of the spatial mode shape; one is not limited to

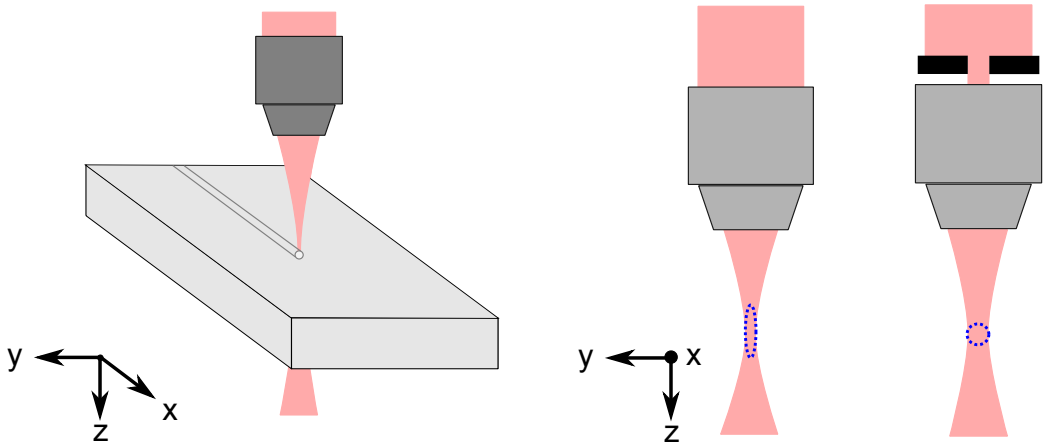


Figure 3.1: A femtosecond laser can permanently increase the index of refraction of a substrate. The modified region is localized to the focus, allowing a waveguide to be fabricated by translating the sample (left). Focusing a Gaussian beam through a typical objective leads to an elliptical focal volume where the blue dashes indicate a contour of constant writing laser intensity (center). Since material modification only occurs above a certain intensity threshold, the blue dashed lines also indicate the resulting waveguide cross section. Inserting a slit before the objective leads to a broader beam waist along the y-axis, allowing waveguides with circular cross sections to be fabricated (right).

circular cores.

We use a low PRF writing laser to take advantage of this added flexibility, though beam-shaping is generally required. Since the fabricated region matches the focal shape, then the waveguide cross section in the yz plane of Figure 3.1 produced by a focused Gaussian beam will be elliptical with axes defined by the Rayleigh range and beam waist. The Rayleigh range is given by $z_R = \frac{\pi w_0^2}{\lambda}$, where w_0 is the radius of the beam waist. Achieving a circular cross-section, a useful fabrication benchmark even if we eventually wish to fabricate other shapes, then requires $z_R = w_0$. Using the common femtosecond laser wavelength $\lambda = 800$ nm, this condition implies $w_0 = z_R = 254$ nm. Achieving this small a focal spot would be difficult and, even if achieved, would be smaller than desired to make a single mode well-matched to standard optical fibers. Typical focusing conditions instead lead to $z_R > w_0$, which leads to the highly elliptical focal volumes shown in the middle of Figure 3.1.

One can shape the beam to produce a circular focus by inserting a slit before the focusing objective. The fields in the front and back focal planes of the objective are related by the Fourier transform. The right portion of Figure 3.1 then shows that the smaller spatial extent of the beam before the objective in the y direction leads to a broadening of the waist so that circular cross sections can be achieved. This simple beam shaping approach has been successfully implemented with physical slits in the way depicted in Figure 3.1 [89].

We use a spatial light modulator (SLM) to reproduce this slit beam shaping effect. By replacing a mechanical slit with this adaptive element, one adds the capability to change the focal shape during fabrication to create a variety of more elaborate structures. Figure 3.2 shows the specific femtosecond laser writing apparatus used to fabricate the SFWM sources discussed here [90]. A regeneratively amplified titanium sapphire laser was attenuated using

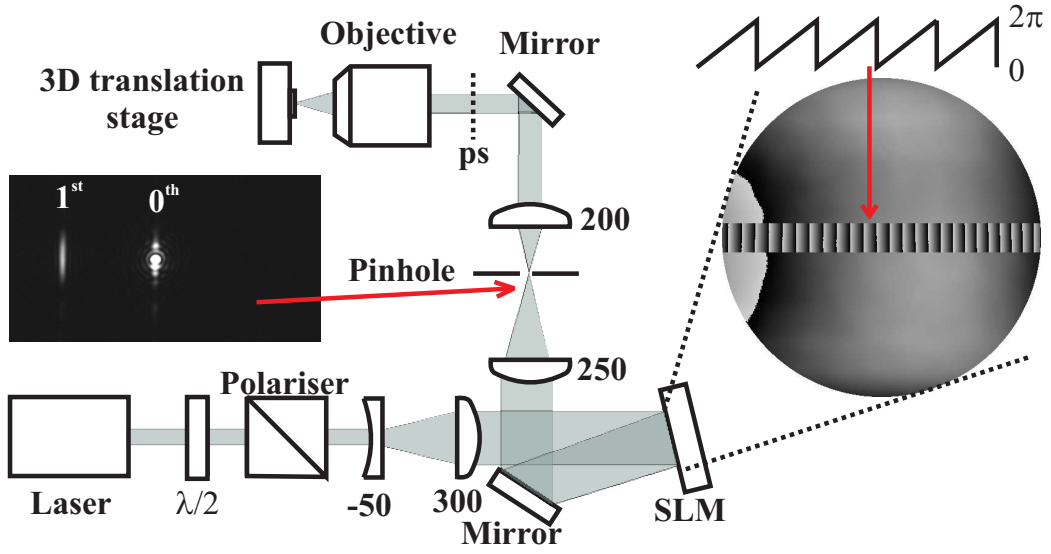


Figure 3.2: Experimental setup for the adaptive slit beam shaping. The inset on the right shows the phase pattern displayed on the SLM, while the left shows the zeroth and first diffracted orders of the beam in the pinhole plane.

a rotatable half-wave plate and polarizer. The beam was then expanded to fill a reflective phase-only SLM. The right inset shows the phase pattern displayed on the SLM; a blazed

slit across the center is added to a background phase profile that corrects for any system aberrations including the flatness of the SLM. Two lenses are used in a $4f$ arrangement to image the SLM onto the entrance face of the objective. A pinhole is placed at the intervening Fourier plane that transmits the first diffracted order from the blazed grating on the SLM but blocks the zeroth order. The area of the blazed region thus defines the size of the beam-shaping slit. A silica chip is placed on a three-axis translation stage (Aerotech ABL10100, ANT95-3-V) and translated through the shaped focus to fabricate a waveguide.

The efficacy of this fabrication method is benchmarked by writing low-loss waveguides whose spatial modes are well-matched to optical fibers. Waveguides were written with the SLM beam-shaping method described above and also with a physical slit (PS) of equivalent size inserted just before the focusing objective in Figure 3.2. The spatial modes were imaged with a beam-profiling CCD, and the fabricated regions inside the substrate were probed with a third-harmonic generation (THG) microscope[91,92]. This latter method is sensitive to variations in $\chi^{(3)}$ which typically coincide with regions of varying refractive index. The results are shown in Figure 3.3 and demonstrate that SLM beam-shaping can shape the writing beam similarly to the PS. Both produce circularly symmetric guides that are very well-matched to standard optical fibers. In addition, the propagation loss for these guides was < 0.4 dB/cm, which compares favorably with other femtosecond laser writing results[28,89,93–95].

We require a birefringent waveguide to phasematch SFWM with a large pump-idler detuning to minimize Raman noise. This birefringence is necessarily achieved by introducing some asymmetry into the fabricated region. Many fibers achieve this by varying the composition of the glass across the fiber to create a stress which, through the strain-optic

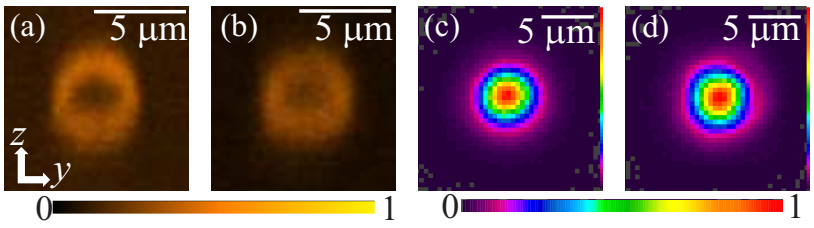


Figure 3.3: THG images showing cross sections through waveguides written in fused silica using (a) SLM and (b) physical slit beam shaping. (c)-(d) Near field profiles of the guided mode associated with (a) and (b) for an 825 nm source.

effect, translates into a birefringence [96]. Since we employ femtosecond writing in a uniform silica substrate, relying on inhomogeneity in the bulk dielectric in this way is not an option.

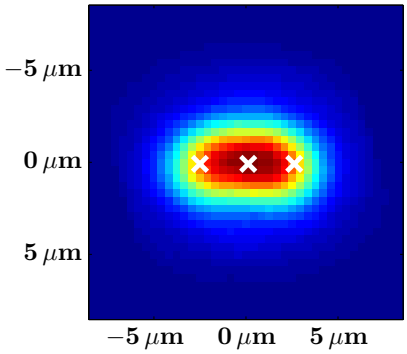


Figure 3.4: Near field image of the spatial mode from an elliptical waveguide fabricated from three femtosecond laser fabrication runs, each separated by 2.5 μm shift in-between (white marks).

We achieve a high birefringence by fabricating an elliptical core. This can be achieved by using the SLM slit to shape the writing focus into an ellipse. However, we achieved the lowest loss by using a circular focus and making multiple writing passes with a transverse shift of the writing beam between each pass. The SFWM sources presented here used the same beam shaping that led to the waveguide in Figure 3.3. Three fabrication runs are performed, laterally separated by 2.5 microns, which leads to the elliptical mode shown in Figure 3.4.

3.1.2 Observation of SFWM

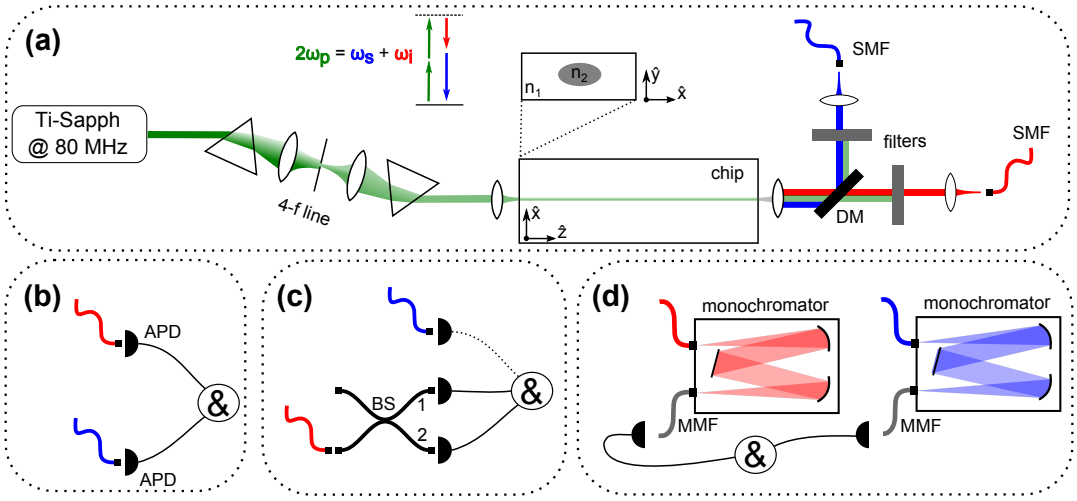


Figure 3.5: The experimental layout consists of (a) state preparation and (b-d) characterization. (a) A 4-f line with an adjustable slit is used to control the bandwidth of the pump ($\Delta\lambda_p$). After the chip, the pump is filtered out while the signal and idler fields are separated by a dichroic mirror (DM) and coupled into separate single-mode fibers (SMFs). (b) Source count rate and cross-correlation $g^{(2)}_{si}(0)$ are measured with the SMFs coupled directly into APDs. (c) The autocorrelation $g^{(2)}_{ii}(0)$ is measured by connecting the idler fiber to a fiber beam splitter (BS) with a reflectivity of 50% and ignoring the signal field, while the signal field is used as a herald when determining $g^{(2)}_{H}(0)$. We measure $g^{(2)}_{ss}(0)$ by inserting the signal fiber into the BS and ignoring the idler field (not shown). (d) To measure joint spectral intensities, the source SMFs are connected to separate monochromators with multi-mode fibers (MMFs) at the output performing a raster scan over the joint spectral range while APDs count in coincidence.

We operate our SFWM sources using the configuration shown in Figure 3.5(a). A single ultrashort pulsed pump laser with central wavelength of $\lambda_p = 729$ nm is generated by a Ti-Sapphire oscillator (Spectra-Physics Mai Tai) whose spectral bandwidth $\Delta\lambda_p$ is adjusted with a mechanical slit in a $4f$ line external to the laser cavity. A dichroic beam splitter (Semrock FF740) separates the signal and idler photon modes while spectral filters (Semrock 684/24 and 800/12) remove the pump. The signal and idler are coupled into separate single-mode fibers which are connected to a variety of detector configurations as shown in Figure 3.5(b)-(d).

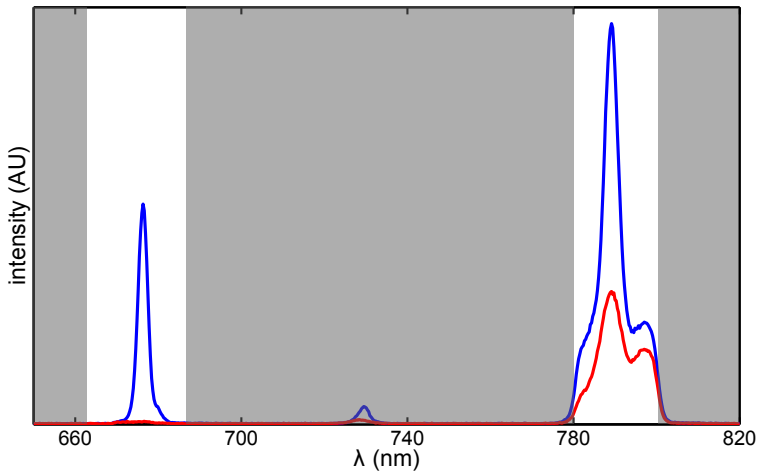


Figure 3.6: A single photon spectrum shows the presence of the signal and idler at 676 nm and 790 nm respectively. The pump field, centered at 729 nm, has its polarization oriented along the slow axis of the waveguide to phasematch SFWM (blue). Background measurements are taken by rotating the pump polarization to the fast axis (red). Spectral filters block frequencies lying in the gray shaded area.

We first measure the signal and idler modes using a single-photon spectrometer (Andor SR163 and DV420A), with the results shown in Figure 3.6. The signal and idler are clearly visible at 676 nm and 790 nm respectively. The Raman scattering noise overlapping the idler is clearly visible, though it has been greatly reduced due to the significant spectral detuning from the pump.

Ideally, one would independently measure the index of refraction and birefringence of the source over the spectral region of interest. These parameters and Equation (2.27) then allow calculation of the expected phasematched signal and idler wavelengths. Experimental limitations, particularly the determination of the refractive index of our short samples, made this analysis infeasible. We instead assume the index of bulk fused silica and infer a birefringence, using Equation (2.27) and the phasematched wavelengths in Figure 3.6, of $\Delta n = 1{\cdot}10^{-4}$.

3.1.3 Low-loss source of single photons

We confirm the low-loss operation of our source by measuring the heralding efficiency η_H given in Equation (2.25). We calculate $\eta_H > 40\%$ for pump powers of 75 to 150 mW as seen in Figure 3.7. The measured η_H is limited primarily by the detector efficiency, which can be readily improved [70]. We thus estimate the preparation efficiency η_P, which corresponds to the probability that the idler photon arrives at its detector given detection of the heralding signal photon. Due to the difficulty of precisely measuring absolute detector efficiencies, we follow the practice [76,97] of using the manufacturer's specifications which results in an estimated $\eta_P = 80\%$ for our source. The remaining inefficiency is due to loss in the heralded idler mode; losses in the signal mode reduce the count rate but do not affect η_P.

We quantify the suppression of undesired higher order photon emission by measuring $g_H^{(2)}(0)$, the heralded second order coherence of the idler mode, using the arrangement shown in Figure 3.5. We calculate

$$g_H^{(2)}(0) = \frac{N_{i_1 i_2 s} N_s}{N_{i_1 s} N_{i_2 s}}, \tag{3.1}$$

where the number of two-and three-fold coincident detection events between the signal and the two arms of the idler after the beam splitter are given by $N_{i_1 s}$,$N_{i_2 s}$, and $N_{i_1 i_2 s}$. An ideal single-photon source would give $g_H^{(2)}(0) = 0$. We measure $g_H^{(2)}(0) = 0.0092 \pm 0.0004$ for a pump power of 25 mW. Spontaneous Raman scattering is the primary noise source, which explains the linear increase in $g_H^{(2)}(0)$ as the source is pumped harder. Our low $g_H^{(2)}(0)$ values are comparable to other SFWM sources that take advantage of birefringent phase-matching [73,98].

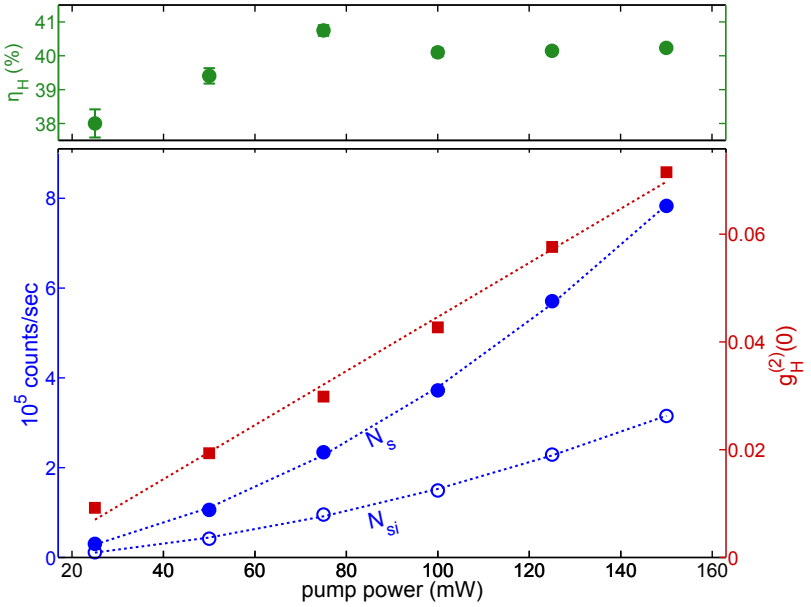


Figure 3.7: The count rate for the signal mode alone (N_s) and in temporal coincidence with the idler mode (N_{si}) is shown as a function of the pump power (blue circles). The heralding efficiency, which includes APD detector losses, is calculated according to $\eta_H = N_{si}/N_s$ (green). At low pump powers, η_H decreases due to the heightened importance of detector dark counts, which cause false herald events. The heralded autocorrelation of the idler photon with the signal field as herald, $g_H^{(2)}(0)$, increases linearly due to spontaneous Raman scattering (red squares). Quadratic (blue) and linear (red) fits to the data are shown with dotted lines. Error bars are smaller than the markers for the count data and $g_H^{(2)}(0)$.

3.1.4 Heralded-state purity

Low-loss heralding of single photons is not sufficient for many applications; these photons must also be in pure states. Undesired entanglement between the signal and idler fields that is not resolved by the heralding detector will result in heralding of mixed states and reduced interference visibility as shown in Equation (1.22). For our sources, these correlations appear in the spectral degree of freedom and are conveniently measured via the number statistics. Section 2.2.4 showed that the spectrally factorable sources we wish to operate are distinguished by an autocorrelation that is thermal, while undesired entanglement leads to increasingly Poissonian statistics. Equation (2.44) indicates that we can control the level

of factorability by adjusting $\Delta\lambda_p$.

We demonstrate this control by measuring $g_{ss}^{(2)}(0)$ as $\Delta\lambda_p$ is adjusted to produce the results shown in Figure 3.8. For $\Delta\lambda_p = 3.1$ nm, we measure our optimal $g_{ss}^{(2)}(0) = 1.86 \pm 0.02$ which is close to the ideal thermal result $g^{(2)}(0) = 2$. Since the source count rate decreases for shorter pump bandwidths, the detector dark counts become increasingly important in these configurations. We accordingly subtract these false events from the data used to produce Figure 3.8.

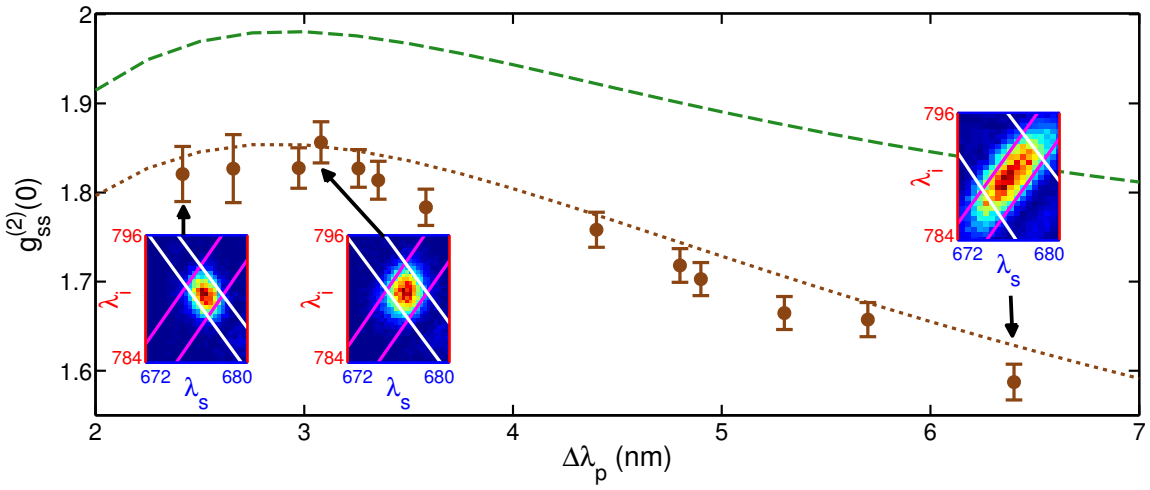


Figure 3.8: Unheralded $g_{ss}^{(2)}(0)$ of the signal field shows control over the amount of spectral entanglement between the signal and idler fields as $\Delta\lambda_p$ is adjusted (brown). Theoretical curve (dotted line) and data (points) are shown. Filtering out the peripheral lobes in the joint spectral intensity (JSI) with a 4.5 nm filter on the signal field is calculated to give $g_{ss}^{(2)}(0) = 1.98$ at $\Delta\lambda_p = 3$ nm (green dashed line). Insets: JSI measurements demonstrate spectral entanglement control. The FWHM of the pump envelope $|\alpha|^4$ (white) and phase-matching function $|\phi|^2$ (purple) accurately predict JSI orientation.

The near single-mode emission of our source is further supported by joint spectral intensity measurements. A spectrally uncorrelated state has a factorable joint spectral amplitude, and thus a factorable joint spectral intensity. In the middle inset of Figure 3.8, we find that the major and minor axes of the central ellipse are parallel to the $\lambda_{i,s}$ axes

which shows a factorable intensity $|F(\lambda_i, \lambda_s)|^2 = I(\lambda_i)S(\lambda_s)$. As $\Delta\lambda_p$ is adjusted away from this optimal value, the joint spectral intensities in the left and right insets of Figure 3.8 indicate entanglement with an increasingly tilted central lobe. These spectral correlations are in agreement with the measured decrease in $g_{ss}^{(2)}(0)$.

At the optimal $\Delta\lambda_p$, the slight deviation from ideal thermal statistics, and corresponding reduction in heralded state purity, is principally due to peripheral lobes in the phasematching function. These arise from the hard-edge boundaries of the waveguide which lead to a sinc function in Equation (2.17). Figure 2.8 showed that a filter on the herald photon, which removes only these small lobes, achieves a significant increase in heralded purity for only a minor decrease in count rate. For our femtosecond-laser written source, a 4.5 nm filter on λ_s would yield $g_{ss}^{(2)}(0) = 1.98$, and corresponding $P = 0.98$, while still passing 90% of photons. Such a filter leaves η_P unchanged, as it would only be applied to the heralding signal photon.

Our source offers large spectral tunability of the signal and idler fields. Figure 3.9 shows that SFWM is phase-matched over a wide range of pump wavelengths. The inset figures illustrate that over this entire range $\Delta\lambda_p$ can be adjusted to achieve factorable output. For larger λ_p the group velocities of the pump and idler photons become comparable, which results in a large spectral bandwidth for the idler. This relatively simple route to broadband emission could be useful, for example in quantum optical coherence tomography [99]. In contrast, many applications, such as quantum memories, frequently require narrowband emission [100]. A theory for cavity-enhanced SFWM has been developed that allows emission down to MHz bandwidths [101]. Cavities can be implemented using the refined Bragg grating technology available in silica [102,103].

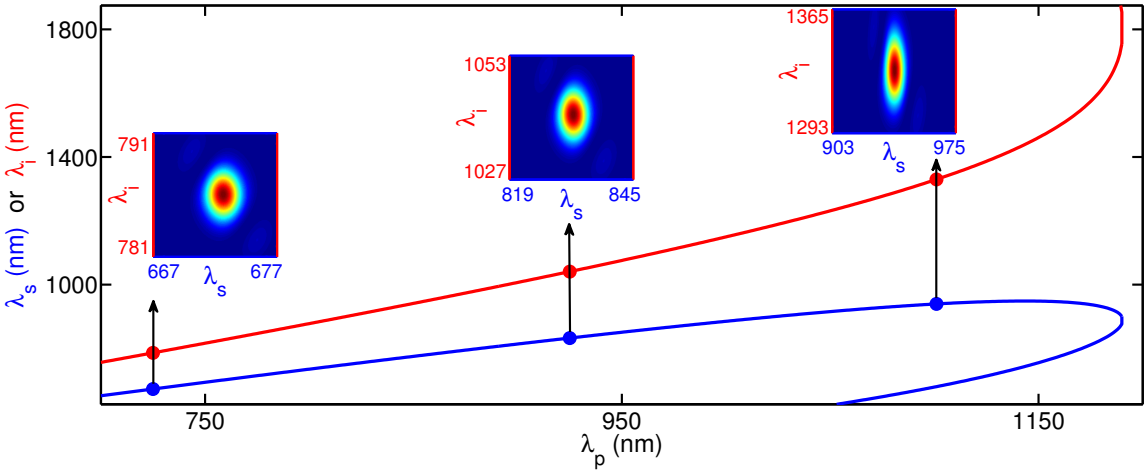


Figure 3.9: Theoretical phasematching curves show the signal (blue) and idler (red) wavelengths that satisfy the condition $\Delta k = 0$ for a range of pump wavelengths. Insets show predicted SFWM joint spectral intensities. A factorable output suitable for heralded production of pure photons can be produced over a wide spectral range, from visible to telecom wavelengths, by adjusting λ_p and $\Delta\lambda_p$ (left to right: $\Delta\lambda_p = 3.1$, 7.5, 20 nm). These calculations assume $L = 4$ cm and $\Delta n = 10^{-4}$.

3.1.5 Single-source homogeneity

Our analysis so far assumes the waveguide, and thus the wavevectors $k_{p,s,i}$, are constant throughout the source. Variation of these wavevectors along the length of the guide can diminish the purity of heralded photons. Due to our operation of a solid-clad guide far from its zero dispersion wavelength along with the excellent material uniformity of commercial silica[98,104], the dominant source of inhomogeneity is believed to be the birefringence. The spectral output is extremely sensitive to variation in the birefringence since the phase-matched wavelengths depend critically on this parameter[73,82].

We consider two models of birefringence inhomogeneity and their effect on heralded state purity in Figure 3.10. The effects of gradual changes during fabrication, for example variation in the writing laser power or local environment, are modeled as a birefringence that varies linearly along the length of the guide. Rapid fluctuations, on the other hand,

are described as a random variation in the birefringence of a specified mean and standard deviation. For both models, the resulting phase-matching function is found by numerically integrating $\phi(\omega_s, \omega_i) = \int_0^L e^{iz\Delta k(z)} dz$, where $\Delta k(z)$ has an explicit spatial dependence. The corresponding joint spectral amplitude is then used to determine both the heralded state purity via the Schmidt decomposition and the joint spectral intensities in the insets of Figure 3.10. These simple models suggest our measured purity (gray band, Figure 3.10) corresponds to a birefringence inhomogeneity of $\delta_{max}(\Delta n) \leq 3 \cdot 10^{-6}$.

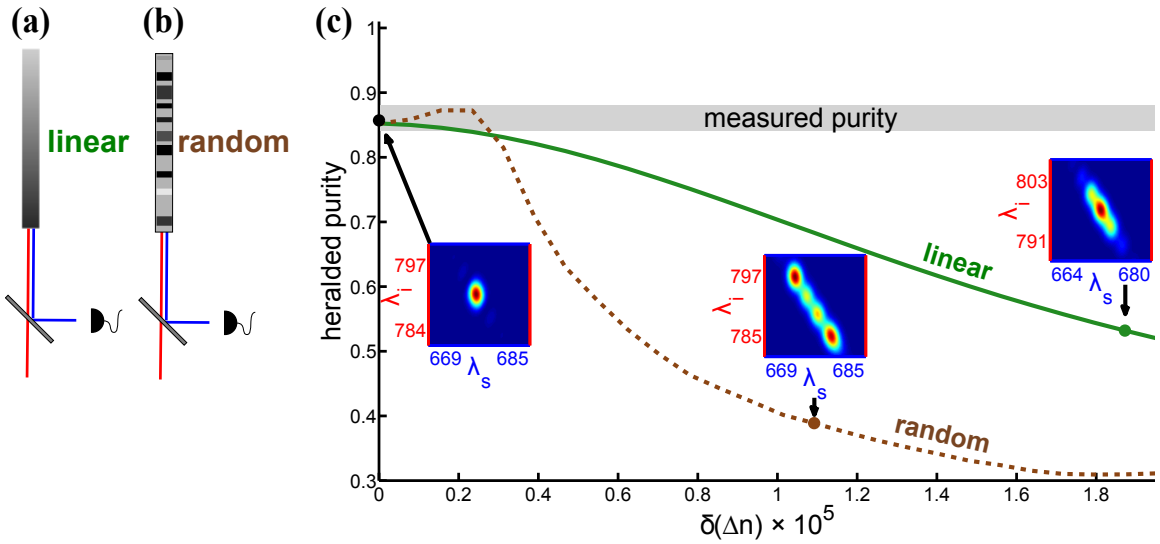


Figure 3.10: The heralded state purity from two models of birefringence variation are shown. (a) Gradual variation in the waveguide is approximated by a linear dependence $\Delta n(z) = \Delta n_0 + \delta(\Delta n) \times z/L$. (b) Rapid variation is modelled by randomly fluctuating birefringence of mean $\langle \Delta n \rangle = \Delta n_0$ and standard deviation $\delta(\Delta n)$. Random inhomogeneity increases the heralded purity for small $\delta(\Delta n)$ due to suppression of the peripheral lobes in the joint spectral amplitude. All results assume $L = 4$ cm and $\Delta n_0 = 10^{-4}$. Insets show representative joint spectral intensities. The gray shaded region indicates the heralded purity of our source, $P = 0.86 \pm 0.02$.

Equation (1.22) showed that the HOM interference visibility obtained from two heralded photons from identical sources is equal to the photon purities. Undesirable birefringence variations that reduce heralded purity, as shown in Figure 3.10, will therefore reduce interference visibility. In the ideal case of $\delta(\Delta n) = 0$, using a 4.5 nm filter on the heralding

signal photon increases the heralded purity from 0.86 (left side of Figure 3.10) to 0.98 while still transmitting 90% of photons. For an inhomogeneity of $\delta(\Delta n) = 3 \cdot 10^{-6}$, the heralded purity, and thus interference visibility, remains high at 0.95 while the filter transmission is now 84%. Accounting for source inhomogeneity, current fabrication methods thus appear sufficient to obtain high visibility interference from heralded sources.

3.1.6 Difficulty achieving multi-source homogeneity

While a single SFWM source in a femtosecond laser written waveguide was demonstrated, efforts to fabricate multiple identical sources proved difficult. This difficulty in achieving repeatable fabrication results can possibly be explained by the highly nonlinear nature of the femtosecond laser writing process[87]. Small fluctuations in the writing parameters, for example the laser power, could then cause significant changes in the fabricated waveguide. Demonstrating uniformity in the phasematched wavelengths across multiple SFWM sources proved prohibitively difficult. Even more fundamentally, consistently demonstrating low optical loss was also challenging.

Figure 3.11 shows the three different strategies that were employed to introduce asymmetry, and potentially significant birefringence, into the fabricated waveguide. This asymmetry can be introduced in the guided mode or in the surrounding silica to introduce stress. The first approach, and the one that successfully produced the birefringent source discussed above, was to fabricate an elliptical waveguide. Secondly, circular waveguides were fabricated close to the top surface of the silica chip where the air-silica interface introduces an asymmetric boundary condition. Finally, a circular waveguide can be fabricated in proximity to intentionally damaged regions of silica.

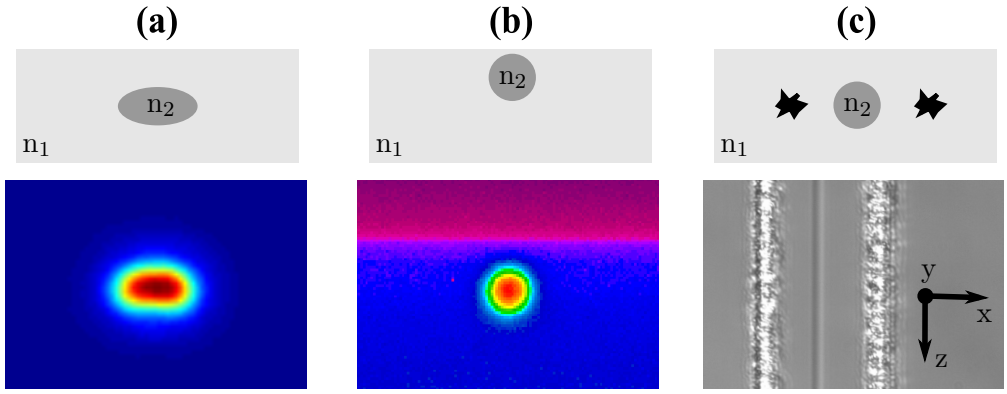


Figure 3.11: Three different methods were used in an attempt to create a birefringent waveguide with the femtosecond laser writing technique. (a) One can fabricate an elliptical core, achieved here with multiple shifted passes of the writing laser, that leads to an elliptical spatial mode (bottom). (b) A circular core can be fabricated very close to the top surface of the silica chip. Coupling a laser to the waveguide shows that it is circular, while white light illumination of the entire chip allows the top surface of the chip to be seen. (c) The silica chip can be intentionally damaged in close proximity to the waveguide. A microscope image reveals the two obvious damage regions with a faint waveguide running in-between. The coordinate system for the bottom right subfigure is indicated; all other subfigures are in the xy plane.

Only the elliptical waveguides proved successful in achieving a birefringence which led to observable SFWM. Three waveguides were birefringent enough to phasematch SFWM sufficiently detuned from the pump to allow measurement of the signal and idler modes. The single-photon spectra for these three sources are shown in Figure 3.12. The spectra from these three sources are clearly distinguishable, which is unsurprising since the three sources were fabricated with different combinations of the fabrication strategies shown in Figure 3.11. The goal was to fabricate all three sources with identical elliptical cores. However, one of the sources also was fabricated close to damage regions (green), one was also fabricated close to the top surface of the chip (red), while the source characterized above only had the elliptical core (blue).

Over a hundred waveguides were fabricated using the strategies illustrated in Figure 3.11 in an attempt to make more SFWM sources. Only the three elliptical-core sources shown

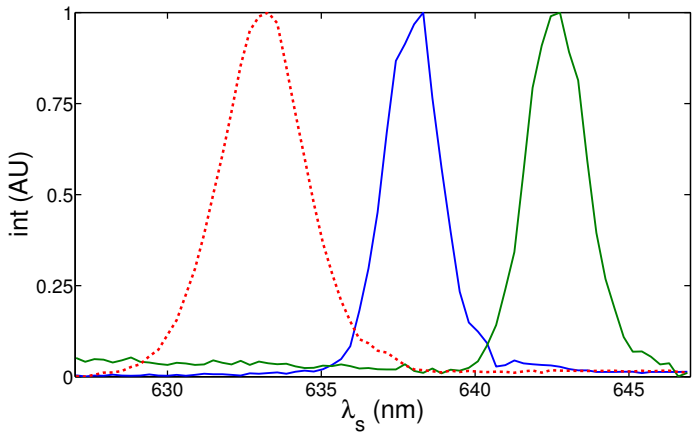


Figure 3.12: The spectra of the signal mode is shown for the three SFWM sources successfully made with the femtosecond laser writing technique. All have elliptical cores and were fabricated with multiple passes of a transversely-shifted writing beam. Two of the sources (blue, green) are pumped at $\lambda_p = 690$ nm, while technical difficulties forced the third to be pumped at 701 nm. To compare it with the other sources, we use this data to infer its spectra (red dotted line) for a 690 nm pump.

in Figure 3.12 succeeded, and these were written on the same day and in the same chip. The source near damaged regions (green) and near the top surface (red) are only slightly detuned from the source that only has the elliptical core. This observation that fabricating near the surface or close to damaged regions only introduces a slight perturbation to the birefringence perhaps explains why these approaches fail when used with a circular core.

The common trait the three SFWM sources share is an elliptical core. To investigate the origin and magnitude of the birefringence, we model asymmetrical core waveguides in silica using the commercial numerical mode solver FIMMWAVE, which is based on film mode matching[105]. The results in Figure 3.13 indicate that the birefringence from an asymmetrical core alone, $\Delta n \approx 3{\cdot}10^{-6}$, is far too low to account for the inferred birefringence of our source, $\Delta n = 1{\cdot}10^{-4}$. This model only accounts for the optical properties shown in the insets of Figure 3.13, not stress-induced birefringence. Therefore, we hypothesize that our larger-than-expected birefringence is due to such a stress when an elliptical core is

fabricated. Surprisingly, follow-on attempts to fabricate more elliptical core SFWM sources did not succeed. Imaging of the spatial modes confirmed modal ellipticity, but no signal or idler were observed in spectra.

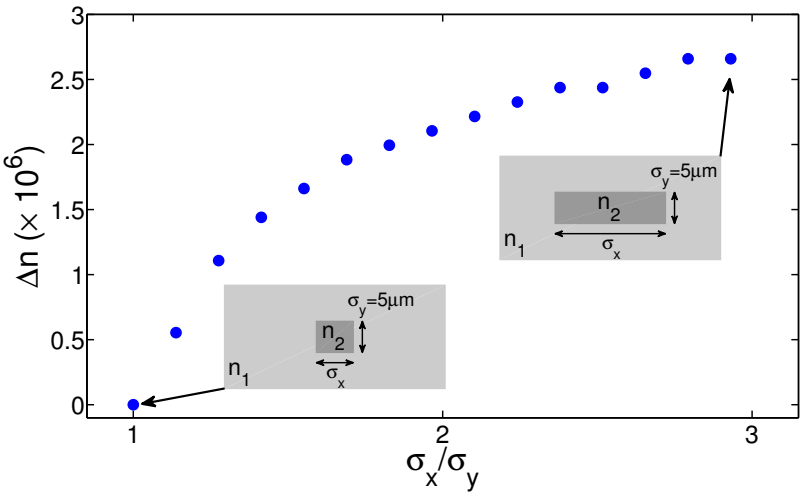


Figure 3.13: The birefringence of a waveguide is modeled as the core becomes increasingly rectangular. The height of the waveguide is kept constant at $\sigma_y = 5\,\mu$m, while the width is scanned from 5 to 15 μm. The core-clad index contrast and wavelength are $n_2 - n_1 = 4.5 \cdot 10^{-3}$ and $\lambda = 730$ nm throughout. While the birefringence does increase as the core becomes increasingly asymmetric, it still remains well below the inferred value of $\Delta n = 1 \cdot 10^{-4}$ for the SFWM source characterized here.

After encountering these erratic fabrication results and with no clear route to quickly improving repeatability, femtosecond laser writing of SFWM sources was suspended. A key obstacle was the absence of a method to quantify the quality of fabrication *in situ*. Waveguides were fabricated and qualitatively judged by visual inspection of a microscope image. Unfortunately, optical performance parameters, such as loss and birefringence, did not correlate well with observations of these microscope images. A newly-added ability to perform refractive-index profiling while writing a waveguide [106] may remedy this issue with repeatability in future work.

3.2 Silica-on-silicon UV-written SFWM sources

There are alternate silica waveguide fabrication techniques that offer the potential for more repeatable results. Waveguide formation via exposure to UV light has been successfully used to produce a variety of photonic circuits while still avoiding the need for expensive lithography[21,24,25]. Figure 3.14 illustrates how a localized index change is created by focusing a UV laser into a photosensitive silica layer. By translating the chip through the focus, a photonic circuit can be fabricated within this layer.

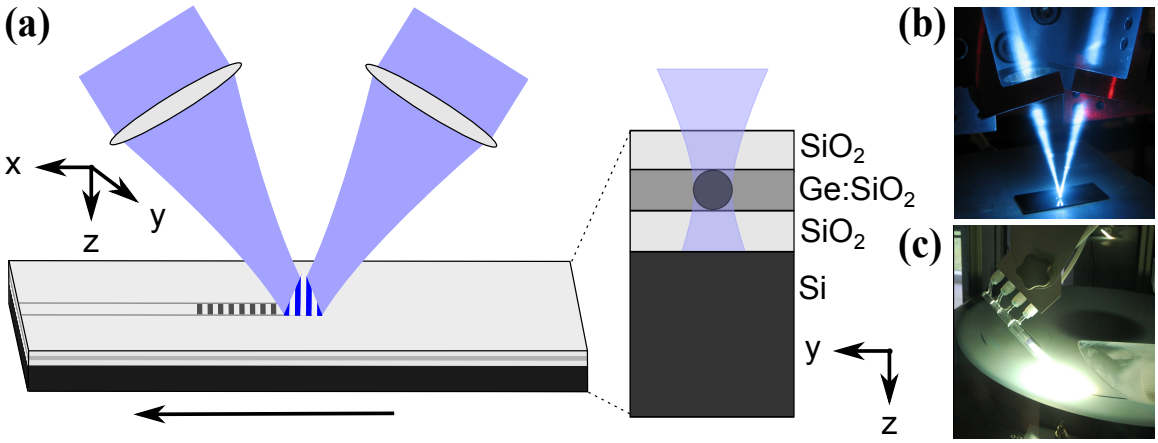


Figure 3.14: (a) Waveguides are fabricated by focusing a continuous-wave UV laser ($\lambda = 244$ nm) into a silica-on-silicon wafer. Dopants are added to the middle silica layer to make it photosensitive, causing a permanent refractive index increase when exposed to the UV writing beam. The interference pattern from a second UV beam allows the fabrication of Bragg gratings, when fabricating waveguides without gratings only a single UV beam is required. (b) A photograph of the UV writing system fabricating a Bragg grating. (c) The wafers are fabricated by flame hydrolysis deposition. A turntable with silicon wafers rotates underneath a burner that deposits a fine layer of silica which is later consolidated by heating in an oven. The process is repeated for each of the three silica layers.

The flame hydrolysis deposition technique, by which the silica-on-silicon wafers are constructed, leads to significant birefringence. Each of the three silica layers is constructed in a two-step process. First, a silica particulate layer is deposited by a burner as shown in Figure 3.14(c). Second, this 'soot' layer is consolidated via heating in an oven. These

thermal treatments lead to a permanent stress between the planar layers due to their differing coefficients of thermal expansion. Through the strain-optic tensor of silica, this permanent stress manifests optically as a birefringence.

One significant benefit of this technique is the ability to write Bragg gratings at arbitrary points within the circuit. This is accomplished by employing a second UV writing beam, as shown in Figure 3.14, to produce an interference pattern at the focal spot that leads to a spatially varying index change. The angular orientation and power in these two writing beams allows one to fabricate gratings with a wide variety of spectral responses. While we do not use this grating fabrication capability in the SFWM results discussed below, it does offer many promising avenues for future work.

3.2.1 Single SFWM source

We apply a strong pulsed pump at $\lambda_p = 731$ nm along the slow axis of our birefringent waveguide and observe the SFWM spectra shown in Figure 3.15. From the location of the signal and idler, we infer a birefringence for this waveguide of $\Delta n = 1.65 \cdot 10^{-4}$ which is higher than the value of $\Delta n = 1{\cdot}10^{-4}$ obtained in the femtosecond laser written source discussed above.

There is significant background evident for both the signal and idler modes shown in Figure 3.15. This noise is not unique to this particular waveguide, but is representative of all the UV-written SFWM sources. By comparison, the femtosecond laser-written source spectra shown in Figure 3.6 shows almost no background on the signal and less background on the idler. Fluorescence is the likely cause of this additional noise in our UV-written waveguides.

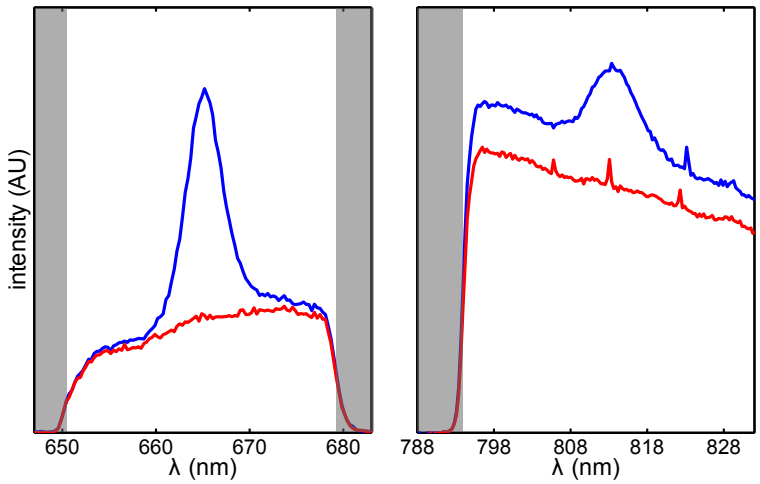


Figure 3.15: A single photon spectrum shows the SFWM signal and idler at $\lambda_s = 665$ nm and $\lambda_i = 813$ nm respectively (blue) for a pump at 731 nm (not shown). There is significant background (red) on both the signal and idler modes. Spectral filters attenuate frequencies in the gray shaded areas.

The noise overlaying the signal mode is potentially caused by the non-bridging oxygen-hole center (NBOHC)[107,108]. This widely-studied silica defect is comprised of a dangling bond in the lattice ($\equiv \mathrm{Si} - \mathrm{O}\bullet$) and can be caused by a number of physical mechanisms, including UV illumination[109]. This defect exhibits a broad luminescence from $600 - 700$ nm with a peak at 655 nm which lies close to the SFWM signal. When our waveguides are pumped with a CW laser, this NBOHC fluorescence disappears. This implies that the high pump intensities when using an ultrashort pulsed pump allows for a multiphoton excitation of an NBOHC defect, evident by the broad luminescence we see in Figure 3.15 overlapping the signal mode.

The majority of the background surrounding the idler mode in Figure 3.15 is fluorescence. The > 40 THz detuning of the idler from the pump implies that Raman scattering is minor. This dominance of the noise by fluorescence is confirmed in Figure 3.16 where the noise spectrum is measured at two different pump wavelengths. If the noise is due to

a fixed fluorescence feature, then the noise spectrum should not change shape. However, Raman emission exhibits a constant frequency detuning from the pump and thus should shift along with the pump. The right side of Figure 3.16 demonstrates almost no change in the background at the two widely separated pump wavelengths, confirming the majority of the noise in these UV-written guides is due to fluorescence.

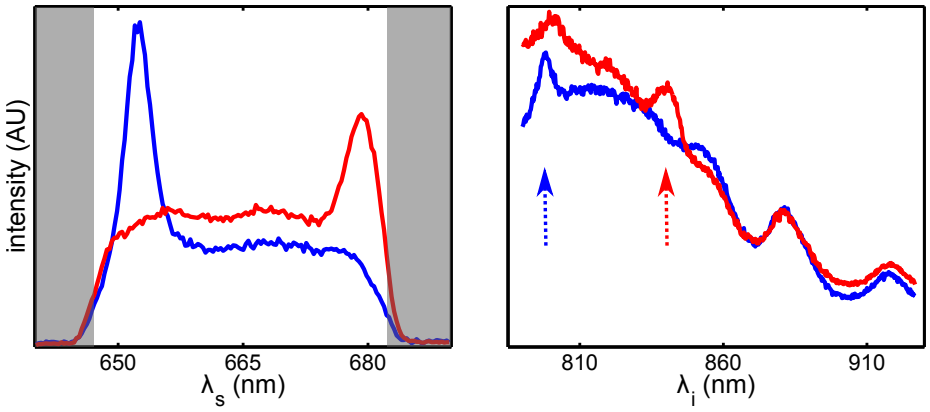


Figure 3.16: The spectra from a UV-written SFWM source is measured for two different pump wavelengths, 715 (blue) and 750 nm (red). The signal mode is clearly visible on the left plot at the edges of the bandpass filter. In the right plot, the idler is evident (arrows) on top of a broad background. The shape of this background changes very little with the pump, which is consistent with fluorescence.

Figure 3.9 shows that SFWM is phasematched over a wide frequency range which, in principle, should allow one to move the signal and idler frequencies to a range where this fluorescence background is minimized. In practice, the currently available pump laser and detectors significantly constrain adjustability. Our Ti-Sapphire pump laser cannot operate below 700 nm while our silicon APDs monitoring the signal and idler can only operate in the range $400-1000$ nm. For the multi-source experiments of interest here, we are further constrained to the most efficient region of the detectors from $500-850$ nm. We therefore consistently operate with wavelengths $\lambda_p \approx 730$ nm, $\lambda_s \approx 665$ nm, and $\lambda_i \approx 815$ nm despite the presence of the fluorescence noise shown in Figure 3.16.

The significant background evident in Figure 3.15 could significantly deteriorate source performance. Noise photons in the heralding mode reduce η_H by introducing false herald events that are not correlated with any photons in the idler mode. This increases the denominator of Equation (2.25) without a corresponding increase in coincident photon emission. On the other hand, noise in the heralded mode causes undesired heralding of more than one photon that is reflected in an increased $g_H^{(2)}(0)$. In both cases, noise reduces the photon number correlations between the signal and idler modes upon which our source performance relies.

We analyze the number statistics of a single source to quantify η_H and $g_H^{(2)}(0)$. To maximize source performance, we filter as much of the background as possible without significantly attenuating the signal and idler modes. We angle-tune our spectral filters to implement a narrower bandpass about the signal and idler modes. The results are shown in Figure 3.17, which is plotted on the same scale as the spectra using a larger bandpass in Figure 3.15. In addition, we use a home-built APD coincidence counting program, optimized for high temporal resolution and described in Appendix C.3.2, to temporally filter the noise. Fluorescence features in silica generally have decay times of microseconds which, for our 80 MHz pulsed pump, implies a near temporally constant background. However, the signal and idler are sub-picosecond wavepackets. We therefore modify the experimental arrangement shown in Figure 3.5(b)-(c) for measuring η_H and $g_H^{(2)}(0)$ respectively. To filter the temporally constant noise from the ultrashort pulsed SFWM, we count in coincidence with a pickoff from the pump laser. In this arrangement, a herald is a signal photon temporally coincident with a pump pulse.

Table 3.1 reveals that this combination of spectral and temporal filtering allows simul-

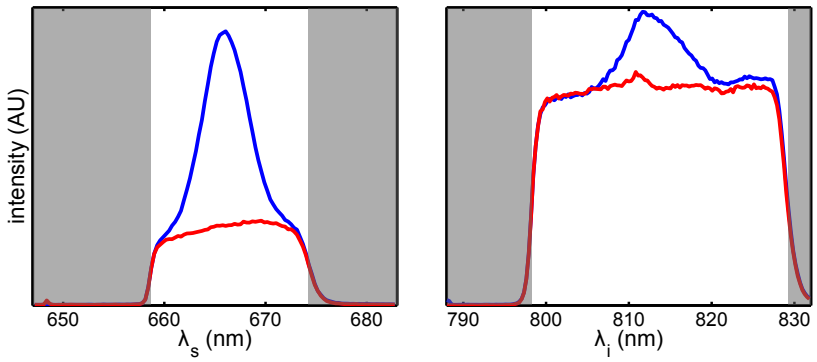


Figure 3.17: The spectra of the SFWM signal and idler (blue) and corresponding background (red) are recorded before measuring the η_H and $g_H^{(2)}(0)$ with APDs. The spectral filters are angle-tuned to narrow the bandpass to minimize the background while introducing negligible loss on signal and idler.

taneously high η_H and low $g_H^{(2)}(0)$. Table 3.1(b) shows the expected behavior that increasing either the pump power or the width of the temporal filter increases $g_H^{(2)}(0)$ and thus increases the contribution of higher order photon emission. The dependence of η_H on these paramaters in Table 3.1(a) is more complicated. The highest value is obtained for the lowest pump power and tightest temporal filtering. There isn't a monotonic decrease in η_H for increasing powers and temporal windows since more noise photons in the heralded idler mode actually increase η_H. However, the heralded state is not the desired mode given in Equation (2.20) but rather a mixed state over the SFWM idler and noise. In our experiments, we therefore focus on implementing the tightest temporal filtering and moderate pump powers to simultaneously attain high η_H and low $g_H^{(2)}(0)$.

We moved to UV-written silica waveguides to achieve better source homogeneity, and so we measure the SFWM spectra from many guides which are fabricated on the same UV-written chip. Figure 3.12 showed that source variation manifests as a change in the spectral position of the signal and idler. Comparing just λ_s or λ_i between sources introduces error into our estimation of chip homogeneity due to small shifts in λ_p. We therefore measure

Table 3.1: The heralding efficiency (η_H) and heralded second-order correlation $g_H^{(2)}(0)$ are shown for a single SFWM source fabricated via UV-writing. Each row corresponds to an average pump power, while the columns indicate the temporal window applied with the FPGA coincidence counting program.

(a) Heralding efficiency-η_H

	0.8 ns	1.8 ns	3.3 ns	4.9 ns
80 mW	30.3%	28.3%	24.7%	22.7%
120 mW	28.3%	28.6%	25.5%	23.8%
160 mW	21.5%	28.1%	26.0%	24.3%

(b) Heralded second-order correlation-$g_H^{(2)}(0)$

	0.8 ns	1.8 ns	3.3 ns	4.9 ns
80 mW	0.032	0.065	0.118	0.154
120 mW	0.062	0.099	0.156	0.196
160 mW	0.105	0.144	0.197	0.248

$\lambda_i - \lambda_s$ which is insensitive to these small pump instabilities.

Figure 3.18 shows the results of this SFWM source homogeneity test. In general, the homogeneity between sources is vastly improved from what was observed in the femtosecond laser written waveguides. The waveguides were UV-written in the order shown in Figure 3.18, with guide 1 written first. Waveguides $1-4$ have a significantly higher birefringence, leading to a larger $\lambda_i - \lambda_s$. While this abrupt shift has not been observed in other UV-written source chips, the HOM interference experiments described below used waveguides drawn from the set in Figure 3.18. For clarity, we use data taken from this single chip for all analysis in this section. The remainder of the sources are well-matched with a slowly decreasing birefringence farther into the writing process. There are at least two possible causes for this gradual drift. Non-uniformity in the wafer from which this chip was diced could gradually change the stress, and thus the birefringence, across the chip. Secondly, the chips gradually outgas hydrogen during the writing process. Since this hydrogen makes the chips photosensitive, then the magnitude of the UV-induced index change also decreases for later writing times.

Pumping our UV-written SFWM sources for extended periods of time causes a gradual decrease in the signal and idler detuning from the pump. Figure 3.19 shows this spectral

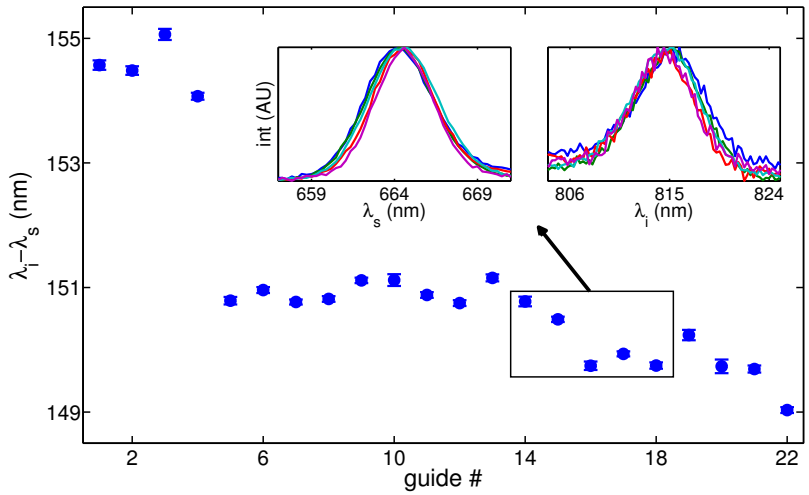


Figure 3.18: The uniformity of UV-written SFWM sources is quantified by measuring $\lambda_i - \lambda_s$ for 22 different guides on the same chip. The phasematched wavelengths are determined by a Gaussian fit to the background subtracted data. A representative set of background-subtracted data is shown in the insets for five of the sources. All spectra are measured shortly after first exposure to the pulsed pump laser and are thus independent of the spectral shift effect discussed below.

shift for two waveguides, each of which are pumped with the same average power over the course of several days. Investigation into the physical cause of this gradual shift is ongoing.

Before each interference experiment, we separately pump all involved sources until they spectrally match one another. During subsequent periods of extended use, Figure 3.19 demonstrates that the signal and idler spectra remain well-matched between multiple sources despite gradually shifting in time. Importantly, we have noted that this spectral shift is not the limiting factor in stability when running sources for extended periods of time. For example, spatial mode coupling and the optical path length matching degrade more quickly than the spectral match. We later show it is possible to perform high-visibility HOM interference experiments despite this gradual spectral shift.

The main motivation for pursuing UV-written SFWM sources was to achieve excellent inter-source homogeneity to allow high-visibility HOM interference. The excellent initial

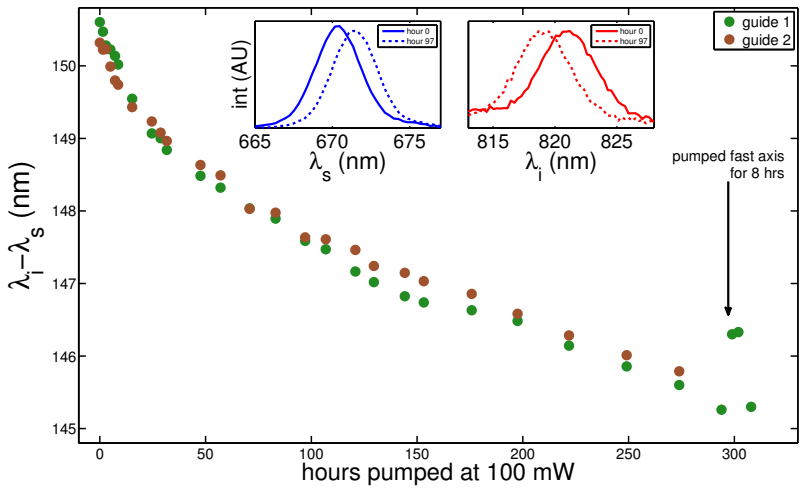


Figure 3.19: The UV-written SFWM sources demonstrate a gradual shift in the signal and idler wavelength after being pumped for an extended period of time. We quantify this shift by periodically measuring $\lambda_i - \lambda_s$ for two sources (brown/green) that are each pumped with the same average power ultrashort pulsed pump. The insets show representative signal (blue) and idler (red) spectra at the start (solid) and finish (dotted) of the annealing run. This spectral variation shows a polarization dependence; pumping the fast axis of the guide, where SFWM is not phase-matched, leads to an increase in $\lambda_i - \lambda_s$.

source uniformity in Figure 3.18 and the repeatable annealing demonstrated in Figure 3.19 show that this goal is within reach.

3.2.2 Two-source interference

We perform a series of two-source HOM interference experiments to explicitly demonstrate this operation of many identical heralded sources of pure single photons. The experimental arrangement is shown in Figure 3.20, and is similar to the single-source schematic shown in Figure 3.5. We perform multiple two-source HOM experiments to demonstrate operation of many matched sources. We select a subset of five neighboring sources on the chip and interfere four of these five with a fifth refererence source in pairwise fashion. For example, we first interfere source 1 and 2, then 1 and 3 and so on.

For each of these two-source interference experiments, we measure two HOM dips. We

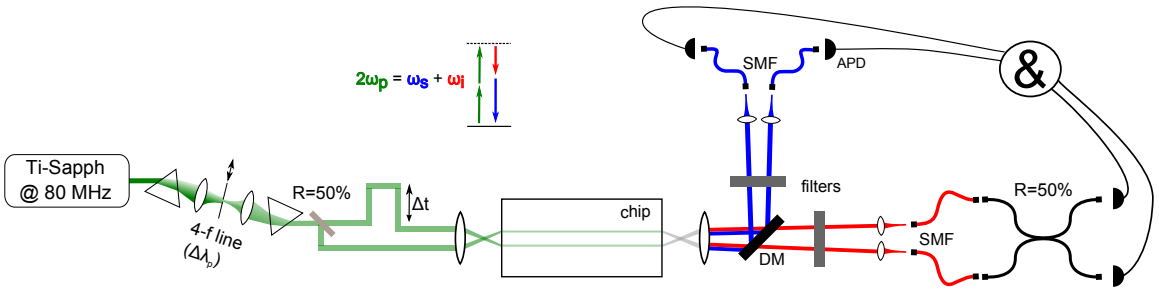


Figure 3.20: A $4f$ line with an adjustable slit is used to control the bandwidth ($\Delta\lambda_p$) of an ultrashort pulsed pump which is split by a $R = 50\%$ beam splitter. A free-space optical delay line applies a programmable delay Δt to one of these two pump beams, which are subsequently coupled into two SFWM sources. After the chip, a dichroic mirror (DM) separates signal and idler beams while spectral filters remove the pump. Each of the four modes is then coupled into a separate single-mode fiber. The signal photons act as heralds and are coupled directly to APD detectors, while the idlers propagate through an $R = 50\%$ fiber beam splitter before detection. An FPGA counts all possible coincident detection events amongst the detectors.

first measure the HOM dip visibility with only broad spectral filtering on the signal and idler. Residual spectral entanglement between the signal and idler, evident from the sinc lobes in the phasematching function, reduces the HOM dip visibility. Despite this reduced interference quality, this experiment is vital to demonstrating operation of identical sources, as applying identical narrow filters can hide significant distinguishability amongst the sources. Secondly, we measure the HOM dip visibility when narrower spectral filtering is applied to the herald photon to remove the sinc lobes. This is the intended configuration of the source if it was to supply high purity photons for an experiment.

A summary of the two-source interference results is shown in Figure 3.21. In all cases, the measured HOM interference visibilities are close to their maximum possible values. For the broadly filtered configuration, we expect a maximum visibility of $V^b_{max} = 0.88$ and observe a mean visibility, across all the interference experiments shown in Figure 3.21, of $V^b_{raw} = 0.807\pm0.018$ from the raw data and $V^b_{bg} = 0.836\pm0.019$ with background subtraction. With a narrower bandpass filter about the herald photon, we obtain $V^n_{max} = 0.97$, $V^n_{raw} =$

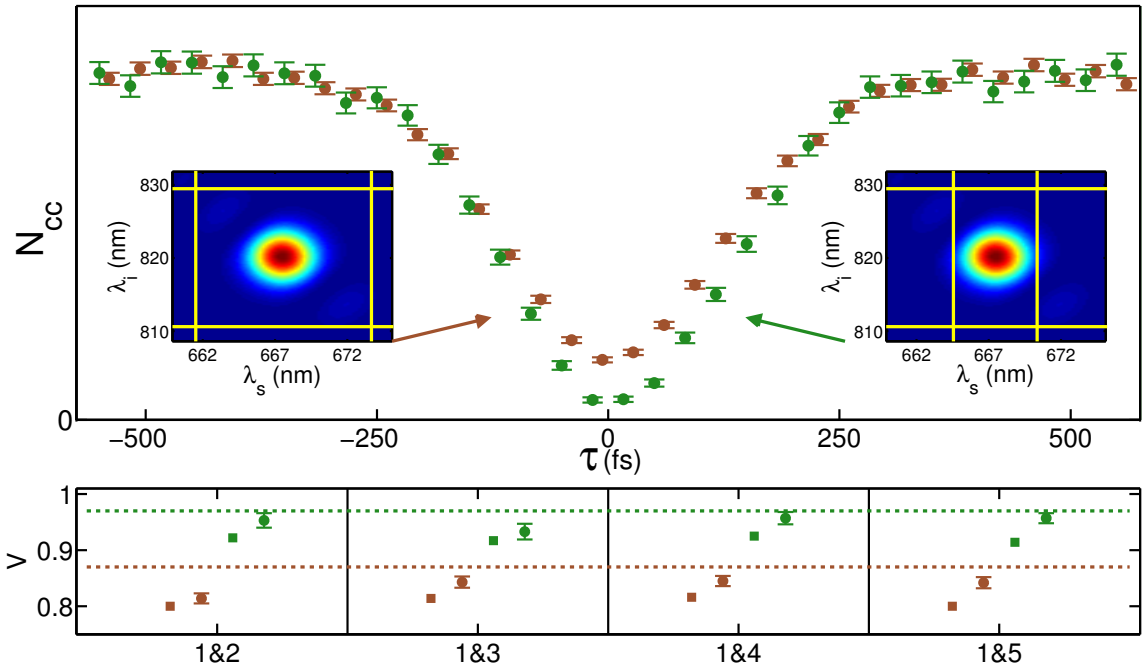


Figure 3.21: We perform a series of HOM interference experiments amongst five waveguides, interfering the heralded emission from two sources at a time. For each pair of sources we perform two interference experiments; the first configuration employs a broad spectral bandpass on signal and idler (top, left inset) while the second narrows the bandpass about the heralding signal photon (top, right inset). The yellow lines indicate the bandpass boundaries. This broad and narrow filtering configuration leads to the brown and green interference results shown (top). Narrower filtering removes residual spectral correlations between signal and idler and thus improves dip visibility. A summary of the interference visibility results amongst the set of five waveguides are summarized at the bottom. Brown and green marks correspond to broad and narrow filtering respectively. Points with error bars indicate background subtracted results, while points show raw visibility results. The maximum expected visibilities for each configuration are shown (dashed lines). In the narrow filtered case, the maximum expected visibility is reduced to $V^n_{max} = 0.97$ (green dashed line) due to the constraints of our bulk optics and is not intrinsic to the source. The slight angular offset between the two herald photons after the chip implies a slightly different bandpass from the shared spectral filters; using dedicated filters for each individual source would restore $V^n_{max} > 0.99$.

0.919 ± 0.023, and $V^n_{bg} = 0.95 \pm 0.024$. These values indicate we have successfully operated nearly identical heralded sources of highly pure states on the same chip. Furthermore, our spectral and temporal filters have significantly mitigated the impact of background noise.

3.2.3 Source performance comparison

We first demonstrated a SFWM source in a silica waveguide fabricated by femtosecond laser writing in an undoped substrate. This source successfully demonstrated the low-loss heralding of high purity single photons, as shown in Figures 3.7 and 3.8, that originally motivated the choice of SFWM in silica as a source architecture. A key drawback of this approach was the inability to fabricate multiple identical sources which precluded interfering the heralded emission between two or more sources. Fabrication of hundreds of waveguides only led to three sources whose SFWM spectra were successfully observed and Figure 3.12 shows that these sources were clearly spectrally distinguishable from one another. We hypothesize that the highly nonlinear dependence of the fabrication results on the writing laser power [87] leads to inhomogeneity between guides.

Fabrication of many identical silica SFWM sources was achieved by switching to UV-writing of waveguides in a silica-on-silicon photonic chip. Table 3.1 showed these sources still allow for low-loss heralding of single-photons, with Figure 3.18 demonstrating that the SFWM spectra for different sources across a chip are now nearly identical. This allowed us to directly demonstrate high heralded state purity and inter-source homogeneity through the set of HOM interference experiments shown in Figure 3.21. Unfortunately, these sources also exhibited significant fluorescence that is evident from the marginal spectra shown in Figure 3.16. This undesired fluorescence can reduce the number correlations between the signal and idler modes upon which heralded photon sources rely. However, Table 3.1 and our interference results in Figure 3.21 show that temporal and broadband spectral filtering significantly mitigate the deleterious effects of this fluorescence and still allow for excellent

source performance.

The slow shift in the spectral output of the SFWM sources in UV-written waveguides shown in Figure 3.19 is an interesting feature with unclear impacts for future, more advanced source designs. A best-case scenario involves first pumping each guide individually and utilizing this slow spectral shift to fine-tune the SFWM spectral emission between sources post-fabrication so that even better inter-source homogeneity is achieved. The experimenter then requires some way to stop, or at least slow, future shifting so many sources can be operated for extended periods at a fixed wavelength. In our experiments, we were able to tune, using our pump laser, each pair of sources to spectral homogeneity before interference experiments but were unable to find a way to stop the spectral shift. Instead, we relied on the two sources in an interference experiment shifting together, as shown in Figure 3.19, so that homogeneity is maintained throughout. Research into the physical cause of this spectral shift, in particular finding ways to slow or stop it, is ongoing.

Fair comparison of our silica SFWM architecture to alternatives is complicated by both the wide variety of photon source designs and the many performance parameters of interest[34]. Single-emitter sources, including trapped atoms, ions, quantum dots, and NV-centers in diamond, are far more complicated to operate than the source described here and generally exhibit higher losses with significant difficulties in achieving inter-source homogeneity[34]. While SPDC sources in bulk crystal are far easier to operate with repeatable crystal fabrication widely available, these sources still emit into many spatial-spectral modes.[1] Heralding of high purity states requires detection of a single pair of spatial-spectral

[1] Extensive source engineering can be used to limit signal-idler emission to a single pair of spectral modes, but even then the multimode spatial structure persists[57].

modes which is generally accomplished with heavy spatial-spectral filtering, but the added loss from this strategy reduces source brightness and can preclude quantum-enhanced operation. In the last several years, waveguide-based SPDC [59,76,110,111] and fiber-based SFWM sources [67–69,73,97] have been successfully used to emit into a single pair of spatial-spectral modes. These sources face one or more of the following challenges: excellent homogeneity between multiple sources has not been demonstrated, the single spatial mode of the sources is frequently not well-matched to the optical fibers that are commonly found in optics experiments which increases loss, and many of these sources are restricted to certain wavelength regions by phasematching constraints.

We have demonstrated many identical low-loss sources of high purity single photons in an exceptionally simple architecture capable of emitting from visible to telecom wavelengths. The main drawback of our source is the low $\chi^{(3)}$ nonlinearity in silica, which forces the use of a high average-power ultrashort pulsed pump laser. Waveguide SPDC sources can operate with pump powers orders of magnitude lower than described here [59,76,110,111]. This is the key reason we have focused on a heralded single-photon source application, where the nonlinear optical interaction must deliberately be kept small to avoid multi-pair emission. In this regime, Ti-Sapphire lasers can still output sufficient power to simultaneously pump many silica SFWM sources. We anticipate that in this application silica SFWM sources on photonic chips can become a workhorse source for future quantum optics experiments.

Chapter 4

Boson sampling theory

Can quantum systems solve problems more efficiently than classical computers? While many believe the answer is 'yes', we still lack clear experimental evidence. This is partly because the quantum machine that has attracted the most attention, the universal quantum computer, is also very hard to build[2,3]. There is increasing interest in intermediate quantum computers that operate with less stringent experimental requirements, yet can nonetheless demonstrate a quantum speed-up for a limited class of problems.

The boson sampling problem is a computational task particularly well-suited to such non-universal quantum computers. We begin this chapter by introducing an abstract model of computation so that we can more specifically define what we mean by classical and universal quantum computers. With this framework, it becomes clear how intermediate devices straddle the space between these two extremes. We then introduce a simple quantum system that solves the boson sampling problem, summarize why it's likely classically hard, and explain why photonics provides a natural physical implementation. The chapter concludes by summarizing the key points from the paper by Aaronson and Arkiphov (AA) that ori-

ginally formulated the boson sampling problem and gave strong evidence for its classical hardness[112]. Their work necessarily involved a significant amount of computational complexity theory with which many physicists – a group very interested in building boson sampling machines – are unfamiliar. The goal here is to present a self-contained summary of their arguments, including an introduction to the pertinent aspects of complexity theory.

4.1 Models of computation

4.1.1 Classical Computers

Anything that executes an algorithm, a finite step-by-step procedure to solve a problem, is a computer. These machines can take many physical forms ranging from an integrated circuit to a person with pen and paper. The Turing Machine (TM), shown in Figure 4.1, is an abstract model that allows us to describe the capabilities of a computer without referencing the details of its physical construction[113]. The Church-Turing thesis states that every function that is algorithmically computable is computable by a TM[2]. In other words, if a TM can't solve a problem then it can't be solved by *any* type of computer, including a quantum machine.

The Church-Turing thesis does not comment on how efficiently the TM will solve a problem. We follow the computer science convention of equating computational *efficiency* with a polynomial scaling of the required resources, space or time, with the size of the problem. By contrast, *inefficiency* corresponds to an exponential scaling. These two categories of problems are also frequently labeled *easy* and *hard* respectively. Classical computers

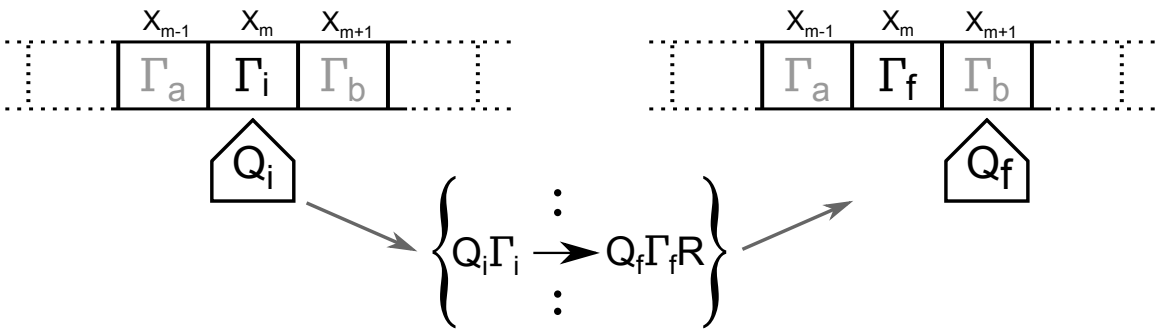


Figure 4.1: The TM consists of several parts: an infinite length of tape divided into cells whose entries Γ_m each hold a symbol from some finite alphabet Γ, a head (pentagon) that can write and read symbols to tape cells and move between them, a memory register that holds the current state of the system Q_j drawn from a finite list of possibilities Q, and a table of instructions. This table contains entries of the form $Q_i\Gamma_i \rightarrow Q_f\Gamma_f(\mathrm{L},\mathrm{R})$ which tells the TM, given the TM is in state Q_i and the head reads Γ_i in its current tape cell, to write symbol Γ_f to the cell, update its state to Q_f and move the head left or right. Given some input state of the tape $\Sigma \subseteq \Gamma$, the TM proceeds deterministically through a computation until reaching a final state $F \subseteq Q$ at which point it outputs all or part of its tape. At any point of the computation, the current state of a TM can be summarized by the triplet (Q_i, T, X_k) where Q_i gives the current state of the machine, T is a vector giving the current configuration of the tape, while X_k gives the position of the head. Only a single step of the computation is shown here.

are those class of machines whose operation can be efficiently simulated by a TM[1]. The resources consumed by a classical computer to solve a problem will scale similarly, for example polynomially or exponentially, to that required by a TM solving the same problem.

4.1.2 A Universal Quantum Computer

There are many problems that are hard for a classical computer to solve. Feynman observed that these exponential computational scalings are commonly encountered when classical computers simulate quantum mechanical systems[114]. Consider a group of N two level systems. If these particles are classical, and accordingly occupy a definite state, a TM modeling their time evolution only needs to store a single N-bit number. Accurate modeling of quantum particles requires the computer to store up to 2^N numbers corresponding to

[1]Sometimes classical computers are equated with a TM that has access to a random number generator, though this detail is not relevant to our discussions.

the dimensionality of the resulting superposition state. This exponential scaling in the quantum case implies prohibitively large computational overhead for even modestly sized systems. Quantum mechanics presents a wide variety of systems that are unlikely to be efficiently simulated by a TM[114].

The TM shown in Figure 4.1 is a classical device. Physical implementations of a TM can obey the principles of local realism; all portions of the machine shown in Figure 4.1 have definite values independent of measurement and these values can only be altered by interactions with their local environment. It is perhaps unsurprising that these classical devices cannot efficiently simulate quantum systems that are not required to obey these restrictions.

Building a computer that operates quantum-mechanically, a quantum Turing Machine (QTM), overcomes these problems in simulating quantum systems[114,115]. The state of this *universal quantum computer* is now described, not by a single triplet (Q_i, T, X_k), but by a superposition of triplets $\sum_\gamma \alpha_\gamma (Q_{i_\gamma}, T_\gamma, X_{k_\gamma})$. This superposition state can exhibit entanglement, and thus the QTM is no longer constrained to local realistic operation.

The QTM is believed to be more powerful than a TM since the state of the latter machine is a subset of the superposition states allowed by the quantum device. For example, it is known that a QTM can efficiently simulate quantum systems[116]. However, the advantages of a QTM extend beyond direct physical simulation to more abstract mathematical problems. Most famously, they enable an exponential speed-up in finding the prime factors of an integer through the implementation of Shor's algorithm[117]. Many additional quantum algorithms have been developed for a wide variety of problems[2,118–120].

Despite widespread interest in building a QTM, there still has not been a clear demon-

stration of a quantum-enhanced computation. While small proof-of-principle devices have been constructed using a wide variety of physical systems [3], they do not surpass the abilities of classical computers. Universal quantum computers require physical systems that are well-isolated from the decohering effects of their environment, while at the same time allowing precise manipulation during computation. Simultaneously meeting these two goals is challenging, and explains this slow progress in experimental efforts to build such a machine. Despite significant research investment, no physical architecture has yet demonstrated all the criteria for a scalable universal quantum computer [121].

4.1.3 Boson Sampling Machine: An Intermediate Quantum Computer

One does not need to build a QTM to demonstrate a computational advantage using quantum mechanics. Intermediate quantum computers bridge the gap between classical and universal quantum machines [122–124]. They can efficiently solve specific problems believed to be hard for a TM, but only a subset of those problems which are efficiently solvable by a QTM. Ideally this reduction in computing power comes with the benefit of significantly relaxed requirements for physical construction. Such 'purpose-built' quantum computers offer a promising route to quantum-enhanced computation far sooner than a universal machine. Boson sampling is a recently-introduced intermediate model of quantum computation [112].

There are currently no known applications for a boson sampling machine, besides demonstrating the principle that quantum systems can provide a computational advantage over classical machines. This characteristic of the boson sampling problem polarizes researchers. A common criticism is that boson sampling is not 'useful for anything'; a boson

sampling machine does not solve problems with widespread application, unlike a QTM.

A counterargument is that demonstrating quantum-enhanced computation is *interesting in its own right* because it furthers our understanding of what can be efficiently computed in the physical world. To proponents of this position, the level of interest in a problem is proportional to the strength of the evidence for its classical hardness and the feasibility in building a physical machine to solve it. The presence or absence of wider applications for the problem are irrelevant.

Adherents of this viewpoint likely view boson sampling as an extremely interesting problem. A machine that efficiently solves this problem has a very natural photonic implementation that may be significantly easier to construct than universal quantum computers. In addition, boson sampling presents a problem that is strongly believed to be classically hard; if it were classically easy then a broad range of computational problems would be easier than is widely believed (a collapse of complexity classes). This is far stronger evidence for the computational power of quantum mechanics than the well-popularized Shor's factoring algorithm. While factoring has obvious practical applications, an efficient classical factoring algorithm would not have drastic consequences in the field of complexity theory [112]. This simultaneously strong evidence of classical hardness and straightforward physical construction differentiates boson sampling, not just from Shor's algorithm, but also from previous intermediate quantum computational models [122–124].

4.2 Boson Sampling Overview

4.2.1 A Boson Sampling Machine

A quantum boson sampling machine (QBSM) consists of a source of N indistinguishable bosons, random unitary evolution of these bosons amongst M modes, and single boson sensitive detection on each mode. Figure 4.2 shows a photonic implementation of this design where photons are injected into an interferometric network with detectors monitoring all outputs. Given a description of the input bosonic state and the unitary, it is likely hard for a classical computer to output samples drawn from the correct probability distribution; the computation time required by a classical computer is believed to scale exponentially in N [112].

Figure 4.3 shows that the QBSM is the 'bosonic analogue' of the classical Galton board [112]. In the latter device, balls dropped into an array of pegs perform a classical random walk before finally setting into one of the bins at the bottom. The binomial distribution describes the probability of a ball landing in a specific bin and these statistics are confirmed after a sufficient number of balls are dropped. In our bosonic version, the photons explore all possible paths through the interferometric network in each trial. The probability amplitudes of these possible paths interfere with one another, leading to an entangled state at the output.

The number statistics of this entangled output are hard to classically compute because bosonic scattering amplitudes are given by the matrix permanent [125,126]. The specific form of this functional dependence on the permanent is illustrated in Figure 4.2(b). The matrix permanent is identical to the well-known determinant except, when expanded as

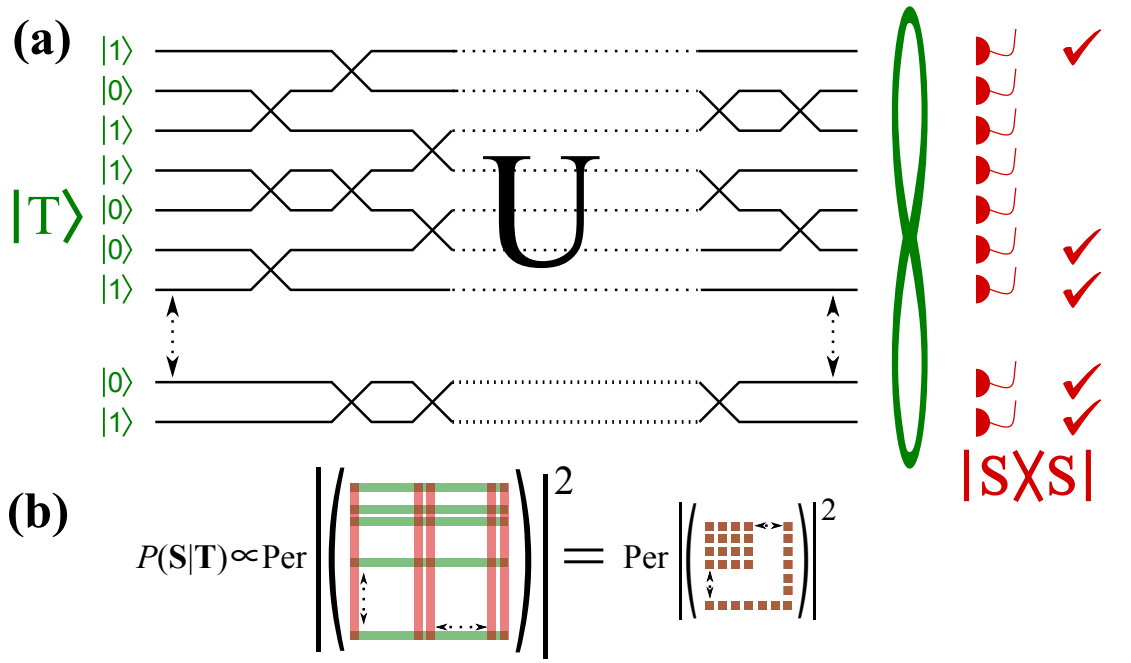


Figure 4.2: (a) A photonic boson sampling machine inputs N photons into an interferometric network comprised of beam splitters with randomly assigned reflectivities. The input state $|\mathbf{T}\rangle$ has no more than one photon per input mode. The resulting unitary transformation over the M modes is given by the matrix $\mathbf{U}$. Single photon detectors monitor the number statistics of the entangled output. (b) The probability of observing a particular output $|\mathbf{S}\rangle$ given the input $|\mathbf{T}\rangle$ is given by the permanent of a subsection of the unitary matrix describing the optical network $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$.

a polynomial in the matrix elements, the permanent has only positive terms while the determinant alternates positive and negative. Despite this minor difference, the determinant is easy to compute while there is strong evidence the permanent is classically hard [127].

The dependence of the boson number distribution on the matrix permanent takes a simple form. Following the notation in Figure 4.2(a), we specify the input state to our unitary by

$$|\mathbf{T}\rangle = |T_1 ... T_M\rangle \propto \prod_{i=1}^{M} (\hat{a}_i^\dagger)^{T_i} |0\rangle \,, \tag{4.1}$$

where $\hat{a}_i^\dagger$ is a creation operator for input mode i. There are a total of $N = \sum_{i=1}^{M} T_i$ photons distributed in the M modes. The output state $|\psi_{out}\rangle$ is generated according to the unitary

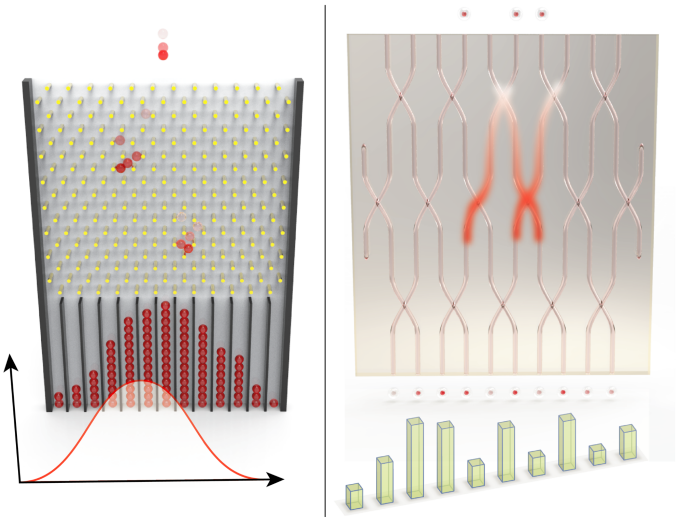


Figure 4.3: The classical Galton board illustrates the binomial distribution by having many balls perform a random walk through a series of pegs (left). A photonic implementation of boson sampling using a waveguide chip is shown (right). In this analogue to the Galton board, indistinguishable photons injected at the top explore all possible paths through the interferometric network. Resulting interference leads to an entangled output whose number statistics are hard to sample.

map

$$\hat{a}_i^\dagger = \sum_{j=1}^{M} U_{ij} \hat{b}_j^\dagger, \tag{4.2}$$

where $\hat{b}_j^\dagger$ is the creation operator for output mode j and U_{ij} is an element of the unitary matrix $\mathbf{U}$ describing the interferometric network. Finally, we measure the number of photons in each mode. The probability of observing a particular measurement outcome $|\mathbf{S}\rangle = |S_1 ... S_M\rangle$ is given by

$$P(\mathbf{S}|\mathbf{T}) = |\langle \mathbf{S}|\psi_{out}\rangle|^2 = \frac{|\mathrm{Per}(\mathbf{U}^{(\mathbf{S},\mathbf{T})})|^2}{\prod_{j=1}^{M} S_j! \prod_{i=1}^{M} T_i!}, \tag{4.3}$$

where the $N \times N$ submatrix $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$ is obtained by keeping S_j copies of the j^{th} column (and T_i copies of the i^{th} row) of $\mathbf{U}$ as shown in Figure 4.2(b). One can envision a race between a classical and a quantum machine to sample the boson distribution given $|\mathbf{T}\rangle$ and

U. The classical machine must evaluate at least part of the probability distribution, which requires the analysis of matrix permanents. A QBSM instead creates indistinguishable bosons, physically implements **U**, and records the outputs. For sufficiently large N, the QBSM is expected to win [112].

It likely remains hard for a classical computer to even approximately sample from the boson distribution [112]. This is a critical point as all experiments will inevitably have some noise and thus a QBSM will be unable to sample exactly from the distribution in Equation (4.3). We quantify the distance between two probability distributions, $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$, using the L_1 measure

$$d(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}) = \frac{1}{2} \sum_k |P_k^{(1)} - P_k^{(2)}|, \tag{4.4}$$

where bold font distinguishes a probability distribution $\mathbf{P}$ from a single element of that distribution P_k. This L_1 measure is used to formalize the approximate boson sampling claim. Let $\boldsymbol{\mathcal{D}}$ be the true boson distribution and $\boldsymbol{\mathcal{D}'}$ be the approximate distribution sampled by the classical machine that has additive error given by $d(\boldsymbol{\mathcal{D}'}, \boldsymbol{\mathcal{D}}) \leq \epsilon$. The time required by a classical computer to approximately sample from the boson distribution is unlikely to scale poly$(N, 1/\epsilon)$, that is polynomially in N and $1/\epsilon$. Quantum mechanics provides a computational advantage if one could operate a QBSM that required only polynomially more resources for increasing numbers of photons and sampling accuracy [112].

Photonics is a natural platform to implement boson sampling because sources of indistinguishable photons are well-developed [34,59,77], integrated optics offers a scalable route to low decoherence linear transformations over many modes [21,23,24,26], and a variety of high-efficiency photon detectors are available [70,71,128]. Importantly, boson sampling re-

quires neither the nonlinearities nor on-demand entanglement that are needed for universal quantum computation and are challenging to implement in photonics [129]. Building large QBSMs with existing photonic technology in the not-distant future may be feasible.

4.2.2 Bosonic scattering amplitudes given by the matrix permanent

We show how the matrix permanent in Equation (4.3) arises when predicting the number statistics at the output of the QBSM. This connection to the permanent is critical to the analysis in the next section that shows that boson sampling is likely classically hard [112]. In particular, we prove Equation (4.3) in the 'collisionless' case where all modes, at both the input and output, have at most one photon. This instance is particularly relevant here because the original boson sampling problem [112] assumed an input state with at most one boson per mode. Furthermore, it required that the number of modes be much larger than the number of bosons so that it was unlikely to observe more than one boson per mode at the output.[2]

We assume N bosons are injected into a network that performs a unitary transformation over M modes. Without loss of generality, let modes 1 to N contain an input boson, while modes $N+1$ to M have vacuum inputs. The input state can then be described by

$$|\mathbf{T}\rangle = \prod_{i=1}^{N} \hat{a}_i^\dagger |0\rangle . \tag{4.5}$$

[2]This bound on the number of modes, which implies we are unlikely to measure more than one boson per mode, is needed in the computational hardness analysis of boson sampling. However, it provides a significant secondary benefit since this output no longer requires number-resolving detectors, which are challenging to implement [70,71], for accurate sampling [112].

Using Equation (4.2), we then obtain the output state

$$|\Psi_{\text{out}}\rangle = \prod_{i=1}^{N} \left(\sum_{j=1}^{M} U_{ij} \hat{b}_j^\dagger \right) |0\rangle \,. \tag{4.6}$$

To find the boson distribution, we project our output onto a state $|\mathbf{S}\rangle$ which, in the number basis, we describe by an N element vector $\underline{\mathbf{S}}$, where an element of this vector $\underline{S}_j$ gives the mode of the j-th boson. We stress that this vector is different from the M-element vector $\mathbf{S}$, where an element S_j gives the occupation number of the j-th mode.[3] Returning to our example, the probability of measuring $|\mathbf{S}\rangle$ is then

$$\begin{aligned} P_S &= |\, \langle \mathbf{S} | \psi_{\text{out}} \rangle \,|^2 \\ &= \left| \langle 0 | \prod_{i=1}^{N} \hat{b}_{\underline{S}_i} \left[\prod_{j=1}^{N} \left(\sum_{k=1}^{M} U_{jk} \hat{b}_k^\dagger \right) \right] |0\rangle \right|^2 . \end{aligned} \tag{4.7}$$

We rewrite $|\psi_{out}\rangle$ in Equation (4.7) to facilitate later comparison with the matrix permanent. Expanding this term in square brackets leads to a summation over M^N terms, corresponding to all the ways one can put N bosons into M modes. We rewrite Equation (4.7) accordingly

$$P_S = \left| \langle 0 | \prod_{i=1}^{N} \hat{b}_{\underline{S}_i} \left[\sum_{j=1}^{M^N} \left(\prod_{k=1}^{N} U_{k,\widetilde{V}_k^j} \hat{b}_{\widetilde{V}_k^j}^\dagger \right) \right] |0\rangle \right|^2 , \tag{4.8}$$

where $\widetilde{\mathbf{V}}$ is the set of M^N permutations of N photons amongst M modes. The tilde notation will be used throughout to represent a set of permutations. The expression $\widetilde{V}_k^j$ indicates

[3] An example helps to clarify this dual notation. Assume we inject $N = 3$ photons into an $M = 6$ mode unitary. One of the possible outputs is $|\mathbf{S}\rangle = |010101\rangle$ which can be equivalently described by either the M-element vector $\mathbf{S} = (0, 1, 0, 1, 0, 1)$ or the N-element vector $\underline{\mathbf{S}} = (2, 4, 6)$. Both forms are useful; for example, the former notation appears in Equation (4.3) while the latter facilitates the analysis in this section and also appears in a boson sampling certification proposal [130].

the mode of the k-th boson in the j-th permutation. As an example, consider the case with $M = 3$ modes and $N = 2$ input bosons, then

$$\begin{aligned} \widetilde{V}^1 &= [1,1] \quad \widetilde{V}^2 = [1,2] \quad \widetilde{V}^3 = [1,3] \\ \widetilde{V}^4 &= [2,1] \quad \widetilde{V}^5 = [2,2] \quad \widetilde{V}^6 = [2,3] \\ \widetilde{V}^7 &= [3,1] \quad \widetilde{V}^8 = [3,2] \quad \widetilde{V}^9 = [3,3] \end{aligned} \tag{4.9}$$

We use the bosonic commutation relations to simplify Equation (4.8). It is helpful to first rewrite this equation

$$P_S = \left| \langle 0| \sum_{j=1}^{M^N} \prod_{k=1}^{N} U_{k,\widetilde{V}_k^j} \left(\prod_{i=1}^{N} \hat{b}_{\underline{S}_i} \right) \hat{b}^\dagger_{\widetilde{V}_k^j} |0\rangle \right|^2 . \tag{4.10}$$

We use several steps to show the effect of commuting the product of operators in Equation (4.10). We first re-express the standard bosonic commutation relation which, with our current notation, is written

$$\begin{aligned} \left[\hat{b}_{\underline{S}_i}, \hat{b}^\dagger_{\widetilde{V}_k^j} \right] &= 1, \quad \text{if } \widetilde{V}_k^j = \underline{S}_i \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{4.11}$$

This leads to the intermediate result

$$\begin{aligned} \left[\prod_{i=1}^{N} \hat{b}_{\underline{S}_i}, \hat{b}^\dagger_{\widetilde{V}_k^j} \right] &= 1, \quad \text{if } \widetilde{V}_k^j \in \underline{\mathbf{S}} \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{4.12}$$

Finally, we multiply the commutators in Equation (4.12)

$$\prod_{k=1}^{N}\left[\prod_{i=1}^{N}\hat{b}_{\underline{S}_i},\hat{b}^{\dagger}_{\widetilde{V}_k^j}\right]=1,\quad \text{if } \widetilde{V}_k^j\in\underline{\mathbf{S}}\ \forall\ k$$
$$=0,\quad \text{otherwise.} \tag{4.13}$$

We summarize this result: $\widetilde{V}^j$ and $\underline{\mathbf{S}}$ must share the same values, though ordering of their elements is irrelevant, otherwise the product of commutators in Equation (4.13) is zero. As a consequence, only a subset of the terms in the summation in Equation (4.10) contribute to P_S. Specifically, if $\widetilde{V}^j$ matches any of the $N!$ permutations of $\underline{\mathbf{S}}$ then the product of commutators in Equation (4.13) is nonzero and contributes to our final result. We label the j-th permuation of the output state $\underline{\widetilde{S}}^j$. Keeping only these $N!$ terms allows us to simplify Equation (4.10)

$$P_S=\left|\sum_{j=1}^{N!}\prod_{k=1}^{N}U_{k,\underline{\widetilde{S}}_k^j}\right|^2. \tag{4.14}$$

The formula for the permanent of an $n\times n$ matrix A with elements a_{ij} is

$$\mathrm{Per}(A)=\sum_{i=1}^{n!}\prod_{j=1}^{n}a_{j,\widetilde{\sigma}_j^i} \tag{4.15}$$

where $\widetilde{\sigma}_j^i$ gives the j-th element of the i-th permutation of the numbers $1,2,....,n$. The term inside the modulus of Equation (4.14) has the same form as the matrix permanent in Equation (4.15). Our original unitary $\mathbf{U}$ can be described by an $M\times M$ matrix. However, it is obvious from Equation (4.14) that we take the permanent of a subsection of $\mathbf{U}$. Specifically, we only keep the first N rows which correspond to modes with input photons. In addition,

we only retain columns corresponding to the elements of $\underline{\mathbf{S}}$. This modified subsection of our original unitary is written $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$ and is illustrated in Figure 4.2(b). Using the definition of the permanent, we can then rewrite Equation (4.14)

$$P(\mathbf{S}|\mathbf{T}) = \left|\mathrm{Per}(\mathbf{U}^{(\mathbf{S},\mathbf{T})})\right|^2. \tag{4.16}$$

If one allows the possibility of more than one photon per input or output mode, then a similar analysis yields Equation (4.3) [131].

We note that the above treatment assumes the bosonic commutation relation $[\hat{b}_i, \hat{b}_j] = [\hat{b}_i^\dagger, \hat{b}_j^\dagger] = 0 \; \forall \; i, j$. If the system consisted of indistinguishable fermions, then the corresponding anticommutation relation $\{\hat{b}_i, \hat{b}_j\} = \{\hat{b}_i^\dagger, \hat{b}_j^\dagger\} = 0 \; \forall \; i, j$ would be used, leading to alternating plus and minus signs introduced in the summation in Equation (4.14), yielding the easily classically computable determinant.

This fermion scattering problem conveniently illustrates a common misconception about what constitutes sufficient proof for the classical hardness of a quantum simulation. Observing that the dimensionality of a quantum system's Hilbert space increases exponentially in system size does *not* imply that predicting the results of measurements on this system is classically hard. Consider an N fermion sampling experiment that is analogous to Figure 4.2. The size of the output state increases exponentially in N, however a classical machine can efficiently sample from the fermion distribution because fermionic scattering amplitudes are given by the efficiently calculable determinant [112].

4.3 Computational Complexity of Boson Sampling

We review the computational hardness arguments for exact and approximate boson sampling[112] on a level accessible to experimental physicists. A satisfactory presentation of this topic requires an introduction to a variety of complexity theory concepts. While this results in a section that reads markedly different than the rest of the thesis, it still comprises an important part of the overall work. In a highly multidisciplinary field like quantum information, it is important that good communication exists between the complexity theorists formalizing interesting problems and the experimentalists attempting to build physical devices. This section is one contribution to that effort.

We first present an overview of a variety of concepts from complexity theory that are indispensable to later discussion. The bulk of this section is then spent reviewing the computational hardness of exact and approximate boson sampling[112].

4.3.1 Complexity Class Zoo

Computational problems can be grouped into distinct families based on the necessary *resources* consumed by a *machine* to solve them. These groups of problems are called complexity classes.

Most of the time we will discuss classes of decision problems. In these cases, the machine receives an arbitrarily long input x but outputs $y = 0$ or 1. This single bit output is frequently described as the machine 'rejecting/accepting' or outputting 'no/yes'. This focus on decision problems is potentially surprising since we are often more interested in search problems where the machine is not restricted to a single bit output. However, many search

problems, for example factoring, are efficiently reducible to decision problems.[4] In addition, classifying the computational complexity of search problems can be more complicated and has received less attention in the literature.

The machines solving these problems will often be variants on the TM described above. It is frequently useful to classify a problem by what additional capability a Turing-like machine requires to efficiently solve that problem. The QTM provides a convenient example; efficiently simulating quantum systems requires that a TM be given the ability to take on a coherent superposition of states. Other variations on the TM are introduced below to classify other sets of problems.

Classes of problems are categorized by the resources, time or space, required to solve them. None of the decision problems we discuss require significant space to solve using a TM. For an N bit input, the size of the TM will be, at worst, a polynomial of degree N expressed as poly(N). We therefore distinguish classes by the computation time required by the machine. This asymmetry between time and space resources can be generalized. If the physical space required by a TM to solve a problem scales like $g(N)$, then the time consumed will scale at least as fast as $g(N)$. This is because a TM head can only access one tape location at a time, but may need to revisit the same location multiple times.

The first class we introduce is P, those problems which can be solved in polynomial time by a TM. More specifically, for an N-bit input the computation time required by a TM to output the correct answer, 0 or 1, scales at worst as poly(N). These problems are conventionally termed 'easy' for classical computers to solve due to this polynomial scaling.

[4]The decision form of factoring asks the following: Given a composite integer z and $y < z$, does z have a non-trivial factor less than y?[2] The better known factoring search problem, where we ask for the prime factors of z, can be answered by solving the decision problem a number of times that scales polynomially in z. In this sense, the decision and search factoring problems are equivalent.

The addition of a random number generator to a TM may add additional computational power. The bounded-error probabilistic polynomial time class (BPP) contains those problems that can be solved by a probabilistic Turing machine (PTM), a TM with access to a random number generator, in polynomial time with an error probability $\leq 1/3$. Since this variant of a TM can access random numbers, it is no longer guaranteed to return the same output when run multiple times with a constant input. The error probability of $1/3$ given above is arbitrary; as long as the error rate is $< 1/2$ than one can run polynomially many trials and output the majority from this ensemble of answers. The Chernoff bound guarantees this procedure will result in an exponentially small error rate [2]. Since deterministic algorithms are a special case of probabilistic ones, then $\mathsf{P} \subseteq \mathsf{BPP}$.

The nondeterministic polynomial time class (NP) consists of those decision problems where 'yes' solutions can be verified in polynomial time by a TM. This efficient verification of a candidate solution does not generally imply an efficient way to find the solution. The name of this class arises because these problems can be efficiently solved by a nondeterministic Turing machine (NTM). A regular TM always has a single machine state, head location, and tape status. The NTM can instead branch into multiple states by performing more than one action in response to a given situation. As the machine executes multiple lines of its program, Figure 4.4 shows that it can repeatedly branch where each strand describes a single possible machine state. If any of these strands leads to an answer of 'yes' to the decision problem, then the machine returns 'yes'. The family coNP is the opposite of this family in that 'no' solutions can be verified efficiently.

A NTM is *not* the same as a PTM nor a QTM. A PTM is simply a regular TM with access to a random number generator. While it's operation is probabilistic, at any given

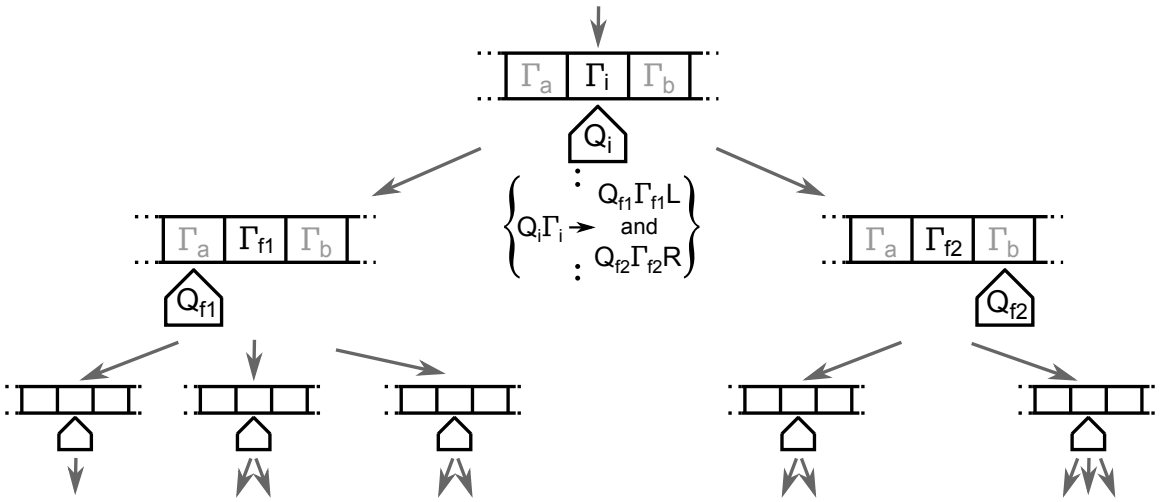


Figure 4.4: At every step in an algorithm, a NTM can branch into multiple parallel states. For example, given the current machine state and tape symbol $Q_i\Gamma_i$ the instruction table can tell the NTM to update to $Q_{f1}\Gamma_{f1}$ with the head moving left and simultaneously update to $Q_{f2}\Gamma_{f2}$ while the head moves right. As computation proceeds, the NTM can assume an exponentially increasing number of states. If any of these strands accepts when the machine reaches the halt state, then the NTM accepts. Otherwise, the machine rejects.

computational step the machine's tape, internal state, and head location have a single value. In contrast, Figure 4.4 shows that a NTM can assume an exponentially increasing number of definite machine values. This also distinguishes a NTM from a QTM; the latter device can be in a coherent superposition over exponentially many possible machine values. However, a theoretical measurement of the QTM state would yield a single value.

While a useful analytical tool, there isn't a known way to build a NTM. When we resort to solving an NP problem on a TM, the computation time frequently scales exponentially because these machines can't do much better than a brute force search through the branches shown in Figure 4.4[132]. Each possible strand can be computed efficiently by the TM, which explains why solutions to NP problems can be verified in polynomial time. However, solving the original problem involves searching through the exponentially many strands to see if any accept.

The archetypical example of an NP problem is factoring a number into its constituent primes. For convenience we discuss factoring in the better-known search problem form given above, though it can be recast as a decision problem. Factoring is in NP because it is easy to verify a candidate solution, by multiplying the numbers together, however there is no known efficient way to solve for the factors in the first place.

The wide variety of problems in NP are not equally difficult, the most challenging problems in this class are called NP-complete. An NP problem C is called NP-complete if any problem L in NP is *reducible* to C; solving L requires only a polynomial number of calls to a subroutine that solves C[2]. Consequently, if C can be efficiently solved then one can efficiently solve every problem in NP and P = NP. This explains why NP-complete problems are described as the hardest in NP. Proving, or more likely disproving, this last equality is the greatest unsolved problem in computational complexity theory.

We introduce a specific example of an NP-complete problem, circuit satisfiability (SAT), to make these concepts more concrete. The SAT problem is simple to articulate: given the Boolean equation $f(\mathbf{x})$, is there an $\mathbf{x_0}$ such that $f(\mathbf{x_0}) = 1$? For example, can we assign the elements of the 3-bit number $\mathbf{x}$ such that $f(\mathbf{x}) = (x_1 \vee x_2) \wedge (\neg x_2 \wedge x_3)$ evaluates to one? For this small instance of SAT, it's straightforward to find the answer is yes ($\mathbf{x_0} = [1, 0, 1]$). However, our intuition is to solve this problem by guess-and-check which leads to an exponentially increasing computation time as the problem size grows. However, it remains classically easy to evaluate $f(\mathbf{x_0})$ for any candidate solution $\mathbf{x_0}$, therefore SAT is in NP. While the details are not presented here, any problem in NP can be reduced to SAT so the latter is NP-complete. In contrast, if we replace 'is there an $\mathbf{x_0}$...?' with 'do all $\mathbf{x_0}$...?' in the problem statement then we obtain a coNP problem because then rejecting

answers are easy to verify.

These 'complete' problems exist in other complexity classes. In general, a problem C is X-complete if $C \in \mathsf{X}$ and any other problem in X is *reducible* to C. A problem H is described as X-hard if C is reducible to H. However, H does not actually need to be a member of the class X. Colloquially, H is 'at least as hard as all problems in X', while C is similar but must also be a member of X. Linking an entire complexity class to the specifics of a single problem, either via completeness or hardness, will later prove useful in the hardness analysis of boson sampling.

It is useful to have a methodology to isolate independent sources of complexity in a problem. For example, suppose we know a problem L is classically hard, however part of this problem requires solving SAT. Is the problem classically hard just because SAT is hard? Or is it even harder than this?

The concept of an *oracle* is useful for studying these situations. An oracle is a black-box that instantly solves a problem from a specified complexity class. A TM with access to an oracle can query it for answers in a single operation. The notation X^{Y} is used to describe those problems solvable by a TM for class X with access to an oracle for a complete problem for class Y. For example, P^{NP} indicates those problems that can be efficiently solved by a TM with an oracle for an NP-complete problem. Using our example in the previous paragraph, if $L \in \mathsf{P}^{\mathsf{NP}}$ then SAT is the only source of hardness in this problem. If we showed $L \in \mathsf{NP}^{\mathsf{NP}}$ then this problem is classically hard independent of SAT.

Oracle machines exhibit some counterintuitive properties. For example, while we know $\mathsf{NP} \subseteq \mathsf{P}^{\mathsf{NP}}$ it is strongly believed that actually $\mathsf{NP} \subset \mathsf{P}^{\mathsf{NP}}$. In other words, the class of problems efficiently solvable by a TM with an NP oracle is likely larger than the family of

NP problems! Care must be taken when relating classes of problems defined in terms of oracles.

The classes P, NP, and coNP form the lowest levels of a hierarchy of complexity classes shown in Figure 4.5 that are related to each other by oracles[133]. This polynomial hierarchy (PH) is central to the later complexity claims of the boson sampling problem. Figure 4.5 shows that P is at the base of PH since it contains the easiest computational problems. Anticipating the split of the PH into three strands, we label this lowest level $\mathsf{P} \equiv \Delta_0^{\mathsf{P}} \equiv \Sigma_0^{\mathsf{P}} \equiv \Pi_0^{\mathsf{P}}$. Recursive relations generate the higher levels of the hierarchy

$$\begin{aligned} \Delta_{i+1}^{\mathsf{P}} &\equiv \mathsf{P}^{\Sigma_i^{\mathsf{P}}} \\ \Sigma_{i+1}^{\mathsf{P}} &\equiv \mathsf{NP}^{\Sigma_i^{\mathsf{P}}} \\ \Pi_{i+1}^{\mathsf{P}} &\equiv \mathsf{coNP}^{\Sigma_i^{\mathsf{P}}}, \end{aligned} \tag{4.17}$$

for $i \geq 0$. If $\Sigma_{i+1}^{\mathsf{P}} = \Sigma_i^{\mathsf{P}}$ then PH is said to collapse to the i-th level. If $\mathsf{P} = \mathsf{NP}$ then the PH collapses entirely (to the zeroth level). It is strongly believed that all levels of PH are distinct and therefore any assertion that implies a collapse of PH at any level is very unlikely.

4.3.2 Fast classical solution of exact boson sampling collapses PH

If a classical computer could efficiently solve the boson sampling problem without any error, then PH collapses to the third level[112]. While this presumed hardness of boson sampling is related to the matrix permanent, PH collapse is not a trivial consequence of Equation (4.3). One must connect the complexity of boson *sampling* to the hardness of search and decision

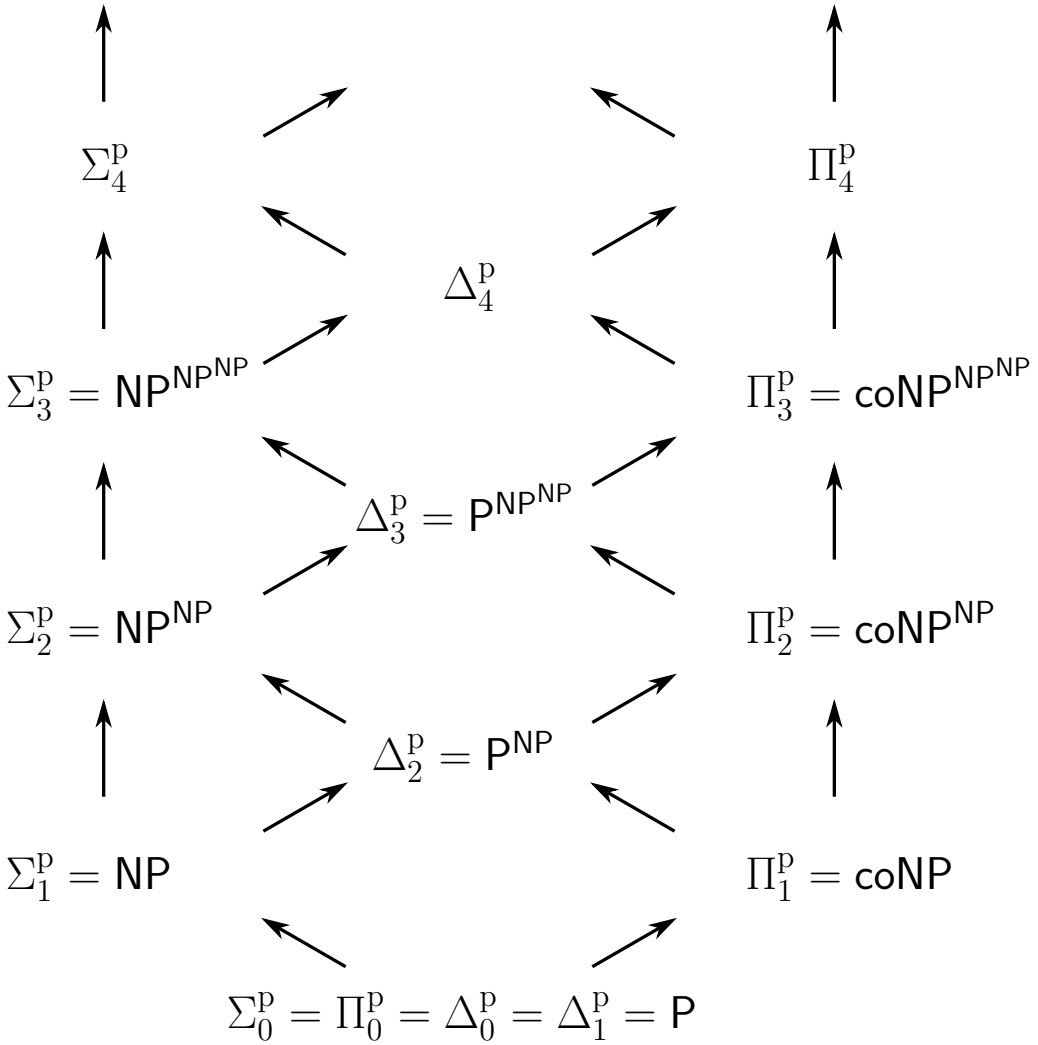


Figure 4.5: Starting with P,NP, and coNP at the base, one introduces subsequent tiers of the polynomial hierarchy (PH) by defining classes of problems solvable by oracle machines of increasing power. The recursive relations that generate higher levels of the hierarchy are defined in the text. Arrows indicate a subset relationship between classes, with P being the smallest set.

problems. This requires a more careful analysis that shows access to an efficient classical boson sampling computer would allow one to estimate permanents surprisingly fast [112]. In this section, and the following section on approximate boson sampling, we deliberately reuse the notation used by AA whenever possible [112].

Computing the permanent of an $n \times n$ matrix is a #P-hard problem [127]. Even when the matrix A takes the special form where all of its entries are either zero or one ($A \in \{0,1\}^{n\times n}$), computing $\mathrm{Per}(A)$ remains #P-complete. Problems in #P require counting the number of accepting paths in the NTM diagrammed in Figure 4.4. This is more computationally demanding than the class NP, which merely required finding if any of these paths accepted. While many of the other complexity classes introduced so far have been for decision problems, #P is a family of search problems because the output is not necessarily a single bit. Toda's theorem famously states that PH is contained within the class $\mathsf{P}^{\#\mathsf{P}}$ [134]. If there were an efficient classical algorithm for computing $\mathrm{Per}(A)$, a #P-complete problem, then $\#\mathsf{P} = \mathsf{P}$. As a result of Toda's theorem, we then find $\mathsf{PH} \subseteq \mathsf{P}^{\#\mathsf{P}} = \mathsf{P}^{\mathsf{P}} = \mathsf{P}$; in other words, PH collapses entirely. Since such a collapse is unlikely, it is widely believed that computing permanents is classically hard, even harder than NP-complete problems.

We cannot use a boson sampling machine to compute permanents, since the device outputs samples and not the distribution itself. However, we can use an ensemble of samples to estimate a permanent which, for certain families of matrices, remains classically hard. There exists an algorithm that efficiently estimates $\mathrm{Per}(X)$ to multiplicative error if all the elements of X are real and positive-valued [135]. If we drop the positivity requirement, such that $X \in \mathbb{R}^{n\times n}$, then approximating $\mathrm{Per}(X)^2 = |\mathrm{Per}(X)|^2$ to multiplicative error is #P-hard. We focus on $\mathrm{Per}(X)^2$ rather than $\mathrm{Per}(X)$ because observable probabilities in quantum

mechanics are given by the square of the scattering amplitude. One can repeatedly query a boson sampling machine to estimate $\mathrm{Per}(X)^2$. What must be clarified is the following: How quickly can this estimation be performed and with what error?

First we must resolve an incompatibility in problem statements, as boson sampling is framed in terms of a unitary transformation U while X is real-valued. However, given X, one can always efficiently find a larger U_x that contains X, within a global factor, as a submatrix. More formally, for $X \in \mathbb{R}^{n\times n}$ there exists an $m \times m$ unitary matrix U_x that contains ϵX as a submatrix, where $m \geq 2n$, $\epsilon \leq 1/\left\|X\right\|$, where $\left\|X\right\|$ is the largest singular value of X [112]. Given X, this U_x can be found in polynomial time.

Having efficiently reframed our problem in terms of unitary operations, we input U_x to a boson sampling machine which we repeatedly query for samples. Equation (4.3) shows that the boson distribution is given by permanents of submatrices of this unitary $U^{(\mathbf{S},\mathbf{T})}$. Our matrix X corresponds to one of these submatrices for a particular input and output state. It simplifies the analysis to assume, without loss of generality, that X lies in the top-left of U_x. In this case, $U^{(\mathbf{S},\mathbf{T})} = X$ when the input and output states both correspond to a single photon on the first n modes, written $|1_n\rangle$. Since we have control over the input state, we choose $|\mathbf{T}\rangle = |1_n\rangle$. Using Equation (4.3), we find the probability that the output state is also $|1_n\rangle$

$$p_x = \epsilon^{2n}|\mathrm{Per}(X)|^2. \tag{4.18}$$

As we draw an increasing number of samples, we obtain an increasingly accurate estimate of p_x by observing how often we observe the output $|1_n\rangle$.

We want to relate the computational complexity of boson sampling to PH, a hierarchy defined in terms of oracles. We therefore introduce the oracle for the boson sampling problem $\mathcal{O}(U, \mathbf{T}, r)$ which accepts three inputs: a unitary matrix U, the input bosonic state $\mathbf{T}$, and a random number r that allows the oracle to output a random sample drawn from the boson distribution. We have dropped the ket notation from $\mathbf{T}$ because we no longer assume quantum operation, $\mathbf{T}$ is merely a binary string specifying the input state. We can re-express p_x from Equation (4.18) as the probability that our oracle outputs 1_n given a unitary U_x and input 1_n

$$p_x = \Pr_r \left[\mathcal{O}(U_x, 1_n, r) = 1_n\right], \tag{4.19}$$

where the probability over all r is denoted by the subscript.

To conclude this complexity analysis we require a theorem regarding Boolean functions. We follow AA's restatement [112] of a famous result from Stockmeyer [136].

Theorem 1. *Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, let*

$$p = \Pr_{x \in \{0,1\}^n} [f(x) = 1] = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x). \tag{4.20}$$

Then for all $g \geq 1 + \frac{1}{\text{poly}(n)}$, there exists an $\mathsf{FBPP}^{\mathsf{NP}^f}$ machine that approximates p to within a multiplicative factor of g.

We paraphrase this result. The function f has 2^n possible n-bit inputs and outputs either 0 or 1. The overall probability p that the function outputs 1 is equal to the arithmetic average over all 2^n possible inputs. We introduce the class FBPP of those search problems

efficiently solvable by a PTM.[5] The probability p can be efficiently estimated to within a factor g by an $\mathsf{FBPP}^{\mathsf{NP}^f}$ machine. This is a search, not decision, problem since in general p cannot be expressed as a single bit.

We observe that Equations (4.19) and (4.20) are similar; both express the probability of a Boolean function accepting. To make the relationship more explicit, we observe that $\mathcal{O}(U_x, 1_n, r)$ can be treated as a Boolean function whose outputs are either 1_n (accepting) or 'not 1_n' (rejecting). As a consequence, estimating p_x is in the class $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$. However, Equation (4.18) shows that this implies estimating $\mathrm{Per}(X)^2$ is also in the class $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$.

We have observed that estimating $\mathrm{Per}(X)^2$ is in both the class $\#\mathsf{P}-$hard and $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$. We outline how this dual-membership implies the collapse of PH. Consider a machine that is performing a $\mathsf{P}^{\#\mathsf{P}}$ computation. Every query the machine makes to the $\#\mathsf{P}$ oracle can be reduced, in polynomial time, to the $\#\mathsf{P}-$hard problem of estimating $\mathrm{Per}(X)^2$. This is true by definition; $\#\mathsf{P}$ problems are efficiently reducible to $\#\mathsf{P}-$hard problems. However, we know that estimating $\mathrm{Per}(X)^2$ can be done in $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$ from our earlier analysis. In addition, a BPP machine can simulate an FBPP machine at intermediate stages of computation. Both are TMs with a random number generator, they only differ in the form of the final output. We accordingly can estimate $\mathrm{Per}(X)^2$ in $\mathsf{BPP}^{\mathsf{NP}^{\mathcal{O}}}$ which leads to the final result $\mathsf{P}^{\#\mathsf{P}} \subseteq \mathsf{BPP}^{\mathsf{NP}^{\mathcal{O}}}$.

There are surprising consequences if exact boson sampling could be efficiently performed by a TM. This assumption implies $\mathsf{P}^{\#\mathsf{P}} \subseteq \mathsf{BPP}^{\mathsf{NP}}$. We then relate both sides of this equation to PH and note that Toda's theorem says that $\mathsf{PH} \subseteq \mathsf{P}^{\#\mathsf{P}}$ [134] while it's known

[5] Search problems are frequently, and misleadingly [137], described as function problems. This explains why the notation for classes of search problems frequently begin with F. For example, the class FBPP contains search problems, while BPP contains decision problems, that are efficiently solvable with a PTM.

that $\mathsf{BPP} \subseteq \Sigma_2^P$ [138]. Assembling these relations together yields

$$\mathsf{PH} \subseteq \mathsf{P}^{\#\mathsf{P}} \subseteq \mathsf{BPP}^{\mathsf{NP}} \subseteq (\Sigma_2^{\mathsf{P}})^{\mathsf{NP}} \equiv \Sigma_3^{\mathsf{P}}. \tag{4.21}$$

However, by definition, $\Sigma_3^{\mathsf{P}} \subseteq \mathsf{PH}$ and therefore equality holds throughout Equation (4.21) so that $\mathsf{PH} = \Sigma_3^{\mathsf{P}}$. Thus, if boson sampling was efficiently solvable by a TM then PH collapses to the third level.

4.3.3 Fast classical solution of approximate boson sampling likely collapses PH

A realistic quantum device will inevitably be imperfect and thus only sample approximately from the true boson distribution. For a fair comparison, we hold a classical computer to the same standard and allow it some error. Since the sampling oracle above was assumed to exactly sample from the boson distribution, a new approach is required to show the collapse of PH. While AA took a significant step towards proving PH collapse in this approximate case, the proof relies on unproven, but plausible, conjectures.

We first introduce some notation that is essential to outlining the arguments which follow.

One of the inputs to the boson sampling problem is a random unitary matrix. This randomness is essential in preventing a classical machine from potentially exploiting structure in the matrix to efficiently sample from the boson distribution. This randomness was not required in the hardness proof of exact boson sampling, though it will prove crucial in the approximate case here. An $n \times n$ unitary matrix can be parameterized by n^2 real numbers

θ_i where $1 \leq i \leq n^2$. We define a measure, $d\mu(g)$ such that the probability of drawing a unitary matrix from some subset S of all possible $n \times n$ unitary matrices $U_{n,n}$ is given by

$$\mu(S) = \int d\mu(S) = \int .... \int f(\theta_1, ..., \theta_{n^2}) d\theta_1 ... d\theta_{n^2}, \tag{4.22}$$

where $f(\theta_1, ..., \theta_{n^2})$ is the probability density function and the limits of integration (not shown) define the subset S. To obtain random unitaries, we seek a probability measure that is uniform. As a consequence, the probability of selecting a unitary from a specific subset S should only be a function of the volume of S, not its specific location within the set $U_{n,n}$. This shift-invariance of the probability is captured by the Haar measure, which imposes the condition

$$d\mu(U_0 S) = d\mu(S U_0) = d\mu(S), \tag{4.23}$$

where $U_0 \in U_{n,n}$ and produces an arbitrary shift within the group. A random unitary matrix *drawn from the Haar measure* is chosen from a probability distribution taking the form of Equation (4.22), where $d\mu(S)$ satisfies Equation (4.23). These Haar random matrices are interesting not only for the lack of structure that precludes algorithmic shortcuts in permanent estimation but because they naturally arise in linear optics. When a large linear optical network is constructed out of beam splitters and phase offsets with randomly chosen values, the resulting linear transformation resembles a Haar-random unitary [139].

We illustrate these concepts by reviewing the Haar measure in the simplest case of $U_{1,1}$ unitaries. In this case, a single number, θ_1, parametrizes elements of this group, which are written $e^{i\theta_1}$. The group $U_{1,1}$ describes a circle in the complex plane and the most obvious

measure is the infinitesimal arc segment $d\mu(S) = d\phi$. Translations in this group are produced with the operation $U_0 = e^{i\phi_0}$ yielding the new element $U_0 U = e^{i(\phi+\phi_0)}$. The Haar condition is satisfied since $d(U_0 S) = d(\phi + \phi_0) = d\phi = d\mu(S)$. For higher dimensional systems $U_{n,n}$, there are systematic methods for determining the Haar measure[139,140]. They are not reviewed here because their specific functional form is not important to the arguments below.

Haar-random unitaries have the convenient property that sufficiently small subsections resemble identical and independently distributed (iid) Gaussian matrices. We introduce the notation for these two distributions of matrices. The distribution of $n \times n$ submatrices of $m \times m$ Haar-random unitaries is written $\mathcal{S}_{m,n}$. These submatrices are rescaled by a factor of $\sqrt{m}$ so that they have a mean of zero and variance of one. The distribution of iid Gaussian matrices with a mean of zero and variance of one is written $\mathcal{G}^{n\times n}$. If n is sufficiently small relative to m, then $\mathcal{S}_{m,n}$ is close in L_1 distance to $\mathcal{G}^{n\times n}$. An algorithm can take advantage of this similarity to hide a nearly iid Gaussian matrix in a larger Haar-random matrix. The notation $a \sim b$, which indicates that a is randomly drawn from the distribution b, allows us to formalize these claims. There exists a $\mathsf{BPP}^{\mathsf{NP}}$ algorithm that takes as input $X \sim \mathcal{G}^{n\times n}$ and, with high probability, outputs a Haar-random matrix $A \sim \mathcal{S}_{m,n}$ that contains $X/\sqrt{m}$ as a random submatrix.

Gaussian Permanent Estimation

The key result of AA was their ability to link the ability of classical computers to approximately sample from boson distributions with their ability to estimate permanents of Gaussian matrices[112]. They refer to the latter as the Gaussian Permanent Estimation

(GPE) problem.

Problem 1($|\mathbf{GPE}|^2_{\pm}$) *Given as input a matrix $X \sim \mathcal{N}(0,1)^{n\times n}_{\mathbb{C}}$ of i.i.d. Gaussians, together with error bounds $\epsilon, \delta > 0$, estimate $|\mathrm{Per}(X)|^2$ to within additive error $\pm\epsilon \cdot n!$, with probability at least $1-\delta$ over X, in $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ time.*

Their main result then links $|\mathrm{GPE}|^2_{\pm}$ to approximate boson sampling.

Theorem 2. *Let $\mathcal{D}_A$ be the probability distribution sampled by a boson computer A. Suppose there exists a classical algorithm $\mathcal{C}$ that takes as input a description of A as well as an error bound ϵ, and that samples from a probability distribution $\mathcal{D}'_A$ such that $\|\mathcal{D}'_A - \mathcal{D}_A\| \leq \epsilon$ in $\mathrm{poly}(|A|, 1/\epsilon)$ time [$|A|$=number of bosons]. Then the $|\mathrm{GPE}|^2_{\pm}$ problem is solvable in $\mathsf{BPP}^{\mathsf{NP}}$. Indeed, if we treat $\mathcal{C}$ as a black box, then $|\mathrm{GPE}|^2_{\pm} \in \mathsf{BPP}^{\mathsf{NP}^{\mathcal{C}}}$.*

We can estimate Gaussian permanents by inputting a large Haar random unitary matrix A to an *approximate* boson sampling oracle and then counting the fraction of times that it outputs a sample corresponding to the desired submatrix X. This strategy is very similar to that employed in Section 4.3.2 to prove the hardness of exact boson sampling. In the current approximate case, the boson sampling oracle receives an $m \times m$ matrix A and an input state of n bosons that we choose to target the rows of A that contain X. We write $A \in \mathcal{U}^{n\times m}$ to express that A can be truncated to an $n \times m$ matrix given this knowledge of the input state, where $\mathcal{U}^{n\times m}$ is the group of $n \times m$ row-orthonormal matrices with complex entries. We repeatedly draw samples from the oracle and track how often these outputs correspond to the columns of A that contain X. Using Equation (4.3), this gives an estimate of $|\mathrm{Per}(X)|^2$.

This reliance on an approximate boson sampling oracle introduces a complication in our attempt to link boson sampling with $|\mathrm{GPE}|^2_{\pm}$. While the oracle is guaranteed to sample with some accuracy over the entire distribution, it might not exhibit uniform error. The distribution from which it samples might be very accurate for some output states and inaccurate for others, which is problematic if the latter corresponds to our desired output! This can be rephrased in terms of an adversarial oracle that does not want PH to collapse and therefore attempts to spoil our efforts to estimate $|\mathrm{Per}(X)|^2$. Given A as an input, this oracle attempts to find which submatrix is X and then samples very accurately from the boson distribution except for the single element of that distribution corresponding to X.

Resolving this issue relies on our ability to hide an iid Gaussian matrix X inside the unitary A. When the oracle receives A and a description of the N boson input, it gains little information about which of the M columns of A contain X because small subsections of Haar-random unitaries closely resemble iid Gaussians[112]. In other words, the adversarial oracle knows our iid Gaussian X is in A somewhere. However, every subsection of A looks close to an iid Gaussian matrix and the oracle cannot quickly find X. To satisfy its global sampling accuracy requirement, this ignorance forces the oracle to sample accurately across the entire distribution including that element corresponding to $|\mathrm{Per}(X)|^2$. AA showed that this is sufficient to prove the link between boson sampling and $|\mathrm{GPE}|^2_{\pm}$ in Theorem 2.

We outline in more detail the steps reviewed above. The individual who wishes to solve $|\mathrm{GPE}|^2_{\pm}$ knows X. They can generate $A \in U^{n\times m}$ that contains $X' \equiv X/\sqrt{m}$ with high probability in $\mathsf{BPP}^{\mathsf{NP}}$. The original $|\mathrm{GPE}|^2_{\pm}$ problem can then be recast as estimating $|\mathrm{Per}(X)|^2 = m^n|\mathrm{Per}(X')|^2$. We consolidate the $|\mathrm{GPE}|^2_{\pm}$ tolerance parameters into $\beta \propto \epsilon\delta$. The boson sampling oracle is then given the inputs A, β, along with a random variable

r. By definition, it then samples from a distribution $\mathcal{D}'_A$ such that $\|\mathcal{D}_A - \mathcal{D}'_A\| \leq \beta$. The probability that our oracle outputs the sample that corresponds to $|\mathrm{Per}(X)|^2$ is labeled q_{S^*} and takes the form

$$q_{S^*} = \Pr_r \left[\mathcal{O}(A, \beta, r) = S^*\right]. \tag{4.24}$$

We can again use Theorem 1 to approximate q_{S^*} in $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$; this approximation is labeled $\tilde{q}_{S^*}$. However, Equation (4.3) shows that the probability of sampling the output state S^* is also given by the matrix permanent

$$p_{S^*} = |\mathrm{Per}(A_{S^*})|^2 = |\mathrm{Per}(X')|^2. \tag{4.25}$$

The $|\mathrm{GPE}|^2_{\pm}$ proof concludes by showing that $\tilde{q}_{S^*}$ and p_{S^*} are close in the sense that

$$\Pr\left[|\tilde{q}_{S^*} - p_{S^*}| > \epsilon \cdot \frac{n!}{m^n}\right] < \delta. \tag{4.26}$$

Therefore, the probability that our $\mathsf{FBPP}^{\mathsf{NP}^{\mathcal{O}}}$ approximation ($\tilde{q}_{S^*}$) to the permanent of an iid Gaussian matrix (p_{S^*}) has large error ($> \epsilon \cdot \frac{n!}{m^n}$) is minimal ($< \delta$). This concludes the proof of Theorem 2.

Conjectured Collapse of PH

Efficient classical solution of the approximate boson sampling problem implies $|\mathrm{GPE}|^2_{\pm}$ can be solved in $\mathsf{FBPP}^{\mathsf{NP}}$. As noted in 4.3.2, a BPP can efficiently simulate an FBPP machine at intermediate stages of computation which indicates that $|\mathrm{GPE}|^2_{\pm}$ can be solved in $\mathsf{BPP}^{\mathsf{NP}}$. This does not collapse PH. If two further conjectures prove true, then this same permanent

estimation problem also belongs to the class $\mathsf{P}^{\#\mathsf{P}}$ and PH again collapses to the third level by reasoning similar to that presented at the end of Section 4.3.2.

We introduce a variant on the permanent estimation problem that later proves useful as it may be more naturally related to the class $\#\mathsf{P}$ than the $|\mathrm{GPE}|^2_{\pm}$ problem discussed so far.

Problem 2 ($\mathbf{GPE}_{\times}$) *Given as input a matrix $X \sim \mathcal{N}(0,1)^{n\times n}_{\mathbb{C}}$ of i.i.d. Gaussians, together with error bounds $\epsilon, \delta > 0$, estimate $\mathrm{Per}(X)$ to within error $\pm\epsilon \cdot |\mathrm{Per}(X)|$, with probability at least $1-\delta$ over X, in $\mathrm{poly}(n, 1/\epsilon, 1/\delta))$ time.*

There are two key differences between $\mathrm{GPE}_{\times}$ and $|\mathrm{GPE}|^2_{\pm}$. The former gives a multiplicative rather than additive estimate of the permanent of a Gaussian matrix. In addition, it estimates $\mathrm{Per}(X)$, a complex number, rather than the real-valued $|\mathrm{Per}(X)|^2$.

We want to show that $\mathrm{GPE}_{\times}$ and $|\mathrm{GPE}|^2_{\pm}$ problems are polynomial-time equivalent. This means that a problem in either class can be efficiently solved, given access to an oracle from the other class. One can straightforwardly prove that any problem in $|\mathrm{GPE}|^2_{\pm}$ can be efficiently solved given access to a $\mathrm{GPE}_{\times}$ oracle. Proving the reverse, that $\mathrm{GPE}_{\times}$ instances can be solved in polynomial time given a $|\mathrm{GPE}|^2_{\pm}$ oracle, is more involved. One obvious complication is that $\mathrm{GPE}_{\times}$ requires the estimation of a complex number but our oracle only outputs real values. If one can show that access to a $|\mathrm{GPE}|^2_{\pm}$ oracle allows a sufficiently accurate estimation of the phase $\theta = \mathrm{Per}(X)/|\mathrm{Per}(X)|$, then $\mathrm{GPE}_{\times}$ is reducible to $|\mathrm{GPE}|^2_{\pm}$.

An accurate estimation of θ is possible if $|\mathrm{Per}(X)|$ is generally not too close to zero for $X \sim \mathcal{G}^{n\times n}$. This statement is formalized in the Permanent Anti-Concentration Conjecture (PACC).

Conjecture 1 (Permanent Anti-Concentration Conjecture-PACC) *For all n and $\epsilon > 0$,*

$$\Pr_{X \sim \mathcal{N}(0,1)_{\mathbb{C}}^{n \times n}} \left[|\mathrm{Per}(X)| < \epsilon \sqrt{n!} \right] < C n^D \epsilon^\beta, \tag{4.27}$$

where C,D, and β are positive constants.

There is strong evidence that PACC is true. An analogous anti-concentration result with the determinant replacing the permanent has been proved. Numerical comparison of the probability density functions of $|\mathrm{Det}(X)|$ and $|\mathrm{Per}(X)|$ reveal that the permanent appears even less concentrated about zero than the determinant. It is therefore plausible that PACC is true.

Assuming PACC is true, then a $|\mathrm{GPE}|_{\pm}^2$ oracle can be used to efficiently solve $\mathrm{GPE}_\times$ problems, though this still does not collapse PH. For this, we require a final conjecture that $\mathrm{GPE}_\times$ is $\#\mathsf{P}$-hard.

Conjecture 2 (Permanent-of-Gaussians Conjecture-PGC) $\mathrm{GPE}_\times$ *is $\#\mathsf{P}$-hard. In other words, if $\mathcal{O}$ is any oracle that solves* $\mathrm{GPE}_\times$*, then* $\mathsf{P}^{\#\mathsf{P}} \subseteq \mathsf{BPP}^{\mathcal{O}}$.

It has been proved that exactly computing $\mathrm{Per}(X)$ for most matrices $X \sim \mathcal{N}(0,1)_{\mathbb{C}}^{n \times n}$ is $\#\mathsf{P}$-hard. However, it is unknown how to extend this to the approximation of these permanents, hence the hardness of $\mathrm{GPE}_\times$ remains a conjecture.

If Conjectures 1 and 2 prove true, then an efficient classical solution of the approximate boson sampling problem would imply the collapse of PH to the third level. The line of reasoning is presented in Figure 4.6.

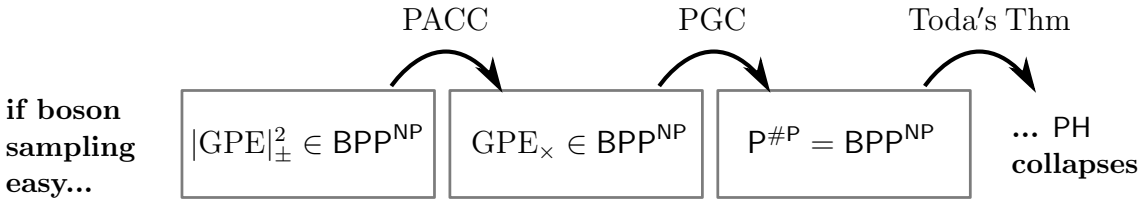


Figure 4.6: There would be surprising consequences if we assume boson sampling could be solved in polynomial time by a classical computer. Starting with this assumption, one can prove that $|\mathrm{GPE}|^2_{\pm} \in \mathsf{BPP}^{\mathsf{NP}}$. If the PACC and PGC are true, then $\mathsf{P}^{\#\mathsf{P}} = \mathsf{BPP}^{\mathsf{NP}}$ and, by Toda's theorem, this collapses PH to the third level. Therefore, approximate boson sampling is likely classically hard.

Chapter 5

Demonstration of boson sampling

We performed one of the first proof-of-principle demonstrations of boson sampling [25,141–143]. Losses distributed throughout any boson sampling experiment play a pivotal role in our results and are the subject of the first half of this chapter. In Section 5.1, we formalize how these losses affect our boson distribution. A simple case study is included to make these ideas more tangible. Section 5.2 presents the circuit characterization method used in our experiment that correctly accounts for these losses.

In the latter half of this chapter, we present our boson sampling results. While our QBSM does not surpass the limits of classical computers, in Section 5.3 we successfully demonstrate that matrix permanents correctly predict the observed boson distribution. This is important since this assumption underlies the computational hardness proofs of boson sampling presented in the previous chapter. In Section 5.4, we identify the source and approximate magnitude of errors in our experiment. Diagnosing these imperfections in QBSMs will be critical in scaling to larger systems.

5.1 Effect of loss

5.1.1 Lossy boson distribution still given by matrix permanent

The boson sampling problem shown in Figure 4.2 calls for a unitary transformation [112]. Due to inevitable experimental losses, QBSMs won't satisfy this requirement. Generalizing the computational hardness proofs in Sections 4.3.2-4.3.3 to this case of nonzero loss remains an important open problem for complexity theorists. However, in this section we demonstrate that boson distributions are still a function of matrix permanents in this case of non-unitary evolution.

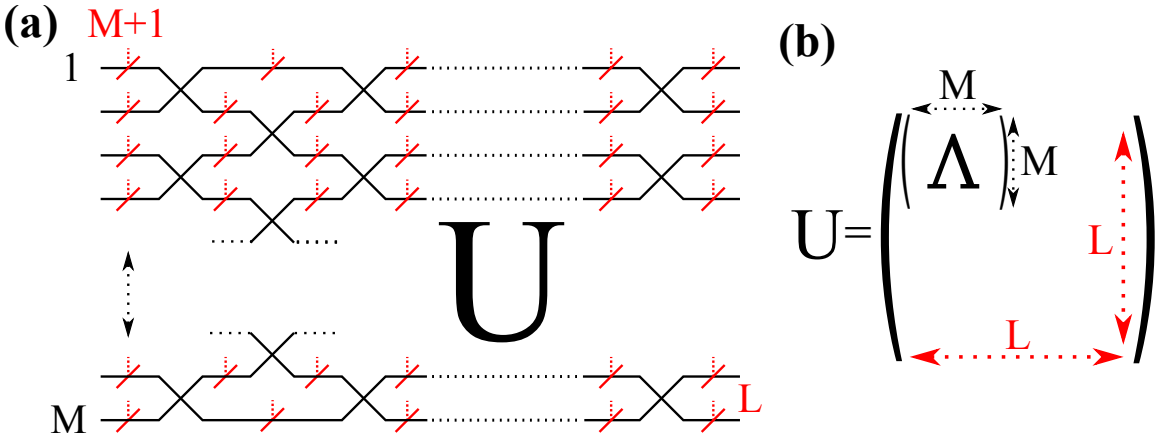


Figure 5.1: (a) A general linear optical network is modeled by a series of loss beam splitters (red) linking accessible modes (black lines) to loss modes (red dashed). We label the accessible modes $1 \leq i \leq M$, while loss modes are labeled $M+1 \leq i \leq L$. (b) The $L \times L$ matrix describing this transformation is unitary. However, the top-left $M \times M$ submatrix Λ describing the transformation over the accessible modes observed by the experimenter is not unitary.

We wish to find the analogue of Equation (4.3) for this lossy boson sampling scenario. Figure 2.3 showed that optical loss can be modeled by a beam splitter where $(1-\eta^2)$ is equivalent to the power loss. Figure 5.1 illustrates that losses throughout a linear optical circuit can be modeled by adding many such loss beam splitters with, in general, different

reflectivities. There are a total of L modes in this system. Without loss of generality, we label the first $1 \leq x \leq M$ modes as accessible to the experimenter with the remainder describing loss modes. Figure 5.1 then shows that the observed non-unitary transformation is the top-left $M \times M$ sub-matrix of an $L \times L$ unitary transformation. We input N photons into modes $1 \leq x \leq N$. We can then rewrite Equation (4.8)

$$P_S = \left| \langle 0| \prod_{i=1}^{N} \hat{b}_{\underline{S}_i} \left[\sum_{j=1}^{L^N} \left(\prod_{k=1}^{N} U_{k,\widetilde{V}_k^j} \hat{b}^{\dagger}_{\widetilde{V}_k^j} \right) \right] |0\rangle \right|^2, \tag{5.1}$$

where $1 \leq \widetilde{V}_k^j \leq L$ because photons can emerge on any of the L modes. Despite this difference, using the commutation relation in Equation (4.13) to simplify Equation (5.1) still yields

$$P_S = \left| \sum_{j=1}^{N!} \prod_{k=1}^{N} U_{k,\underline{\widetilde{S}}_k^j} \right|^2 = |\mathrm{Per}(\mathbf{U}^{(\mathbf{S},\mathbf{T})})|^2, \tag{5.2}$$

where we have merely rewritten Equation (4.14). Our detectors only monitor the M accessible modes so $1 \leq \underline{\widetilde{S}}_k^j \leq M$, which restricts $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$ to the first M columns of U. In addition, because we input photons on the first N modes, then $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$ is restricted to the first N rows of $\mathbf{U}$. It is convenient to actually keep the first M rows of $\mathbf{U}$ corresponding to the accessible modes, rather than just the subset of N modes where photons are input. Thus while $\mathbf{U}$ is an $L \times L$ matrix, we can retain only the $M \times M$ submatrix Λ in the top-left of $\mathbf{U}$ which describes the transformation over the accessible modes. We use this notation to rewrite Equation (5.2)

$$P_S = |\mathrm{Per}(\Lambda^{(\mathbf{S},\mathbf{T})})|^2. \tag{5.3}$$

The possibility of more than one photon per mode on the input or output is addressed by

including the denominator in Equation (4.3) in Equation (5.3).

5.1.2 Postselected boson sampling with loss: a case study

Equation (5.3) shows that the boson distribution is a function of the non-unitary matrix Λ. However, all boson sampling experiments, including the one detailed below, have used post-selection; if we seek to perform an N-photon sampling experiment then we require that N detectors fire for a trial to be included in the data set. Do we still use Equation (5.3) to predict this boson distribution? Or does postselection allow us to work in terms of an effectively unitary (lossless) transformation so that Equation (4.3) applies? In other words, does post-selecting on trials where no photons are lost remove the need to characterize these losses?

We present a case-study of the post-selected boson distribution observed from a lossy circuit to address these questions. While we simplify the algebra by restricting the losses to a special form, the analysis remains sufficiently general to draw important conclusions that apply to more realistic scenarios. We first introduce notation for the decomposition of a linear optical circuit before proceeding to the details of the case study.

Decomposition of a linear optical network

Any linear optical network can be assembled from two types of elements: beam splitters and phaseshifters. A lossy linear optical network is comprised of these same elements, with losses represented by a subset of the beam splitters which couple to loss modes. We retain the modal notation in Figure 5.1 such that there are M accessible modes in the system with L total modes.

We first define the action of these basic elements on the L-dimensional space. By generalizing the transformation in Equation (1.12), we can express the elements of the matrix describing a beam splitter on this larger space

$$\begin{aligned}(U_{\rm bs}(i,j,\eta))_{k,l} = \delta_{k,l}(1-\delta_{k,i})(1-\delta_{k,j}) + \eta\,(\delta_{k,i}\delta_{l,i} - \delta_{k,j}\delta_{l,j}) \\ + \sqrt{1-\eta^2}\,(\delta_{k,i}\delta_{l,j} + \delta_{k,j}\delta_{l.i}),\end{aligned} \tag{5.4}$$

where the beam splitter links modes i, j with reflectivity η[144]. Equation (5.4) describes the identity matrix where the elements (i,i) and (j,j) are replaced by η and $-\eta$ respectively, while the (i,j) and (j,i) elements are replaced by $\sqrt{1-\eta^2}$. The form of this matrix is sketched in Equation (5.5) with the nontrivial rows and columns labeled by i and j

$$U_{\rm bs}(i,j,\eta) = \begin{array}{c} \\ \\ i \\ \\ \\ j \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} & & & i & & j & & \end{array} \\ \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & \eta & & \sqrt{1-\eta^2} & & \vdots \\ \vdots & & & & & & \vdots \\ \vdots & & \sqrt{1-\eta^2} & & -\eta & & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} \end{array}. \tag{5.5}$$

The analogous expression for a phaseshift of ϕ on mode i takes the simpler form

$$(U_{\rm ps}(i,\phi))_{k,l} = e^{i\phi}\delta_{k,i}\delta_{l,i} + \delta_{k,l}(1-\delta_{k,i}). \tag{5.6}$$

Equation (5.6) describes the identity matrix but with the (i,i) element replaced by $e^{i\phi}$ [144]. The form of this matrix is sketched in Equation (5.7) with the i-th row and column labeled

$$U_{\mathrm{ps}}(i,\phi) = \begin{array}{c} \\ i \end{array} \begin{array}{c} \begin{array}{ccccc} & & i & & \end{array} \\ \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & e^{i\phi_{ii}} & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \end{array}. \tag{5.7}$$

The overall $L \times L$ unitary transformation U is calculated from the matrix product of its K constituent elements. This allows us to write

$$\mathbf{U} = \prod_{z=1}^{K} U_{\alpha_z}(x_z), \tag{5.8}$$

where α_z is either PS or BS while x_z contains the arguments of Equation (5.4) or (5.6) appropriate to this z-th element [144].

Case study

If we know all the individual elements in our photonic circuit including the losses, then we can use Equation (5.8) to construct $\mathbf{U}$. We can then isolate the submatrix Λ describing the transformation over the accessible modes, and use Equation (5.3) to predict the boson distribution.

It is difficult to characterize optical losses embedded within an interferometric network.

As a result, one might have accurately characterized all the beam splitters and phaseshifters over the accessible modes but remain ignorant about the losses. We restate this assertion using the notation of Equation (5.5); one may know $U(x_z)$ if the modal labels $1 \leq i \leq M$ and $1 \leq j \leq M$. However, the optical losses, described by beam splitters with a mode label in the range $M+1 \leq j \leq L$, are unknown.

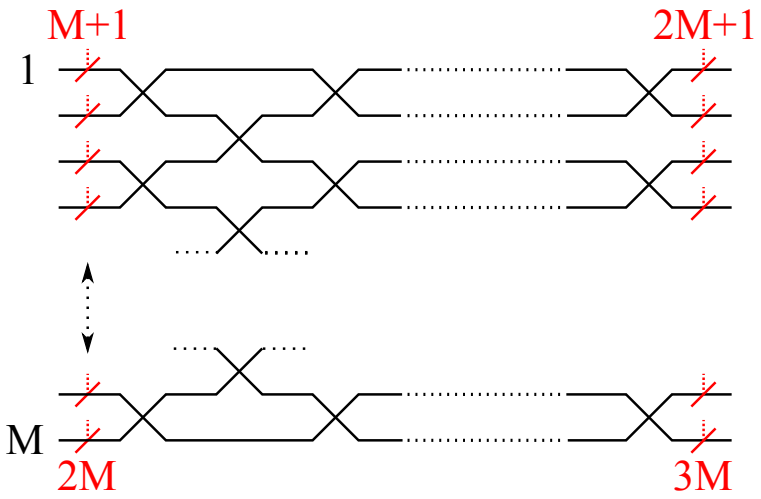


Figure 5.2: A simplified model allows for losses (red) before and after an interferometric network (black). There are no losses within the network itself. The numbering convention for the M accessible modes (black) and $2M$ total loss modes (red) are indicated.

We present a simple case-study to analyze the effect of these unbalanced losses on our postselected experiment. In particular we answer the question: Does postselection remove the need to characterize loss in our system? The answer informs our later choice of circuit characterization method. Our case-study consists of an M-mode linear network, shown in Figure 5.2, with losses at the input and output but not within the interferometric network itself. This special case of a lossy network allows us to rewrite Equation (5.8)

$$U = \prod_{\beta=1}^{M} U_{\mathrm{BS}}(x_\beta) \prod_{\gamma=1}^{K} U_{\alpha_\gamma}(x_\gamma) \prod_{\delta=1}^{M} U_{\mathrm{BS}}(x_\delta), \tag{5.9}$$

where the product over β and δ indicates losses at the input and output of the linear network respectively. This network is represented by the middle product over γ and is lossless.

Our boson sampling measurements are only affected by the $M \times M$ submatrix Λ in the top-left of U that represents the transformation over the accessible modes. The effect of the losses in this subspace takes a simple form. Expressing transformations as matrices in block form provides a convenient method of isolating subspaces. There are $3M$ modes in total with the following numbering convention: accessible modes are labeled $1 \leq n \leq M$, the loss modes on the input are $M{+}1 \leq n \leq 2M$, and the loss modes on the output are $2M{+}1 \leq n \leq 3M$. These mode labels are illustrated in Figure 5.2. We can then express each of the products in Equation (5.9) as a 3×3 matrix where each element represents a transformation amongst these M-dimensional spaces

$$U = \begin{pmatrix} \overline{\eta}_{11} & \overline{\eta}_{12} & 0 \\ \overline{\eta}_{21} & \overline{\eta}_{22} & 0 \\ 0 & 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \overline{U}_c & 0 & 0 \\ 0 & \mathbb{1} & 0 \\ 0 & 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \overline{\eta}'_{11} & 0 & \overline{\eta}'_{13} \\ 0 & \mathbb{1} & 0 \\ \overline{\eta}'_{31} & 0 & \overline{\eta}'_{33} \end{pmatrix}. \tag{5.10}$$

The middle matrix describes the lossless transformation amongst the accessible modes and therefore only has the non-trivial element $\overline{U}_c$. The form of the matrices on the left and right, describing the losses on the input and output respectively, are a consequence of Equation (5.4) and our mode numbering convention described above. By multiplying the matrices in Equation (5.10) and retaining the top-left element, we find $\Lambda = \overline{\eta}_{11}\overline{U}_c\overline{\eta}'_{11}$.

We determine the effect of loss on our boson distribution by expanding Λ into its full form. Equation (5.4) shows that $\overline{\eta}_{11}$ and $\overline{\eta}'_{11}$ are diagonal matrices. This simplifies the form

of Λ

$$\Lambda = \begin{pmatrix} \eta_1^{\text{in}} & 0 & \cdots & 0 \\ 0 & \eta_2^{\text{in}} & & \\ \vdots & & \ddots & \\ 0 & & & \eta_{\text{M}}^{\text{in}} \end{pmatrix} \begin{pmatrix} \text{U}_{11} & \text{U}_{12} & \dots & \text{U}_{1\text{M}} \\ \text{U}_{21} & \text{U}_{22} & & \\ \vdots & & \ddots & \\ \text{U}_{\text{M}1} & & & \text{U}_{\text{MM}} \end{pmatrix} \begin{pmatrix} \eta_1^{\text{out}} & 0 & \cdots & 0 \\ 0 & \eta_2^{\text{out}} & & \\ \vdots & & \ddots & \\ 0 & & & \eta_{\text{M}}^{\text{out}} \end{pmatrix}$$

$$= \begin{pmatrix} \eta_1^{\text{in}}\eta_1^{\text{out}}\text{U}_{11} & \eta_1^{\text{in}}\eta_2^{\text{out}}\text{U}_{12} & \dots & \eta_1^{\text{in}}\eta_{\text{M}}^{\text{out}}\text{U}_{1\text{M}} \\ \eta_2^{\text{in}}\eta_1^{\text{out}}\text{U}_{21} & \eta_2^{\text{in}}\eta_2^{\text{out}}\text{U}_{22} & & \\ \vdots & & \ddots & \\ \eta_{\text{M}}^{\text{in}}\eta_1^{\text{out}}\text{U}_{\text{M}1} & & & \eta_{\text{M}}^{\text{in}}\eta_{\text{M}}^{\text{out}}\text{U}_{\text{MM}} \end{pmatrix}. \tag{5.11}$$

This is the final expression for Λ we will use to observe the effect of loss on our boson distribution.

We use postselection in our boson sampling experiment which implies we don't attempt to estimate the absolute probability of a given outcome given by Equation (5.3). We instead select a subset of trials and compute the relative probability of measuring the state $|\mathbf{S}\rangle$ compared to $|\mathbf{S}'\rangle$

$$P_{\mathbf{S},\mathbf{S}'} = \frac{\left|\text{Per}(\Lambda^{(\mathbf{S},\mathbf{T})})\right|^2}{\left|\text{Per}(\Lambda^{(\mathbf{S}',\mathbf{T})})\right|^2}. \tag{5.12}$$

We note that the two submatrices, $\Lambda^{(\mathbf{S},\mathbf{T})}$ and $\Lambda^{(\mathbf{S}',\mathbf{T})}$, are both drawn from the first N rows of Λ since the input state $|\mathbf{T}\rangle$ is constant throughout the experiment. Using the matrix permanent expansion from Equation (4.14) and the form of Λ from Equation (5.11), we

rewrite Equation (5.12)

$$P_{\mathbf{S},\mathbf{S}'} = \frac{\left|\sum_{j=1}^{N!}\prod_{k=1}^{N}\eta_{S_k}^{\text{out}}U_{k,\underline{\widetilde{S}}_k^j}\right|^2}{\left|\sum_{j=1}^{N!}\prod_{k=1}^{N}\eta_{S'_k}^{\text{out}}U_{k,\underline{\widetilde{S}}'_k{}^j}\right|^2}. \tag{5.13}$$

We summarize several conclusions from Equation (5.13).

- Since the input state is constant for all sampling trials, losses before the interferometric network (η^{in}) do not affect the postselected boson distribution.[1]
- If the losses are balanced ($\eta_k^{\text{out}} = \text{constant}\ \forall\, k$) or zero ($\eta^{\text{out}} = 1$), we can accurately predict the postselected boson distribution from the unitary transformation alone.
- Unbalanced losses after the interferometric network (η^{out}) do affect the postselected boson distribution. These losses must be accurately characterized.

The simplified case presented above, with losses only at the input and output, provided a concise way of introducing some of the salient features of lossy boson sampling. In a more general scenario, there will be losses distributed throughout the linear network. We do not investigate this case in detail here. However, the circuit characterization method used must accurately describe these losses, including those embedded within the interferometric network, as they also affect the postselected boson distribution.

[1] This is strictly true only assuming true single photon inputs. The absolute system loss, which depends on η^{in}, does affect the contribution of errors from higher-order photon emission from our SPDC sources. This analysis is described in Section 5.4.

5.2 Circuit characterization

Before attempting to sample the boson distribution we must first experimentally determine the non-unitary transformation Λ. We thus briefly review photonic circuit characterization in the presence of loss. The relevance of this topic extends beyond boson sampling experiments. Accurately and efficiently characterizing large lossy linear transformations will become increasingly important for a variety of applications.

5.2.1 Overview

There are two strategies to characterizing a linear optical transformation. They differ in the amount of information they give about the decomposition in Equation (5.8).

In an *element-wise* approach, one individually characterizes every U_{α_z} comprising the circuit. The phase differences in every interferometer and the reflectivities of all beam splitters are identified. This method has been successfully used to characterize photonic circuits used in quantum optics experiments assuming zero loss[24]. New experimental methods, which allow losses to be measured at an arbitrary point within a circuit, generalize this approach[145].

In a *blackbox* strategy, one characterizes $\mathbf{U}$ from Equation (5.8) without isolating the individual contributions of its constituent elements. All we know about the circuit is that it has M input and output ports and this approach quantifies the linear transformation between them. Blackbox characterization strategies using classical[146] and nonclassical[147] light have been developed.

The element-wise approach gives more information than a blackbox characterization. It

is particularly useful for programmable circuits where adjustable phase shifters allow one to change the linear transformation. More specifically, the decomposition of such a circuit includes terms $U_{\mathrm{ps}}(i,\phi)$ where, ideally, the phase can be set anywhere in the range $0 \leq \phi \leq 2\pi$. With knowledge of this decomposition, one can efficiently determine the required phases to implement a desired $\mathbf{U}$[148]. A blackbox characterization of a programmable circuit does not provide this predictive power if a phase adjustment is made. One can verify that $\mathbf{U}$ changes as ϕ is adjusted, but there is not a clear route to setting phases in a deliberate fashion to implement a desired transformation.

Despite these drawbacks, all the boson sampling experiments performed to-date[25,141–143], including the one described in this thesis, have implemented a blackbox characterization. These characterizations are relatively straightforward to perform; light is injected in one or two of the input ports and detectors monitor all outputs. A series of such measurements are sufficient to determine the complex-valued transformation including loss. Element-wise characterization of a lossy circuit, while possible, can be more involved and benefits from technologies, like intra-circuit Bragg gratings[145], that increase experimental complexity.

5.2.2 Blackbox characterization with single photons

We outline the blackbox method used to characterize the photonic chips used in our boson sampling experiment. These chips perform a complex-valued linear mapping $\Lambda_{ij} = \tau_{ij}e^{i\phi_{ij}}$. We outline the two sets of experiments that are performed to determine τ and ϕ[147].

The moduli of this complex matrix are determined by injecting single photons into mode

i_1 and monitoring the probability P_{j_1} of the photon emerging from output j_1. We find that

$$P_{j_1} = \left|\langle 0|\, \hat{b}_{j_1} \hat{a}^\dagger_{i_1}\, |0\rangle\right|^2 = \left|\langle 0|\, \hat{b}_{j_1} \sum_{j=1}^{M} \Lambda_{i_1} \hat{b}^\dagger_j\, |0\rangle\right|^2$$
$$= |\Lambda_{i_1 j_1}|^2 = |\tau_{i_1 j_1}|^2 . \tag{5.14}$$

We do not need single photons to measure $\tau_{i_1 j_1}$, a source of classical light would suffice. However, since a heralded source of single-photons is already coupled to our circuit for the boson sampling experiment, using the same source for the preceding characterization is convenient.

We then perform a series of HOM interference experiments to determine ϕ_{ij}. We input two photons into modes i_1, i_2 and detect in modes j_1, j_2. Equation (1.19) shows that the intereference visibility is a function of $P_{cc}(0)$ and $P_{cc}(\infty)$, the probability of coincident photon detection at the outputs for indistinguishable and distinguishable photons respectively. We use Equation (5.3) to find

$$P_{cc}(0) = \left|\mathrm{Per}(\Lambda^{(\mathbf{S},\mathbf{T})})\right|^2 = \left|\mathrm{Per}\begin{pmatrix} \tau_{i_1 j_1} e^{i\phi_{i_1 j_1}} & \tau_{i_1 j_2} e^{i\phi_{i_1 j_2}} \\ \tau_{i_2 j_1} e^{i\phi_{i_2 j_1}} & \tau_{i_2 j_2} e^{i\phi_{i_2 j_2}} \end{pmatrix}\right|^2$$
$$= (\tau_{i_1 j_1}\tau_{i_2 j_2})^2 + (\tau_{i_1 j_2}\tau_{i_2 j_1})^2 + 2\tau_{i_1 j_1}\tau_{i_1 j_2}\tau_{i_2 j_1}\tau_{i_2 j_2} \cos(\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} + \phi_{i_2 j_2}). \tag{5.15}$$

If the photons are distinguishable, we observe the incoherent sum of their individual scat-

tering probabilities. We therefore obtain

$$P_{cc}(\infty) = (\tau_{i_1 j_1}\tau_{i_2 j_2})^2 + (\tau_{i_1 j_2}\tau_{i_2 j_1})^2, \tag{5.16}$$

which is the same as $P_{cc}(0)$ in Equation (5.15) but without the sinusoid term from interference. We use Equation (1.19) to find the HOM visibility

$$\begin{aligned} V &= \frac{P_{cc}(\infty) - P_{cc}(0)}{P_{cc}(\infty)} \\ &= -\frac{2\tau_{i_1 j_1}\tau_{i_1 j_2}\tau_{i_2 j_1}\tau_{i_2 j_2}}{(\tau_{i_1 j_1}\tau_{i_2 j_2})^2 + (\tau_{i_1 j_2}\tau_{i_2 j_1})^2} \cos(\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} + \phi_{i_2 j_2}). \end{aligned} \tag{5.17}$$

We then solve for the phase terms as a function of the observed HOM visibility and previously determined τ_{ij}

$$\cos(\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} + \phi_{i_2 j_2}) = -\left(\frac{V}{2}\right) \frac{(\tau_{i_1 j_1}\tau_{i_2 j_2})^2 + (\tau_{i_1 j_2}\tau_{i_2 j_1})^2}{\tau_{i_1 j_1}\tau_{i_1 j_2}\tau_{i_2 j_1}\tau_{i_2 j_2}}. \tag{5.18}$$

We have a single equation with four unknown phase terms; more information is required to solve for the individual ϕ_{ij}.

The phase insensitivity of our single photon inputs and detectors allows us to set $\phi_{ij} = 0$ for several of the elements in Λ. We introduce the diagonal matrix D_ϕ whose k-th term along the diagonal is given by $e^{i\phi_k}$. The phase freedom at the input and output of our circuit implies that Λ and $D_{\phi_{\text{in}}}\Lambda D_{\phi_{\text{out}}}$ give equivalent results when inserted into Equation (5.3). We are thus free to adjust $\phi_{\text{in,out}}$, the set of M phases along the accessible input or output modes, to make any row and column of our matrix real-valued. In this section, we

choose to make the first row and column real.

When inverting the cosine function in Equation (5.18), there will be a sign ambiguity in the calculated phase. We introduce two relevant properties of the signs of these phases. First, we note that the complex conjugate of Λ in Equation (5.3) does not change our results. We use this invariance to ensure that the top-left nonzero phase, ϕ_{22}, is positive. Second, we can determine the sign of one of the four phases in Equation (5.3), say $\phi_{i_2 j_2}$, if we know: the magnitude $|\phi_{i_2 j_2}|$, the magnitude $\beta_{i_1 i_2 j_1 j_2} \equiv |\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} + \phi_{i_2 j_2}|$, and the values of the other three phases ($\phi_{i_1 j_1}$,$\phi_{i_1 j_2}$,$\phi_{i_2 j_1}$) including their sign. Using this information, one finds the sign from

$$\begin{aligned}\text{sign}\left[\phi_{i_2 j_2}\right] = \text{sign}\Bigg[&\Big|\beta_{i_1 i_2 j_1 j_2} - \big|\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} - |\phi_{i_2 j_2}|\big|\Big| \\ &- \Big|\beta_{i_1 i_2 j_1 j_2} - \big|\phi_{i_1 j_1} - \phi_{i_1 j_2} - \phi_{i_2 j_1} + |\phi_{i_2 j_2}|\big|\Big|\Bigg]. \end{aligned} \tag{5.19}$$

We can now outline the strategy to find all the phases in Λ. We inject single photons into modes $i_1 = 1$ and $i_2 = 2$ and observe the HOM interference visibility at outputs $j_1 = 1$ and $j_2 = 2$. Since the first row and column of Λ are real, then the only nonzero phase in Equation (5.18) is ϕ_{22}. We solve for this phase which we know is positive. We then repeat this experiment but change i_2 to mode γ and use Equation (5.18) to solve for $|\phi_{\gamma 2}|$. To determine the sign of this phase, we first measure the HOM visibility for $i_1 = 2$ and $i_2 = \gamma$ which allows us to determine $|-\phi_{22} + \phi_{\gamma 2}|$. We now have all the necessary information to use Equation (5.19) to determine the sign of $\phi_{\gamma 2}$. One repeats this strategy to solve for all ϕ_{ij} in Λ.

This strategy for calculating the phase terms is always overconstrained. Since the first

row and column of the M dimensional matrix Λ are real, there are $M^2-(2M-1)$ nonzero phases. By comparison, there are $\binom{M}{2}$ independent assignments of (i_1, i_2) or (j_1, j_2) amongst M modes. There are thus $\binom{M}{2}^2$ different HOM experiments one can perform, which is greater than or equal to the number of phases. The number of HOM experiments grows exponentially with M and thus, in practice, it may be infeasible to perform all of them even for modest circuit sizes. Regardless, one always has the option of performing additional interference experiments and implementing a least-squares minimization to improve the estimate of the phases.

5.3 Experimental results

5.3.1 Scope of the experiment

While we perform a proof-of-principle boson sampling experiment, our QBSM does not outperform a classical computer. It is instructive to outline why a classical device can quickly solve this small boson sampling problem. In our experiment, we sample the $N = 3$ or $N = 4$ photon distribution over $M = 6$ modes. Due to our use of non-number resolving detectors, we only sample from the $\binom{M}{N}$ elements of the distribution where all photons emerge on separate modes. Each element of this distribution corresponds to the permanent of an $N \times N$ submatrix of the $M \times M$ linear transformation describing our chip. The entire boson distribution can thus be exactly obtained by calculating $\binom{M}{N}$ permanents of $N \times N$ matrices. For $N = 3$ and $M = 6$, this amounts to calculating the permanents of 20 different 3×3 matrices which can be done with 340 floating point operations. A desktop computer performs this computation in a small fraction of a second for this small system,

while our photonic QBSM only outputs a few samples drawn from this distribution every hour. Clearly, our quantum machine is slower than its classical counterpart.

We instead demonstrate the central tenet of boson sampling that matrix permanents correctly predict the boson distribution. By repeatedly injecting photons into a circuit and detecting where they emerge, we experimentally infer the underlying distribution $\mathbf{P}^{\text{exp}}$. Figure 4.3 shows that this is analogous to dropping balls into a Galton board to infer the binomial distribution. We also characterize the linear transformation enacted by our photonic circuit and use Equation (4.3) to predict the boson distribution $\mathbf{P}^{\text{th}}$. Using Equation (4.4), we show that the distance $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ between these two distributions is small which is consistent with the assertion that bosonic scattering amplitudes are a function of the matrix permanent. Sections 4.3.2-4.3.3 showed that this relationship is crucial to demonstrating the computational hardness of boson sampling.

5.3.2 Experimental layout

The geometry of the chip used for the boson sampling experiments is shown in Figure 5.4(a). The circuit was fabricated on a silica-on-silicon chip using the UV-writing process illustrated in Figure 3.14. beam splitters of a desired reflectivity are fabricated by crossing waveguides at a specific angle. Such beam splitters take up less space than traditional evanescent couplers, allowing reduced chip size and lower loss[149]. While there are eight spatial modes in the central section of the chip, the outer waveguides are not written to the edge and are thus not accessible. These outer modes can thus be treated as losses, leaving a six (accessible) mode interferometric network.

We use a bulk crystal SPDC source to generate the photons used for characterization

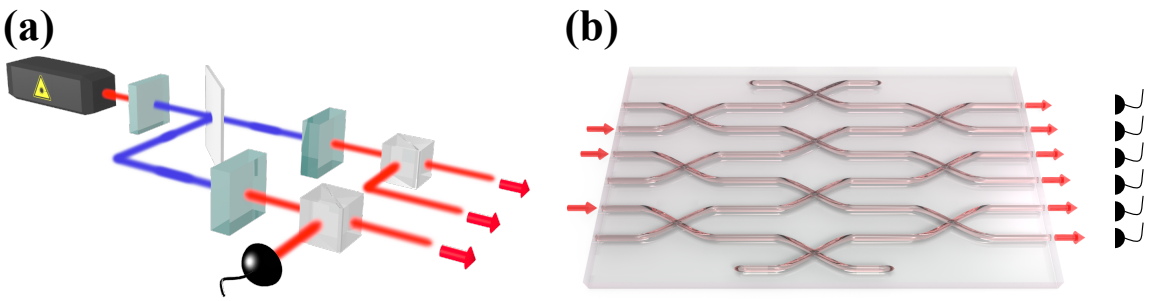


Figure 5.3: (a) An 80 MHz Ti-Sapphire oscillator (SpectraPhysics Mai Tai) outputs 100 fs pulses at 830 nm, which undergo type-I second harmonic generation in a $700\mu m$ $\beta - BaB_2O_4$ (BBO) crystal. This 415 nm light is then used to pump type-II collinear SPDC in two 8-mm long Potassium Dihydrogen Phosphate (KDP) crystals[57]. All modes of the resulting squeezed state are passed through spectral filters (Semrock $\Delta\lambda$=3 nm, not shown) to maximize indistinguishability before being coupled into PM fibers. (b) An $M = 6$ mode linear optical circuit fabricated in a UV-written silica-on-silicon substrate is used for the boson sampling experiment. Photons from the SPDC source are coupled to the photonic chip from the left and APDs monitor all outputs on the right. This figure represents the configuration for the $N = 3$ photon boson sampling experiment. For the four-photon experiment, which was implemented on a different chip of identical geometry, we coupled a double photon pair from a single SPDC crystal into two of the six chip inputs (not shown).

and in the boson sampling experiment. While the on-chip SFWM sources discussed in earlier chapters would have been well-suited to this experiment, the successful fabrication of multiple matched sources occured only after boson sampling data collection. The photon source at the input and the detectors at the output are both coupled to the chip via PM fiber V-groove arrays.

5.3.3 Characterization results

Figure 5.4 shows the result of the characterization of our photonic circuit using the method outlined in Section 5.2.2. While we later present both three- and four-photon boson sampling experiments, performed on different chips, here we present only the characterization relevant to the three-photon case.

We clarify several steps used to produce Figure 5.4. The photons used for the three-

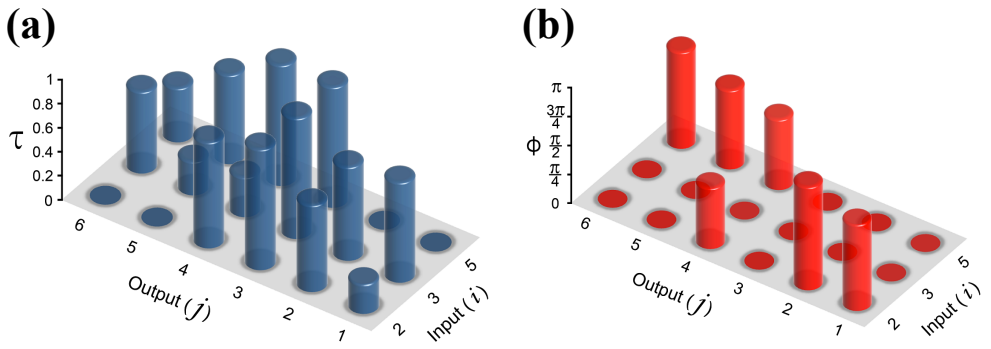


Figure 5.4: We characterize the (a) modulus and (b) phase terms of the complex linear transformation $\Lambda_{ij} = \tau_{ij} e^{i\phi_{ij}}$ linking the input mode i to the output mode j of our three-photon apparatus. The circuit geometry dictates that several τ_{ij} are zero, and our phase-insensitive input states and detection methods imply only six nonzero ϕ_{ij}. Only relative values were needed because of postselection, so we rescaled each row of τ so that its maximum value is unity.

photon boson sampling are coupled into modes 2, 3, and 5 of the photonic chip while APDs monitor all 6 outputs. We therefore only need to characterize a 3×6 element matrix. Several τ_{ij} are zero because photons injected into some modes cannot access all outputs. In addition, since Equation (5.13) showed the postselected boson distribution is insensitive to losses before the interferometric network, we rescale each row of τ so that its maximum value is unity. Finally, we can use the phase freedom on the input and output modes to set any row and column of Λ to be real. We choose the $i = 3$ row and $j = 3$ column in Figure 5.4 because this eliminates the most phases given that several of the τ_{ij} are known to be zero.

5.3.4 Boson sampling results

We perform three- and four-photon boson sampling experiments. For the three-photon experiment, we use two SPDC sources, one in a heralded configuration, to inject the state $|\mathbf{T}\rangle = |011010\rangle$ as shown in Figure 5.3. We only keep trials when three detectors at the chip output and the herald detector fire. For the four-photon experiment, the higher order

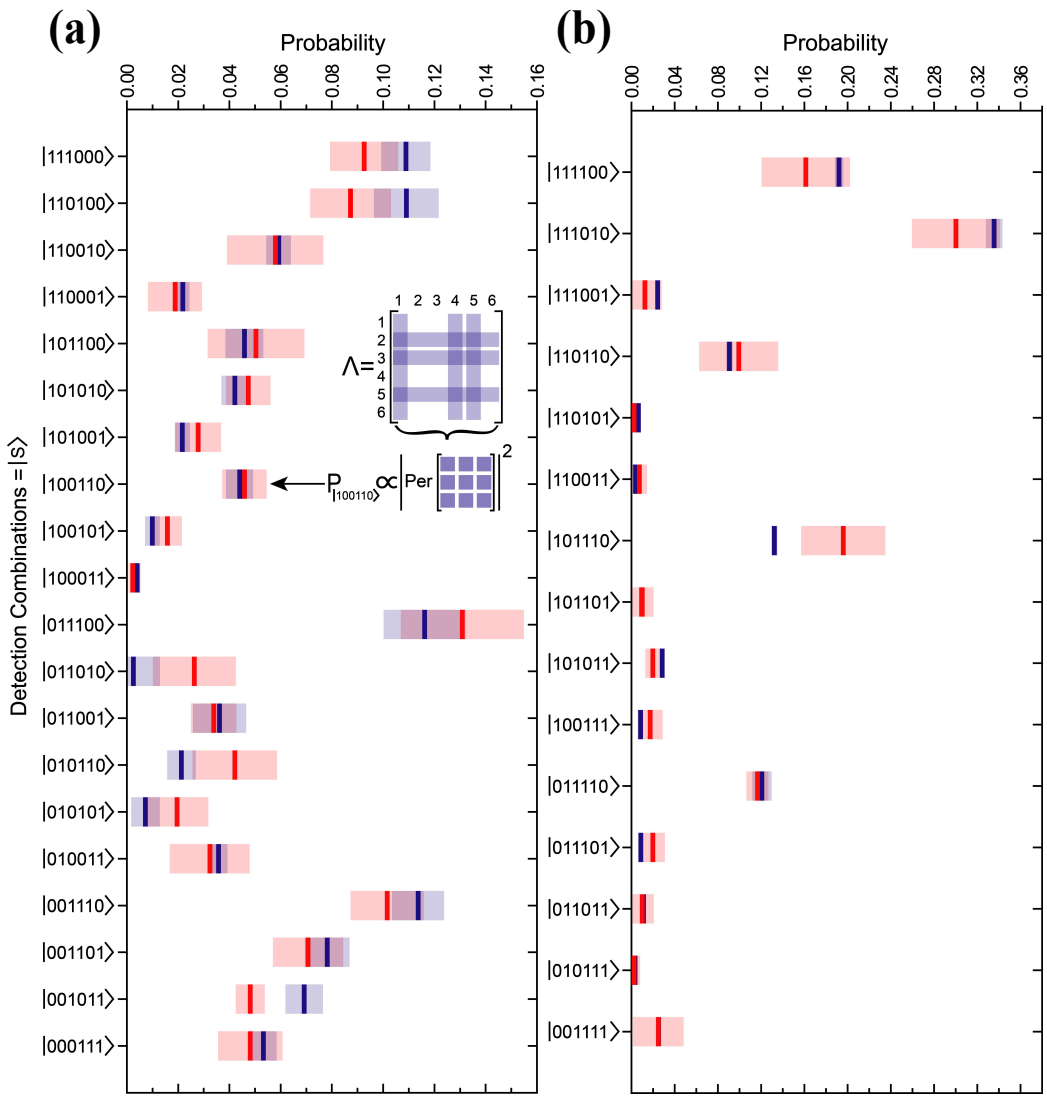


Figure 5.5: The relative frequencies $\mathbf{P}^{\mathbf{exp}}$ of outcomes in which the photons were detected in distinct modes are shown in red for (a) three- and (b) four-photon experiments. Each data set was collected over 160 hours, and statistical variations in counts are shown by the red shaded bars. The theoretical distributions $\mathbf{P}^{\mathbf{th}}$ (blue) are obtained from the permanents of submatrices constructed from the full transformation Λ, as depicted in the inset. The blue error bars arise from uncertainties in the characterization of Λ.

emission from a single source is injected into a different chip of identical geometry. No heralding detectors are used in this case. The input state is given by $|202000\rangle$ and we postselect on four detectors after the chip firing. In both experiments, data is collected for 160 hours. The SPDC sources are pumped as lightly as possible while maintaining an

acceptable count rate. This mitigates errors originating from higher-order pair emission in a single source which is discussed in more detail in Section 5.4.1. In the three-photon case, a total of 1421 samples are collected, implying a data point is collected on average every 7 minutes. In the four photon case, a total of 405 samples are collected, implying a data point is recorded on average every 21 minutes.

Figure 5.5 shows our main boson sampling results. The observed samples are used to infer the relative frequencies $\mathbf{P}^{\text{exp}}$ shown by solid red marks in Figure 5.5. The red error bars are calculated by sub-dividing the data set and finding the variation in the observed counts between these subsets. This gives information about the variation in our observed boson distribution without assuming an error model. In particular this analysis does not assume Poissonian errors; due to fluctuations in our experimental apparatus over time, most of the error bars are notably larger than the shot noise limit.

The inset of Figure 5.5(a) shows that our theoretical prediction $\mathbf{P}^{\text{th}}$ indicated by blue marks is obtained by calculating the permanent of the appropriate subsection of our characterized Λ. The blue error bars arise from uncertainties in the characterization of Λ. In the three-photon experiment, this error is dominated by variation in the alignment of our chip to the fiber-coupled detectors which causes the effective detector efficiency to change with time. The chip used in the four photon experiment was glued to fiber pigtails which eliminated this alignment variation and explains the significantly smaller blue error bars in Figure 5.5(b).

5.4 Error analysis

To analyze the performance of our QBSM, we compare our results to an ideal machine. We quantify the match of two sets of relative frequencies $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ by calculating the L_1 distance $d^{(N)}(\mathbf{P}^{(1)}, \mathbf{P}^{(2)})$, originally defined in Equation (4.4), where N denotes the number of photons in a sample. Identical and maximally dissimilar distributions correspond to $d = 1$ and $d = 0$ respectively. For our experiments, we calculated $d^{(N)}(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ and obtained $d^{(3)} = 0.094 \pm 0.014$ and $d^{(4)} = 0.097 \pm 0.004$, where errors arise solely from the experimental uncertainty represented by red shaded bars in Figure 5.5. Even in an ideal QBSM with perfect state preparation and detection, statistical variations result in nonzero d. When we substitute for our data a Monte Carlo sampling of $\mathbf{P}^{\text{th}}$ with sample size equivalent to our experiments, we instead calculated $d^{(3)} = 0.043 \pm 0.012$ and $d^{(4)} = 0.059 \pm 0.022$. This suggests there are contributions to $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ beyond statistical deviation which represent sources of error in our QBSM.

Analyzing the causes of these errors is an integral part of the boson sampling problem[112]. Theorem 2 in Section 4.3.3, which states $|\text{GPE}|^2_{\pm} \in \mathsf{BPP}^{\mathsf{NP}^{\mathcal{C}}}$, explicitly requires that the classical algorithm $\mathcal{C}$ be given an error estimate ϵ. If we are to devise a fair race between a quantum and classical machine, then we must know the accuracy of our QBSM because this acts as an upper bound for ϵ. While QBSMs have only been built on a small scale, it is worthwhile to continually investigate sources of sampling error as the size of these devices is increased.

We analyze the effect of two significant sources of error in our experiment: higher order pair emission from our SPDC sources and photon distinguishability. This was the first,

and remains the only, experimental investigation of boson sampling error to-date[25]. A new theoretical boson distribution that accounts for these effects, $\mathbf{P}^{\text{mod}}$, is calculated and we show that $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{mod}}) < d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$. In particular, $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{mod}})$ is explained by the finite number of experimental counts used to infer $\mathbf{P}^{\text{exp}}$. These observations lead us to conclude that we have accurately diagnosed the source and approximate magnitude of the principal error terms in our experiment.

5.4.1 Higher-order SPDC pair emission

SPDC sources do not emit a single pair of photons, as we have so far assumed, but rather the superposition of pair states described by Equation (2.26). Due to loss in our system and our use of non-number resolving detectors, this higher-order emission contaminates our boson distribution. For example, when one of the SPDC sources in our experiment shown in Figure 5.3 occasionally emits two pairs of photons we inject $N = 4$ or 5 photons into our circuit. Due to system imperfections, these trials contribute to the $N = 3$ boson statistics. The new boson distribution that includes the effect of this higher order emission can no longer be calculated from Λ using Equation (5.3). This formula assumed an experiment where N photons are injected and N photons are detected.

Figure 5.6 presents a schematic for our lossy circuit that is helpful in analyzing the effect of these higher order terms. There are $M = 6$ accessible modes on the input and output of our chip and one can insert 24 loss beam splitters throughout the circuit to model the effect of unbalanced losses. Appendix A shows that this system is overspecified; one can accurately model the effect of these losses using the alternate arrangement of 13 loss modes shown in Figure 5.6(b). Assuming that an element-wise characterization has been

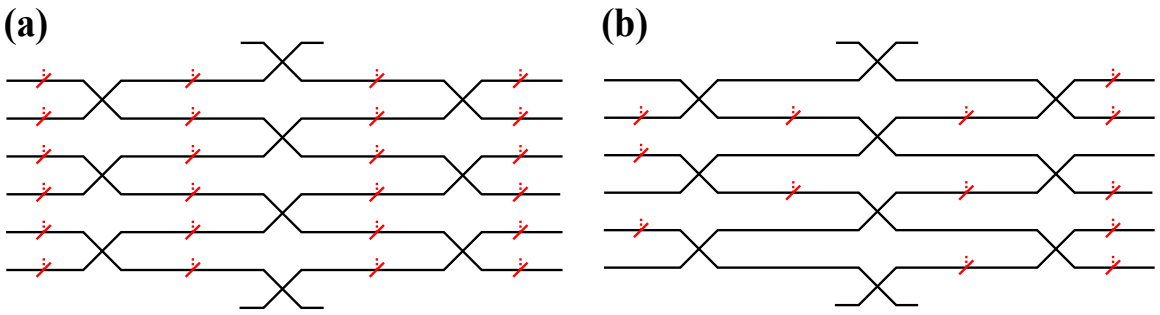


Figure 5.6: (a) One can represent the effect of nonuniform loss in our experiment with a set of 24 loss modes (red). Inefficiency in our SPDC sources and APD detectors are accounted for with the first and last column of losses respectively. (b) For our three-photon boson sampling experiment, 13 losses are sufficient for an accurate model.

performed on this lossy circuit, how can we model the effect of the higher order SPDC emission on our boson distribution?

Calculate full output state

One approach is to calculate the full state output from our circuit, $|\psi_{out}\rangle$. An element-wise characterization gives us knowledge of all the beam splitter reflectivities and phase offsets, so we can construct the $L \times L$ unitary U. We use this to find $|\psi_{out}\rangle$ and then apply detector operators similar to Equation (1.8) to N of the accessible spatial modes to accurately predict the corresponding element of the boson distribution P_S.

We do not use this strategy because it is computationally intensive since $|\psi_{out}\rangle$ occupies a high-dimensional Hilbert space. The Hilbert space of the overall system is the tensor product of the spaces of the subsystems, $H = H_1 \bigotimes H_2 \bigotimes ... \bigotimes H_L$, where H_j indicates the Hilbert space of the j-th spatial mode. Therefore, the dimensionality of the full system is given by $D = d_1 d_2 ... d_L$ where D and d_j are the dimensions of the full system and j-th subsystem respectively. We note that $d_j = n_j + 1$, where n_j is the maximum number of photons that could possibly reach a certain output and the '+1' accounts for the vacuum

term.

It is now straightforward to calculate D for our experiment. We choose the input state $|\mathbf{T}\rangle = |022010\rangle$, one of the higher-order SPDC terms when we're attempting to couple $|011010\rangle$ to perform a three-photon experiment, for a numerical example. We find n_j for all 19 modes, 6 accessible modes in black and 13 losses in red, shown in Figure 5.6(b). For example, $n_4 = 5$ because all five input photons can potentially access the fourth accessible output mode. Multiplying the resulting d_j together gives $D \approx 10^{12}$. Calculating $|\psi_{out}\rangle$ in this large a space is computationally demanding, even when using sparse matrix representations.

Direct calculation with matrix permanent

An alternate approach uses the matrix permanent to determine the effect of these higher order terms. This strategy does not require an explicit calculation of $|\psi_{out}\rangle$ and thus avoids the problems highlighted above. To analyze the effect of higher order terms, we revert to Equation (5.2) which shows that boson distributions in lossy systems are given by $\mathrm{Per}(\mathbf{U}^{(\mathbf{S},\mathbf{T})})$. Figure 5.1 shows that $\mathbf{U}$ acts over the full L-dimensional system which includes accessible and loss modes. We consider the case when our input state $|\mathbf{T}\rangle$ contains N' photons, but we are attempting to perform an N-photon boson sampling experiment where $N' > N$. The state $|\mathbf{T}\rangle$ is still restricted to the first M modes shown in Figure 5.1. However, our output state $|\mathbf{S}\rangle$ can range over the full L dimensional space. Accordingly, the sub-matrix $\mathbf{U}^{(\mathbf{S},\mathbf{T})}$ is drawn from the first M rows but any of the columns of $\mathbf{U}$.

We do not know where photons are lost in the circuit, nor can our detectors resolve photon number. We account for this ignorance by summing over all the possible N'-photon

output states that could be mistaken for a desired N-photon component of the boson distribution. The notation $\underline{\mathbf{S}}$ was previously used to represent an N element vector whose entries list the spatial modes of all the photons in the state and are limited to the accessible modes so that $1 \leq \underline{S}_j \leq M$. The new notation $\widetilde{\underline{\mathbf{S}}}'$ is introduced to represent the set of N'-photon outputs that could be mistaken for $\underline{\mathbf{S}}$. Every member of this set, written $\widetilde{\underline{\mathbf{S}}}'^k$, is an N' element vector whose entries indicate the modal positions of the photons and take values in the range $1 \leq \widetilde{\underline{S}}_j'^k \leq L$. By definition, if we take any of these vectors $\widetilde{\underline{\mathbf{S}}}'^k$ and remove values $\widetilde{\underline{S}}_j'^k > M$ (photons in loss modes) and any duplicate entries (two or more photons in the same output) then we are left with $\underline{\mathbf{S}}$. This final property specifies what is meant when we say $\widetilde{\underline{\mathbf{S}}}'^k$ could be mistaken in our experiment for $\underline{\mathbf{S}}$.

We enumerate how many of these N'-photon states can be mistaken for $\underline{\mathbf{S}}$. More specifically, we find γ which is the upper range of $1 \leq k \leq \gamma$ in the set of states $\widetilde{\underline{\mathbf{S}}}'^k$. We know that N of the N' photons in our input state must occupy the modes defining $\underline{\mathbf{S}}$. We are therefore left with $N' - N$ photons to distribute. These photons can only be placed in $L - (M - N)$ of the modes because $(M - N)$ of the accessible modes must remain unoccupied to match $\underline{\mathbf{S}}$. Our problem thus reduces to placing $N - N'$ photons (indistinguishable balls) in $L - (M - N)$ spatial modes (distinguishable bins) where repetitions are allowed because multiple photons are allowed to occupy the same spatial mode. This well-known combinatoric problem yields the result

$$\gamma = \binom{(L - M + N) + (N' - N) - 1}{N' - N} = \binom{L - M + N' - 1}{N' - N}. \tag{5.20}$$

We obtain our final result by summing up over these γ possible output states that could

be mistaken for $|\mathbf{S}\rangle$

$$P_S = \sum_{k=1}^{\gamma} \left(\frac{|\mathrm{Per}(\mathbf{U}^{(\widetilde{\mathbf{S}}'^k,\mathbf{T})}|^2}{\prod_{j=1}^{L} \widetilde{S}_j'^k! \prod_{i=1}^{L} T_i!} \right). \tag{5.21}$$

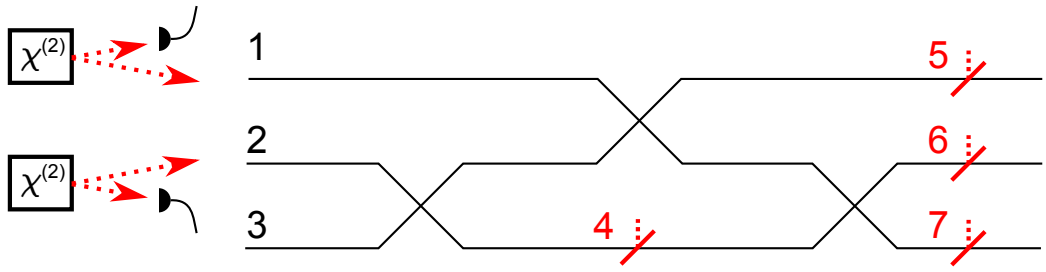


Figure 5.7: Two SPDC sources in a heralded configuration are coupled to a simple $M = 3$ mode interferometer with 4 losses (red) implying there are $L = 7$ modes in total. We intend to perform an $N = 2$ photon boson sampling experiment with the two heralded modes, however multipair emission from a single source leads to errors in our observed boson distribution. The largest of these errors is when $N' = 3$ photons are coupled into the circuit.

A simple case-study helps illustrate the use of Equation (5.21). Figure 5.7 shows an $L = 7$ mode system, with $M = 3$ accessible modes. We use two heralded SPDC sources in an attempt to perform an $N = 2$ photon boson sampling experiment. However, occasionally one of our sources emits two pairs of photons and we actually have an input state with $N' = 3$ photons which we describe with the vector $\underline{\mathbf{T}} = [1, 1, 2]$. In this example, we attempt to measure the element of the N boson distribution corresponding to a single photon in the first two outputs which corresponds to $\underline{\mathbf{S}} = [1, 2]$. There are $\gamma = \binom{7-3+3-1}{3-2} = 6$ terms in the summation shown in Equation (5.21). The set of outputs that can be mistaken for $\mathbf{S}$ are

$$\begin{aligned} \underline{\widetilde{\mathbf{S}}}^{\prime 1} &= [1, 2, 4] & \underline{\widetilde{\mathbf{S}}}^{\prime 2} &= [1, 2, 5] & \underline{\widetilde{\mathbf{S}}}^{\prime 3} &= [1, 2, 6] \\ \underline{\widetilde{\mathbf{S}}}^{\prime 4} &= [1, 2, 7] & \underline{\widetilde{\mathbf{S}}}^{\prime 5} &= [1, 1, 2] & \underline{\widetilde{\mathbf{S}}}^{\prime 6} &= [1, 2, 2]. \end{aligned} \tag{5.22}$$

If one can characterize all the elements of the circuit in Figure 5.7, then Equation (5.8) allows one to calculate $\mathbf{U}$. Equations (5.21)-(5.22) show that we calculate P_S by summing $\gamma = 6$ terms each of which is proportional to the permanent of an $N' \times N' = 3 \times 3$ matrix.

Characterization of circuit elements

We need to perform an element-wise characterization of our lossy circuit to find $\mathbf{U}$ and implement Equation (5.21). Our previous characterization outlined in Section 5.2.2 is not sufficient; that blackbox method only provided Λ, the $M \times M$ submatrix of $\mathbf{U}$ over the accessible modes.

Figure 5.8 illustrates the $K = 26$ individual elements we require to model our lossy system. Appendix A explains why all losses in our system, including inefficiency in our SPDC sources and APD detectors, can be represented by η_1^ℓ-η_{13}^ℓ.

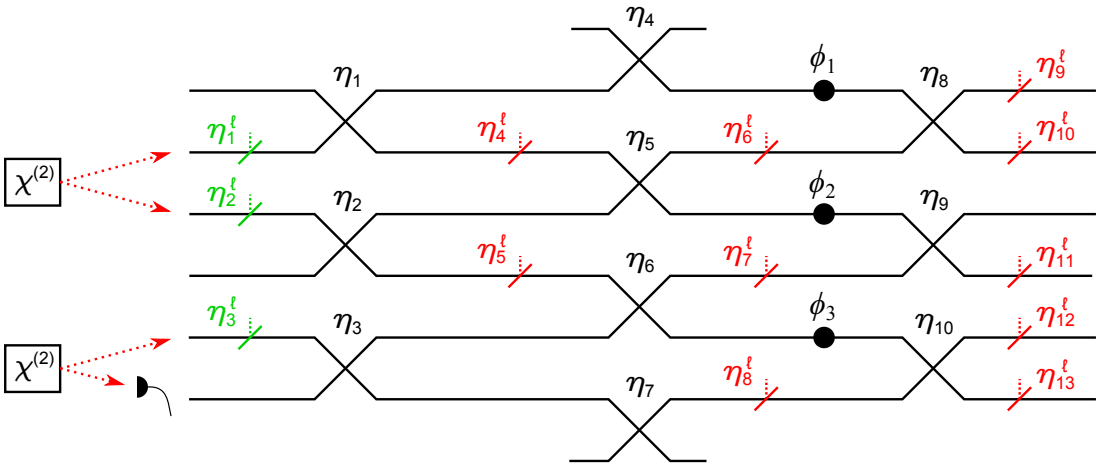


Figure 5.8: A schematic of our 3-photon boson sampling experiment highlights the role of losses distributed throughout the system. There are ten fabricated beam splitters (η_1-η_{10}) comprising three interferometers whose relative phases are shown (ϕ_1-ϕ_3). We can model the role of unbalanced loss with thirteen beam splitters η_1^ℓ-η_{13}^ℓ. These 26 total elements are marked black, red, or green according to how they were characterized. The SPDC source configuration is indicated on the left. We note that a loss mode labeled η_x^ℓ actually has reflectivity $\sqrt{1-(\eta_x^\ell)^2}$ which is consistent with our earlier definition in Figure 2.3; we use the abridged labeling to keep the figures clearer.

Three characterization methods, corresponding to the three colors used in the circuit shown in Figure 5.8, are used to characterize these 26 elements. The details are presented in Appendix A. We use a loss-independent characterization strategy to determine the UV-written beam splitter reflectivities, η_1-η_{10}, and relative phases, ϕ_1-ϕ_3, drawn in black in Figure 5.8. We perform a numerical search for the relative losses η_4^ℓ-η_{13}^ℓ drawn in red that is constrained by our earlier blackbox characterization of Λ. Finally, we can use the number correlations between modes of our SPDC sources to determine the overall transmission from source to detector which allows η_1^ℓ-η_3^ℓ drawn in green to be calculated.

The results of this element-wise characterization along with Equation (5.8) then determine the overall unitary operation $\mathbf{U}$ describing our entire system, including loss modes. We can then use Equation (5.21) to determine the boson distribution for any input state, including those with higher order photon numbers. The dominant error term corresponds to one of the SPDC sources emiting two pairs of photons while the other still emits a single pair. Figure 5.3 illustrates that this corresponds to the input states $|\mathbf{T}\rangle = |022010\rangle$ and $|\mathbf{T}\rangle = |011020\rangle$ depending on which crystal emits two pairs. We label the boson distributions obtained from Equation (5.21) for these inputs $\mathbf{P}^{221}$ and $\mathbf{P}^{112}$ respectively.

We must determine how often our SPDC source emits these higher order terms to accurately weight the contribution of $\mathbf{P}^{221}$ and $\mathbf{P}^{112}$ relative to our original distribution $\mathbf{P}^{\text{th}}$. Using the number-state expansion of a two-mode squeezer in Equation (2.48), we write the state emitted by our two sources

$$|\psi\rangle = \mathcal{N}\big(\,|1111\rangle + \lambda(|2211\rangle + |1122\rangle)\big), \tag{5.23}$$

where $\mathcal{N}$ is a normalization constant. The $|1111\rangle$ term is the desired term for our 3-photon boson sampling experiment, where one of the photons is used as a herald. The latter two terms are the lowest order errors from higher order SPDC emission and lead to $\mathbf{P}^{221}$ and $\mathbf{P}^{112}$ respectively. Equation (5.23) shows that λ is the appropriate weighting factor for these error terms.

We measure λ^2 by relating it to the heralded second-order correlation $g_H^{(2)}(0)$ from Equation (2.30). We re-write this latter quantity

$$g_H^{(2)}(0) = \frac{N_{34H}N_H}{N_{3H}N_{4H}} = \frac{P_{34H}P_H}{P_{3H}P_{4H}}, \tag{5.24}$$

where P is the probability, on any experimental trial, of observing coincident events on the subscripted modes. Accurate determination of λ^2 from $g_H^{(2)}(0)$ requires the inclusion of loss into our model. We outline how to calculate one of the terms in Equation (5.24), specifically P_{34H}. Detectors on modes 3, 4, and the herald only fire simultaneously if two or more pairs of photons are emitted from the SPDC source. The probability of this occuring, since $\lambda^2 \ll 1$, is approximately λ^4. Two of these photons are coupled to a 50% reflectivity beam splitter, as shown in Figure 2.5, and the probability of them splitting so that there is a photon in both modes is 1/2. The probability that neither of these photons in modes 3 and 4 are lost is $\eta_3^2\eta_4^2$. The corresponding probability that at least one of the two photons in the herald mode is not lost is $1-(1-\eta_H^2)^2$. Multiplying these terms together gives $P_{34H} = \lambda^4(1-(1-\eta_H^2)^2)\eta_3^2\eta_4^2/2$. Using similar logic to determine the other three terms, we find

$$g_H^{(2)}(0) = \frac{\frac{\lambda^4(1-(1-\eta_H^2)^2)\eta_3^2\eta_4^2}{2}\lambda^2\eta_H^2}{\frac{\lambda^2\eta_H^2\eta_3^2}{2}\frac{\lambda^2\eta_H^2\eta_4^2}{2}}. \tag{5.25}$$

We solve Equation (5.25) for λ^2

$$\lambda^2 = g_H^{(2)}(0)\frac{\eta_H^2}{2(1-(1-\eta_H^2)^2)}. \tag{5.26}$$

We measure $g_H^{(2)}(0)$ using Equation (5.24) and the heralding efficiency from $\eta_H = (N_{H3} + N_{H4})/N_H$. This then allows us to calculate λ^2 using Equation (5.26). While our SPDC source at full power is capable of $\lambda^2 \approx 0.03$ we deliberately adjust the pump as low as possible while maintaining an acceptable boson sampling data rate. This leads to our final result of $\lambda^2 = 0.011$ and 0.022 in the three- and four-photon boson sampling experiments respectively.

There is finally sufficent information to predict the effect of higher order SPDC emission on our boson distribution. Using our element-wise characterization, we can form new distributions $\mathbf{P}^{221}$ and $\mathbf{P}^{112}$ that correspond to when one of our two crystals emits two pairs of photons. We then used Equation (5.26) to calculate λ^2 which tells us how often this higher-order emission occurs. We then form the new distribution

$$\mathbf{P}^{\mathrm{h}} = \mathcal{N}\big[\mathbf{P}^{\mathrm{th}} + \lambda^2(\mathbf{P}^{221} + \mathbf{P}^{112})\big] \tag{5.27}$$

where $\mathcal{N}$ ensures $\mathbf{P}^{\mathrm{h}}$ is normalized.

5.4.2 Photon distinguishability

The analysis which showed boson sampling is computationally hard assumed that the bosons were indistinguishable so that their scattering amplitudes were given by the matrix

permanent. In practice, any QBSM will use bosons that are at least slightly distinguishable from one another. This distinguishability introduces error because Equation (4.3) no longer exactly predicts the observed boson distribution. If we perform an N-photon sampling experiment with distinguishable particles then the underlying distribution is no longer given by the permanents of $N \times N$ matrices.

Our model of photon distinguishability assumes that the SPDC sources are spectrally factorable which allows us to treat each input photonic mode as a separate system. This is a good approximation as our SPDC sources were engineered to be nearly spectrally factorable without any spectral filtering[57]. In addition, we employ narrowband spectral filters on all photons that are coupled to our photonic circuit. These two factors eliminate nearly all spectral correlations between the photon pairs emitted by our sources.

We can express each of these N single photons as a superposition of a desired mode and an undesired distinguishable mode[150]. This decomposition is convenient for predicting the effect of photon distinguishability within the matrix permanent formalism. Our input state is then written

$$|\mathbf{T}\rangle = \prod_{i=1}^{N} \left(\alpha_i \hat{A}_i^\dagger(\xi_0) + \sqrt{1-\alpha_i^2}\, \hat{A}_i^\dagger(\xi_i) \right) |0\rangle \,, \tag{5.28}$$

where ξ_0 and ξ_i describe the desired and undesired mode with relative weightings quantified by the distinguishability parameter α_i. In the extreme cases of totally indistinguishable or distinguishable photons, $\alpha = 1$ and 0 respectively. Our SPDC sources emit a spectrally narrowband and broadband photon whose spectra have a bandwidth of ≈ 3.5 and 11 nm (FWHM) respectively. While all photons used in these boson sampling experiments are passed through a 3 nm filter, slight spectral distinguishability persists. We assume that dif-

ferences between α_1, α_2,..., α_N primarily arise due to these slight variations in the spectra between photons. Consequently, we do not need to determine N distinguishability parameters but rather just α_{nb} and α_{bb} corresponding to the filtered narrowband and broadband SPDC photons. Appendix B describes how we estimate these two parameters, resulting in $\alpha_{nb} = 0.984$ and $\alpha_{bb} = 0.969$.

We use the fact that $\alpha_{nb,bb} \approx 1$ to only consider the lowest order contribution to error from photon distinguishability when one photon is completely distinguishable from the other $N-1$ indistinguishable photons. In this case, the resulting boson distribution is no longer a function of permanents of $N \times N$ matrices but rather an incoherent summation of permanents of $(N-1)\times(N-1)$ matrices. For example, assume the state $|\mathbf{T}\rangle = |1\rangle^{\tau} |11000\rangle$ is input to a circuit where τ labels a distinguishable photon. Then for a unitary transformation $\mathbf{U}$ the probability of obtaining an output $|\mathbf{S}\rangle$ is

$$\begin{aligned} P_S = &\left|U_{11}^{(\mathbf{S},\mathbf{T})}\right|^2 \left|\left(U_{22}^{(\mathbf{S},\mathbf{T})}U_{33}^{(\mathbf{S},\mathbf{T})} + U_{23}^{(\mathbf{S},\mathbf{T})}U_{32}^{(\mathbf{S},\mathbf{T})}\right)\right|^2 \\ &+ \left|U_{12}^{(\mathbf{S},\mathbf{T})}\right|^2 \left|\left(U_{21}^{(\mathbf{S},\mathbf{T})}U_{33}^{(\mathbf{S},\mathbf{T})} + U_{23}^{(\mathbf{S},\mathbf{T})}U_{31}^{(\mathbf{S},\mathbf{T})}\right)\right|^2 \\ &+ \left|U_{13}^{(\mathbf{S},\mathbf{T})}\right|^2 \left|\left(U_{21}^{(\mathbf{S},\mathbf{T})}U_{32}^{(\mathbf{S},\mathbf{T})} + U_{22}^{(\mathbf{S},\mathbf{T})}U_{31}^{(\mathbf{S},\mathbf{T})}\right)\right|^2, \end{aligned} \tag{5.29}$$

where the terms in parentheses are permanents of 2×2 matrices. Repeating the calculation in Equation 5.29 for all possible outputs yields the probability distribution $\mathbf{P}^{\mathrm{d_i}}$ when the i-th input photon is distinguishable from the other photons.

We add $\mathbf{P}^{\mathrm{d_i}}$, with appropriate weighting, to our ideal distribution $\mathbf{P}^{\mathrm{th}}$ to estimate the effect of photon distinguishability on our boson distribution. We accordingly form a new

probability distribution

$$\mathbf{P}^{\mathrm{d}} = \mathcal{N}\left[\left(\alpha_{nb}^2\alpha_{bb}\right)^2\mathbf{P}^{\mathrm{th}} + \left(\alpha_{nb}\alpha_{bb}\sqrt{1-\alpha_{nb}^2}\right)^2\left(\mathbf{P}^{\mathrm{d}_1}+\mathbf{P}^{\mathrm{d}_2}\right) + \left(\alpha_{nb}^2\sqrt{1-\alpha_{bb}^2}\right)^2\mathbf{P}^{\mathrm{d}_3}\right], \tag{5.30}$$

where $\left(\alpha_{nb}^2\alpha_{bb}\right)^2$ describes the probability that all three photons are in the desired mode, $\left(\alpha_{nb}\alpha_{bb}\sqrt{1-\alpha_{nb}^2}\right)^2$ is the probability that one of the narrowband photons is distinguishable from the other photons, and $\left(\alpha_{nb}^2\sqrt{1-\alpha_{bb}^2}\right)^2$ is the probability that the narrowband photon is distinguishable from the other photons. We multiply by a normalization constant $\mathcal{N}$ that accounts for the fact that we've only considered the case when a single photon is distinguishable.

5.4.3 Final error analysis results

We combine the probability distributions $\mathbf{P}^{\mathrm{h}}$ and $\mathbf{P}^{\mathrm{d}}$, which account for higher order emission and photon distinguishability respectively, to conclude this analysis of error in our boson sampling results. We use Equations (5.27) and (5.30) to form the new distribution

$$\begin{aligned}\mathbf{P}^{\mathrm{mod}} = \mathcal{N}\Bigg[&\left(\alpha_{nb}^2\alpha_{bb}\right)^2\mathbf{P}^{\mathrm{th}} + \left(\alpha_{nb}\alpha_{bb}\sqrt{1-\alpha_{nb}^2}\right)^2\left(\mathbf{P}^{\mathrm{d}_1}+\mathbf{P}^{\mathrm{d}_2}\right) + \left(\alpha_{nb}^2\sqrt{1-\alpha_{bb}^2}\right)^2\mathbf{P}^{\mathrm{d}_3} \\ &+ \lambda^2\left(\left(\alpha_{nb}^3\alpha_{bb}^2\right)^2\mathbf{P}^{221} + \left(\alpha_{nb}^3\alpha_{bb}\right)^2\mathbf{P}^{112}\right)\Bigg].\end{aligned} \tag{5.31}$$

We have ignored the case when there is a distinguishable photon and higher order emission simultaneously. These terms would be of the order $\lambda^2(1-\alpha^2)$ and are thus ignored because both λ^2 and $(1-\alpha^2)$ are close to zero.

Figure 5.9 presents the results of this boson sampling error analysis. The distance

$d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{mod}})$ shown by the green point (Figure 5.9(a)-(b) insets) was found to be consistent with the statistical variation resulting from a finite sample size, for both the three- and four-photon experiments. This suggests we have correctly identified and modeled the largest sources of inaccuracy. To investigate how the performance of our QBSM depends on λ and photon distinguishability, we calculated $d(\mathbf{P}^{\text{mod}}, \mathbf{P}^{\text{th}})$ for a range of operating parameters shown in Figure 5.9(c)-(d). In terms of λ, a clear trade-off is presented between data rate and inaccuracy because of multiphoton emission, which is an intrinsic consequence of using SPDC sources. Improvement in photon indistinguishability increases the fidelity to the ideal machine by reducing the contribution of lower-order matrix permanents to the observed distribution. Minimizing these effects allows the sampling error of the QBSM to be reduced which increases the computation time of a classical computer attempting to sample with the same error bound.

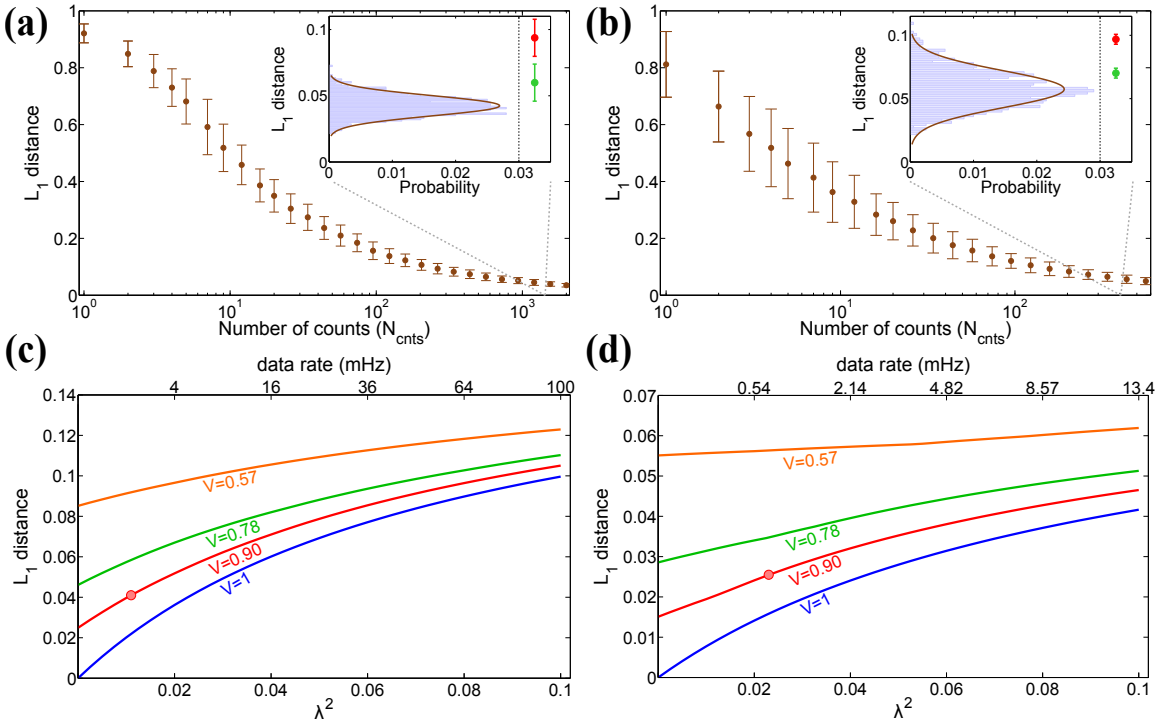


Figure 5.9: We consider several boson distributions: the experimental samples $\mathbf{P}^{\text{exp}}$, the ideal predictions of the matrix permanent $\mathbf{P}^{\text{th}}$, and the predictions of the full model $\mathbf{P}^{\text{mod}}$ that includes higher-order emission and photon distinguishability. **(a)** The L_1 distance d between $\mathbf{P}^{\text{th}}$ and a Monte Carlo simulation of an ideal machine that samples $\mathbf{P}^{\text{th}}$ a finite number of times for our three- and **(b)** four-photon cases. The errors in this case are solely a result of the finite number of samples collected by the ideal machine. The inset histograms show the variation in d expected for a sample size corresponding to the 1421 and 405 counts collected in our three- and four-photon experiments. The distance $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ (red) suggests an underlying systematic inaccuracy as it falls outside the range of outputs of an ideal machine indicated in the histogram. Our full model is validated by the distance $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{mod}})$ (green) which is consistent with statistical variation. The red and green dot positions correspond to the L_1 axis only. **(c)** The predicted variation in $d(\mathbf{P}^{\text{th}}, \mathbf{P}^{\text{mod}})$ is shown as a function of λ and the photon distinguishabilites, represented by the reduction in two-photon interference visibility V, for the three- and **(d)** four-photon cases. Our experimental conditions are marked (red dot).

Chapter 6

Conclusion and outlook

6.1 SFWM sources

With femtosecond-laser fabrication, we have demonstrated an on-chip source of heralded single photons that achieves extremely low loss ($\eta_h = 40\%$) and high output purity ($P = 0.86$) without any narrow filtering. To our knowledge, no previous on-chip source has simultaneously demonstrated such low-loss heralding of high purity states. Achieving this performance relied on spectrally factorable photon emission which was enabled by the mode control allowed by source integration.

UV-writing in a silica-on-silicon chip allowed the repeatable fabrication of SFWM sources on a large scale. Single-source characterization confirmed that low-loss heralding of single photons was still possible with these sources. The purity and indistinguishability of the heralded photons from different sources was confirmed in a series of HOM interference experiments where consistently high visibilities were observed.

Simultaneous operation of many on-chip SFWM sources is a promising avenue for future

research. In the near term, operating $N = 3 - 5$ sources and postselecting on N heralded photon events may be feasible and is facilitated by the low-loss, spectrally factorable source design. In the longer term, one can overcome the probabilistic emission inherent in SFWM and SPDC with a multiplexing strategy where a single near-deterministic emitter is constructed from many sources linked together with an active switchyard [151,152]. Constructing a nearly-deterministic source with this approach would require the simultaneous operation of far more heralded photon sources [55,60] than is currently common in quantum optics experiments.

Achieving this ambitious increase in source complexity will be facilitated by further integration in future SFWM source designs. Bulk optics comprise a significant portion of the interference experiment shown in Figure 3.21 and preclude scaling to more than a few sources. The effort to replace these optics surrounding the source with integrated components will benefit from mature fabrication techniques in silica. Below, we outline potential strategies to reduce the reliance on bulk optics in future designs.

The use of a $1 \times N$ waveguide splitter on the input of the chip allows the pump to be coupled to a single spatial mode which is then equally split, on chip, to pump N different sources in parallel. This removes the need for bulk beam splitters and multiple sets of coupling optics before the chip. However, silica damage thresholds place an upper bound on the allowed optical power; surface damage at the input of the $1 \times N$ splitter could limit this approach. A possible solution is evident from Equation 2.57 which shows that the SFWM efficiency increases as we shrink the transverse mode area. This mitigates damage constraints by lowering the power required by each source. Adiabatic mode tapers [90] at the input and output could allow such small-core sources to still be mode-matched to 'normal'

fiber modes at the input and output. An alternate solution is to use coreless end caps on chip facets which allows larger beam sizes at the air-silica interface which frequently has the lowest damage threshold [153].

The signal, idler, and residual pump light after the source must be separated into different spatial modes to herald single photons without saturating single photon detectors. This is currently done with bulk dichroic mirrors and interference filters which requires a free space beam path between the SFWM source chip and single-photon sensitive detectors. There are a variety of wavelength selective integrated components, in particular Bragg gratings and evanescent couplers, that could separate these modes. Achieving the required $\approx$100 dB pump extinction appears very challenging and may require the concatenation of several integrated filtering elements. However, separation of signal and idler into separate modes is less demanding and is feasible with existing technology [111]. Since the herald modes would be separated on the source chip, this would allow direct connection of SFWM sources to more complicated linear optical circuits with eventual free-space pump filtering just before detection.

Alternate integrated source designs, particularly SPDC in periodically-poled $\chi^{(2)}$ crystals, have also demonstrated impressive progress in recent years. Heralding of high purity states with very low pump powers [59,76] has been demonstrated. This latter quality makes these sources particularly promising for generation of highly-squeezed states which is challenging with silica sources due to their significantly lower $\chi^{(3)}$ nonlinearity. However, fabrication of matched sources has not been demonstrated in these designs, it is unclear how low the scattering loss in some of these substrates can be lowered, and the natural interfacing with external silica waveguides is lost.

Our sources offer a simple route to low-loss production of high purity single photons over a broad spectral range. Fabrication of identical sources on a large scale has been demonstrated through a series of HOM interference experiments. Increasing the number of SFWM sources operated simultaneously is facilitated by the wide-range of ancillary components that are already well developed in silica. We anticipate that on-chip silica sources can have immediate application in a wide range of single-photon quantum optics applications which employ silica photonics [23–29].

6.2 Boson sampling

We have performed one of the first experimental demonstrations of boson sampling. A simple case-study showed unbalanced losses in the system affect the boson distribution, even with postselection. We account for these losses with a blackbox characterization method that, while not resolving the contribution of individual optical elements, reveals the overall linear transformation illustrated in Figure 5.4 describing our lossy system. Both three- and four-photon boson sampling experiments were conducted and Figure 5.5 shows that the resulting data closely matches the predictions of the matrix permanent. While our QBSM was far slower than a classical computer, this provides supporting evidence for the central tenet of boson sampling that permanents accurately predict bosonic scattering amplitudes.

Two significant sources of error in our sampling experiment were investigated. First, the effect of higher-order SPDC pair emission was estimated by performing an element-wise characterization of our lossy circuit. Knowledge of these elements allowed the unit-

ary operation describing our system, including loss modes, to be specified. The effect of higher-order pair emission was predicted by taking permanents of larger matrices where one or more photons is in a loss mode. Secondly, photon distinguishability also reduces our boson sampling accuracy. We estimated the degree of distinguishability by noting the reduction in HOM interference visibility and the spectral overlap between injected photons. Incoherently summing over the permanents of smaller matrices predicts the effect of this distinguishability on the boson distribution. Figure 5.9(a)-(b) shows that accounting for these effects brings the observed error in our experiment within the range expected from statistical variation due to our finite data set.

Photonic QBSMs could feasibly be the first quantum systems to surpass the limits of classical computation. Two of the principal challenges in building a photonic universal quantum computer, the very large number of photons required and active feed-forward, are no longer necessary. As few as 20 photons may be sufficient [112] to demonstrate a quantum advantage in boson sampling and a completely non-adaptive circuit is adequate for such a device. In addition, all of the components for a QBSM, including single-photon sources, large low-loss photonic circuits, and high-efficiency single-photon detectors, are all active research areas due to their utility in a wide variety of experiments.

Scaling to larger systems appears to require integration, an effort which has witnessed significant advances in recent years. Traditionally reliant on SPDC in a bulk crystal, singe-photon sources are now often integrated on-chip [59,77,111,154] which, in addition to potentially lower-loss and higher state purity, offers the hope of scaling to larger numbers of sources in a single experiment. While many early integrated quantum optics experiments were feasible with bulk optics, there has been a consistent trend to scale to larger and lower-loss circuits

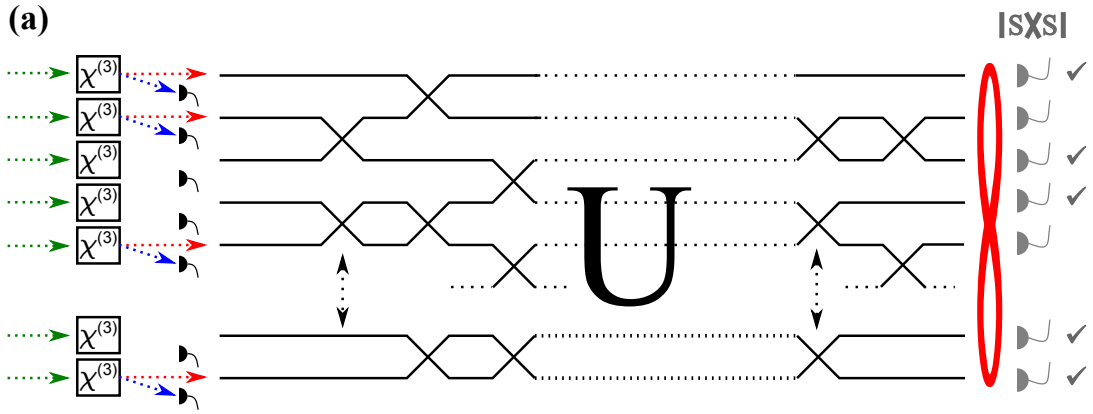
such that current experiments can often only be realistically performed on-chip. Detectors have also recently been integrated by evanescently coupling waveguides to superconducting films deposited directly onto chips [71,155,156]; particularly relevant to boson sampling are the recent demonstration of high-efficiency versions of such detectors which facilitates scaling to larger photon number.

Formalizing protocols that plausibly show quantum-enhanced operation in future, larger QBSMs is an important area for future work. Our choice of the words 'plausibly show' is deliberate; because boson sampling is not an NP problem then we cannot efficiently verify correct operation with a classical computer. For relatively small instances, one can retain the approach used in our experiment where operation is inefficiently verified by classically computing permanents to predict the entire boson distribution which is then compared with QBSM output. This approach will become untenable for larger systems that are necessary to demonstrate a quantum advantage. In this case, a reasonable strategy is to draw samples from the QBSM and subsequently use a classical computer to calculate the corresponding matrix permanents. This (hopefully) shows the user that the QBSM tends to output samples that correspond to relatively large matrix permanents. Implementation of this strategy relies on choosing a problem size that allows the quantum-machine to sample faster than the classical computer but small enough that the latter device can still, eventually, calculate the relevant matrix permanents. Interestingly, while we cannot efficiently verify correct operation with certainty, one can often efficiently discount alternate distributions, for example the uniform distribution [157].

The greatest impediment to scaling to larger QBSMs is the quality of single-photon sources. Both low-loss circuits and high-efficiency detectors are currently available, the

technical challenge is in replicating these high-quality components on a larger scale. In contrast, current source architectures are insufficient for straightforward scaling to larger photon numbers due to their highly probabilistic nature; the probability of pair emission on any trial $p \ll 1$. Scaling to larger photon numbers by waiting for N sources to fire simultaneously has an exponentially decreasing success rate given by p^N. This problem is not unique to boson sampling and has impeded progress in a wide variety of quantum optics experiments.

A variant on this computational model, dubbed scattershot boson sampling, obviates the need for near-deterministic photon sources before scaling to larger systems. By removing this key experimental challenge, scattershot boson sampling offers an especially promising route to quantum-enhanced computation. Rather than using a fixed input state and drawing different output samples on each trial, Figure 6.1 illustrates that in scattershot boson sampling both the input state and the output samples vary on each iteration. Given $\mathbf{U}$ and the squeezing parameter λ^2 of the sources, a classical computer proceeds by randomly choosing a heralded input state appropriate to λ^2 and then solves the corresponding boson sampling problem. A quantum device instead couples the squeezed sources to each of M modes and uses the observed statistics on the herald and post-circuit modes to deduce the input and output states respectively. In this new problem, both the classical and quantum machines are tasked with sampling, not just over a single boson distribution, but over a distribution of boson distributions. In other words, rather than just outputting a list of samples $|\mathbf{S}\rangle$, the output is now both $|\mathbf{T}\rangle$ and $|\mathbf{S}\rangle$. Since every experimental trial is accordingly an instance of the original boson sampling problem, the classical hardness analysis described in Section 4.3 still applies.



(b)

Trial	|T⟩	|S⟩
X-1	|01000...11⟩	|10111...00⟩
X	|11001...01⟩	|10110...11⟩
X+1	|00111...00⟩	|01100...01⟩

Figure 6.1: (a) A scattershot boson sampling trial begins with the pumping of M two-mode squeezed vacuum sources (here using SFWM). One mode of each source is coupled to a random unitary $\mathbf{U}$ while the other leads directly to a heralding detector. The heralded state $|\mathbf{T}\rangle$ is known on every individual trial, though it can vary between trials due to the spontaneous nature of the photon sources. Note that only a subset of the heralded sources emit on any given trial. Similar to the original boson sampling problem, detectors after $\mathbf{U}$ monitor the number statistics of the entangled output. (b) A scattershot boson sampling data set consists of a list of input $|\mathbf{T}\rangle$ and output $|\mathbf{S}\rangle$ states for each trial. Sub-figure (a) is the X-th trial in this example.

Scattershot boson sampling no longer requires nearly-deterministic sources to scale to large photon numbers. Assuming factorable sources, we can use Equation (2.48) to write the state output by each source as

$$|\psi\rangle = \sqrt{1-\lambda^2}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}|nn\rangle\,, \tag{6.1}$$

where the final ket indicates the occupation number in the two-modes of the squeezed state. The probability of any given source heralding a single photon is thus $p_1 = (1-\lambda^2)\lambda^2$ and the probability of simultaneous emission of photons from a *fixed* set of N sources is $P_N = p_1^N$

which decreases exponentially. However, the probability of *any* set of N sources, amongst the M total sources in Figure 6.1, having non-vacuum output is given by the binomial distribution

$$P_N = (\lambda^2)^N (1-\lambda^2)^{(M-N)} \binom{M}{N}, \tag{6.2}$$

which has a mean photon number $\overline{N} = M\lambda^2$ that no longer decays exponentially with N. On every experimental trial we perform a boson sampling experiment, on average with a large number of photons $\overline{N}$. In the original boson sampling problem, one lowers the probability of observing output states with more than one photon in a mode by having many more modes than bosons. In scattershot boson sampling, the analogous condition is to lower λ^2 as M increases which allows for states $|\mathbf{T}\rangle$ and $|\mathbf{S}\rangle$ that are unlikely to have more than one photon per mode despite a potentially large $\overline{N}$. Importantly, this can be achieved with existing and widely-available two-mode squeezed vacuum sources.

By merging the two topics discussed in this thesis, scattershot boson sampling provides an instructive conclusion. While this model of quantum-enhanced computation alleviates the need for deterministic single-photon sources, arguably the greatest impediment to QBSM scaling, Figure 6.1 illustrates that it requires identical sources on an even larger scale. If we wish to perform an N-photon experiment, we no longer need just N but $M \gg N$ sources. In the first half of this thesis, we have demonstrated that SFWM in silica waveguides provides a promising architecture for exactly this type of application. It is possible to fabricate identical, low-loss sources of high purity photons on a large scale. A key drawback of such sources is their probabilistic nature, however this is not necessary in scattershot boson sampling. Inclusion of such sources in future boson sampling experiments

could feasibly provide strong evidence for the computational power of quantum mechanics in the not-distant future.

Appendix A

Circuit element characterization

We perform an element-wise characterization of our circuit, including losses, to predict the boson distribution observed from higher order SPDC emsision. Quantifying the performance of every element in our system allows one to calculate the $L \times L$ unitary matrix $\mathbf{U}$. This transformation is sufficient, using Equation (5.21), to determine our desired boson distribution.

Characterization of all the necessary elements in our circuit proceeds in several steps. Figure A.1 shows a schematic of our experiment with loss modes inserted between every UV-written beamsplitter to model non-uniform loss. We first perform a loss-independent characterization of the intentionally fabricated circuit drawn in black. Determining the losses illustrated in red is less straightforward and proceeds in two steps. The majority of the losses are determined from a numerical minimization routine; we eliminate redundant loss beamsplitters from Figure A.1 to aid this search. Several of the losses cannot be determined from this numerical analysis and are instead measured using the number-correlated emission from our SPDC source.

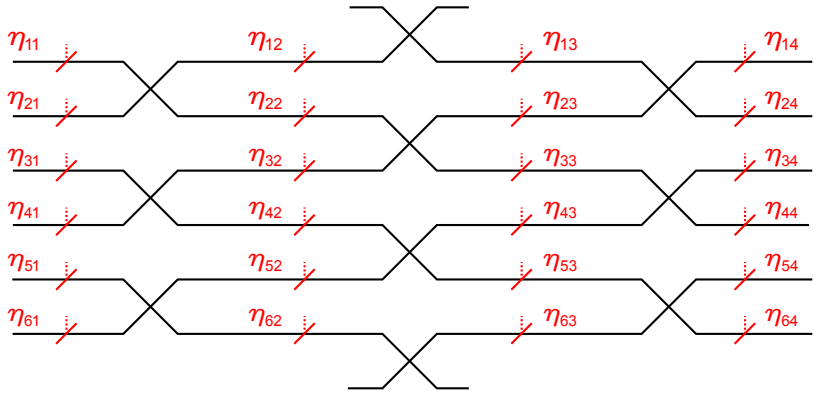


Figure A.1: Inefficiencies in our system can be represented by beamsplitters (red) that couple our experiment to loss modes. Accounting for potentially unbalanced losses throughout our system initially leads to 24 of these beamsplitters in our model. Losses before and after the chip, for example from coupling our SPDC source to the circuit and the imperfect detectors, are included in the first and last column of these loss elements. Many of these losses are redundant. We note that a loss mode labeled η_x actually has reflectivity $\sqrt{1-\eta_x^2}$ which is consistent with our earlier definition in Figure 2.3; we use the abridged labeling to keep the figures clearer.

A.1 Characterization of UV-written elements

One can characterize the reflectivity of a beamsplitter even in the presence of unbalanced losses[20]. We use the labeling convention described in Figure A.2. Coupling a light source with power P into input a leads to the following powers observed at outputs c and d

$$\begin{aligned} P_{ac} &= \eta_a^2\eta^2\eta_c^2 P \\ P_{ad} &= \eta_a^2(1-\eta^2)\eta_d^2 P. \end{aligned} \tag{A.1}$$

This light source is then coupled into port b giving the output powers

$$\begin{aligned} P_{bc} &= \eta_b^2(1-\eta^2)\eta_c^2 P \\ P_{bd} &= \eta_b^2\eta^2\eta_d^2 P. \end{aligned} \tag{A.2}$$

We then calculate the ratio $r = P_{ad}P_{bc}/P_{ac}P_{bd}$ and solve for η. Keeping the solution that gives $0 \leq \eta^2 \leq 1$ gives

$$\eta^2 = \frac{1}{1+\sqrt{r}}. \tag{A.3}$$

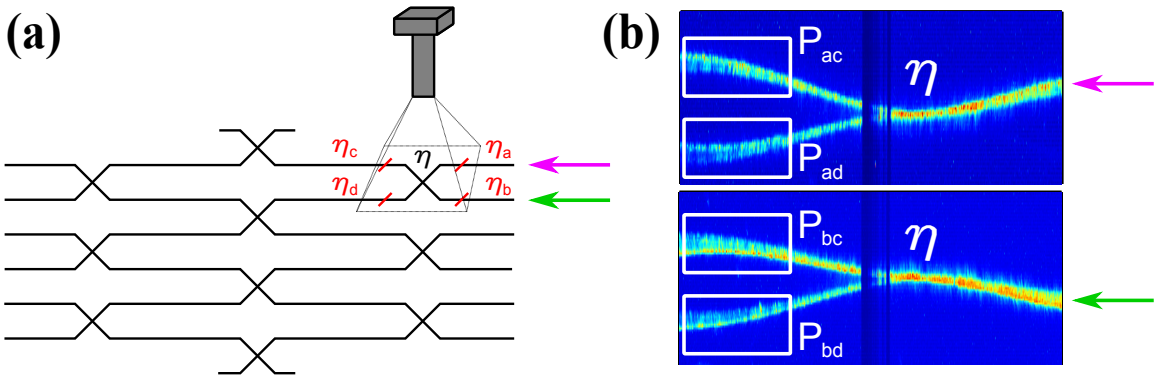


Figure A.2: (a) A light source is coupled into the top output port (purple) of a beamsplitter and the power on the two opposing ports is measured (P_{ac},P_{ad}). Light is then coupled into the second port (green) and the output powers are again measured (P_{bc},P_{bd}). These four powers are sufficient to calculate the reflectivity of the beamsplitter η independent of the losses $\eta_a - \eta_d$. It is not possible to measure these four powers from the left-side of the circuit due to the intervening interferometer. We therefore use a CCD camera to record top-down images of the scatter from the waveguide. (b) Example CCD images are shown. The scattered light from a portion of each image (white box) is integrated and corresponds to one of the four powers necessary to calculate η.

Figure A.2 shows this method relies on the ability to separately couple light into the two input ports and take independent power measurement on the outputs. For a single beamsplitter, meeting these requirements is straightforward. However, obtaining independent access to the four ports of a beamsplitter embedded in a larger photonic circuit, including our chip, is more involved. For example, the output powers from the beamsplitter being characterized in Figure A.2(a) cannot be independently measured from the far side (left side) of the chip due to the middle interferometer. We implement a more flexible power measurement strategy, using top-down CCD images of scatter from the waveguide as shown in Figure A.2(b), to overcome this issue. Due to our loss-independent character-

ization method, our CCD measurements do not assume the scatter is constant along the length of the guide.

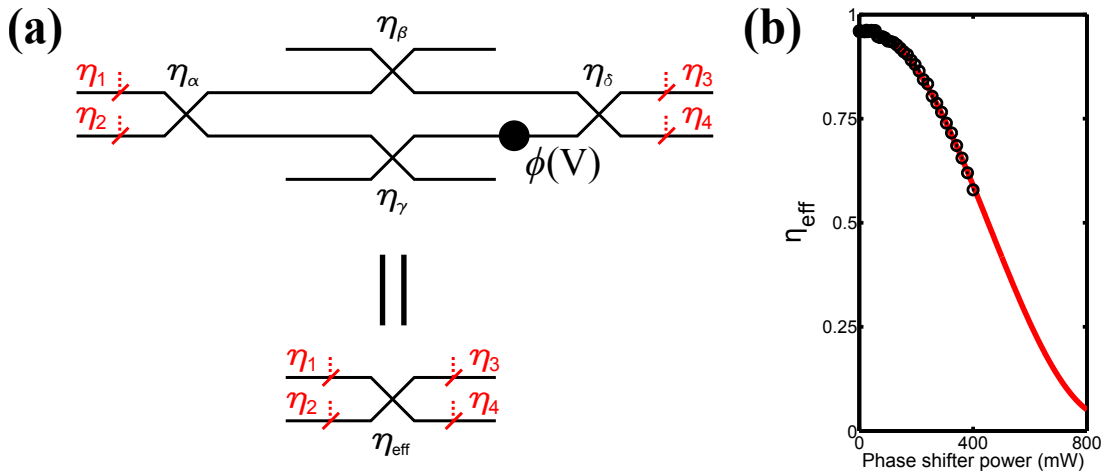


Figure A.3: (a) Our circuit contains interferometers comprised of four beamsplitters η_α-η_δ and a voltage-adjustable phase $\phi(V)$. There are unbalanced losses $\eta_1 - \eta_4$ on the input and output ports. This interferometer is equivalent to a single lossy beamsplitter whose reflectivity is a function of the constituent parts of the interferometer: $\eta_{\text{eff}}(\eta_\alpha, \eta_\beta, \eta_\gamma, \eta_\delta, \phi(V))$. (b) One can perform a loss-independent characterization of η_{eff} as ϕ is adjusted. Fitting a sinusoid to the resulting curve yields the phase-offset when $V = 0$ which is relevant to the boson sampling experiment.

We also need to characterize the three relative phases in the middle interferometers of our circuit shown in Figure A.1. These phases are adjustable via resistive heaters bonded to the top of the chip above each of the interferometers. The phase $\phi = \phi_0 + bV^2$ is determined by the zero-voltage offset ϕ_0 and the power dissipated in the resistor which is proportional to V^2, where V is the voltage across the heater. Adjusting this phase changes the effective reflectivity of the overall interferometer $\eta_{\text{eff}}\big(\eta_\alpha, \eta_\beta, \eta_\gamma, \eta_\delta, \phi(\text{V})\big)$. Any losses internal to the interferometer are described by $\eta_{\beta,\gamma}$ and are thus included in this model.

Repeated application of the loss-independent beamsplitter characterization method allows these relative phases to be determined[20]. One uses this strategy to first separately find η_α, η_β, η_γ, and η_δ. The reflectivity of the overall beamsplitter η_{eff} is then re-

peatedly determined as the voltage applied to the adjustable phase shifter is adjusted. We fit $\eta_{\text{eff}}\big(\eta_\alpha, \eta_\beta, \eta_\gamma, \eta_\delta, \phi(\text{V})\big)$ to the resulting data, where ϕ_0 and b are fitting parameters. Since the phase-shifters are not used during this boson sampling experiment, then ϕ_0 is the relevant phase used in our later models.

A.2 Characterization of losses

We rely on a numerical search to find the majority of the losses in our experiment. To minimize the size of the parameter space, we seek to eliminate redundant loss terms from Figure A.1. Specifically, we show why the 13 losses shown in Figure 5.8 are sufficient to describe the losses in our experiment. We only consider the geometry for the three-photon boson sampling experiment, though a similar approach was used in the four-photon case.

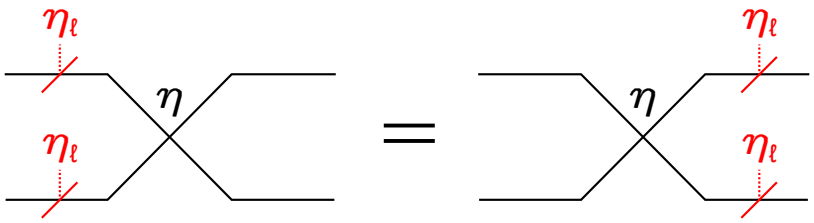


Figure A.4: A beamsplitter with balanced loss can be modeled with loss modes at the input (left) or output (right).

Our analysis relies on the fact that balanced losses commute with a beamsplitter operation as shown in Figure A.4. We use this property repeatedly to move from a very general model of our experiment with 24 loss modes drawn in Figure A.1 to the 13 losses shown in Figure A.8. These steps are shown graphically in Figures A.5-A.8.

A numerical search is performed to determine the losses shown in red in Figure A.9. While these losses are unknown, section A.1 described how we characterize the elements

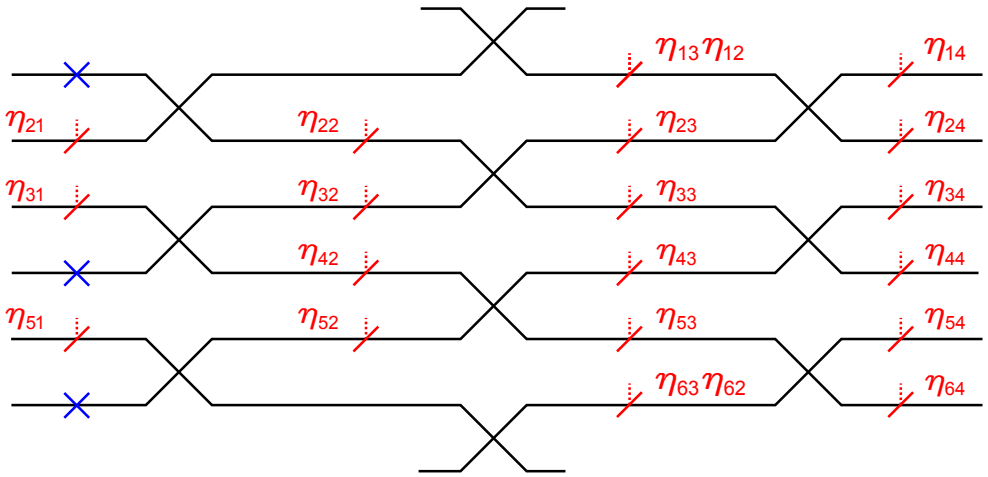


Figure A.5: We always inject photons on input modes 2,3, and 5 during the 3-photon boson sampling experiment. Consequently, all other losses in the first column can be ignored (blue crosses). In addition, the top and bottom fabricated beamsplitters are not written to the edge of the circuit and are thus effectively losses. They therefore commute with η_{13} and η_{63} and this property is used to combine terms.

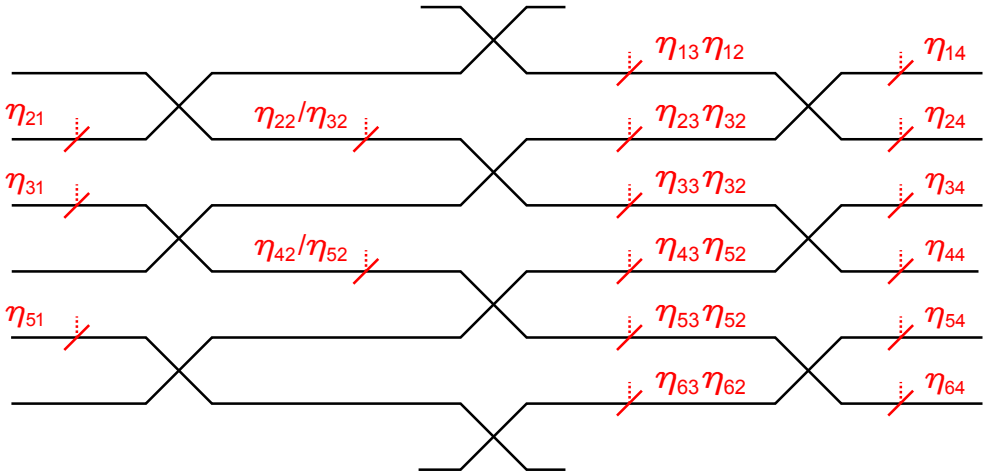


Figure A.6: We eliminate two losses from the second column. We consider losses η_{22} and η_{32} from Figure A.5 where we assume η_{22} is smaller. We introduce the relative loss η_{22}/η_{32} between these two modes and move the shared loss η_{32} to the right past the black beamsplitter. This same strategy is used to define the relative loss η_{42}/η_{52}.

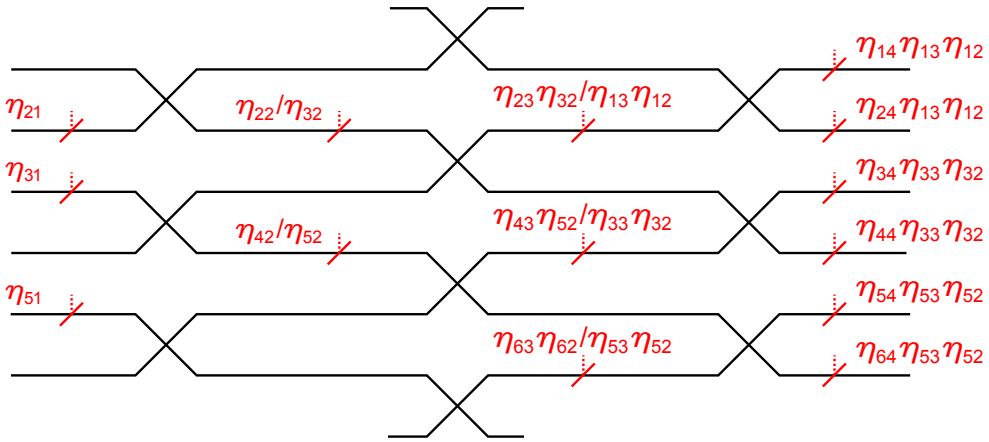


Figure A.7: Three new relative losses are defined in the third column with shared losses again moved to the right back the relevant black beamsplitter.

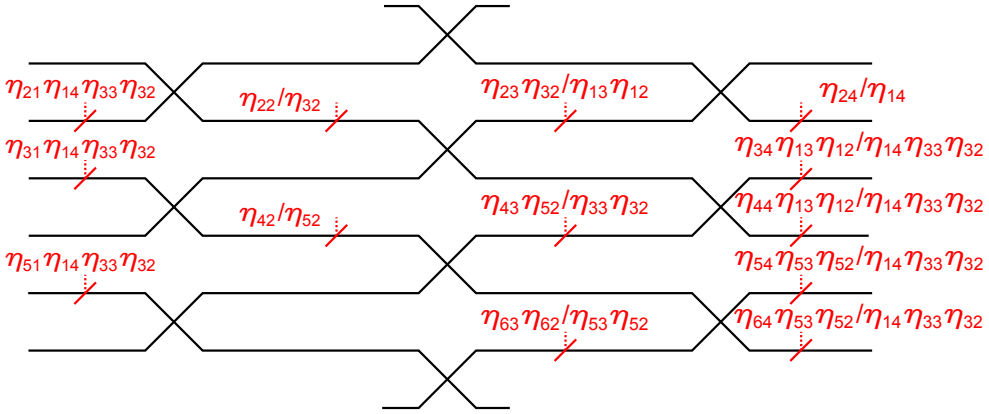


Figure A.8: We introduce five new relative losses at the output of our circuit, leaving a loss $\eta_{14}\eta_{33}\eta_{32}$ that is common to all modes. This shared loss can therefore be placed along any column of losses and we choose to place it along the input.

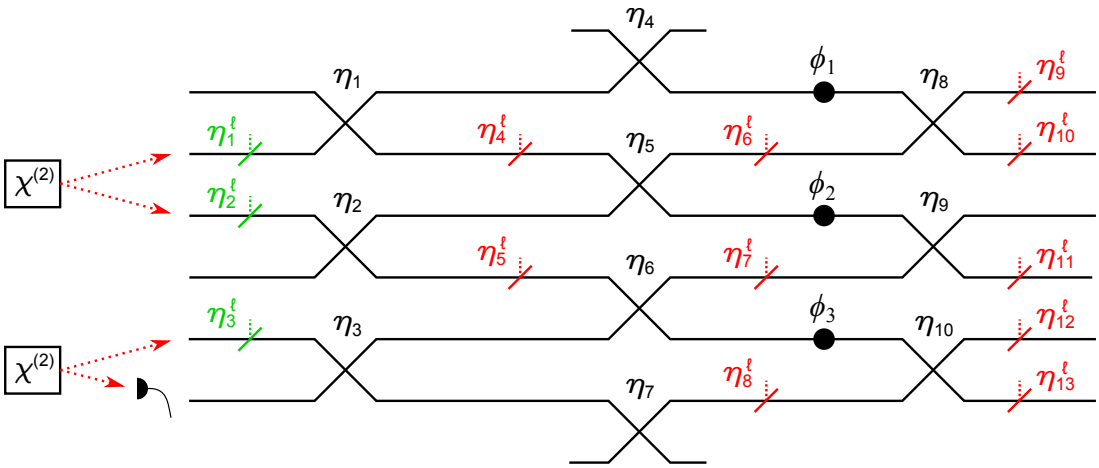


Figure A.9: We rename the 13 loss terms (red, green) according to the convention used in the main text. The loss terms in red are found through a numerical search that seeks to minimize the difference between our independently characterized transformation Λ and that inferred from the linear elements shown in black and red. The losses in green are obtained by using the number correlations between the modes of our SPDC sources.

drawn in black. Using Equation (5.8), we find an expression for the transformation given by the 23 total elements drawn in black and red in Figure A.9; this matrix is a function of the 10 unknown terms η_4^ℓ-η_{13}^ℓ drawn in red. The relevant 3×6 submatrix of this transformation, Λ^{th}, should match Λ which was already found through the blackbox characterization approach with the results shown in Figure 5.4. We thus perform a numerical optimization that searches for the set of losses that minimizes $\sum_{ij}|\Lambda_{ij}^{\text{th}}-\Lambda_{ij}|/|\Lambda_{ij}|$.

This search returns consistent results across a variety of initial conditions and search parameters. As another test of accuracy, when the search algorithm was allowed to vary the values of independently characterized elements, for example ϕ_1-ϕ_3, it returned self-consistent values. We quantify this accuracy by calculating $\Delta\tau_{ij}=|\tau_{ij}^{\text{num}}-\tau_{ij}^{\text{bb}}|$ and $\Delta\phi_{ij}=|\phi_{ij}^{\text{num}}-\phi_{ij}^{\text{bb}}|$ for all elements that are not zero or one in Figure 5.4. The superscripts refer to values obtained through this numerical search or from our blackbox characterization shown in Figure 5.4. We observe mean values $\overline{\Delta\tau}=0.028$ and $\overline{\Delta\phi}=8.0^\circ$ along with standard

deviations of $\sigma(\Delta\tau) = 0.037$ and $\sigma(\Delta\phi) = 6.3°$. The similarity of our inferred Λ^{num} with our blackbox characterization of Λ is consistent with an accurate modeling of η_4^ℓ-η_{13}^ℓ.

A different strategy is required to find the losses η_1^ℓ-η_3^ℓ in Figure 5.8 (green). This is due to our rescaling of the rows of Λ such that the maximum value in each row in Figure 5.4(b) is one. This implies there is an unknown loss on each input to the photonic circuit. Thus, while we can use Λ to determine the losses drawn in red, we still have the three unknown terms drawn in green in Figure 5.8.

We use the number correlations from the two modes of our SPDC source to estimate these three losses. For this measurement one of our sources is used in a heralded configuration with the heralded mode coupled to one of the circuit inputs; in this example we use the mode corresponding to η_3^ℓ. The fraction of counts that make it through our circuit is given by $\eta^{\text{tot}} = \sum_{m=1}^{6} N_{mH}/N_H$, where N_H is the number of counts at the herald detector and N_{mH} is the number of coincidence counts between circuit output mode m and the herald detector. Using Equation (5.14) then indicates

$$\eta^{\text{tot}} = \eta_3^\ell \sum_{j=1}^{6} |\tau_{5j}|^2, \tag{A.4}$$

where we use the elements τ_{5j} because the loss η_3^ℓ occurs on input $i = 5$ in Figure A.9. We thus measure η^{tot} and use our blackbox characterized values of τ_{5j} to determine η_3^ℓ. A similar strategy allows η_1^ℓ and η_2^ℓ to be determined.

This completes our elementwise characterization of the system shown in Figure A.9. A loss-independent analysis allowed the elements drawn in black to be determined. Independently, a blackbox characterization of our circuit was performed leading to Λ in Figure 5.4.

This matrix was used to both constrain a numerical search for the relative losses drawn in red. Finally, we solve for the global losses in green using our number-correlated SPDC source.

Appendix B

Photon distinguishability characterization

The photons coupled into our photonic circuit are slightly distinguishable from one another. This introduces error into our sampling because the matrix permanent no longer exactly predicts the bosonic scattering amplitudes. We introduce a rudimentary model that allows us to estimate the effect of this distinguishability. The three-photon boson sampling experiment is considered here, though a similar approach was used in the four-photon experiment.

B.1 Overview of model

We assume that the emission from our SPDC sources is factorable which allows us to accurately describe each of the three photons coupled to our circuit as a separate pure state. Spectral correlations between the photon pairs emitted from many SPDC sources often render this approximation invalid. However, our SPDC sources were engineered to emit the signal and idler into a near-spectrally factorable state[57] and we also use narrow spectral filters. These factors eliminate nearly all spectral correlations between the photons

used in our experiment.

One can decompose any pure state into a superposition of two modes, in our case a desired mode with creation operator $\hat{A}_j^\dagger(\xi_0)$ and an orthogonal mode with operator $\hat{A}_j^\dagger(\xi_i)$ [150]. This approach is useful in modeling distinguishability in boson sampling within the matrix permanent formalism. One can then write the three-photon state used in our boson sampling experiment as

$$\begin{aligned} |\mathbf{T}\rangle =&(\alpha_1 \hat{A}_1^\dagger(\xi_0) + \sqrt{1-\alpha_1^2}\hat{A}_1^\dagger(\xi_1)) \\ &(\alpha_2 \hat{A}_2^\dagger(\xi_0) + \sqrt{1-\alpha_2^2}\hat{A}_2^\dagger(\xi_2)) \\ &(\alpha_3 \hat{A}_3^\dagger(\xi_0) + \sqrt{1-\alpha_3^2}\hat{A}_3^\dagger(\xi_3))\,|0\rangle\,, \end{aligned} \tag{B.1}$$

where each photon has a different overlap, quantified by the distinguishability parameter α_i, with the desired mode ξ_0. We have used the subscripts $1, 2, 3$ on the operators to indicate coupling of photons into three different ports of our circuit, though technically we use inputs $2, 3, 5$ as illustrated in Figure 5.8.

Ideally, our experiment would use indistinguishable photons which implies there exists a mode ξ_0 such that $\alpha_{1,2,3} = 1$ in Equation (B.2). Variation between photons in any degree of freedom decreases α. One can perform a series of three HOM interference experiments between every pair of photons in this set, with the reduction in observed dip visibility constraining the relevant terms α_i and α_j. Three interference experiments then allow calculation of α_1, α_2, and α_3. Technical difficulties after boson sampling data collection prevented this series of experiments from being completed. Only one of the three requisite HOM interference experiments was performed.

We introduce the further approximation that variation in the spectra of our photons dominates differences between α_1, α_2, and α_3. The photon pairs emitted by our SPDC sources have different bandwidths; one is narrowband (≈ 3.5 nm) while the other is broadband (≈ 11 nm). While both photons are passed through a 3 nm filter, Figure B.1 shows that the observed spectra are still slightly distinguishable. Two narrowband photons and one broadband photon, all filtered, are used in the three-photon sampling experiment. We use our assumption that spectral variation dominates differences in α to rewrite Equation (B.2) as

$$\begin{aligned} |\mathbf{T}\rangle =&(\alpha_{nb}\hat{A}_1^\dagger(\xi_0) + \sqrt{1-\alpha_{nb}^2}\hat{A}_1^\dagger(\xi_1)) \\ &(\alpha_{nb}\hat{A}_2^\dagger(\xi_0) + \sqrt{1-\alpha_{nb}^2}\hat{A}_2^\dagger(\xi_2)) \\ &(\alpha_{bb}\hat{A}_3^\dagger(\xi_0) + \sqrt{1-\alpha_{bb}^2}\hat{A}_3^\dagger(\xi_3))\,|0\rangle\,. \end{aligned} \tag{B.2}$$

where we introduce two distinguishability parameters, α_{nb} and α_{bb}, corresponding to the narrowband and broadband SPDC photons.

B.2 Determination of α_{nb} from interference visibility

We interfere the two narrowband photons from our SPDC sources to determine α_{nb}. We first use the methods described in Appendix A to perform a loss-independent characterization of the reflectivity of a beamsplitter. This reflectivity allows us to predict a HOM interference visibility, assuming indistinguishable photons. Our two narrowband photons are then interfered on this beamsplitter and we use the reduction in dip visibility to estimate

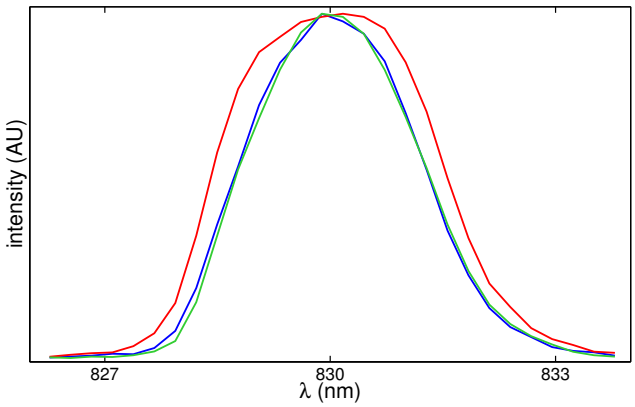


Figure B.1: We measure the spectra of the three SPDC modes, two narrowband (blue, green) and one broadband (red), coupled to our circuit for three-photon boson sampling. All modes are spectrally filtered but the broadband mode is still clearly distinguishable from the narrowband photons.

α_{nb}.

A beamsplitter within our circuit is chosen for this HOM experiment since it gives us the best estimate of the distinguishability α_{nb} during the interference relevant to boson sampling. Given the complex geometry of our circuit, we are forced to use the beamsplitter shown in Figure B.2 for this experiment. It is the only choice that allows us to predict a HOM dip visibility from a single beamsplitter reflectivity. A loss-independent characterization of this beamsplitter allows us to calculate its reflectivity $\eta^2 = 0.57$.

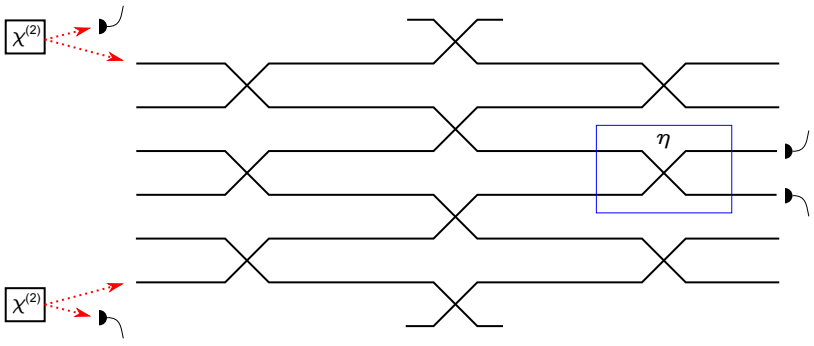


Figure B.2: Two SPDC sources, each in a heralded configuration, are used to perform a HOM interference experiment on a beamsplitter with reflectivity η (blue box) in our photonic circuit. Preceding beamsplitters in the circuit act as loss and thus reduce the count rate of our experiment but do not affect the interference visibility.

Our two narrowband SPDC photons are now used to perform a HOM interference experiment using the configuration shown in Figure B.2. The background-subtracted visibility of the resulting dip is $V = 0.903$. We next find the dependence of this visibility on the distinguishability parameter. The input state to the beamsplitter is

$$|\psi_{in}\rangle = (\alpha_{nb}\hat{A}_1^\dagger(\xi_0) + \sqrt{1-\alpha_{nb}^2}\hat{A}_1^\dagger(\xi_1))(\alpha_{nb}\hat{A}_6^\dagger(\xi_0) + \sqrt{1-\alpha_{nb}^2}\hat{A}_6^\dagger(\xi_2))\,|0\rangle\,. \tag{B.3}$$

The beamsplitter transformation in Equation (1.12) is then used to find the output state after the beamsplitter $|\psi_{out}\rangle$. The resulting interference visibility is then

$$V(\eta,\alpha) = \\ 1 - \frac{\alpha^4(1-2\eta^2)^2 + 2\alpha^2(1-\alpha^2)(\eta^4+(1-\eta^2)^2) + (1-\alpha^2)^2(\eta^4+(1-\eta^2)^2)}{\eta^4+(1-\eta^2)^2}. \tag{B.4}$$

Inserting our characterized beamsplitter reflectivity of $\eta^2 = 0.57$ and measured dip visibility of $V = 0.903$ into Equation (B.4) yields $\alpha_{nb} = 0.984$.

B.3 Determination of α_{bb} from spectral overlap

In principle, we could repeat the above HOM interference experiment with a narrowband and broadband photon to calculate α_{bb}. Technical constraints at the time of the experiment made this measurement impractical. Instead, here we estimate α_{bb} from the spectra shown in Figure B.2.

Following the form of the narrowband photons in Equation (B.2), we express the broad-

band photon as

$$\left(\alpha_{nb}\hat{A}^{\dagger}(\xi_{bb}) + \sqrt{1-\alpha_{nb}^2}\hat{A}^{\dagger}(\xi_3')\right)|0\rangle\,, \tag{B.5}$$

where ξ_{bb} captures the spectral differences from the desired mode ξ_0. We focus on the spectral degree of freedom and express the wavepacket operator as

$$\hat{A}^{\dagger}(\xi_i) = \int g_i(\omega)\hat{a}^{\dagger}(\omega)d\omega, \tag{B.6}$$

where $g_i(\omega)$ is the photon spectra and we assume flat spectral phase so that this function is real. The amplitude α_{bb} in Equation (B.2) can then be estimated by the spectral overlap

$$\alpha_{bb} = \alpha_{nb}\int g_{nb}(\omega)g_{bb}(\omega)d\omega. \tag{B.7}$$

Performing this overlap integral leads to $\alpha_{bb} = 0.969$.

Appendix C

FPGA coincidence counting

The author has written a variety of FPGA coincidence counting programs. The combination of low-cost hardware, straightforward Matlab-based interfacing, and flexible data collection has led these FPGA-based counters to be used in a wide variety of experiments across several research groups [24,25,77,158–164].

We begin this Appendix by outlining some of the common challenges when collecting data in quantum optics experiments. We then transition to describing two principal variants on our FPGA coincidence counter. The first is a general purpose program that has been widely used over the last two years, including in the boson sampling and fs-laser-written SFWM source experiments described in this thesis. The second is a counter optimized for high temporal resolution coincidence counting that was used in the UV-written SFWM source experiments to improve temporal filtering of fluorescence noise.

C.1 Data collection challenge

How much memory is required to store all the relevant information in a quantum optics experiment? Answering this question highlights some of the challenges in collecting and analyzing data and motivates the FPGA coincidence counter described in the latter part of this appendix.

Our analysis is simplified by making assumptions that are applicable to many experiments, including the ones described in this thesis. First, we assume discrete-variable operation so that our data set includes the occupation number of a finite number of photonic modes. A significant portion of current experiments operate in this regime, encoding information in either polarization, spatial, or frequency degrees of freedom. Second, since most experiments currently employ APDs, which cannot resolve photon number, we further assume that the occupation number of each mode is at most one. Third, we assume pulsed operation so that our experiment consists of a finite number of well-defined trials.

A complete data set requires storing *when* and *where* each detector clicks. This can take the form of a list where each entry describes at which time step and in which mode that detection event occured. If we collect data for a time τ with a temporal resolution w then there are τ/w total time-steps. These entries will require approximately $\log_2(\tau/w)$ bits to store on average throughout the data-set. If there are a total of M modes then on average we require approximately $\log_2(M)$ bits to record where each detection occurs. Assuming an average detection rate of r photons/sec then the size of our data set is

$$b_t \approx r\,\tau\left[\log_2\left(\frac{\tau}{w}\right)+\log_2\left(M\right)\right]=r\,\tau\log_2\left(\frac{\tau}{w}M\right), \tag{C.1}$$

where b_t is the memory required in bits.

Collecting a complete data set is often difficult because b_t can be very large. This occurs because many experiments require simultaneous detection of a specific number of photons. The probability of this occuring on any given trial is frequently small due to inevitable optical loss and the highly probabilistic nature of many single-photon sources. A frequently-employed solution is to collect an extremely large number of data points ($r\tau$) and postselect the few of interest.

The three-photon boson sampling experiment described in this thesis provides a convenient case-study. The combined rate at which any of the $M = 7$ APDs, a herald detector and the six detectors after the chip, fired was $r \approx 4 \cdot 10^5\,\mathrm{sec}^{-1}$. Data was collected for $\tau = 1$ week with a resolution of $w = 5$ ns. Inserting these values into Equation (C.1) gives $b_t = 11.9$ terabits. Storing and analyzing this large a data set would have been challenging with the computing resources available at the time of the experiment.

These difficulties in handling a complete data set will increase in many future, larger-scale experiments. We re-express Equation (C.1) to illustrate this point. We assign p to be the probability of a single photon being emitted by a heralded source and successfully being detected at the end of the experiment. The probability of an N-fold coincidence is then p^N. If we need to collect D total data points, each of which is an N-fold coincidence count, then we must perform D/p^N experimental trials which replaces the product $r\tau$ in Equation (C.1)

$$b_t \approx \frac{D}{p^N} \log_2 \left(\frac{\tau}{w} M \right). \tag{C.2}$$

The limiting factor in Equation (C.2) is the exponential scaling in N; future experiments

will have to allocate exponentially increasing resources if they wish to postselect on larger N-fold events while maintaining a complete data set.

These challenges have led many experimentalists to use alternate strategies that throw away unnecessary information and thus simplify the storage and analysis of data. Frequently, one does not care on what experimental trial each detector fires. It is instead sufficient to count the total number of times specific combinations of detectors click simultaneously over the course of the experiment. Using our three-photon boson sampling as an example, the seven detectors used in the experiment lead to

$$C = \sum_{i=1}^{7} \binom{7}{i} = 2^7 - 1 = 127, \tag{C.3}$$

combinations of detectors that can potentially fire simultaneously. For example, labeling the detectors alphabetically we can count the total number of clicks on detector B or simultaneous events between EFG. Our final boson sampling data set could then be reduced from the terabits mentioned above to 127 numbers, each of which are ≤ 38 bits since the number of counts on any detector must be less than the total number of photons detected $r\tau$. This strategy reduces the size of the boson sampling data set by nine orders of magnitude by not time-tagging every individual detection event.

This *direct coincidence counting* approach still encounters undesired exponential scalings. As Equation (C.3) suggests, counting all possible combinations of coincidence counts amongst M detectors requires storing approximately 2^M numbers which scales exponentially in the number of modes. Even counting only the N-fold coincidences requires $\binom{M}{N}$ numbers which, using our earlier assumption that $M \gg N$ so that most modes are unoccu-

pied, increases exponentially with N. These examples highlight that blindly applying this strategy to larger experiments still eventually results in untenable data collection requirements.

With no single approach that can clearly be scaled to larger systems, future work may rely on data collection customized to each individual experiment. For example, one may perform conditional time-tagging; a complete data set is collected conditional on certain combinations of detectors firing. Alternatively, one might do direct coincidence counting but only on a relevant subset of the possible detector combinations.

C.2 FPGA introduction

Field-programmable gate arrays (FPGAs) have emerged as a valuable data collection tool for many quantum optics researchers. FPGAs are integrated circuits that can be configured by the end-user to perform desired logic operations. They offer several benefits to researchers.

1. Large numbers of logical operations can be performed in parallel on dedicated circuity. In contrast, relying on traditional computers can lead to bottlenecks when many logical operations must queue for a small number of serial processors.

2. As a consequence of this parallelism, logic operations have predictable timing. Given certain inputs to the FPGA, one can accurately predict the time required to perform the desired logic irrespective of what is happening on the other inputs. This can be vital in some experiments, for example when feed-forward operation is necessary. In contrast, the aforementioned queuing for serial processor time can cause erratic

timing.

3. FPGAs can operate at an impressively high clock rate so that the delay incurred to perform logic is minimal[1]. While serial processors also operate at high clock rates, the parallelism of FPGAs frequently leads to a performance advantage.

Importantly, FPGAs offer these benefits in a very flexible platform. Rather than constructing a fixed application-specific integrated circuit (ASIC), FPGAs allow the user to repeatedly configure their circuit in a variety of different ways without purchasing new hardware. This facilitates data collection programs customized to particular experiments with minimal cost.

C.3 Coincidence counters

We introduce two direct coincidence counting programs written by this author and loaded onto the Xilinx SP605 development board. The first is a general-purpose program that has been in use for several years with minor variants used in a variety of experiments spanning several research groups[24,25,77,158–164]. The second program was developed more recently and was designed to optimize temporal resolution; it can coincidence count with a resolution that surpasses its own clock period.

The goal of this section is to provide a top-level summary; specifically we describe how the program turns input pulses from detectors into output coincidence count data. It is not intended as a user's manual, therefore many of the details required for actually using these FPGA programs are omitted.

[1]Importantly, low-cost FPGAs operate at clock frequencies far higher than the 80 MHz repetition rate of widely-used Ti-Sapphire lasers that pump many quantum optics experiments.

C.3.1 General purpose coincidence counter

We introduce the general purpose program in the simplest possible terms by describing how the FPGA counts coincident events from two detectors. The program is capable of registering higher-order coincidences, though this serves as a useful starting point. Figure C.1 shows that we consider a pair of photons produced via SPDC. These photons are coupled to detectors A and B which produce electronic pulses that are subsequently read in on separate FPGA input pins. When the voltage on these FPGA inputs crosses a user-defined threshold, the digital FPGA inputs (A_{in} and B_{in}) switch from 0 to 1. Since the two input pulses are from the same physical event, our program should eventually register these signals as coincident.

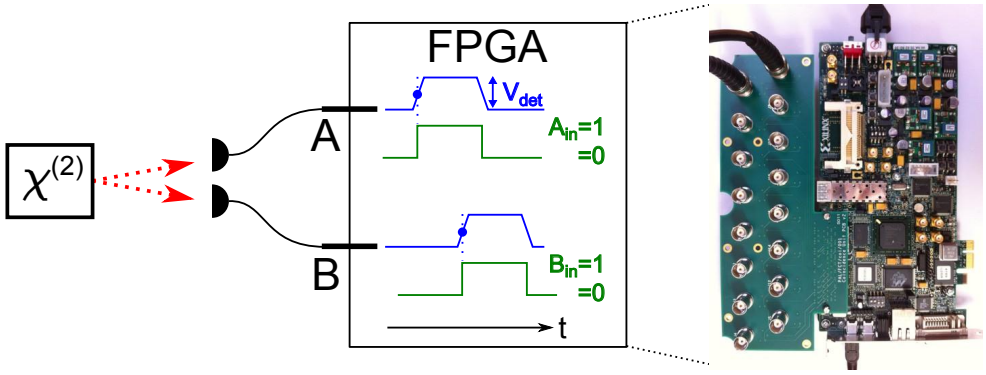


Figure C.1: We illustrate operation of our FPGA coincidence counting program with a SPDC pair source coupled to two detectors (A and B). Upon detecting a photon, a detector produces an electronic pulse with voltage V_{det} (blue). When these signals cross a threshold voltage (blue dot), a digital FPGA signal switches from 0 to 1 (green). Path length differences and different detector response times can cause a temporal offset between the signals from detector A and B. The photo on the right shows the hardware: the circuit board on the right contains the FPGA while the board on the left was developed internally to allow connection of many coaxial detector signals to the FPGA. Two detectors are in use in this picture.

The first step in successfully resolving coincident events is adding an adjustable electronic delay to each input pulse. This compensates for timing offsets between each input that can cause signals originating from simultaneous events to be read in by the FPGA at

different times. Such offsets can arise from path length mismatches, in either the optical or electronic domains, or different detector response times. If the pulses from detectors A and B are read in by the FPGA at times separated by Δt then we seek to add an electronic delay of $-\Delta t$ so that the signals are again coincident when they arrive at the counting logic.

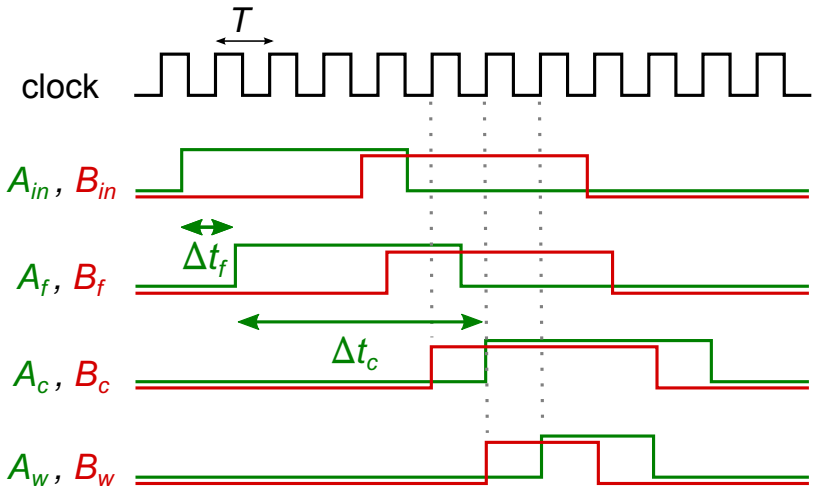


Figure C.2: Several signal processing steps are done on each input (green = detector A, red = detector B) before attempting to count coincidences. The raw input signals (A_{in}, B_{in}) are temporally offset due to different detector response times and path length differences before the FPGA. These signals are resynchronized with programmable electronic delays inside the FPGA. These delays are accomplished with the definition of two new signals, a fine (A_f, B_f) and coarse (A_c, B_c) delay, for each input. These new signals implement temporal delays Δt_f and Δt_c; these delays are only labeled for input A. The input and fine delay signals are asynchronous; they have no defined timing relationship relative to the FPGA's internal clock. All subsequent signals are synchronous; logic levels always change on the positive edge of the FPGA's clock. Note that if the user programs $\Delta t_c = x$ then the actual delay can be as large as $\Delta t_c = x + T$ as the FPGA clock is generally not synchronized to the pulsed pump laser. The temporal resolution of the coincidence counting is equivalent to the duration w of a final signal (A_w, B_w); in this example $w = 2T$. Detector jitter and imperfect setting of electronic delays can lead to the synchronous signals from different detectors (ex. A_c and B_c) being offset; here this offset is one clock period.

This electronic delay is added internally in the FPGA in two steps shown in Figure C.2: an asynchronous fine delay and a synchronous coarse delay. The fine delay uses dedicated hardware provided by the manufacturer to delay signals on each input pin in steps of approximately 30 ps over an 8-bit step range. Therefore, this element allows the user to program delays Δt_f of up to 7.6 ns with a resolution of 30 ps. This delay is asynchronous; if

a signal arrives at time t then it leaves the fine delay step at time $t + \Delta t_{\mathrm{f}}$ irrespective of the FPGA's internal clock. After this delay, all remaining logical steps occur on the rising edge of the FPGAs clock; this is synchronous logic. After the fine delay, an 8-bit synchronous coarse delay is applied yielding a programmable delay of up 255 FPGA clock periods. The most recent version of the FPGA coincidence counter has a clock period of 350 MHz and so the total coarse delay is $\Delta t_{\mathrm{c}} \approx 728$ ns with a resolution of 2.9 ns. Both coarse and fine delays can be set to individual values for each FPGA input.

This combination of fine and coarse delays is sufficient for many quantum optics experiments. First, the step size of the electronic delays is sufficiently small to resolve individual experimental trials. Ti-Sapphire pump lasers frequently set the period between experimental trials. These lasers often have an 80 MHz repetition rate so that the time between trials is 12.5 ns. The delays described above have step sizes well below this value. Second, the range of the electronic delays is sufficient to account for typical experimental conditions. Optical path-length mismatch, and different APD response times, commonly introduce timing mismatch of 100 ns or less, which can easily be accomodated with the above coarse delay.

The FPGA next modifies these delayed signals to account for the user-defined coincidence counting resolution. We define this temporal resolution as the coincidence window w. Since this program uses synchronous logic, w is a multiple of the FPGA clock period. The relative time of arrival of signals from the two detectors is $\Delta t = t_A - t_B$. If $|\Delta t| \leq w$ then the FPGA should count these events as coincident. This is accomplished by changing the temporal duration of the signals inside the FPGA to w. If these modified signals (A_w, B_w) overlap then the FPGA will register a coincidence.

It is not always desirable to maximize the temporal resolution by using the smallest coincidence window. For example, searching for the appropriate electronic delay on each input would proceed by scanning these delays until one observes a significant increase in coincidences. Since a photon-pair source is being used, this signifies that the correct electronic delay is being used so that signals from each detector reach the counting logic at the same time. This scan must use a step size no larger than the coincidence window and is thus aided by initially using a relatively large w. Detector jitter also implies using $w = T$ will lead to missed coincidence counts. Even if the electronic delays are set perfectly, jitter can cause the two detector signals to be read in by the FPGA on different clock cycles and thus not registered as coincident. For most data collection in this thesis, $w = 2T$ was used to avoid this issue while preserving reasonable temporal resolution. Sacrificing temporal resolution by increasing w does not significantly negatively impact many experiments. For example, when using SPDC photon sources and silicon APDs the time uncorrelated background, possibly originating from fluorescence or detector dark counts, is often negligible.

After these signal processing steps, the FPGA is now ready to count coincidences. While our example only illustrates functionality with detectors A and B, in reality there are up to M allowed inputs. The processed signals from each of these inputs are each fanned out to C identical sets of counting logic. The FPGA thus effectively has C copies of an M-fold coincidence counter. While this is redundant, it allows the user to count coincidences on any subset of these M detectors by instructing the FPGA to ignore irrelevant detectors. For example, Figure C.3 shows that, in our two detector example, we can instruct the first counter to ignore all inputs except detector A, the second counter to ignore all except detector B, and the third counter to ignore all except both detectors A and B. This collects

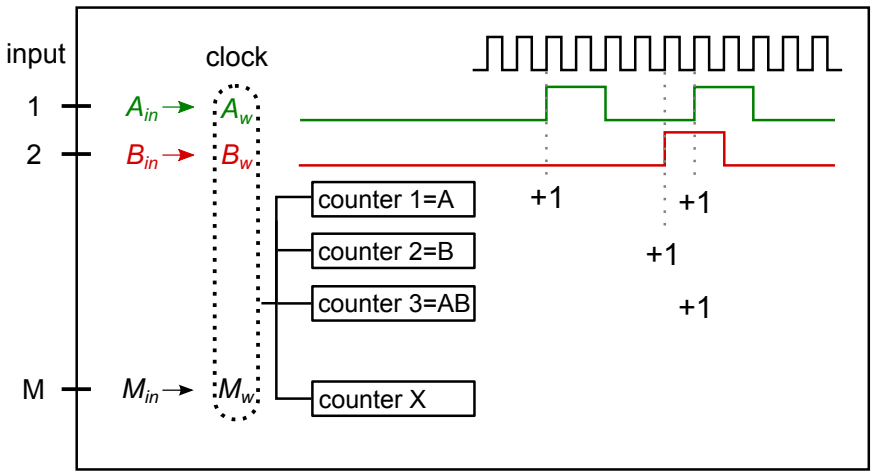


Figure C.3: We illustrate how the FPGA counts coincident events. The program accepts up to M possible signal inputs; in our example we have assumed two detectors, A and B, are plugged into inputs 1 and 2 respectively. Each of these inputs undergoes separate electronic delays and temporal windowing leading to outputs A_w and B_w. Representative forms for these two signals and the FPGA clock are illustrated. These signals are fanned out to X duplicate copies of counting logic, each of which is capable of counting an M-fold coincidence. Each counter can be separated programmed to ignore certain inputs to count lower-order events; for example counter 1 counts single events from detector A because it ignores all other inputs. Incrementing of counters are designated by $+1$ notation at the appropriate clock cycle.

a full set of statistics: single events on A and B, and coincidences between A and B.

This coincidence counting strategy has the advantage of allowing a large number of detector inputs while giving the user the flexibility to collect data on only the relevant combinations of detectors. For example, in the latest program version $M = 18$ inputs are possible but counting coincidences on all $2^{18} - 1 = 262143$ possible combinations is generally not necessary. Our FPGA is currently programmed to accomodate $C = 70$ counters which has been sufficient for our experiments to-date. Increasing M and C further is possible without new hardware.

A suite of Matlab commands were developed to allow easy control of the FPGA program. Through this interface, the user sets all the required parameters for counting and then reads the results after data collection is complete. A summary of the available commands is given

Table C.1: A set of Matlab commands is used to set relevant coincidence counting parameters and collect data. Some entries are blank because they are not defined in the text, though their functionality is apparent from context.

parameter	file	description
Δt_f	setFineDelay.m	asynchronous fine delay
Δt_c	setCoarseDelay.m	synchronous coarse delay
w	setWindow.m	coincidence window
-	setChannels.m	inputs for each counter
-	setCntTime.m	data collection time
-	takeData.m	collect data/send results to PC

in Table C.1. This functionality was used to construct a graphical user interface (GUI), shown in Figure C.4, which can continually refresh data and is thus a valuable tool for optimizing experimental alignment. For longer data-taking sessions, for example the week-long boson sampling data collection, one writes a specific Matlab script that calls the FPGA commands in Table C.1.

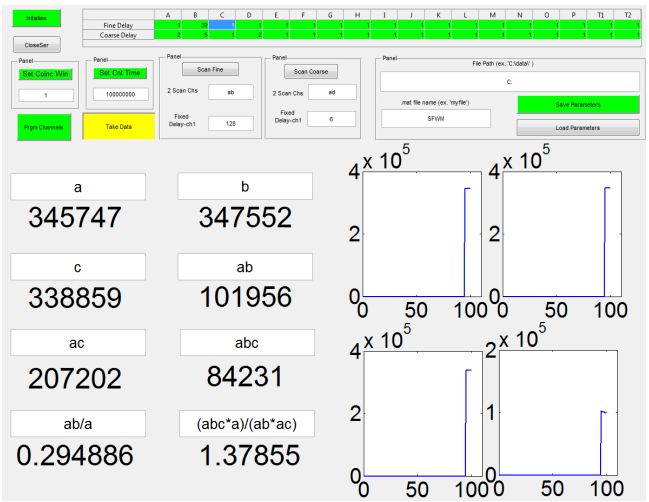


Figure C.4: An example screenshot from the coincidence counter GUI is shown. Counting parameters, specifically the electronic delays, coincidence window, and counting time, are set in the upper section of the GUI. Data from 8 of the counters is shown on the left side, with graphical plots on the right showing counts vs time for 4 of these counters. Algebraic formulas are also allowed inputs, as shown in the lower left, which allows one to adjust an experiment to optimize parameters such as η_H or $g_H^{(2)}(0)$. Scanning of the fine and coarse electronic delays, in the top center, allow automation of delay setting.

C.3.2 High temporal resolution

Some experiments would significantly benefit from counting coincident events with the smallest possible coincidence window w. One application of this improved temporal resolution is improved temporal filtering of noise. Given a continuous noise background but pulsed signal, the signal to noise ratio increases as the coincidence window w is decreased. No 'true' coincidences are lost if w remains larger than both the optical pulse duration and detector jitter.

The synchronous counting logic detailed in Figure C.3 is incapable of counting coincidences with a temporal resolution w below its clock period. The lowest possible clock period for our FPGA is 2.5 ns and we must use a resolution of $w = 5$ ns to avoid losing data due to detector jitter. Since our optical pulses are sub-picosecond and the jitter of our APDs is a few hundred picoseconds, the synchronous FPGA coincidence counter described above only applies a broad temporal filter. One can significantly reduce w in many existing experiments without losing data.

We developed a new asynchronous coincidence counter that implements a smaller w to aid data collection in experiments with a large time uncorrelated background. The key asynchronous element that makes this program possible is the fine delay element first depicted in Figure C.2. Previously this element was used once on each input to provide an electronic delay. Figure C.5 illustrates that this new program uses these fine delays many times; the raw signal on each input is fanned out to multiple fine delay elements to create a sequence of S signals with temporal spacing Δt_s. In principle, this allows us to implement a coincidence window less than the clock period: $w < T$. We introduce the variable n_x

which describes the number of elements from input x that are high on a given clock cycle and is illustrated in Figure C.5. We register a coincidence between detectors A and B if the following three conditions are satisfied: $0 < n_a < S$, $0 < n_b < S$, and $|n_a - n_b| = x$. The first two conditions merely indicate that both detectors fired, while the latter indicate they fired within a time $w = x\Delta t_s$ of each other where x is a user-defined integer.

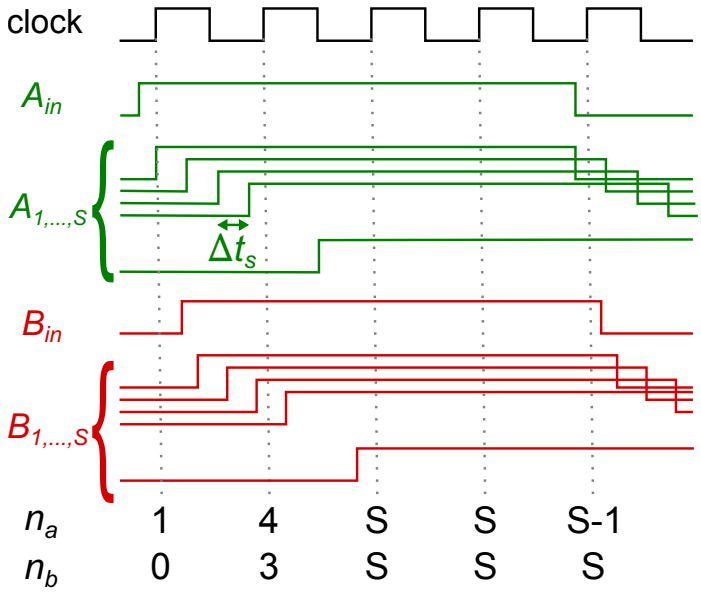


Figure C.5: Repeated use of an asynchronous delay allows counting of coincident events with a temporal resolution below the FPGA clock period. Each input signal (A_{in}, B_{in}) is fanned out to create multiple signals ($A_{1,2,...,S}$) each of which are programmed with a different delay to create a series with temporal spacing Δt_s. On the positive edge of every clock cycle, the FPGA logic counts how many of these signals, n_x for input x, are high. We ensure that the spacing from the first to last signal $(S-1)\Delta t_s$ is greater than the FPGA clock period so that each detector firing leads to $0 < n_x < S$ on at least one clock cycle. One can then count coincidence events with a resolution $w = \Delta t_s$ which can be significantly less than the clock period.

A calibration step is essential in producing the evenly spaced set of signals shown in Figure C.5 upon which this counting strategy relies. Ideally, if a detector signal arrives at time t then signal m in the sequence goes high at time $t_m = t + m\Delta t_s$. In reality, each of

these fine delay elements has a different fixed timing offset t_m^0 leading to

$$t_m = t + t_m^0 + \Delta t_s m. \tag{C.4}$$

This offset can be relatively large and vary significantly between elements, which is unsurprising since this hardware was not intended to be used in sub-nanosecond timing performance applications. However, the offsets between two elements of the series, for example $t_{A_1}^0$ and $t_{A_2}^0$, are easily determined. A signal is applied to input A, for example from a function generator, and coincidences are counted using the synchronous method described in Section C.3.1 between these two elements of the series as the delay on one of the elements is scanned. Since these two internal FPGA signals originate from the same input pin, the peak of these coincidences occurs at time $t_{A_1} - t_{A_2} = 0$. Inserting Equation (C.4), gives

$$t_2^0 - t_1^0 = (m_1 - m_2)\Delta t_s. \tag{C.5}$$

We repeat this calibration step for all elements of the series as depicted in Figure C.6 which gives all the timing offsets relative to input 1. Once these offsets are known, we can compensate for them when setting delays on signals $A_{2,\ldots,S}$ to produce the evenly spaced set shown in Figure C.5.

In principle, we can set $w = \Delta t_s$ which can be as low as 30 ps, the resolution of the fine delay hardware. In practice, this isn't optimal. Our APD detectors have jitter on the order of hundreds of picoseconds; we will miss true coincidences if w is set below this value. In addition, the number of fine delay elements S must be large enough to span an FPGA

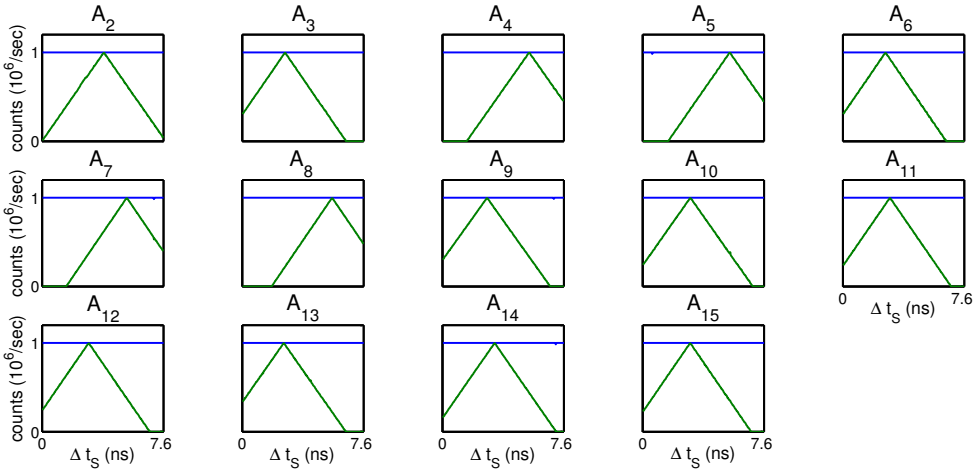


Figure C.6: A calibration step is required before using the asynchronous coincidence counter. A function generator applies a 1 MHz square-wave to FPGA input A. In each subplot, a delay Δt_{A_m} is incremented while the number of individual events (blue) and two-fold coincidences (green) using signals A_1 and A_m are counted. Fully synchronous counting logic is used in this calibration with $w = T \approx 3.3$ ns. If there were no fixed temporal offsets between channels, $t_s^0 = 0$, then the peak of the green curve would be on the far left of each subplot because in this example we implement a fixed $\Delta t_{A_1} = 0$ ns. The nonzero positions of the green peaks allows us to determine $t_S^0 - t_1^0$ for all signals, which can later be compensated when counting coincidences.

clock period, $(S-1)\Delta t_s > T$, or we again omit true coincidences. As we lower Δt_s, then S must increase with an ultimate upper bound set by the number of fine delay elements on the FPGA. In practice, calibrating an extremely large number of fine delays becomes increasingly cumbersome and places a practical upper bound that is below that allowed by the hardware. The current program iteration uses $S = 15$ signals in the temporal sequence for each input, with a typical spacing of $\Delta t_s = 400$ ps.

We present representative results showing successful operation of this high-temporal-resolution coincidence counter. A function generator is applied to input pins A and B, the calibration routine illustrated in Figure C.6 is performed, and we set $\Delta t_s = 400$ ps. In this example, two signals are coincident if $|n_a - n_b| \leq 1$. The asynchronous delay on one of the signals is scanned while coincidences are counted, with the results shown in Figure

C.7. For a window w, such a scan should produce coincidences in a temporal range $2w$. From Figure C.7, we estimate $w = 1$ ns and note that, at the optimal delay point, there are no lost two-fold coincidences. Using a synchronous counting program we would need to set $w = 2T$ to collect all true coincidences. Since the minimum clock period specified by the manufacturer is $T = 2.5$ ns, we would need to use $w = 5$ ns with a fully synchronous counting program. We have thus significantly shortened w, and thus improved temporal resolution, through this asynchronous strategy.

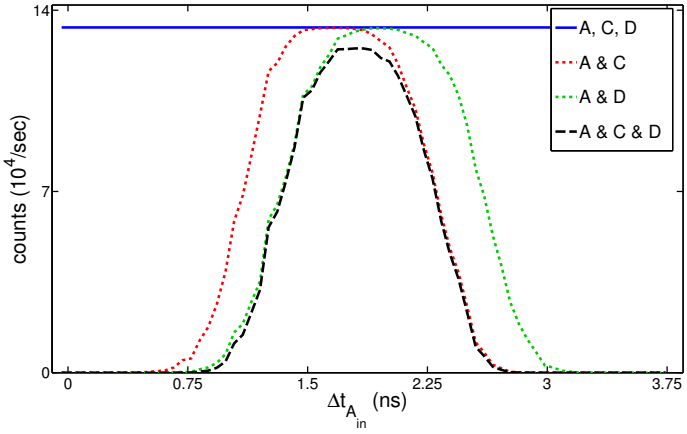


Figure C.7: Asynchronous coincidence counting allows high temporal resolution. A function generator is connected to FPGA inputs A, B, and C. After all three channels undergo calibration, the fine delay on input A ($\Delta t_{A_{in}}$) is scanned while single (solid), two-fold coincident (dotted), and three-fold coincident (dashed) events are counted. For illustrative purposes, $\Delta t_{B_{in}}$ and $\Delta t_{C_{in}}$ are programmed with a slight offset; for real data collection this would be fixed so that the red and green curves overlaid one another. The total duration of the two-fold coincidences is, by definition, $2w$; we thus estimate $w \approx 1$ ns. At appropriate $\Delta t_{A_{in}}$, the two-fold coincidences are equivalent to the single events indicating no data is lost. A synchronous coincidence counter on this FPGA would require $2w > 5$ ns.

The improved temporal filtering allowed by this FPGA program was used in the on-chip SFWM experiments using UV-written waveguides. Figure 3.16 showed that these waveguides exhibit significant fluorescence and Table 3.1 showed that decreasing the coincidence window, which was only possible with the novel counting method described in

this section, produced a desirable increase in η_H and decrease in $g_H^{(2)}(0)$. Both observations are consistent with improved temporal filtering of fluorescence increasing the number correlations between signal and idler modes.

Bibliography

[1] E. Schrödinger. Are there Quantum Jumps? *The British Journal for the Philosophy of Science*, 3(11):233–242, 1952.

[2] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[3] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. Quantum computers. *Nature*, 464(7285):45–53, 2010.

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

[5] J. W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger. Experimental Entanglement Swapping: Entangling Photons That Never Interacted. *Physical Review Letters*, 80(18):3891–3894, 1998.

[6] D. Bouwmeester, J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. Experimental quantum teleportation. *Nature*, 390:575–579, 1997.

[7] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik. Unconditional Quantum Teleportation. *Science*, 282:706–709, 1998.

[8] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat. Detection-Loophole-Free Test of Quantum Nonlocality, and Applications. *Physical Review Letters*, 111(13):130406, 2013.

[9] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger. Bell violation using entangled photons without the fair-sampling assumption. *Nature*, 497(7448):227–230, 2013.

[10] C. H. Bennett. Quantum Cryptography Using Any Two Nonorthogonal States. *Physical Review Letters*, 68(21):3121–3124, 1992.

[11] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning. Demonstration of an all-optical quantum controlled-NOT gate. *Nature*, 426(6964):264–267, 2003.

[12] S. Gasparoni, J. W. Pan, P. Walther, T. Rudolph, and A. Zeilinger. Realization of a Photonic Controlled-NOT Gate Sufficient for Quantum Computation. *Physical Review Letters*, 93(2):020504, 2004.

[13] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg. Super-resolving Phase Measurements with a Multiphoton Entangled State. *Nature*, 429:161–164, 2004.

[14] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi. Beating the Standard Quantum Limit with Four-Entangled Photons. *Science*, 316(5825):726–729, 2007.

[15] H. de Riedmatten, I. Marcikic, J. A. W. van Houwelingen, W. Tittel, H. Zbinden, and N. Gisin. Long-distance entanglement swapping with photons from separated sources. *Physical Review A*, 71(5):050302, 2005.

[16] D. Salart, A. Baas, J. A. W. van Houwelingen, N. Gisin, and H. Zbinden. Spacelike Separation in a Bell Test Assuming Gravitationally Induced Collapses. *Physical Review Letters*, 100(22):220404, 2008.

[17] N. Gisin and R. Thew. Quantum communication. *Nature Photonics*, 1(3):165–171, 2007.

[18] X. S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin, and A. Zeilinger. Quantum teleportation over 143 kilometres using active feed-forward. *Nature*, 489(7415):269–273, 2012.

[19] J. L. O'Brien, A. Furusawa, and J. Vučković. Photonic quantum technologies. *Nature Photonics*, 3(12):687–695, 2009.

[20] N. Thomas-Peter, N. K. Langford, A. Datta, L. Zhang, B. J. Smith, J. B. Spring, B. J. Metcalf, H. B. Coldenstrodt-Ronge, M. Hu, J. Nunn, and I. A. Walmsley. Integrated photonic sensing. *New Journal of Physics*, 13:055024, 2011.

[21] B. J. Smith, D. Kundys, N. Thomas-Peter, and I. A. Walmsley. Phase-controlled integrated photonic quantum circuits. *Optics Express*, 17(16):264–267, 2009.

[22] A. Peruzzo, M. Lobino, J. C. F. Matthews, N. Matsuda, A. Politi, K. Poulios, X. Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, Y. Bromberg, Y. Silberberg, M. G. Thompson, and J. L. O'Brien. Quantum Walks of Correlated Photons. *Science*, 329(5998):1500–1503, 2010.

[23] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. F. Matthews, M. G. Thompson, and J. L. O'Brien. Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit. *Nature Photonics*, 6(1):45–49, 2011.

[24] B. J. Metcalf, N. Thomas-Peter, J. B. Spring, D. Kundys, M. A. Broome, P. C. Humphreys, X. M. Jin, M. Barbieri, W. S. Kolthammer, J. C. Gates, B. J. Smith, N. K. Langford, P. G. R. Smith, and I. A. Walmsley. Multiphoton quantum interference in a multiport integrated photonic device. *Nature Communications*, 4:1356, 2013.

[25] J. B Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X. M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley. Boson Sampling on a Photonic Chip. *Science*, 339:798–801, 2013.

[26] A. Crespi, R. Osellame, R. Ramponi, V. Giovannetti, R. Fazio, L. Sansoni, F. D. Nicola, F. Sciarrino, and P. Mataloni. Anderson localization of entangled photons in an integrated quantum walk. *Nature Photonics*, 7(4):322–328, 2013.

[27] G. D. Marshall, A. Politi, J. C. F. Matthews, P. Dekker, M. Ams, M. J. Withford, and J. L. O'Brien. Laser written waveguide photonic quantum circuits. *Optics Express*, 17(15):12546–12554, 2009.

[28] A. Crespi, R. Ramponi, R. Osellame, L. Sansoni, I. Bongioanni, F. Sciarrino, G. Vallone, and P. Mataloni. Integrated photonic quantum gates for polarization qubits. *Nature Communications*, 2:566, 2011.

[29] J. O. Owens, M. A. Broome, D. N. Biggerstaff, M. E. Goggin, A. Fedrizzi, T. Linjordet, M. Ams, G. D. Marshall, J. Twamley, M. J. Withford, and A. G. White. Two-photon quantum walks in an elliptical direct-write waveguide array. *New Journal of Physics*, 13(7):075003, 2011.

[30] R. Loudon. *The Quantum Theory of Light*. Oxford University Press, Oxford, 1973.

[31] A. Zeilinger. General properties of lossless beam splitters in interferometry. *American Journal of Physics*, 49(9):882–883, 1981.

[32] M. O. T. Cunha, J. A. Dunningham, and V. Vedral. Entanglement in single-particle systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2085):2277–2286, 2007.

[33] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of Subpicosecond Time Intervals between Two Photons by Interference. *Physical Review Letters*, 59(18):2044–2046, 1987.

[34] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov. Invited review article: Single-photon sources and detectors. *The Review of Scientific Instruments*, 82(7):071101, 2011.

[35] A. Kuhn, M. Hennrich, and G. Rempe. Deterministic Single-Photon Source for Distributed Quantum Networking. *Physical Review Letters*, 89(6):4–7, 2002.

[36] J. McKeever, A. Boca, A. D. Boozer, R. Miller, J. R. Buck, A. Kuzmich, and H. J. Kimble. Deterministic generation of single photons from one atom trapped in a cavity. *Science*, 303(5666):1992–4, 2004.

[37] P. B. R. Nisbet-Jones, J. Dilley, D. Ljunggren, and A. Kuhn. Highly efficient source for indistinguishable single photons of controlled shape. *New Journal of Physics*, 13(10):103036, 2011.

[38] M. Keller, B. Lange, K. Hayasaka, W. Lange, and H. Walther. Continuous generation of single photons with controlled waveform in an ion-trap cavity system. *Nature*, 431:1075–1078, 2004.

[39] H. G. Barros, A. Stute, T. E. Northup, C. Russo, P. O. Schmidt, and R. Blatt. Deterministic single-photon source from a single ion. *New Journal of Physics*, 11(10):103004, 2009.

[40] P. Michler, A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, L. Zhang, E. Hu, and A. Imamoglu. A quantum dot single-photon turnstile device. *Science*, 290(5500):2282–2285, 2000.

[41] M. Pelton, C. Santori, J. Vucković, B. Zhang, G. Solomon, J. Plant, and Y. Yamamoto. Efficient Source of Single Photons: A Single Quantum Dot in a Micropost Microcavity. *Physical Review Letters*, 89(23):233602, 2002.

[42] Z. Yuan, B. E. Kardynal, R. M. Stevenson, A. J. Shields, C. J. Lobo, K. Cooper, N. S. Beattie, D. A. Ritchie, and M. Pepper. Electrically driven single-photon source. *Science*, 295(5552):102–105, 2002.

[43] C. Kurtsiefer, S. Mayer, P. Zarda, and H. Weinfurter. Stable solid-state source of single photons. *Physical Review Letters*, 85(2):290–293, 2000.

[44] P. Tamarat, T. Gaebel, J. Rabeau, M. Khan, A. Greentree, H. Wilson, L. Hollenberg, S. Prawer, P. Hemmer, F. Jelezko, and J. Wrachtrup. Stark Shift Control of Single Optical Centers in Diamond. *Physical Review Letters*, 97(8):083002, 2006.

[45] S. Pezzagna, D. Rogalla, D. Wildanger, J. Meijer, and A. Zaitsev. Creation and nature of optical centres in diamond for single-photon emission-overview and critical remarks. *New Journal of Physics*, 13(3):035024, 2011.

[46] A. Kuzmich, W. P. Bowen, A. D. Boozer, A. Boca, C. W. Chou, L. M. Duan, and H. J. Kimble. Generation of nonclassical photon pairs for scalable quantum communication with atomic ensembles. *Nature*, 423(6941):731–734, 2003.

[47] C. H. van der Wal, M. D. Eisaman, A. André, R. L. Walsworth, D. F. Phillips, A. S. Zibrov, and M. D. Lukin. Atomic memory for correlated photon states. *Science (New York, N.Y.)*, 301(5630):196–200, 2003.

[48] C. Chou, S. Polyakov, A. Kuzmich, and H. J. Kimble. Single-Photon Generation from Stored Excitation in an Atomic Ensemble. *Physical Review Letters*, 92(21):213601, 2004.

[49] D. Matsukevich, T. Chanelière, S. Jenkins, S. Y. Lan, T. A. B. Kennedy, and A. Kuzmich. Deterministic Single Photons via Conditional Quantum Evolution. *Physical Review Letters*, 97:013601, 2006.

[50] R. Boyd. *Nonlinear optics*. Academic Press, 2008.

[51] W. P. Grice, A. B. U'Ren, and I. A. Walmsley. Eliminating frequency and space-time correlations in multiphoton states. *Physical Review A*, 64(6):063815, 2001.

[52] A. B. U'Ren, C. Silberhorn, K. Banaszek, I. A. Walmsley, R. Erdmann, W. P. Grice, and M. G. Raymer. Generation of Pure-State Single-Photon Wavepackets by Conditional Preparation Based on Spontaneous Parametric Downconversion. *Laser Physics*, 15(1):146–161, 2005.

[53] M. Varnava, D. Browne, and T. Rudolph. How Good Must Single Photon Sources and Detectors Be for Efficient Linear Optical Quantum Computation? *Physical Review Letters*, 100(6):060502, 2008.

[54] A. Datta, L. Zhang, N. Thomas-Peter, U. Dorner, B. Smith, and I. A. Walmsley. Quantum metrology with imperfect states and detectors. *Physical Review A*, 83(6):063836, 2011.

[55] T. Jennewein, M. Barbieri, and A. G. White. Single-photon device requirements for operating linear optics quantum computing outside the post-selection basis. *Journal of Modern Optics*, 58(3-4):276–287, 2011.

[56] M. Lucamarini, G. Vallone, I. Gianani, P. Mataloni, and G. Di Giuseppe. Device-independent entanglement-based Bennett 1992 protocol. *Physical Review A*, 86(3):032325, 2012.

[57] P. J. Mosley, J. S. Lundeen, B. J. Smith, P. Wasylczyk, A. B. U'Ren, C. Silberhorn, and I. A. Walmsley. Heralded Generation of Ultrafast Single Photons in Pure Quantum States. *Physical Review Letters*, 100:133601, 2008.

[58] T. Gerrits, M. J. Stevens, B. Baek, B. Calkins, A. Lita, S. Glancy, E. Knill, S. W. Nam, R. P. Mirin, R. H. Hadfield, R. S. Bennink, W. P. Grice, S. Dorenbos, T. Zijlstra, T. Klapwijk, and V. Zwiller. Generation of degenerate, factorizable, pulsed squeezed light at telecom wavelengths. *Optics Express*, 19(24):24434–24447, 2011.

[59] G. Harder, V. Ansari, B. Brecht, T. Dirmeier, C. Marquardt, and C. Silberhorn. An optimized photon pair source for quantum circuits. *Optics Express*, 21(12):13975–13985, 2013.

[60] A. Christ and C. Silberhorn. Limits on the deterministic creation of pure single-photon states using parametric down-conversion. *Physical Review A*, 85(2):023829, 2012.

[61] G. P. Agrawal. *Nonlinear Fiber Optics*. Academic Press, 2007.

[62] M. Fiorentino, P. L. Voss, J. E. Sharping, and P. Kumar. All-fiber photon-pair source for quantum communications. *IEEE Photonics Technology Letters*, 14(7):983–985, 2002.

[63] K. Inoue and K. Shimizu. Generation of quantum-correlated photon pairs in optical fiber: Influence of spontaneous Raman scattering. *Japanese Journal of Applied Physics*, 43:8048–8052, 2004.

[64] J. G. Rarity, J. Fulconis, J. Duligall, W. J. Wadsworth, and P. St. J. Russell. Photonic crystal fiber source of correlated photon pairs. *Optics Express*, 13(2), 2005.

[65] O. Alibart, J. Fulconis, G. K. L. Wong, S. G. Murdoch, W. J. Wadsworth, and J. G. Rarity. Photon pair generation using four-wave mixing in a microstructured fibre: theory versus experiment. *New Journal of Physics*, 8(5):1–20, 2006.

[66] K. Garay-Palmett, H. J. McGuinness, O. Cohen, J. S. Lundeen, R. Rangel-Rojo, A. B. U'Ren, M. G. Raymer, C. J. McKinstrie, S. Radic, and I. A. Walmsley. Photon pair-state preparation with tailored spectral properties by spontaneous four-wave mixing in photonic-crystal fiber. *Optics Express*, 15(22):14870, 2007.

[67] O. Cohen, J. S. Lundeen, B. J. Smith, G. Puentes, P. J. Mosley, and I. A. Walmsley. Tailored Photon-Pair Generation in Optical Fibers. *Physical Review Letters*, 102(12):123603, 2009.

[68] M. Halder, J. Fulconis, B. Cemlyn, A. Clark, C. Xiong, W. J. Wadsworth, and J. G. Rarity. Nonclassical 2-photon interference with separate intrinsically narrowband fibre sources. *Optics Express*, 17(6):4670–4676, 2009.

[69] A. Clark, B. Bell, J. Fulconis, M. M. Halder, B. Cemlyn, O. Alibart, C. Xiong, W. J. Wadsworth, and J. G. Rarity. Intrinsically narrowband pair photon generation in microstructured fibres. *New Journal of Physics*, 13(6):065009, 2011.

[70] D. Fukuda, G. Fujii, T. Numata, K. Amemiya, A. Yoshizawa, H. Tsuchida, H. Fujino, H. Ishii, T. Itatani, S. Inoue, and T. Zama. Titanium-based transition-edge photon number resolving detector with 98% detection efficiency with index-matched small-gap fiber coupling. *Optics Express*, 19(2):870–875, 2011.

[71] B. Calkins, P. L. Mennea, A. E. Lita, B. J. Metcalf, W. S. Kolthammer, A. Lamas-Linares, J. B. Spring, P. C. Humphreys, R. P. Mirin, J. C. Gates, P. G. R. Smith, I. A. Walmsley, T. Gerrits, and S. W. Nam. High quantum-efficiency photon-number-resolving detector for photonic on-chip information processing. *Optics Express*, 21(19):22657–22670, 2013.

[72] Peter J. Mosley. *Generation of Heralded Single Photons in Pure Quantum States*. PhD thesis, University of Oxford, 2007.

[73] C. Söller, O. Cohen, B. J. Smith, I. A. Walmsley, and C. Silberhorn. High-performance single-photon generation with commercial-grade optical fiber. *Physical Review A*, 83:031806, 2011.

[74] A. M. Branczyk, A. Fedrizzi, T. M. Stace, T. C. Ralph, and A. G. White. Engineered optical nonlinearity for quantum light sources. *Optics Express*, 19(1):55–65, 2011.

[75] P. B. Dixon, J. H. Shapiro, and F. N. C. Wong. Spectral engineering by Gaussian phase-matching for quantum photonics. *Optics Express*, 21(5):5879–5890, 2013.

[76] A. Eckstein, A. Christ, P. J. Mosley, and C. Silberhorn. Highly efficient single-pass source of pulsed single-mode twin beams of light. *Physical Review Letters*, 106:13603, 2011.

[77] J. B. Spring, P. S. Salter, B. J. Metcalf, P. C. Humphreys, M. Moore, N. Thomas-Peter, X. M. Jin, N. K. Langford, W. S. Kolthammer, M. J. Booth, and I. A. Walmsley. On-chip low loss heralded source of pure single photons. *Optics Express*, 21(11):5932–5935, 2013.

[78] C. K. Law, I. A. Walmsley, and J. H. Eberly. Continuous Frequency Entanglement: Effective Finite Hilbert Space and Entropy Control. *Physical Review Letters*, 84(23):5304–5307, 2000.

[79] Andreas Eckstein. *Mastering quantum light pulses with nonlinear waveguide interactions*. PhD thesis, Friedrich-Alexander University Erlangen-Nurnberg, 2012.

[80] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.

[81] A. Christ, K. Laiho, A. Eckstein, K. N. Cassemiro, and C. Silberhorn. Probing multimode squeezing with correlation functions. *New Journal of Physics*, 13(3):033027, 2011.

[82] Christoph Soller. *Optical fiber sources of pulsed single- and multi-photon states for quantum networks*. PhD thesis, Friedrich-Alexander Universitat Erlangen-Nurnberg, 2011.

[83] K. Garay-Palmett, A. B. URen, and R. Rangel-Rojo. Conversion efficiency in the process of copolarized spontaneous four-wave mixing. *Physical Review A*, 82(4):043809, 2010.

[84] F. Zhang and J. W. Y. Lit. Temperature and strain sensitivity measurements of high-birefringent polarization-maintaining fibers. *Applied Optics*, 32(13):2213–2218, 1993.

[85] W. Zou, Z. He, and K. Hotate. Complete discrimination of strain and temperature using Brillouin frequency shift and birefringence in a polarization-maintaining fiber. *Optics Express*, 17(3):1248–1255, 2009.

[86] K. M. Davis, K. Miura, N. Sugimoto, and K. Hirao. Writing waveguides in glass with a femtosecond laser. *Optics Letters*, 21(21):1729–1731, 1996.

[87] M. Ams, G. D. Marshall, P. Dekker, M. Dubov, V. K. Mezentsev, I. Bennion, and M. J. Withford. Investigation of Ultrafast Laser–Photonic Material Interactions: Challenges for Directly Written Glass Photonics. *IEEE Journal of Selected Topics in Quantum Electronics*, 14(5):1370–1381, 2008.

[88] S. M. Eaton, H. Zhang, P. R. Herman, F. Yoshino, L. Shah, J. Bovatsek, and A. Y. Arai. Heat accumulation effects in femtosecond laser-written waveguides with variable repetition rate. *Optics Express*, 13(12):4708–4716, 2005.

[89] M. Ams, G. Marshall, D. Spence, and M. Withford. Slit beam shaping method for femtosecond laser direct-write fabrication of symmetric waveguides in bulk glasses. *Optics Express*, 13(15):5676–5681, 2005.

[90] P. S. Salter, A. Jesacher, J. B. Spring, B. J. Metcalf, N. Thomas-Peter, R. D. Simmonds, N. K. Langford, I. A. Walmsley, and M. J. Booth. Adaptive slit beam shaping for direct laser written waveguides. *Optics Letters*, 37(4):470–2, 2012.

[91] Y. Barad, H. Eisenberg, M. Horowitz, and Y. Silberberg. Nonlinear scanning laser microscopy by third harmonic generation. *Applied Physics Letters*, 70(8):922–924, 1997.

[92] G. D. Marshall, A. Jesacher, A. Thayil, M. J. Withford, and M. Booth. Three-dimensional imaging of direct-written photonic structures. *Optics Letters*, 36(5):695–697, 2011.

[93] Y. Nasu, M. Kohtoku, and Y. Hibino. Low-loss waveguides written with a femtosecond laser for flexible interconnection in a planar light-wave circuit. *Optics Letters*, 30(7):723–725, 2005.

[94] S. M. Eaton, W. Chen, L. Zhang, H. Zhang, R. Iyer, J. S. Aitchison, and P. R. Herman. Telecom-Band Directional Coupler Written With Femtosecond Fiber Laser. *IEEE Photonics Technology Letters*, 18(20):2174–2176, 2006.

[95] M. Pospiech, M. Emons, B. Väckenstedt, G. Palmer, and U. Morgner. Single-sweep laser writing of 3D-waveguide devices. *Optics Express*, 18(7):6994–7001, 2010.

[96] J. Noda, K. Okamoto, and Y. Sasaki. Polarization-maintaining fibers and their applications. *Journal of Lightwave Technology*, 4(8):1071–1089, 1986.

[97] C. Söller, B. Brecht, P. J. Mosley, L. Y. Zang, A. Podlipensky, N. Y. Joly, P. St. J. Russell, and C. Silberhorn. Bridging visible and telecom wavelengths with a single-mode broadband photon pair source. *Physical Review A*, 81:31801, 2010.

[98] B. J. Smith, P. Mahou, O. Cohen, J. S. Lundeen, and I. A. Walmsley. Photon pair generation in birefringent optical fibers. *Optics Express*, 17(26):23589, 2009.

[99] M. B. Nasr, S. Carrasco, B. E. A. Saleh, A. V. Sergienko, M. C. Teich, J. P. Torres, L. Torner, D. S. Hum, and M. M. Fejer. Ultrabroadband Biphotons Generated via Chirped Quasi-Phase-Matched Optical Parametric Down-Conversion. *Physical Review Letters*, 100:183601, 2008.

[100] C. Simon, M. Afzelius, J. Appel, A. Boyer de la Giroday, S. J. Dewhurst, N. Gisin, C. Y. Hu, F. Jelezko, S. Krll, J. H. Mller, J. Nunn, E. S. Polzik, J. G. Rarity, H. De Riedmatten, W. Rosenfeld, A. J. Shields, N. Skld, R. M. Stevenson, R. Thew, I. A. Walmsley, M. C. Weber, H. Weinfurter, J. Wrachtrup, and R. J. Young. Quantum memories. *The European Physical Journal D*, 58(1):1–22, 2010.

[101] K. Garay-Palmett, Y. Jeronimo-Moreno, and A. B. URen. Theory of cavity-enhanced spontaneous four wave mixing. *Laser Physics*, 23:015201, 2013.

[102] G. D. Marshall, M. Ams, and M. J. Withford. Direct laser written waveguide-Bragg gratings in bulk fused silica. *Optics Letters*, 31(18):2690–2691, 2006.

[103] G. Lepert, M. Trupke, E. A. Hinds, H. Rogers, J. C. Gates, and P. G. R. Smith. Demonstration of UV-written waveguides , Bragg gratings and cavities at 780 nm, and an original experimental measurement of group delay. *Optics Express*, 19(25):24933–24943, 2011.

[104] R. H. Stolen, M. A. Bösch, and C. Lin. Phase matching in birefringent fibers. *Optics Letters*, 6(5):213–215, 1981.

[105] A. S. Sudbo. Film mode matching : a versatile numerical method for vector mode field calculations in dielectric waveguides. *Pure and Applied Optics*, 2:211–233, 1993.

[106] A. Jesacher, P. S. Salter, and M. J. Booth. Refractive index profiling of direct laser written waveguides: tomographic phase imaging. *Optical Materials Express*, 3(9):246–253, 2013.

[107] P. D. Dragic, C. G. Carlson, and A. Croteau. Characterization of defect luminescence in Yb doped silica fibers: part I NBOHC. *Optics Express*, 16(7):4688–4697, 2008.

[108] L. Vaccaro, M. Cannas, and R. Boscaino. Luminescence features of nonbridging oxygen hole centres in silica probed by site-selective excitation with tunable laser. *Solid State Communications*, 146(3-4):148–151, 2008.

[109] E. M. Dianov, D. S. Starodubov, and A. A. Frolov. UV argon laser induced luminescence changes in germanosilicate fibre preforms. *Electronics Letters*, 32(3):246–247, 1996.

[110] P. Aboussouan, O. Alibart, D. B. Ostrowsky, P. Baldi, and S. Tanzilli. High-visibility two-photon interference at a telecom wavelength using picosecond-regime separated sources. *Physical Review A*, 81(2), 2010.

[111] S. Krapick, H. Herrmann, V. Quiring, B. Brecht, H. Suche, and C. Silberhorn. An efficient integrated two-color source for heralded single photons. *New Journal of Physics*, 15(3):033010, 2013.

[112] S. Aaronson and A. Arkhipov. The Computational Complexity of Linear Optics. *arXiv*, 1011.3245v1, 2010.

[113] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2):230–265, 1936.

[114] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[115] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London*, 400:97–117, 1985.

[116] S. Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, 1996.

[117] P. W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[118] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Symposium on Theory of Computing (STOC)*, pages 212–219, 1996.

[119] A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum Algorithm for Linear Systems of Equations. *Physical Review Letters*, 103(15):150502, 2009.

[120] A. M. Childs and W. van Dam. Quantum algorithms for algebraic problems. *Reviews of Modern Physics*, 82(1):1–52, 2010.

[121] D. P. DiVincenzo and D. Loss. Quantum information is physical. *Superlattices and Microstructures*, 23(34):419 – 432, 1998.

[122] E. Knill and R. Laflamme. Power of One Bit of Quantum Information. *Physical Review Letters*, 81(25):5672–5675, 1998.

[123] S. P. Jordan. Permutational Quantum Computing. *Quantum Information and Computation*, 10:470, 2010.

[124] D. Shepherd and M. J. Bremner. Temporally unstructured quantum computation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2105):1413–1439, 2009.

[125] E. Caianiello. On quantum field theory-i: explicit solution of Dyson's equation in electrodynamics without use of Feynman graphs. *Il Nuovo Cimento*, 10(12):1634–1652, 1953.

[126] L. Troyansky and N. Tishby. Permanent uncertainty: on the quantum evaluation of the determinant and the permanent of a matrix. In *Proceedings of Physics and Computation (PhysComp 96)*.

[127] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.

[128] F. Marsili, V. B. Verma, J. A. Stern, S. Harrington, A. E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. D. Shaw, R. P. Mirin, and S. W. Nam. Detecting single infrared photons with 93 % system efficiency. *Nature Photonics*, 7(3):210–214, 2013.

[129] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001.

[130] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Mølmer. Stringent and efficient assessment of Boson-Sampling devices. *arXiv*, 1312.3080v1, 2013.

[131] S. Scheel. Permanents in linear optical networks. *arXiv:0406127v1*, 0406127v1, 2004.

[132] S. Aaronson. *Quantum Computing Since Democritus*. Cambridge University Press, 2013.

[133] A. R. Meyer and L. J. Stockmeyer. The equivalence problem for regular expressions with squaring requires exponential space. In *Proceedings 13th Annual Symposium on Switching and Automata Theory*, pages 125–129, 1972.

[134] S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991.

[135] M. Jerrum, A. Sinclair, and E. Vigoda. A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries. *Journal of the ACM*, 51(4):671–697, 2004.

[136] L. Stockmeyer. The Complexity of Approximate Counting. In *Proceedings of ACM STOC*, number 1, pages 118–126, 1983.

[137] S. Aaronson. The equivalence of sampling and searching. In *Proceedings of the 6th International Conference on Computer Science: Theory and Applications*, CSR'11, pages 1–14, Berlin, Heidelberg, 2011. Springer-Verlag.

[138] C. Lautemann. BPPand the Polynomial Hierarchy. *Information Processing Letters*, 17:215–217, 1983.

[139] M. Poźniak, K. Życzkowski, and M. Kuś. Composed ensembles of random unitary matrices. *Journal of Physics A: Mathematical and General*, 31:1059–1071, 1998.

[140] C. Spengler, M. Huber, and B. C. Hiesmayr. Composite parameterization and Haar measure for all unitary and special unitary groups. *Journal of Mathematical Physics*, 53(1):013501, 2012.

[141] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White. Photonic boson sampling in a tunable circuit. *Science*, 339(6121):794–798, 2013.

[142] A. Crespi, R. Osellame, R. Ramponi, D.J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino. Integrated multimode interferometers with arbitrary designs for photonic boson sampling. *Nature Photonics*, 7(7):545–549, 2013.

[143] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther. Experimental boson sampling. *Nature Photonics*, 7(7):540–544, 2013.

[144] Nicholas Thomas-Peter. *Quantum-Enhanced Precision Measurement and Information Processing with Integrated Photonics*. PhD thesis, University of Oxford, 2012.

[145] H. L. Rogers, S. Ambran, C. Holmes, P. G. R. Smith, and J. C. Gates. In situ loss measurement of direct UV-written waveguides using integrated bragg gratings. *Optics Letters*, 35(17):2849–2851, 2010.

[146] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White. Direct characterization of linear-optical networks. *Optics Express*, 21(11):13450–13458, 2013.

[147] A. Laing and J. L. O'Brien. Super-stable tomography of any linear optical device. *arXiv*, 1208.2868v1, 2012.

[148] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental realization of any discrete unitary operator. *Physical Review Letters*, 73:58–61, 1994.

[149] D. O. Kundys, J. C. Gates, S. Dasgupta, C. B. E. Gawith, and P. G. R. Smith. Use of Cross-Couplers to Decrease Size of UV Written Photonic Circuits. *IEEE Photonics Technology Letters*, 21(13):947–949, 2009.

[150] P. Rohde. Optical quantum computing with photons of arbitrarily low fidelity and purity. *Physical Review A*, 86(5):052321, 2012.

[151] A. Migdall, D. Branning, and S. Castelletto. Tailoring single-photon and multiphoton probabilities of a single-photon on-demand source. *Physical Review A*, 66(5):053805, 2002.

[152] J. H. Shapiro and F. N. C. Wong. On-demand single-photon generation using a modular array of parametric downconverters with electro-optic polarization controls. *Optics Letters*, 32(18):2698–2700, 2007.

[153] J. Limpert, F. Röser, T. Schreiber, and A. Tünnermann. High-Power Ultrafast Fiber Laser Systems. *IEEE Journal of Selected Topics in Quantum Electronics*, 12(2):233–244, 2006.

[154] J. W. Silverstone, D. Bonneau, K. Ohira, N. Suzuki, H. Yoshida, N. Iizuka, M. Ezaki, C. M. Natarajan, M. G. Tanner, R. H. Hadfield, V. Zwiller, G. D. Marshall, J. G. Rarity, J. L. O'Brien, and M. G. Thompson. On-chip quantum interference between silicon photon-pair sources. *Nature Photonics*, 8:104–108, 2014.

[155] J. P. Sprengers, A. Gaggero, D. Sahin, S. Jahanmirinejad, G. Frucci, F. Mattioli, R. Leoni, J. Beetz, M. Lermer, M. Kamp, S. Hofling, R. Sanjines, and A. Fiore. Waveguide superconducting single-photon detectors for integrated quantum photonic circuits. *Applied Physics Letters*, 99(18):181110, 2011.

[156] W. H. P. Pernice, C. Schuck, O. Minaeva, M. Li, G. N. Goltsman, A. V. Sergienko, and H. X. Tang. High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits. *Nature Communications*, 3:1325, 2012.

[157] S. Aaronson and A. Arkhipov. BosonSampling Is Far From Uniform. *arXiv*, 1309.7460v2, 2013.

[158] P. C. Humphreys, B. J. Metcalf, J. B. Spring, M. Moore, X. M. Jin, M. Barbieri, W. S. Kolthammer, and I. A. Walmsley. Linear Optical Quantum Computing in a Single Spatial Mode. *Physical Review Letters*, 111(15):150501, 2013.

[159] T. J. Bartley, G. Donati, X. M. Jin, A. Datta, M. Barbieri, and I. A. Walmsley. Direct Observation of Sub-Binomial Light. *Physical Review Letters*, 110(17):173602, 2013.

[160] T. J. Bartley, G. Donati, J. B. Spring, X. M. Jin, M. Barbieri, A. Datta, B. J. Smith, and I. A. Walmsley. Multiphoton state engineering by heralded interference between single photons and coherent states. *Physical Review A*, 86(4):043820, 2012.

[161] M. Cooper, C. Söller, and B. J. Smith. High-stability time-domain balanced homodyne detector for ultrafast optical pulse applications. *Journal of Modern Optics*, 60(8):611–616, 2013.

[162] M. Cooper, C. Söller, and B. J. Smith. Experimental generation of multi-photon Fock states. *Optics Express*, 21(5):5309–5317, 2013.

[163] M. Cooper, M. Karpinski, and B. J. Smith. Quantum state estimation with unknown measurements. *arXiv*, 1306.6431v2, 2013.

[164] H. Defienne, M. Barbieri, B. Chalopin, B. Chatel, I. A. Walmsley, B. J. Smith, and S. Gigan. Single-photon entanglement generation by wavefront shaping in a multiple-scattering medium. *arXiv*, 1401.3239v2, 2014.