OPEN ACCESS

University of BRISTOL

# This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

*Author:*
**Yard, Patrick W**

*Title:*
**Time as a resource in integrated quantum photonics**

**General rights**
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

**Take down policy**
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

# Time as a resource in integrated quantum photonics

Patrick Yard

Doctorate of Philosophy

University of Bristol

March, 2022

Supervisor: Dr Anthony Laing

## ABSTRACT

Photonics, as a platform for quantum information processing, has the unique ability to exploit the qubits temporal degree of freedom. For photonic quantum computing, this can be used to conserve physical resources by encoding photonic qubits directly in the temporal degree of freedom thereby allowing a quantum processor with a modest physical footprint to operate in a large Hilbert space. Additionally, through the conjugate relationship of time and frequency in quantum optics, the bandwidth of photons is related to the timing resolution of detectors, thereby allowing a route to increase the visibility of quantum interference between photons of differing frequency. Conversely, the role of time itself in quantum mechanics can be explored through the mapping of unitary dynamics to photonic quantum simulators where time is a programmable parameter. In this thesis, we explore each of these aspects of time, through a series of quantum photonics experiments including in integrated quantum photonic technologies.

Initially, we see how increasing the timing resolution of quantum interference experiments can lead to new interference effects. We use a silicon chip to perform both HOM interference and interference in a Fourier interferometer. In both cases, we show that decreasing the timing window of the experiment leads to an increase in the visibility of the interference fringes. Following this, we apply this technique to two photon scattershot boson sampling. To combat the inherent reduction in sample rate that accompanies temporal filtering, we perform boson sampling experiments by sampling directly from the photon's arrival time for both two (four detected) and three (six detected) photons. Next, we look at simulating, designing and testing a high-speed all-optical switch, based on inter-modal cross-phase modulation, in silicon nitride. Here, the signal and pump fields propagate in the fundamental and first order transverse electric modes, respectively. We show first devices with a phase shift of up to 0.3 $\pi$ and show how our simulations match the results closely, giving us confidence that we can reach $\pi$ phase shifts with future devices. Finally, we look at how integrated photonics can be used to simulate the dynamics of Hamiltonians that are non-Hermitian, but symmetric under combined parity and time reversal. Using a silica chip, capable

of implementing arbitrary unitary transformations on up to six modes, we simulate the non-unitary dynamics of these systems. Experimentally, we use a time-varying Hamiltonian, implemented by an adaptive feedback loop, to increase the coherence of a target qubit. In simulation, we more rigorously test the algorithm and compare to a fixed Hamiltonian.

These results provide insight into the benefits of exploiting the temporal degree of freedom in quantum photonic experiments and provide evidence for the feasibility and scalability of integrated quantum photonics.

## DECLARATION OF AUTHORSHIP

This dissertation is submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctorate of Philosophy in the Faculty of Science.

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

Date:

Word count: 45,891

# ACKNOWLEDGEMENTS

I think a key to finishing a PhD is accepting that it will never be perfect. A looming deadline with university admin waiting to remove you from the course, should you be even a day late, seems to help this process of acceptance along. This truth not withstanding, I am proud of this document, and even more of the work that I have put into it. This thesis, however, has relied not only on me writing it but also on many people inside QETLabs and out.

First and foremost I would like to thank Anthony, for not only giving me this opportunity but also for the guidance, and of course the many pub trips! I would like to thank Alex Jones, who has played a large role in this thesis, from introducing me to the interesting world of distinguishability, to sharing frustration in the lab and pints in the pub. To Nicola, Stefano and Massimo helping me take my first steps in integrated photonic experiments, any lab skills I have developed owe themselves to your teaching.

To my fellow office inhabitants, past and present, Reece, Owain, Ross, Brian, Nicola, Alex M., Tom, Ed thanks for many distracting but interesting conversations about all manner of things. Special thanks to Rachel and Konstantina - sharing the office, and the writing process, with you two has made writing this thesis much more bearable. At least we all went insane at the same rate.

To those who helped roll this thesis in glitter, along with those who helped prepare me for my viva, Alex J, Alex M, Lawrence, Dave, Jake, Frazer, Dom and Jacob, thanks a lot, both myself and this thesis were in much worse shape before your input.

The best scientific chats often don't happen in an office or at a whiteboard, they happen in the pub. To Brian, Jake, Alex, Dom, Joel, Frazer, Reece, Matt, thanks for all the pub trips and the chats within both scientific and not. Thanks to all of you for laughing at my lazy jokes and for keeping me on my toes by never letting a single slip-up slide. I would have laughed a lot less over these last few years had it not been for you lot.

QETLabs is filled with too many people to individually name, but I think those who need thanking will know who they are. Thanks for being a group of the smartest, most

insightful people I have ever met. Always happy to answer my stupid questions, whilst pretending they weren't stupid questions.

To the Bristol lot, Georgie, Tom, Josh, thanks for being around for park beers, cycling chats, and generally being good eggs. To Alex, Mike, Nat, Char and Pat, thanks for reminding me that I am a student and almost 30 and that no-one wants to hear about physics. In reality though, thanks for being a welcome distraction and making me laugh without fail. I could do a lot worse for friends than you lot.

To Ted, productivity in the office always went up when you were in here and this thesis would have taken a lot longer without you.

Dad, thanks for your unannounced, mysterious but always well-timed and appreciated parcels and spontaneous phone calls asking about some niche part of physics. They meant more than you thought. To mum, thanks for always being there with unwavering, constant, love and support. Your phone calls, visits and parcels always picked me up and none of this would have been possible without you. To Hannah, thanks for always keeping me grounded and reminding me that 'it's only computer science' and 'you can't *see* physics'. I have enjoyed our dinners and random lunches - writing would have much harder without them. Also, your work stories always make what I have to do seem less bad.

Finally, to Zoe, thanks for sharing this journey with me. Your patience has meant the world to me. Here's to everything becoming a little less stressful now (hopefully), let's go on holiday!

To
Granny

”The only person for whom the house was in anyway special was Athur Dent, and that was only because it happened to be the one he lived in.”

Douglas Adams, *Hitchhikers Guide to the Galaxy*

# CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF PUBLICATIONS

The work presented in this thesis will appear in the following places

## PAPERS

*Chapters* 3 *and* 4

1. *Time-resolved multi-photon interference in a silicon chip*, **P. Yard**, A. Jones, S. Paesani, A. Maïnos, J. Bulmer and A. Laing. *In Preparation* (2022);

*Chapter* 6

1. *Photonic quantum simulations of coupled $\mathcal{PT}$-symmetric Hamiltonians*, N. Maraviglia, **P. Yard**, R. Wakefield, J. Carolan, C. Sparrow, L. Chakhmakhchyan, C. Harold, T. Hashimoto, N. Matsuda, A. Harterm Y. Joglekar and A. Laing. *Physical Review Research*, 4:0130051, (2022);
2. *Coherence optimisation with PT-symmetric Hamiltonians*, **P. Yard**, N. Maraviglia, R. Wakefield, J. Carolan, C. Sparrow, L. Chakhmakhchyan, C. Harold, T. Hashimoto, N. Matsuda, A. Harterm Y. Joglekar and A. Laing. *In Preparation* (2022);

# 1 INTRODUCTION

"*The Hitchhikers Guide to the Galaxy*. The reason why it was published in the form of a micro sub-meson electronic component is that if it were printed in normal book form, an interstellar hitchhiker would require several inconveniently large buildings to carry it around in."

Douglas Adams, *Hitchhikers Guide to the Galaxy*

We are moving into a new age of information processing where the best tool for the job may not always be the biggest supercomputer you can find. Since the field's inception in 1982, there has been an open question for quantum computing to answer: can we build a quantum computer that is sufficiently large and complex, and with sufficiently low error rates to solve a problem or perform a calculation substantially faster than a classical computer? This question was answered first hesitantly [1–4], and then emphatically [5, 6], yes. While the problems these rudimentary quantum computers solved do not have much practical significance, the fact that they could be solved is significant in itself.

A key player in the race for quantum supremacy, as it was named by John Preskill [7], was a linear optics approach based on a variation of Aaronson and Arkiphov's boson sampling problem [8]. However, beyond these relatively simple proof-of-principle experiments, linear optics also holds the promise of being a potential platform for realising a universal quantum computer. Initial schemes, such as the original Knill, Laflamme, and Milburn (KLM) scheme [9], had such large overheads that they were rendered functionally impossible. However, modern schemes learned to embrace the probabilistic entangling gates afforded by linear optics, using ideas from percolation theory to show that even incomplete cluster states are universal for quantum computation, provided the gate probability is above the percolation threshold [10–12].

To date, the largest scale demonstrations of even non-universal photonic quantum computers have been performed in bulk optics. While this has turned out to be feasible for computations involving tens or even hundreds of photons [3, 5], it is likely unfeasible when scaled to the millions of photons necessary for even the smallest useful universal quantum computers [13]. Integrated photonics hopes to offer a solution to this scaling

problem by miniaturising the necessary components and combining them on chip scale devices. So far, either silicon or silicon nitride form the basis of the largest integrated quantum photonic circuits [14–18]. However, materials with a $\chi^{(2)}$ nonlinearity are gaining traction due to their ability to integrate fast electro optic modulators [19–27]. One feature of photonic qubits that makes them particularly appealing for measurement-based architectures, usually favoured by linear optical quantum computing (LOQC) proponents, is that, unlike superconducting and matter based systems, photons are travelling qubits, meaning it is possible to exploit their temporal degree of freedom. For static qubits, time only affects the number of gates that can be implemented on a quantum computer. For a photon, however, not only can qubit states be defined in the photon's arrival times [28] but even for photonic qubits encoded in other degrees of freedom, time is useful for fixing the experimental overhead of a quantum computer. By delaying photons and entangling them with future generations of qubits, only small a fraction of the total amount of qubits must be "alive" at a given time. For a platform where propagation loss is the dominant source of error, this is a huge advantage. Beyond this, cluster states that are both universal for measurement based quantum computing (MBQC) and amenable to error correction must be three-dimensional [29–34]. In the way mentioned above, therefore, time is a natural candidate for this third dimension. The photon's temporal degree of freedom also has more subtle effects. The arrival time, or rather the possible arrival times, of a photon are related to its spectral properties and can therefore be used to control the interference quality between photons [35–38].

Despite this, to date, many integrated photonic experiments ignore the importance of the temporal degree of freedom. Most experiments use large coincidence windows, as the decrease in count rates associated with temporal filtering is too high a price to pay for any gain in interference visibility. The lack of fast and low-loss switching in both silicon and silicon nitride platforms has forced users to proceed with experiments using slow thermo-optic phase shifters. One example of an (at least partially) integrated feed forward experiment was demonstrated in Ref. [39]. They perform multiplexing of photons generated on-chip however the delays and switching are all performed off-chip.

In this thesis, we augment the existing work on integrated photonic experiments by performing experiments where we explicitly take the temporal degree of freedom into account. We explore several different aspects, including performing quantum interference experiments in a time-resolved manner, designing and testing a high-speed modulator, and simulating Hamiltonians that are time and parity symmetric.

## THESIS OUTLINE

The remainder of this thesis is outlined as follows. In chapter two, we introduce the background topics required to understand the later results. We start by introducing the formalism of quantum optics. We then move on to describe the fields of linear optics, looking at unitary decompositions in both spatial and temporal degrees of freedom, and nonlinear optics, in both classical and quantum pictures. Following this, we describe the necessary components required for integrated photonic experiments. Finally, we include a brief introduction of linear optical quantum computing.

In chapter three, we look at the effect of performing two-photon interference experiments in a time-resolved manner, that is, using the exact arrival time of the photons rather than integrating over some interval. We perform both Hong-Ou-Mandel interference at a beam-splitter and interference in a complex Hadamard interferometer with photons that are distinguishable in frequency.

In chapter four, we expand the results of chapter three and look at time-resolved boson sampling. We perform two-photon scattershot and look at the effect of changing the timing window. We also perform two-photon scattershot and three-photon standard boson sampling where we sample directly from the photon's arrival time rather than filtering in time.

In chapter five, we simulate, design and test an all-optical Mach-Zehnder switch based on cross-phase modulation. Here, the pump and signal pulses travel in different transverse waveguide modes, allowing easy and low-loss multiplexing and demultiplexing.

In the final chapter, chapter six, we look at how linear optics can be used to simulate a specific class of non-Hermitian Hamiltonians, which are symmetric under combined time and parity reversal. We show how evolution under these Hamiltonians can be used to increase the coherence of an initial qubit state. We compare a fixed Hamiltonian to a time-varying Hamiltonian implemented via an adaptive feedback loop.

# 2

# BACKGROUND

"'I really wish I'd listened to what my mother told me when I was young.'
'Why, what did she tell you?'
'I don't know, I didn't listen.'"

Douglas Adams, *Hitchhikers Guide to the Galaxy*

## 2.1 INTRODUCTION AND MOTIVATION

The study of light has always been a central tenet of humanity's attempt to understand the natural world. Either to understand human vision, bringing us telescopes and microscopes, in order to communicate, or more recently, to process both classical and quantum information.

In this chapter, we introduce the features of electromagnetic radiation required to give the later results context. We start with the classical description of light through Maxwells equations before seeing how this changes in a quantum picture. We move on to describe linear optics from its building blocks to larger scale devices. Following this, we describe photon generation through nonlinear optics before introducing the tools required for integrated photonic experiments. Next, we will talk about the keystone of photonic quantum technologies - quantum interference. Finally, we will put all the pieces together and describe how a universal quantum computer can be constructed from linear optics.

## 2.2 FROM CLASSICAL TO QUANTUM OPTICS

To start our discussion on optics we begin with a description of how one can derive the quantum picture of light [1]. We start from the classical description of electromagnetism, through the famous equations of James Clerk Maxwell. In the absence of free charges these are

$$\nabla \cdot \boldsymbol{E} = 0, \tag{2.1}$$

1 For this discussion we draw heavily from Refs. [40, 41]

$$\nabla \cdot \boldsymbol{B} = 0, \tag{2.2}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{2.3}$$

$$\nabla \times \boldsymbol{B} = \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t}, \tag{2.4}$$

where $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric and magnetic fields respectively and $c$ is the speed of light. We can now define a vector potential, $\boldsymbol{A}$, which is related to the electric and magnetic fields by

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \qquad \boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t}. \tag{2.5}$$

In the presence of no charges, the above equations are gauge invariant. Formally, this means that any choice of the vector potential that can be written in the form $\boldsymbol{A} + \nabla\phi$, where $\phi$ is a function that is twice differentiable, is valid. It is standard in quantum optics to use the Coulomb gauge which imposes the condition

$$\nabla \cdot \boldsymbol{A} = 0. \tag{2.6}$$

To find the dynamics of $\boldsymbol{A}$, we can rewrite Eq. 2.4 in terms of the vector potential, $\boldsymbol{A}$. To do this we substitute the relations in Eq. 2.5 into Eq. 2.4. Using the standard vector calculus identity $\nabla \times \nabla \times \boldsymbol{A} = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$, and noticing that Eq. 2.6 results in the first term being 0, we arrive at

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = 0. \tag{2.7}$$

This is the wave equation for electromagnetic waves. If instead of considering free space, we consider only a finite volume of space, the continuum of solutions to the wave equation becomes discrete. Specifically we look at a volume of space with $V = L^3$. We can then write a general solution to Eq. 2.7 as

$$\boldsymbol{A}(\boldsymbol{r}, t) = \sum_{\boldsymbol{k}} \boldsymbol{u} A_k e^{i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega_k t)} + \boldsymbol{u}^* A_k^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega_k t)}, \tag{2.8}$$

where we have introduced the unit vector along the direction of polarisation $\boldsymbol{u}$. In general $\boldsymbol{A}(\boldsymbol{r},t)$ would also be a sum over both possible polarisation directions but for clarity we chose the vector potential to be polarised along one direction. As it stands, $\boldsymbol{A}(\boldsymbol{r},t)$ describes a superposition of waves with a $k$-vector given by the elements

$$k_i = \frac{2\pi m_i}{L}, \tag{2.9}$$

where $i \in x,y,z$, $m_i \in \mathbb{Z}^+$. Using the relationships in Eq. 2.5, we can derive expressions for the electric and magnetic fields

$$\boldsymbol{E}(\boldsymbol{r},t) = i\sum_{\boldsymbol{k}} \sqrt{\frac{\hbar\omega_k}{2\epsilon_0}} \left(\boldsymbol{u}a_k e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} - \boldsymbol{u}^* a_k^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)}\right), \tag{2.10}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = i\sum_{\boldsymbol{k}} \boldsymbol{k}\sqrt{\frac{\hbar}{2\omega_k\epsilon_0}} \left(\boldsymbol{u}a_k e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} - \boldsymbol{u}^* a_k^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)}\right), \tag{2.11}$$

where we have rewritten the amplitudes $A_k = \sqrt{\frac{\hbar}{2\omega_k\epsilon_0 L^3}}a_k$ in order to leave us with the dimensionless amplitudes $a_k$. The total energy of an electromagnetic field in a given volume is given by the Hamiltonian

$$H = \frac{1}{2}\int_V \left(\epsilon_0 \boldsymbol{E}^2(\boldsymbol{r},t) + \frac{1}{\mu_0}\boldsymbol{B}^2(\boldsymbol{r},t)\right) d\boldsymbol{r}. \tag{2.12}$$

Substituting in the expressions for $\boldsymbol{E}$ and $\boldsymbol{B}$ we arrive at the Hamiltonian

$$H = \frac{1}{2}\sum_{\boldsymbol{k}} \hbar\omega_k \left(a_k a_k^* + a_k^* a_k\right). \tag{2.13}$$

To quantise the electromagnetic field we replace the complex amplitudes $a_k$ and $a_k^*$ with the canonical mutually adjoint operators $\hat{a}_k$ and $\hat{a}_k^\dagger$. Due to the bosonic nature of photons, the fundamental particles of the electromagnetic field, we require these operators to fulfil the commutator relations

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij} \qquad \left[\hat{a}_i, \hat{a}_j\right] = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] = 0. \tag{2.14}$$

Applying this transformation to Eq. 2.13 we arrive at

$$H = \sum_{k} H_k = \sum_{k} \hbar\omega_k \left( \hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \right). \tag{2.15}$$

From this we can see that each of the $k$ modes is described by a quantum harmonic oscillator. We now look at the operators $\hat{a}_k$ and $\hat{a}_k^\dagger$ a bit more closely. Firstly, using the commutation relations above one can show

$$\left[H_k, \hat{a}_{k'}\right] = \delta_{kk'} \hbar\omega_{k'} \hat{a}_{k'} \qquad \left[H_k, \hat{a}_{k'}^\dagger\right] = \delta_{kk'} \hbar\omega_{k'} \hat{a}_{k'}^\dagger. \tag{2.16}$$

To see the effect of these operators we first assume $|n\rangle_k$ is an eigenstate of $H_k$ with eigenvalue $E_{k,n}$. We can then ask what happens when $\hat{a}_k$ and $\hat{a}_k^\dagger$ act on this state? Using the above commutation relations we can see

$$\hat{a}_k^\dagger |n\rangle_k = \frac{1}{\hbar\omega_k} \left[H_k, \hat{a}_k^\dagger\right] |n\rangle_k . \tag{2.17}$$

Expanding the commutator and rearranging we arrive at

$$H_k \hat{a}_k^\dagger |n\rangle_k = \left(E_{k,n} + \hbar\omega_k\right) \hat{a}_k^\dagger |n\rangle_k . \tag{2.18}$$

Here we see that $\hat{a}_k^\dagger |n\rangle_k$ is also an eigenvector of $H_k$ but with an eigenvalue that is larger by $+\hbar\omega_k$. We label this new eigenstate as $|n+1\rangle_k$. Similar analysis can show that $|n-1\rangle_k = \hat{a}_k |n\rangle_k$ has an eigenvalue of $E_{k,n} - \hbar\omega_k$. We define the ground state $|0\rangle_k$ such that $\hat{a}_k |0\rangle_k = 0$, this state has energy $E_{k,0} = \frac{1}{2}\hbar\omega_k$. So the operators $\hat{a}_k^\dagger$ and $\hat{a}_k$ have the effect of adding and removing an excitation from the mode they act on. For this reason they are called the creation and annihilation operators, respectively. We are now primed to discuss some interesting classes of photonic states.

### FOCK STATES

The first class of states are ones that we have already seen, the Fock states. The Fock states are the eigenstates of the Hamiltonian, $H$, and represent a definite amount of excitations in a given mode. For this reason they are also called number states. We define a state with $n$ excitations in the $k^{\text{th}}$ mode as $|n\rangle_k$. The effect of the creation

and annihilation operators, as expected, is to increase and decrease the number of excitations in the mode

$$\hat{a}_k^\dagger \left|n\right\rangle_k = \sqrt{n+1}\left|n+1\right\rangle_k \qquad \hat{a}_k \left|n\right\rangle_k = \sqrt{n}\left|n-1\right\rangle_k . \tag{2.19}$$

Therefore, an arbitrary number Fock state can be acquired by letting the creation operator act on the vacuum $n$ times. With correct normalisation, this is given by

$$\left|n\right\rangle_k = \frac{\left(\hat{a}_k^\dagger\right)^n}{\sqrt{n!}} \left|0\right\rangle_k . \tag{2.20}$$

The energy of a given Fock state is then given by

$$E_{k,n} = \left\langle n\right|_k H \left|n\right\rangle_k = \hbar\omega_k \left(n + \frac{1}{2}\right) \tag{2.21}$$

We can also define a number operator

$$\hat{N}_k = \hat{a}_k^\dagger \hat{a}_k \quad \rightarrow \quad \left\langle n\right|_k \hat{N}_k \left|n\right\rangle_k = n_k, \tag{2.22}$$

which counts the number of excitations in the mode. For the rest of this thesis we refer to these quantised excitations of the electromagnetic field as *photons*. Any photons in the same mode $k$ are indistinguishable, by this we mean there is no physical way to tell them apart. This is an important feature.

### COHERENT STATES

Throughout quantum mechanics there are pairs of non-commuting variables that cannot be simultaneously measured with arbitrary precision. The correlation in the uncertainty of simultaneous measurements is given by a corresponding uncertainty relation. Two such variables are phase and photon number. We can therefore define two classes of states - one with a well-defined photon number and another with a well-defined phase. We have already seen the former, named the Fock state. The latter is called a coherent state and here we will recap its key features. Coherent states are minimum uncertainty states, meaning the product of the uncertainties in their photon number and phase minimises the uncertainty relation. They are defined as eigenstates

of the annihilation operator $\hat{a}_k |\alpha\rangle_k = \alpha_k |\alpha\rangle_k$. They are generated by a unitary operator called the displacement operator

$$\hat{D}_k(\alpha) = \exp\left(\alpha \hat{a}_k^\dagger - \alpha^* \hat{a}_k\right), \tag{2.23}$$

which transforms the creation and annihilation operators as

$$\begin{aligned} \hat{a}_k^\dagger &\to \hat{D}_k^\dagger(\alpha)\, \hat{a}_k^\dagger \hat{D}_k(\alpha) = \hat{a}_k^\dagger + \alpha^*, \\ \hat{a} &\to \hat{D}_k^\dagger(\alpha)\, \hat{a}_k \hat{D}_k(\alpha) = \hat{a}_k + \alpha. \end{aligned} \tag{2.24}$$

We can also write these states in the Fock basis

$$|\alpha\rangle_k = e^{\frac{-|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle_k . \tag{2.25}$$

Using the number operator we see that the average number of photons $\langle\alpha|_k \hat{N}_k |\alpha\rangle_k = |\alpha|^2$. We can also see that the probability of measuring $n$ photons from a given coherent state follows a Poissonian distribution

$$P(n, \alpha) = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n} e^{-|\alpha|^2}}{n!}, \tag{2.26}$$

centred at $|\alpha|^2$. Coherent states represent an important class of quantum optical states. As they minimise the uncertainty relation, follow Poissonian counting statistics, and have a well-defined phase, they most closely mirror the properties of the classical electromagnetic fields.

## 2.3 LINEAR OPTICS

We have seen in the previous sections how we can describe photonic states in a particular and well-defined optical mode. In this section we will see how we can describe transformations on these optical modes. We will restrict ourselves, for the time being, to linear transformations, that is transformations that do not involve light-matter interactions. Any $m$ mode linear optical network can be described by a unique $m \times m$ complex unitary matrix, $\hat{M}$, where the condition of unitarity reflects that the transformation preserves photon number. Each element maps an input mode to an output mode with a

specific amplitude and phase. Given an array of $m$ input modes $|\boldsymbol{a}\rangle = \left(\hat{a}_1^\dagger, \hat{a}_2^\dagger, \ldots, \hat{a}_m^\dagger\right)^T$, the output modes are defined as

$$|\boldsymbol{b}\rangle = \hat{M}\,|\boldsymbol{a}\rangle\,, \tag{2.27}$$

where $|\boldsymbol{b}\rangle = \left(\hat{b}_1^\dagger, \hat{b}_2^\dagger, \ldots, \hat{b}_m^\dagger\right)^T$ is the corresponding array of output modes. Each element of $\boldsymbol{b}$ is therefore given by

$$\hat{b}_i^\dagger = \sum_{j=0}^{m} M_{ij}\hat{a}_j^\dagger. \tag{2.28}$$

Given a single photon input into mode $j$ the corresponding probability of the photon being detected in mode $i$ is then

$$P(i,j) = |\,\langle 0|\,\hat{b}_i^\dagger\hat{a}_j^\dagger\,|0\rangle\,|^2 = |M_{ij}|^2 \tag{2.29}$$

### 2.3.1 *Building blocks for linear optical networks*

Two key components of any linear optical are the beam-splitter, which shifts power from one mode to another and a phase shifter, which imparts a relative phase on one mode relative to all the others in the network. A general beam-splitter acts on two modes and can be described by the mode transformations

$$\begin{aligned} \hat{a}_i^\dagger &\to \sqrt{\eta}\hat{b}_i^\dagger + \sqrt{1-\eta}\hat{b}_j^\dagger, \\ \hat{a}_j^\dagger &\to \sqrt{1-\eta}\hat{b}_i^\dagger - \sqrt{\eta}\hat{b}_j^\dagger, \end{aligned} \tag{2.30}$$

where the parameter $\eta$ describes the power transmission of the beam-splitter. We can then define a beam-splitter where the transmission can be tuned by some parameter $\theta$, $\eta = \sin^2(\theta)$. The beam-splitter action in matrix form can be written as

$$\begin{pmatrix} \hat{b}_i^\dagger \\ \hat{b}_j^\dagger \end{pmatrix} = \begin{pmatrix} \sin(\theta) & \cos(\theta) \\ \cos(\theta) & -\sin(\theta) \end{pmatrix} \begin{pmatrix} \hat{a}_i^\dagger \\ \hat{a}_j^\dagger \end{pmatrix}. \tag{2.31}$$

We note that in the classical field picture a beam-splitter splits the field amplitude between both modes. When a single photon is incident on a beam-splitter it is put into a coherent superposition across the modes with amplitudes set by the $\sqrt{\eta}$. A

second noteworthy component in linear optics is the phase shifter. This is a single mode component that can be described by the mode transformation

$$\hat{a}_i^\dagger \rightarrow e^{i\phi}\hat{b}_i^\dagger. \tag{2.32}$$

Again we can write this in a matrix format

$$\begin{pmatrix} \hat{b}_i^\dagger \\ \hat{b}_j^\dagger \end{pmatrix} = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_i^\dagger \\ \hat{a}_j^\dagger \end{pmatrix}. \tag{2.33}$$

This gives an insight for how we can write these single and two mode components as part of a larger network. For an $m$ mode photonic network these two mode matrices can be written as block diagonal matrix

$$\hat{M} = \begin{pmatrix} \mathbb{I}_{i-1} & \dots & 0 \\ \vdots & T_{\text{ij}} & \vdots \\ 0 & \dots & \mathbb{I}_{m-j} \end{pmatrix}, \tag{2.34}$$

where $T_{ij}$ is the $n \times n$ component matrix, acting on adjacent modes $i$ to $j$ and $\mathbb{I}_k$ is a $k \times k$ identity matrix. With this, larger photonic networks can be built up by matrix multiplication.

### 2.3.2 *Arbitrary unitary transformations*

As we have seen that a general linear optical network is described a unitary transformation. We can, however, ask is every unitary transformation represented by some linear optical circuit? It has been known since Adolf Hurwitz in 1897 that an arbitrary unitary, $U(N)$ can be decomposed into $U(2)$ building blocks [42]. However, it was Reck et al. that first proved this in the case of linear optics [43] [2]. The building block of both these decompositions is a two parameter tuneable component, this can be written as

$$T(\phi, \theta) = \begin{pmatrix} e^{i\phi}\sin(\theta) & \cos(\theta) \\ e^{i\phi}\cos(\theta) & -\sin(\theta) \end{pmatrix}. \tag{2.35}$$

2 The same scheme had already been discovered in the microwave regime [44]

Each element of $T(\phi, \theta)$ is a unitary matrix in $U(2)$ although not all elements of $U(2)$ can be reached with only these two parameters. We can see that, not by accident, this is a combination of the two components we described in the previous section. Specifically, this is achieved by a phase shifter on one mode, followed by a beam-splitter. The aim of a decomposition, such as the one found by Reck, is to find some specific pattern of $T_{ij}$ matrices, that is blocks as in Eq. 2.35 acting on modes $i$ and $j$, which along with a diagonal matrix of phases $D$ gives the target unitary transformation. Formally, this can be expressed as

$$U(N) = D \prod_{i,j \in S} T_{ij}, \tag{2.36}$$

where $S$ describe the specific pattern of beam-splitters required. To date, there exist two main decompositions. Both discover the pattern $S$ by iteratively applying $T_{ij}$ either to the left or the right of $U(N)$ and then selecting the $\theta$ and $\phi$ such that an element of $U(N)$ is set to zero - from Reck this is called nulling the element. Specifically, multiplying from the left we can null an element in column $i$ or $j$ and multiplying from the left we can null an element in row $i$ or $j$. This process can be repeated until all but the diagonal is zero. Inverting this then gives the decomposition in Eq. 2.36. Initially, we describe the decomposition introduced by Reck. For this we only ever null from rows, that is multiplying from the left of $U(N)$. We proceed to null elements from the bottom row of $U(N)$. Due to unitarity, this also nulls from the last column of $U(N)$. We then repeat this step $N-1$ times until all off-diagonal elements are nulled. At each step the resulting unitary becomes

$$U_k(N) = U_{k-1}(N) \underbrace{T_{k,k-1} T_{k,k-2} \cdots T_{k,1}}_{\boldsymbol{T}_k}, \tag{2.37}$$

where swaps can be used to implement beam-splitters between non-adjacent modes. Repeating this until all the elements are nulled we end up with

$$D = U(N) \boldsymbol{T}_k \boldsymbol{T}_{k-1} \boldsymbol{T}_{k-2} \cdots \boldsymbol{T}_2, \tag{2.38}$$

which, when inverted, leaves us with

$$U(N) = D \left( \boldsymbol{T}_k \boldsymbol{T}_{k-1} \boldsymbol{T}_{k-2} \cdots \boldsymbol{T}_2 \right)^{-1}. \tag{2.39}$$

We note that the swaps used in Eq. 2.37 can be absorbed into the reflectivities of other beam-splitters. Due to this, when rearranged the transfer matrix $\boldsymbol{T}_k\boldsymbol{T}_{k-1}\boldsymbol{T}_{k-2}\cdots\boldsymbol{T}_1$ results in a triangular mesh of beam-splitters and phase shifters. This is an interesting result - any linear optical circuit can be decomposed into just two components. The main fault with this decomposition is that the optical depth, defined as the number of components in a given path, is different for photons that traverse different paths, resulting in an imbalance of loss. This has ramifications for, for example, boson sampling where it means that post-selection is no longer a viable route for retaining the complexity [45, 46]. This problem was solved in 2016 when Clements et al. proposed decomposing a unitary into a *square* beam-splitter mesh [47]. The main difference in this decomposition is that rather than nulling elements row by row, one nulls elements along diagonals. This is achieved by sequentially multiplying the unitary by $T_{ij}$ on the left and the right. This results in a mesh that builds up from the edges into the middle.

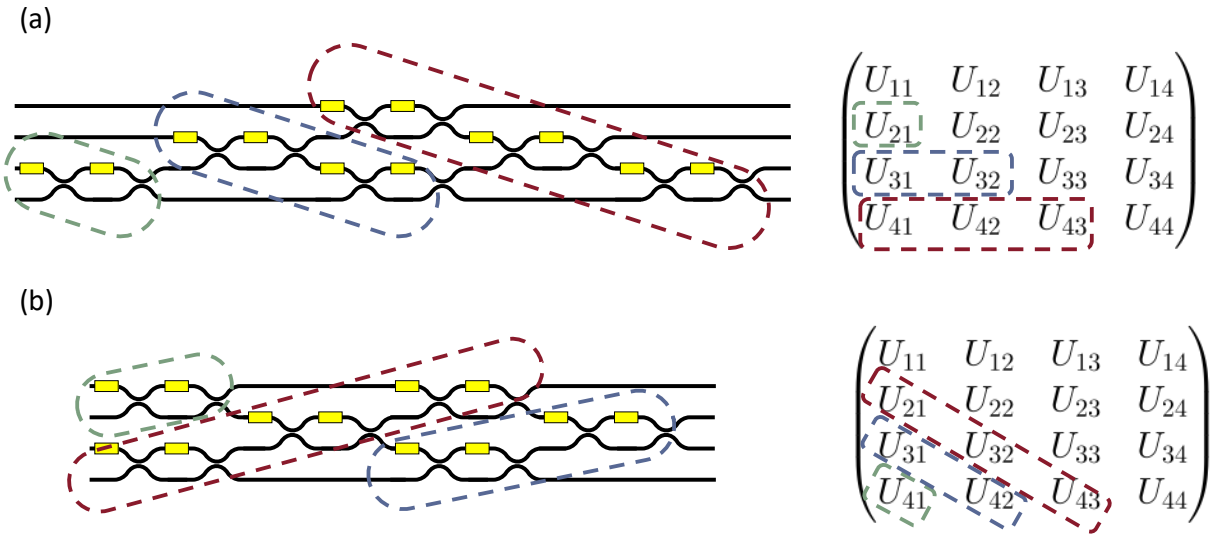


Figure 2.1: Illustrating the unitary decompositions introduced by (a) Reck et al. and (b) Clements et al. Coloured boxes indicate which beam-splitters are used to null which elements of the matrix. End phases have been omitted from the diagram.

It can be seen that, in this decomposition, each mode interferes with its neighbour at the earliest opportunity. This means that the Clements decomposition has the shortest possible optical depth and the optical depth is, on average, the same for all input/output combinations [3]. Both the Reck and Clements decompositions use the minimal number

3 There are some combinations such as $\hat{a}_1^\dagger \rightarrow \hat{b}_1^\dagger$ that have different (in this case shorter) optical depths but the number of these gets proportionally smaller with increasing $N$

of degrees of freedom $\mathcal{O}\left(n^2\right)$. In Fig. 2.1 we illustrate both decompositions and the final design of each. While it may seem that the Clement's decomposition should now always be adopted there are still certain situations where the Reck decomposition performs better. For example, if one has a large reprogrammable unitary, acting on $m$ modes, but would like to perform a $n \times n$ transformation where $n < m$, the optical depth for the Clements unitary will be $\mathcal{O}\left(m\right)$ independent of $n$. Using a Reck style interferometer, and using the top modes, the optical depth is reduced.

### 2.3.3 *Time domain interferometers*

In Fig. 2.1 we represented the unitary decomposition acting on spatial modes. This is how many large scale implementations of reprogrammable unitaries have been made [18, 48]. The primary reason for this is spatial modes lend themselves easily to static components and slow phase shifters. However, as we look to scale up, there are some inherent draw backs to this method. Looking to the future, scaling to large numbers of modes requires thousands of indistinguishable photon sources and expensive single photon detectors. Also, each of the approximately $N^2$ phase shifters for each unitary needs to be regularly calibrated, and can lead to crosstalk inside the unitary, lowering the fidelity. One of the most attractive feature of quantum photonics is that photons have a number of well-defined degrees of freedom and each can be used to encode information. Some, such as polarisation, are limited in the number of modes that we have access to. However, others, such as frequency, transverse mode, or temporal modes do not suffer from this and have already seen experimental implementations. To date, these implementations have either been on small scales [49, 50] or with restricted reprogrammability [51, 52].

Of these, using temporal modes is attractive for a number of reasons. They are continuous, which means that modes can be defined over different lengths of time, depending on the specific experimental setup. It restricts everything to a single spatial mode, and allows components to be reused, reducing the experimental overhead. In this section we will describe how the two decompositions above can be implemented in the time domain. The first time domain interferometer was described in Ref. [53] where they showed that a single photon source, point of interference and detector is sufficient for any sized unitary transformation. Specifically, the Motes scheme used two loops, one fast switch and two slower switches to implement a Reck style decomposition. This

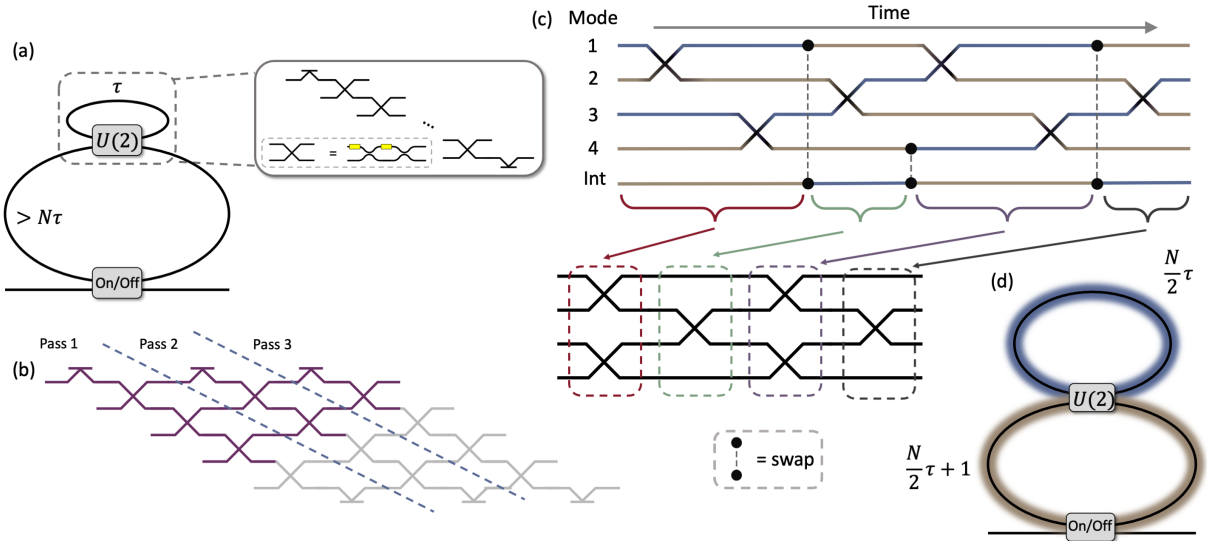


Figure 2.2: Time domain implementations of the Reck and Clements decompositions. (a) Loop based Reck decomposition. Inset shows the spatial mode breakdown of one implementation of the inner loop on a train of time bins. (b) Spatial mode equivalent of (a). If purple beam-splitters are used and the grey beam-splitters are set to identity then this can implement a Reck style decomposition. We also indicate which beam-splitters come from which pass of the interference point. (c) How the Clements style loop system can be thought of in spatial modes. Time is in the horizontal direction. Wires represent the modes and Int represents the interstitial mode. The wire colour represents which loop that mode is in at each time step. Blue is the top loop and brown is the bottom loop. We then show which beam-splitters map to which in the Clements scheme. Dotted lines represent swapping a given mode with the interstitial bin. (d) Clements scheme time loop. Top loop is $N/2$ time bins long and bottom is $N/2+1$.

experimental setup can be simplified slightly, however, by reusing one of the slower switches [54]. This revised scheme is shown in Fig. 2.2(a).

This scheme starts with of a train of time bins, separated by a time $\tau$. An on/off switch capable of switching in a time less than the total length of the train, is used to couple the train into the system, closing after the last bin. The train then sees a reconfigurable $U(2)$ element, with one of the outputs connected to the other input. If this delay is equal to the spacing between the time bins, sequential modes can be interfered. We require that this interference point can be reconfigured on a timescale $\ll \tau$. This allows us to separately tune the interference between pairs of bins and also restricts us to a finite amount of bins, by switching the first mode into and the last mode out of the loop, deterministically. A single pass of this interference point

performs one diagonal of beam-splitters in the Reck scheme. In order to build up a full Reck scheme, a second loop, with a length longer than the entire train of bins, can be used to have the pulse train pass through the interference point an arbitrary amount of times. Figure. 2.2(b) shows the beam-splitter mesh that is built up. From this we can see that $N-1$ passes of the interference point gives sufficient interference to implement a Reck style decomposition. This also demonstrates one of the flaws with this scheme. With $N-1$ outer loops we have twice as many points of interference than we require, increasing the loss of such a device. A second flaw is one mentioned before - the asymmetric optical depth for different input/output combinations. With time domain schemes, this asymmetric path causes photons to spend different amounts of in delay lines. This results in the added effect of photons that experience different levels of dispersion, reducing interference visibility.

Both of these problems were solved in Ref. [54] where they showed that the Clements scheme also has a time domain analogy. Instead of a two very different sized loops, we now have two loops that differ by one time bin. The top loop contains $N/2$ time bins and the bottom has $N/2+1$. The extra interstitial time bin in the bottom is used to change which time bins interfere. As we are free to label the modes as we please, we choose to label the first $N/2$ modes as the odd modes and the final $N/2$ as the even modes. The odd modes are switched into the top loop. As this is left to evolve, it interferes each odd mode, $k$, with the $k+1$ mode. After this first pass, the first mode reaches the beam-splitter at the same time as the interstitial bin, we switch the interstitial bin into the top loop and mode one into the bottom loop. From this point, each time a time bin traverses the loop we switch which loop it is in by adding $\pi$ to the desired beam-splitter phase, letting each time bin map out a figure of 8 path. This alternates interfering each odd mode with its highest neighbour, and shifting the trains with respect to each other so that we now interfere each even mode with its highest neighbour. Figure. 2.2(c) we show how this pattern is mapped out for four modes. The colour of the wire represents which loop that mode is in (colour coded to Figure. 2.2(d)). We also show how this maps to columns of beam-splitters in the Clements scheme.

## 2.4 NONLINEAR OPTICS

In the previous section we looked at *linear* optical transformations. These are transformations that do not mix the creation and annihilation operators and therefore preserve

photon number. In this section we will review *nonlinear* optical process. We will see first how these behave classically and then how we can use these processes to generate another class of quantum optical states. In our derivation of the wave equation in section. 2.2, we assumed a 'free' field, not propagating in a material. When light propagates through a material, the wave equation becomes

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = \mu_0\frac{\partial^2\mathbf{P}}{\partial t^2}, \tag{2.40}$$

where we have introduced the material polarisation, $\boldsymbol{P}$, which describes the response of dipoles in the material to the incoming light field. We have also restricted ourselves to only looking at the $\boldsymbol{E}$ field, this is reasonable as for all the materials in this thesis the magnetic susceptibility is $\approx 0$ and therefore it is only the interaction between the electric field and the material that is of importance. With low field strengths, the polarisation is proportional to the strength of the field $\boldsymbol{P} \simeq \epsilon_0\chi(\omega)\boldsymbol{E}$. Substituting this into Eq. 2.40 and assuming a monochromatic wave, $\boldsymbol{E} = E(\boldsymbol{r}, t)\, e^{-i\omega t}$, we arrive at

$$\nabla^2\mathbf{E} + n^2(\omega)\frac{\omega^2}{c^2}\mathbf{E} = 0, \tag{2.41}$$

where we have introduced $n(\omega) = \sqrt{1+\chi(\omega)}$, called the refractive index of the material and has the effect of slowing the speed of the propagating wave to a velocity of $c/n(\omega)$. This time-independent version of the wave equation is known as the Helmholtz equation. To understand what is happening here, it is useful to think of the potentials which the electrons in the materials experience. In this linear regime, the restoring force seen by the electrons is linear in the electron position and the potential is given by

$$V(x) = -\int F(x)dx. \tag{2.42}$$

This results in a parabolic potential and oscillations of the dipoles at the single frequency, identical to that of the incoming field. While this description is perfectly good for low field strengths, it breaks down if the field strength is sufficiently large. In general the material polarisation is given by

$$\boldsymbol{P} = \epsilon_0\left(\chi^{(1)}\boldsymbol{E} + \chi^{(2)}\boldsymbol{EE} + \chi^{(3)}\boldsymbol{EEE}\ldots\right), \tag{2.43}$$

where $\chi^{(1)}$ now describes the linear dynamics, as we have seen before. The higher order terms add terms to the restoring force seen by dipoles and therefore add features to the potentials. This results in oscillations that are no longer at a single frequency, meaning the energy radiated out by the dipoles contains new frequencies. In general, the strength of $\chi^{(n)}$ decreases as $n$ increases and in most circumstances considering only up to $\chi^{(3)}$ is sufficient. We will now discuss the first two nonlinear terms in a more detail.

### 2.4.1 $\chi^{(2)}$ *processes*

The first non-linear correction to the material polarisation only appears in materials with a specific crystal structure. Specifically, it requires that the crystal is non-isotropic and lacks inversion symmetry. Flipping the crystal direction in an isotropic material should also flip the polarisation direction $\boldsymbol{P}(-\boldsymbol{r}) = -\boldsymbol{P}(\boldsymbol{r})$. However, we see that as the $\chi^{(2)}$ term depends on $\boldsymbol{EE}$ then $\boldsymbol{P}(-\boldsymbol{r}) = \boldsymbol{P}(\boldsymbol{r})$. In this case, the dipole restoring force gains an extra $x^2$ term, resulting in a cubic potential. The asymmetry in this potential reflects the crystal asymmetry. If we consider a bi-chromatic input field $\boldsymbol{E} = E_1 e^{-\omega_1 t} + E_2 e^{-\omega_2 t} + c.c.$, the nonlinear term in the material polarisation becomes

$$\begin{aligned}\boldsymbol{P}_{\mathrm{NL}} = \epsilon_0 \chi^{(2)} \boldsymbol{EE} &= \epsilon_0 \chi^{(2)} \left(E_1 e^{-\omega_1 t} + E_2 e^{-\omega_2 t} + c.c.\right)\left(E_1 e^{-\omega_1 t} + E_2 e^{-\omega_2 t} + c.c\right) \\ &= \epsilon_0 \chi^{(2)} \Bigg( \underbrace{|E_1|^2 + |E_2|^2}_{\text{optical rectification}} + \underbrace{E_1^2 e^{-i2\omega_1 t} + E_2^2 e^{-i2\omega_2 t}}_{\text{second harmonic generation}} \\ &\quad + \underbrace{2E_1 E_2 e^{-i(\omega_1+\omega_2)t}}_{\text{sum frequency generation}} + \underbrace{2E_1 E_2^* e^{-i(\omega_1-\omega_2)t}}_{\text{difference frequency generation}} + c.c. \Bigg).\end{aligned} \tag{2.44}$$

Here, we can see several new frequencies that are generated. Second harmonic generations (SHG) describe the collection of processes that generated new frequencies at twice the original, sum and difference frequency generation are then used to describe processes that generate new frequencies at the sum and difference of the original two, respectively. Interestingly, we also see terms at 0 frequency. These describe the process whereby a DC electric field is generated across the crystal, known as optical rectification. Note here, we have assumed that $\chi^{(2)}$ is a scalar quantity, in general $\chi^{(2)}$ will be a rank 3 tensor with elements $\chi^{(2)}_{ijk}$ where the indices denote the polarisation of the two

mixing fields and the material polarisation. This demonstrates that there are a variety of different frequency generating processes that can in principle occur. However, do they always occur? The answer comes down to a principle known as phase-matching. This is a statement that says the oscillating dipole needs to remain in phase with the propagating fields in order to ensure that the new frequencies constructively interfere. This can be simplified to ensuring conservation of the wave-vectors of the mixing fields, for SHG this means that $2k(\omega_i) = k(2\omega_i)$, for sum and difference frequency generation it is given by $k(\omega_1 \pm \omega_2) = k(\omega_1) \pm k(\omega_1)$. This restricts the frequencies that can be supported and allows tuning of which process are suppressed.

### 2.4.2 $\chi^{(3)}$ *processes*

As mentioned in the previous section, $\chi^{(2)}$ process have strict requirements on the material crystal structures. In materials that do not meet these requirements, the lowest order nonlinear process are $\chi^{(3)}$. These lead to an anti-symmetric polarisation function, $\boldsymbol{P}(-\boldsymbol{r}) = -\boldsymbol{P}(\boldsymbol{r})$, which is true of all materials. As expected, the restoring force for the dipole now becomes cubic, resulting in a symmetric quartic potential. If we now consider a tri-chromatic field $\boldsymbol{E} = E_1 e^{-\omega_1 t} + E_2 e^{-\omega_2 t} + E_3 e^{-\omega_3 t} + c.c.$ The nonlinear polarisation for these materials is can then be written as

$$\boldsymbol{P}_{\mathrm{NL}} = \epsilon_0 \chi^{(3)} \sum_{i,j,k} E_i E_j E_k e^{i(\omega_i+\omega_j+\omega_k)t} + 3E_i E_j^* E_k e^{i(\omega_i-\omega_j+\omega_k)t} + c.c., \tag{2.45}$$

where the indices $i, j, k \in \{1, 2, 3\}$. We have separated terms that oscillate at frequencies $\omega_i + \omega_j + \omega_k$, known as sum frequency generation and terms oscillating at $\omega_i - \omega_j + \omega_k$, typically referred to as four wave mixing. Again, the phase matching conditions restrict the allowed processes. For sum frequency generation this becomes $k(\omega_i + \omega_j + \omega_k) = k(\omega_i) + k(\omega_j) + k(\omega_k)$ and for four wave mixing it becomes $k(\omega_i - \omega_j + \omega_k) = k(\omega_i) - k(\omega_j) + k(\omega_k)$. For the rest of this work we will only be concerned with the four wave mixing terms. Interestingly, unlike $\chi^{(2)}$ processes, for the terms where $\omega_i = \omega_j$ we see that the original frequency can still be retrieved. This will be covered in more detail in chapter 5.

### 2.4.3 *Photon generation through nonlinear optics*

Both of the processes defined above describe how $n$ fields propagating in a material can generate a field oscillating at a new frequency. This can be thought of as $n-1$ pump fields and then the one remaining field acting to seed the new field through stimulated emission. The classical description above says that if the seed field is not there then there cannot be any new frequencies. This is not so in the quantum picture, where vacuum fluctuations, which are always present, can be used to seed the process allowing single photons at multiple different frequencies to be generated. The processes of interest are, for $\chi^{(2)}$, spontaneous parametric down conversion (SPDC) which is difference frequency generation without the seed where one pump photon generates two daughter photons, named signal and idler, by convention. These newly generated photons must conserve energy, $\omega_p = \omega_s + \omega_i$, along with the phase-matching condition described above. For $\chi^{(3)}$ we use the process spontaneous four wave mixing (SFWM). Here, two pump photons are absorbed and signal and idler photons are generated that fulfil $\omega_{p1} + \omega_{p2} = \omega_s + \omega_i$. Generally, two regimes of these processes are defined - that where $\omega_s = \omega_i$, known as the degenerate case and that where $\omega_s \neq \omega_i$, known as non-degenerate. In the following work we employ degenerate SPDC and non-degenerate SFWM. While both interactions are different in nature, the end result is surprisingly similar. This should already be evident - both involve the generation of a pair of single photons via a bright pump field. The brightness of the pump lets us make a semi-classical approximation where the created photons are treated in a quantum picture, with their respective creation and annihilation operators, but the pump is considered to be a classical field with spectral amplitude $\alpha(\omega)$. It is then possible to define a general nonlinear interaction Hamiltonian[4]

$$H_{\text{int}} = -\gamma LP \int\int F(\omega_s, \omega_i)\hat{a}^\dagger(\omega_s)\,\hat{a}^\dagger(\omega_i) + h.c\, d\omega_s d\omega_i, \tag{2.46}$$

where $\gamma$ is proportional to the relevant nonlinear susceptibility tensor element, $P = |\int \alpha(\omega_p) d\omega_p|^2$ describes the total pump power, and $F(\omega_s, \omega_i)$ is the *joint spectral amplitude* (JSA) which includes both the energy conservation and phase-matching conditions described above. For SPDC the JSA is given by $F(\omega_s, \omega_i) = \alpha(\omega_p)\phi(\omega_p, \omega_s, \omega_i)$, where the pump wavelength is set by the energy conservation $\omega_p = \omega_s + \omega_i$. For SFWM the

4 For a more in depth derivation the reader is directed to Refs [55–57] for SPDC and Ref [58] for SFWM

situation is slightly different, as we have two pump photons only one can be fixed by the energy conservation, and so we integrate over the free pump photon

$$F(\omega_s, \omega_i) = \frac{1}{P} \int \alpha(\omega_p)\alpha(\omega_s + \omega_i - \omega_p)\phi(\omega_p, \omega_s, \omega_i)d\omega_p. \tag{2.47}$$

In both cases we have introduced the phase matching condition, $\phi$ given by

$$\phi(\omega_p, \omega_s, \omega_i) = \text{sinc}\left(\frac{\Delta kL}{2}\right) \exp\left(\frac{i\Delta kL}{2}\right), \tag{2.48}$$

where the phase mismatch for both processes are

$$\begin{aligned} \Delta k_{\text{SPDC}} &= k_p(\omega_s + \omega_i) - k_s(\omega_s) - k_i(\omega_i), \\ \Delta k_{\text{SFWM}} &= k_{p1}(\omega_p) + k_{p2}(\omega_s + \omega_i - \omega_p) - k_s(\omega_s) - k_i(\omega_i). \end{aligned} \tag{2.49}$$

In the limit of low pumping power the state generated by this Hamiltonian acting on the vacuum is given by

$$|\psi\rangle \approx |00\rangle + \int\int F(\omega_s, \omega_i)\hat{a}^\dagger(\omega_s)\,\hat{a}^\dagger(\omega_i)\,|00\rangle\; d\omega_s d\omega_i. \tag{2.50}$$

For many applications of single photons, we require that the emitted state is separable in frequency. This is as upon heralding on one of the generated photons, any spectral entanglement results in mixture of the single photon state. The degree of entanglement in the generated state is related to the separability of the joint spectrum. We can write the JSA in its Schmidt decomposition

$$F(\omega_s, \omega_i) = \sum_k \lambda_k \mu_k(\omega_s)\nu_k(\omega_s). \tag{2.51}$$

From this we can see that if only $\lambda_0 \neq 0$ then $F(\omega_s, \omega_i)$ is separable and both signal and idler photons are in pure states after heralding of the other. The purity of the photon can be calculated as

$$\mathcal{P} = \sum_k \lambda_k^4. \tag{2.52}$$

The state in Eq.2.50 only approximates a bi-photon state in the regime of low pump power, where only the first terms of the Taylor expansion dominate. If the pump power

is higher, this Hamiltonian generates further photons into the same modes, always in pairs. This state is called a *squeezed state*. In general, a squeezed state is given by

$$|\xi\rangle = \hat{S}(\xi)\,|0\rangle = e^{\frac{1}{2}(\xi \hat{a}_i^\dagger \hat{a}_j^\dagger + \xi^* \hat{a}_i \hat{a}_j)}\,|0\rangle\,. \tag{2.53}$$

If $i = j$ then this describes single mode squeezing, which we can map to degenerate processes where signal and idler are at the same frequency. If $i \neq j$, which we map to non-degenerate processes, is referred to as two-mode squeezing. For the nonlinear optical processes squeezing parameter, $|\xi| = \gamma LP$. Writing the state $|\xi\rangle$ in the Fock basis we see

$$|\xi\rangle = \mathrm{sech}(|\xi|) \sum_k e^{ik\theta} \tanh^k(|\xi|)\,|n\rangle_i\,|n\rangle_j\,, \tag{2.54}$$

where the angle $\theta$ is related to the phase of the pump field in case where there is a single pump field (or the pump fields are degenerate) and the difference between the pump phases if the pump fields are non-degenerate.

## 2.5 THE INTEGRATED PHOTONIC TOOLKIT

The later chapters in this thesis describe results from experiments performed in the integrated photonics platform, this involves light being routed and mixed in on-chip components. In this section we will discuss the main chip components that are required to perform these experiments.

### 2.5.1 *On-chip components*

#### WAVEGUIDES

Up to now, we have discussed free fields or fields interacting with bulk materials with no constraints placed on the transverse directions. Integrated photonic experiments require that we can confine and guide light across a chip. Creatively, we call these confining structures *waveguides*. Waveguides are created when a guiding material, with refractive index $n_{\text{core}}$ is sandwiched, in at least one of its transverse directions, between a cladding material with refractive index $n_{\text{clad}}$. In the later chapters we exclusively use strip waveguides, these are structures that are confined in both $x$ and $y$ directions. As

we are confining the light, we should expect a discrete set of supported modes. If we assume a set of solutions to the wave equation of the form $\boldsymbol{E} = \sum_m F_m(x,y)e^{i(\beta_m z - \omega t)}$, the supported modes are given by solutions to

$$\nabla^2 F_m(x,y) - \left(\beta_m^2 - \frac{\omega^2}{c^2}n^2(x,y)\right) F_m(x,y) = 0, \tag{2.55}$$

where each supported mode has a spatial distribution given by $F_m(x,y)$ and a propagation constant $\beta_m$. This propagation constant can be written as $\beta_m = \frac{\omega}{c}n_{\text{eff}}$ where $n_{\text{eff}}$ is the effective refractive index which determines the phase velocity of the mode. Many different waveguide geometries exist and have uses throughout integrated photonics, lower loss rib waveguides [59], photonic crystal waveguides [60, 61] and suspended waveguides [62] to name but a few. In the rest of this work we exclusively use strip waveguides. Strip waveguides are rectangular waveguides where the height is usually

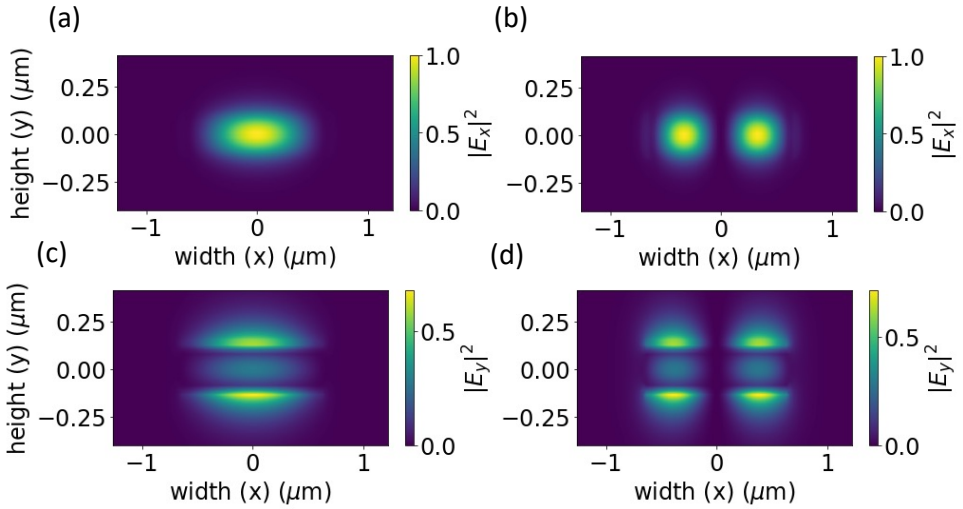


Figure 2.3: Example transverse mode plots for a silicon waveguide with a width of 1.3 µm and a height of 0.22 µm. We show the absolute value squared of the dominant electric field component. (a) TE0 (b) TE1 (c) TM0 (d) TM1.

set by the thickness of the guiding layer offered by the foundry. In these waveguides we can define two classes of propagating modes - those which have the majority of the field in the $E_y$ direction, named transverse electric (TE) and those with the majority of the field in the $E_x$ direction, named transverse magnetic (TM). Note that these modes only approximate transverse modes and will have some field along the propagation direction. A full understanding of how these modes propagate in the waveguide,

requires knowledge of the frequency dependence of $\beta_m$. In general, this cannot be found analytically, however much can be gained from Taylor expanding $\beta_m$ about some central frequency, $\omega_0$ [63]

$$\beta_m(\omega) = \beta_{m,0} + \beta_{m,1}(\omega - \omega_0) + \frac{1}{2}\beta_{m,2}(\omega - \omega_0)^2 ..., \tag{2.56}$$

where the coefficients are given by

$$\beta_{m,k} = \frac{d^k \beta_m}{d\omega^k}. \tag{2.57}$$

The strength of the higher order coefficients decreases, usually only $\beta_{m,1}$ and $\beta_{m,2}$ are considered, while in certain situation terms up to $\beta_{m,4}$ are required [64]. $\beta_{m,1} = 1/v_{m,g}$ is inversely related to the group velocity of the mode - the velocity at which a pulse envelope propagates. The second order coefficient, $\beta_{m,2}$, describes how a pulse propagating broadens or narrows, spectrally, known as the group velocity dispersion (GVD).

### COUPLING BETWEEN WAVEGUIDES

Routing light on a chip is only the first step, we also require the ability to implement beam-splitters. Two main methods are used for coupling light between waveguides. The first, named directional couplers, exploit the evanescent field of modes confined in a waveguide. So long as the effective indices of the modes in each waveguide are the same, moving two waveguides close to each other will result in coupling. The combined system now has joint, rather than separate, eigenmodes known as supermodes with slightly different propagation constants. Beating between these supermodes leads to power transfer between the waveguides. The power transfer is given by

$$P(L) = P_0 \sin^2\left(\frac{\pi L \Delta n_{\text{eff}}}{\lambda}\right), \tag{2.58}$$

where $\Delta n_{\text{eff}}$ describes the difference in effective indices between the supermodes, which depends on the separation between the waveguides. This transmission, for single mode silicon waveguides with different waveguide separations, is shown in Fig. 2.4(b). If the

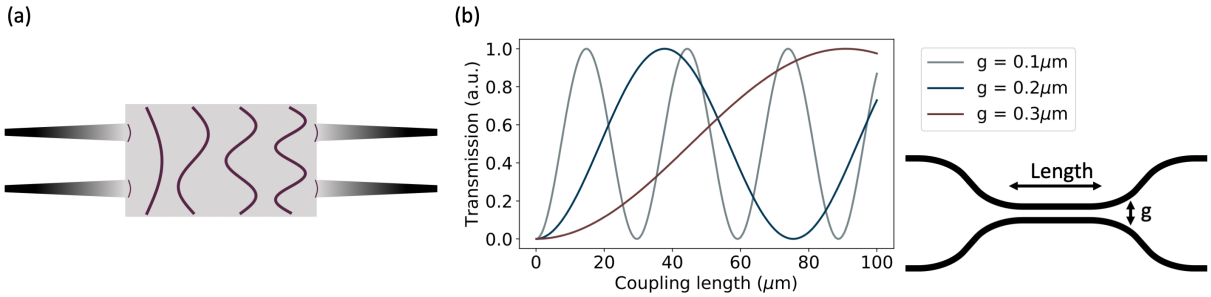


Figure 2.4: (a) Schematic of a multimode interference (MMI) coupler. Single mode waveguides are adiabatically tapered before entering a large multimode region, exciting higher order modes. Interference between these modes can be used to split the input light. (b) Power transfer of a directional coupler as a function of coupling length for different waveguide separations, for single mode silicon waveguides. Schematic of directional coupler is also shown. Here we have shown g in the bent region, this is due to space constraints, g corresponds to the width in the coupling region.

coupling length is selected such that the power transfer is exactly 50% then the unitary transfer matrix of a directional coupler is given by

$$U_{\mathrm{DC}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}. \tag{2.59}$$

Directional couplers can be fabricated with insertion loss that is effectively the waveguide propagation loss. The trade-off, however, is that the coupling coefficient depends very precisely on the waveguide separation, which, in silicon, is usually on the order of a few hundred nanometers. Fabrication imprecision makes reliably fabricating gaps on this scale is hard, often resulting in couplers with undesired coupling coefficients.

Another method for coupling waveguides, that is more robust to fabrication errors, known as the multimode interference coupler. Here, the waveguides are tapered before entering a shared wide section of waveguide. This waveguide can support many transverse modes, each with a different propagation constant $\beta_m$, this leads to interference between the excited modes. If the length is tailored correctly it is possible to pick out a plane where the interference splits the power from each input evenly. The transfer matrix for this component is given by

$$U_{\mathrm{MMI}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{2.60}$$

### COUPLING LIGHT INTO WAVEGUIDES

Another problem is getting light into, and photons off[5], an integrated photonic circuit. There exist many different techniques for this, in this section we will review the two most widely used. The coupling issue arises from the size difference between the waveguide mode and the mode in the optical fibre. One solution, named edge coupling, uses lensed fibre to reduce the mode size of the optical fibre. In order to bring the waveguide mode to the correct size the waveguide can be tapered - either up to a waveguide width comparable to the fibre mode size, or down so that most of the mode is located in the cladding, again increasing the overlap with the fibre mode. The coupling can be calculated with the overlap integral

$$\eta_c = \frac{|\int E_1^* E_2 dxdy|^2}{\int |E_1|^2 dxdy \int |E_2|^2 dxdy}. \tag{2.61}$$

Optimal waveguide taper parameters can then be chosen by optimising $\eta_c$. While these coupler designs are attractive for many reasons, being low loss and broadband, they have some practical limitations. The mode size from lensed fibre is $\approx 10\,\mu m^2$ making the coupling very sensitive to displacements in the fibre position. This makes them less suitable to longer integration times as changes in the laboratory conditions can have large effects on the coupling. Gluing has been used to combat this. The finality of this, which may be undesirable for certain tasks, not withstanding, this can also lead to decreases in the transmission. Edge couplers, as their name suggest, also require that the input and output coupling be at the edge of the chip. Often, due the density of photonic designs on multi-project wafer (MPW) runs, this is not always possible. To combat this another class of couplers, called grating couplers, can be used. These involve using periodic waveguide structures to diffract the light out of the plane of the chip. Tapering the waveguide such that it better matches the fibre mode size and then careful design of the grating, such that waves scattered from all parts of the grating constructively interfere, can result in relatively high efficiency fibre to chip coupling. Often a partial etch layer is also used to break the up/down symmetry of a grating, causing the light to be more preferentially scattered towards the fibre. Other tricks

5 There has been much progress integrated single photon detectors [65–68], but the work in this thesis utilises off chip detectors.

such as reflectors under the waveguide and apodised gratings can lead to collection efficiencies from silicon waveguides as high as 89% [69].

### THERMO-OPTIC PHASE SHIFTERS

Phase shifters represent the main degree of controllability in integrated photonic experiments. While there are many solutions to shifting phase in integrated chips, each with advantages and disadvantages, the current method of choice for large scale experiments are thermo-optic phase shifters. While their speed of reconfigurability is limited to kHz regimes, they are low loss, reliable, and easy to use. These phase shifters are based on the thermo-optic effect whereby a temperature change in a material results in a change of the refractive index. Loosely, this can be thought of as the dipoles in the material moving apart as the temperature increases and therefore the light interacts with them less. In order to change the temperature of the waveguide a metallic layer is patterned on top of the cladding layer. If a voltage is applied across this layer, power is dissipated as $P = IV$, which radiates as heat. The heating layer is around 2 µm above the waveguide, and the temperature radiates through the cladding preferentially, as the thermal conductivity of the cladding material is usually higher than air. This then raises the temperature of the waveguide. The change in refractive index is given by

$$\Delta n = \frac{dn_{\text{eff}}}{dT}\Delta T, \tag{2.62}$$

where $\frac{dn_{\text{eff}}}{dT}$ is the thermo-optic coefficient of the waveguide material. This leads to a phase shift of

$$\phi(L) = \frac{2\pi L}{\lambda}\frac{dn_{\text{eff}}}{dT}\Delta T. \tag{2.63}$$

### MACH-ZEHNDER INTERFEROMETERS

So far we have seen how we can integrate beam-splitters of any reflectivity, however the reflectivity is set at fabrication. In this section we will see how we can fabricate beam-splitters that have a reflectivity that can be changed at will. We do this through via a Mach-Zehnder interferometer (MZI). Here we use two integrated beam-splitters

set to have a transmission of $1/2$, with a phase shifter between. For illustrative purposes we will use the MMI matrix, but the same result can be achieved with directional couplers. The transfer matrix of the MZI is then

$$\begin{aligned} U_{\text{MZI}} &= U_{\text{MMI}} U_{\text{PS}} U_{\text{MMI}} \\ &= \frac{1}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}\begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e^{i\phi}-1 & i\left(e^{i\phi}+1\right) \\ i\left(e^{i\phi}+1\right) & 1-e^{i\phi} \end{pmatrix} \\ &= \frac{e^{i\frac{\phi}{2}}}{2}\begin{pmatrix} e^{i\frac{\phi}{2}}-e^{-i\frac{\phi}{2}} & i\left(e^{i\frac{\phi}{2}}+e^{-i\frac{\phi}{2}}\right) \\ i\left(e^{i\frac{\phi}{2}}+e^{-i\frac{\phi}{2}}\right) & e^{-i\frac{\phi}{2}}-e^{i\frac{\phi}{2}} \end{pmatrix} \\ &= ie^{i\frac{\phi}{2}}\begin{pmatrix} \sin\left(\frac{\phi}{2}\right) & \cos\left(\frac{\phi}{2}\right) \\ \cos\left(\frac{\phi}{2}\right) & -\sin\left(\frac{\phi}{2}\right) \end{pmatrix}, \end{aligned} \tag{2.64}$$

which, up to a global phase, represents a beam-splitter with a reflectivity that can be tuned via the internal phase. This, plus an extra external phase shifter, is precisely the building block of the decompositions described in sec. 2.3.2.

Interestingly, we can introduce frequency dependence to the coupling coefficient by using an MZI with different length arms. To understand this we note that the phase acquired by a waveguide of length $L$ is $\beta L$. We can then write the phase difference between the two arm of an MZI as $\Delta\phi = \beta_{\text{m,wg1}} L_{\text{wg2}} - \beta_{\text{m,wg2}} L_{\text{wg2}}$, where $\beta_{\text{m,wgi}}$ is the propagation constant of the $m^{\text{th}}$ mode in the $i^{\text{th}}$ waveguide. In the previous case we chose to have identical waveguide lengths, but to alter the propagation constant in one arm with a phase shifter. We could, however, also choose to fix the propagation constants and use waveguides of different lengths. The phase is then given by $\Delta\phi = \beta_{\text{m}}\Delta L$. As the phase shift now has a dependence on $\beta_{\text{m}}$ (rather than $\Delta\beta_{\text{m}}$), which is frequency dependent, the relative phase is also frequency dependent. The frequency separation between neighbouring resonant frequencies, known as the free spectral range (FSR), is found to be[6]

$$\Delta\omega_{\text{FSR}} = \frac{2\pi v_g}{\Delta L}, \tag{2.65}$$

where we see that increasing the length mismatch decreases the spacing between resonances. In quantum photonic experiments, devices like this serve two important purposes - routing generated signal and idler photons in different directions, or to sepa-

6 See Appendix C of Ref. [70] for a full derivation

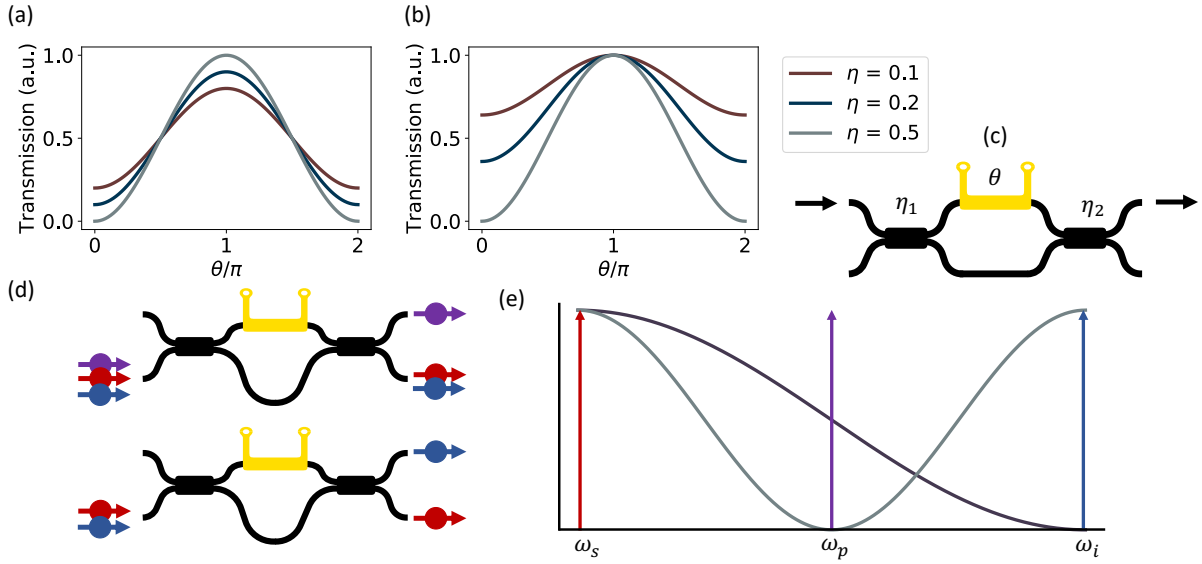


Figure 2.5: (a) and (b) Transmission as a function of Mach-Zehnder interferometer phase as we change the reflectivity of the input and output beam-splitters. (a) $\eta_1 = \eta$ and $\eta_2 = 0.5$. (b) Both reflectivities changed together, $\eta_1 = \eta_2 = \eta$. (c) Schematic of a Mach-Zehnder interferometer also showing input/output ports for the plots in (a) and (b). Reflectivities of the input and output fixed beam-splitters are labelled $\eta_1$ and $\eta_2$, respectively. (d) Both use cases of asymmetric MZI filters. (e) Spectral response of both AMZI cases. Grey line shows spectral response of a pump suppression filter, dark blue shows a splitting filter.

rate them together from the pump. As the signal and idler are generated symmetrically around the pump the FSR required to route the photons is exactly twice that required to suppress the pump, as shown in Fig. 2.5

### 2.5.2 *Comparison of different materials*

Many different materials can be employed to generate and manipulate light in the ways described above, each having different attributes. One considering factor is the index contrast between the core and the cladding. This is responsible for how confined the light is in the waveguide, which therefore dictates the feature size of these integrated devices. Materials such as silicon, in the form of silicon-silicon dioxide waveguides, have a large index contrast, which allows sub-micron scale waveguide profiles. The refractive index of a material often correlates with the material absorption. Higher index waveguides tend to have higher loss, a less than desirable attribute for quantum

applications. As another example, we have seen two methods of generating squeezed states through different nonlinear-processes. However, only certain materials have the necessary crystal structure to exhibit $\chi^{(2)}$ effects. These materials are desirable for two reasons. Generating squeezing through SPDC means that the pump and generated photons are well separated in frequency, making filtering easier. Also, as we have seen the strength of higher order nonlinear susceptibilities decreases, meaning that with SPDC it is often easier to generate high levels of squeezing. Beyond this, if one of the mixing fields in $\chi^{(2)}$ processes is $\omega = 0$, that is DC, then a phase shift is imparted on the second wave which is $\propto \chi^{(2)} E_1(0)$. This allows high speed, low loss, modulators to be integrated, a necessary component for many quantum technologies[7]. While some $\chi^{(2)}$ materials have been demonstrated in materials with large index contrasts [19–21] they suffer from complicated fabrication process that are outside the scope of standard semiconductor foundries - at least for now. In the rest of this thesis we will use

| | Silica ($GeO_2$ doped) | Silicon | Silicon nitride ($Si_3N_4$) |
|---|---|---|---|
| $n$ | 1.460 | 3.478 [71] | 1.985 |
| $n_{\text{core}}/n_{\text{clad}}$ | 0.0045 [48] | 2.41 | 1.37 |
| $\alpha$ ($\text{cm}^{-1}$) | $10^{-6}$ [72] | 0.37 [14] | 0.07 |
| $\chi^{(3)}$ ($\times 10^{-20}$ $\text{m}^2\text{V}^{-2}$ ) | 0.016 [73] | 25.6 [74] | 0.34 [75,76] |
| $\beta_{\text{TPA}}$ (cm $\text{GW}^{-1}$) | - | 0.61 [77] | negligible [78] |
| $E_{\text{bg}}$ (eV) | $> 7.62$ [79] | 1.1 [80] | 5 [71] |
| $\sigma$ ($\text{Wm}^{-1}\text{K}^{-1}$) | 1 [81] | 163 [81] | 90 [82] |
| $\frac{dn}{dT}$ ($\times 10^{-6}\text{K}^{-1}$) | 9.5 [83] | 206 [71] | 24.5 [83] |

Table 2.1: Table showing the typical material parameters for the waveguide materials used throughout this thesis. For doped silica all numbers are at $\lambda$ =800 nm and for the other two $\lambda$ =1550 nm. Index contrast measured with silica cladding, $n_{1550} = 1.444$ ,$n_{800} = 1.453$ [84]. Note, $GeO_2$ doped calculated from $n_{800}$ and the index contrast in Ref. [48]. All other index contrasts are calculated. Propagation loss quoted for average single mode waveguide. Unreferenced SiN parameters are measured at the University of Münster. $\alpha$ is the propagation loss of standard single mode waveguides, $\beta_{\text{TPA}}$ is the two-photon absorption coefficient, $E_{\text{bg}}$ is the bandgap energy, $\sigma$ is the thermal conductivity, and $\frac{dn}{dT}$ is the thermo-optic coefficient.

integrated devices fabricated in three different materials, table 2.1 reviews their main

7 There is an equivalent $\chi^{(3)}$ effect where the phase shift is $\propto \chi^{(3)}|E_1(0)|^2$, but being a $\chi^{(3)}$ effect, this is much weaker.

characteristics. Initially, we use silica waveguides with a germanium doped silica core and silica cladding. Low propagation loss and high coupling efficiency due to the large feature size makes this platform appealing. The low nonlinearity renders on chip sources are impossible. Optical fibre sources are common but either require large lengths of fibre or special fibre with cores engineered to increase the intensity or interaction length of the light [85,86]. Ultimately, the low nonlinear coefficient and large feature size means this platform is unlikely to scale past small proof-of-principle devices. Next, we use a device fabricated from silicon with silica cladding. This material is easily mass manufacturable as these devices fall within the remit of standard complementary metal oxide semiconductor (CMOS) fabrication processes used in the microelectronics industry. These devices have a large index contrast and so can be made with sub-micron sized waveguides. The high nonlinearity is offset by the high two-photon absorption (TPA) at telecom wavelengths. There are approaches to alleviating this problem by moving to higher wavelengths [70,87]. Another approach is to move to a material with a larger bandgap energy, allowing us to retain the current telecom infrastructure. To see this we also study devices where the core is stoichiometric silicon nitride ($Si_3N_4$) waveguides. These devices have a nonlinearity that is two orders of magnitude lower than silicon, however the lack of two photon absorption allows much higher pump powers to be reached.

## 2.6 QUANTUM INTERFERENCE

Interference of quantum amplitudes underpins many quantum technologies. In classical photonics, wave interference is a well known and intuitive phenomenon, in quantum photonics this intuition breaks down. In the following section we will describe two different quantum interference effects.

### HONG-OU-MANDEL INTERFERENCE

The hallmark of quantum photonics, and the backbone for most discrete variable quantum photonic technologies is two photon interference, first discovered by Hong, Ou, and Mandel [88]. This effect is a manifestation of the bosonic nature of photons and is observable when we ask the question, given two indistinguishable photons

incident on a balanced beam-splitter, what is the probability of a coincidence event at beam-splitter outputs? Naively, one would have thought this probability is 0.5 - there are four possible output patterns and two of them would result in a coincidence event. However, as we will now see that this is not the case. Initially we define our input state to be

$$|11\rangle = \hat{a}_1^\dagger \hat{a}_2^\dagger |00\rangle . \tag{2.66}$$

The mode transformation of a balanced beam-splitter is given by Eq. 2.30, with $\eta = 1/2$

$$\begin{aligned} \hat{a}_1^\dagger &\to \frac{1}{\sqrt{2}} \left( \hat{b}_1^\dagger + \hat{b}_2^\dagger \right), \\ \hat{a}_2^\dagger &\to \frac{1}{\sqrt{2}} \left( \hat{b}_1^\dagger - \hat{b}_2^\dagger \right). \end{aligned} \tag{2.67}$$

Applying these to our input state we arrive at

$$\begin{aligned} |11\rangle \to |\psi\rangle_{\text{out}} &= \frac{1}{2} \left( \hat{b}_1^\dagger + \hat{b}_2^\dagger \right) \left( \hat{b}_1^\dagger - \hat{b}_2^\dagger \right) |00\rangle , \\ &= \frac{1}{2} \left( \hat{b}_1^\dagger \hat{b}_1^\dagger + \hat{b}_2^\dagger \hat{b}_1^\dagger - \hat{b}_1^\dagger \hat{b}_2^\dagger + \hat{b}_2^\dagger \hat{b}_2^\dagger \right) |00\rangle , \\ &= \frac{1}{\sqrt{2}} \left( \hat{b}_1^\dagger \hat{b}_1^\dagger + \hat{b}_2^\dagger \hat{b}_2^\dagger \right) |00\rangle = \frac{|20\rangle + |02\rangle}{\sqrt{2}}. \end{aligned} \tag{2.68}$$

where in the last step we have used the commutation relation $\left[ \hat{b}_i^\dagger, \hat{b}_j^\dagger \right] = \delta_{ij}$. From this it is plain to see that we should, in fact, never expect a coincidence measurement. We can see this is an effect of destructive interference between the two output patterns that could give rise to a coincidence. In general this interference effect hinges on photons that are both indistinguishable and in pure states. Adding distinguishability or impurity adds information about which photon came from which source, as with all quantum interference effects, this degrades the interference quality. This means that HOM interference is a useful tool for probing the quality of single photons. To do this experimentally, however, a single measurement of coincidence rates after the beam-splitter is not enough. This rate will fluctuate with the number of photons input or with drifts in the coupling in the setup etc. To really probe the quality of the interference effect we need to measure the contrast of the interference. There are two ways that this is performed. One is to alter the distinguishability of the photons and the other is to change the reflectivity of the beam-splitter. In bulk optics experiments, the first is usually preferred as it is easy to vary the distinguishability with delay lines. The result

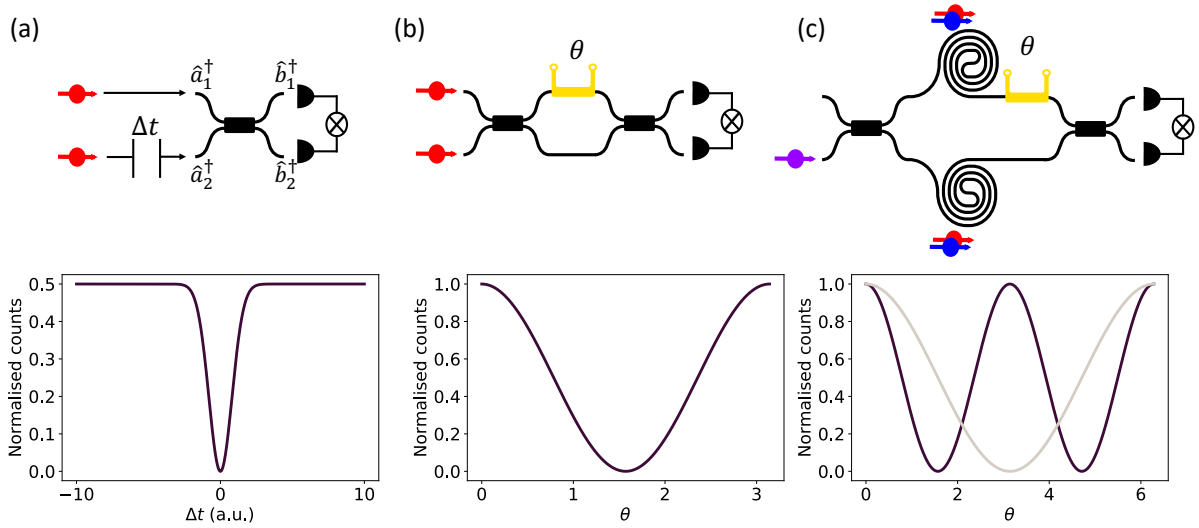


Figure 2.6: Schemes for measuring quantum interference and the resulting measurement outcomes. (a) Standard HOM dip. Coincidences shown as a function of time delay. We also indicate the mode labelling used for all the schemes. (b) HOM fringe where the beam-splitter reflectivity is changed. Fringes as a function of $\theta$ are observed. (c) time-reversed HOM interference. Quantum fringe (purple) with a frequency of twice that of the classical fringe (beige).

of such a measurement is a dip where at large delays, where the photons are completely distinguishable, we retrieve the expected classical coincidence probability of 0.5, and at $\Delta t = 0$ the coincidence probability is a minimum. For this reason, this experiment is referred to as a HOM dip. A schematic of this experiment and an example dip are shown in Fig. 2.6 The visibility of this fringe can be defined as

$$\mathcal{V} = \frac{P_{11}\left(\Delta t \to \infty\right) - P_{11}\left(\Delta t = 0\right)}{P_{11}\left(\Delta t \to \infty\right)} = \frac{C_{\max} - C_{\min}}{C_{\max}}, \tag{2.69}$$

where $C_{\max}$ and $C_{\min}$ are the maximum and minimum measured counts, respectively. A full treatment of this effect, without assuming anything about the spectral properties of the photons shows that this visibility depends on the total overlap of the photons spectra [89, 90]. If we assume that both photons are described by the same density matrix it can be shown that this visibility is equal to the purity of that density matrix. For this reason, the HOM dip is often used to determine the purity of photons generated from different but similar sources. In integrated experiments, it is not so easy to change the distinguishability of photons that are generated on chip. For this reason it is often preferable to alter the beam-splitter reflectivity instead. In this case, we expect

sinusoidal fringes in terms of the beam-splitter phase. Now, if we calculate the visibility as above it no longer maps directly to the purity. This is because at the peak of the fringe, where the beam-splitter either implements an identity or swap, we expect no bunching at all and all photons created should give rise to a coincidence event. In order to link the coincidence measurements we can set $P_{11}\ (\Delta t \to \infty)$ to be $C_{\max}/2$ as in Ref. [91], which gives us the purity to be

$$\mathcal{P} = \frac{C_{\max} - 2C_{\min}}{C_{\max}}. \tag{2.70}$$

In later chapters we will use this definition to determine the purity of our photon sources, however we will define the visibility of our fringe as

$$\mathcal{V} = \frac{C_{\max} - C_{\min}}{C_{\max} + C_{\min}}. \tag{2.71}$$

This quantity is often called the contrast and is used to quantify the quality of classical interference fringes. In the following work we will refer to this kind of interference as a HOM fringe. Beyond this, classically, for the HOM dip experiment, one would expect a visibility of 0. For the fringe, however, with the beam-splitter at the fringe peak we would expect no bunching, however at the minimum we would still expect 50% of the photons to result in bunching. This leads to a classical fringe visibility of $1/3$.

### TIME-REVERSED HOM INTERFERENCE

The visibility of a HOM dip or fringe provides information on the spectral overlap of the individual photons. For spontaneous sources we may also wish to have information on the complete indistinguishability of the source, including the heralding photons. For this we turn to a different type of quantum interference which is the time reversed version of the standard HOM experiment we described above. In the forward case we took the state $|11\rangle$ to $(|20\rangle + |02\rangle)/\sqrt{2}$. To see what happens in the reverse case we will, again, start by ignoring the photon's spectrum. To generate the end state we can pump two sources in superposition, if we post select on only one of the sources firing we are in the state

$$|\psi\rangle_{in} = \frac{1}{\sqrt{2}}\left(\hat{a}_1^\dagger \hat{a}_1^\dagger + e^{i2\phi}\hat{a}_2^\dagger \hat{a}_2^\dagger\right)|00\rangle\,, \tag{2.72}$$

where we have included a phase shift in one of the modes in order to allow us to sweep a fringe. Note that the phase shift is $2\phi$, this is because the phase shift operator adds a $\phi$ phase shift to each photon, see the Eq. 2.32. When we apply the beam-splitter relations to this state we see

$$\begin{aligned}|\psi\rangle_{in} \rightarrow |\psi\rangle_{\text{out}} &= \frac{1}{2\sqrt{2}}\left(\left(\hat{b}_1^\dagger + \hat{b}_2^\dagger\right)\left(\hat{b}_1^\dagger + \hat{b}_2^\dagger\right) + e^{i2\phi}\left(\hat{b}_1^\dagger - \hat{b}_2^\dagger\right)\left(\hat{b}_1^\dagger - \hat{b}_2^\dagger\right)\right)|00\rangle \\ &= \frac{1}{2\sqrt{2}}\left(\left(\hat{b}_1^\dagger\hat{b}_1^\dagger + \hat{b}_2^\dagger\hat{b}_2^\dagger\right)\left(1 + e^{i2\phi}\right) + \left(\hat{b}_1^\dagger\hat{b}_2^\dagger + \hat{b}_2^\dagger\hat{b}_1^\dagger\right)\left(1 - e^{i2\phi}\right)\right)|00\rangle \\ &= \frac{e^{i\phi}}{\sqrt{2}}\left(\left(\hat{b}_1^\dagger\hat{b}_1^\dagger + \hat{b}_2^\dagger\hat{b}_2^\dagger\right)\cos(\phi) + i\left(\hat{b}_1^\dagger\hat{b}_2^\dagger + \hat{b}_2^\dagger\hat{b}_1^\dagger\right)\sin(\phi)\right)|00\rangle\,.\end{aligned} \tag{2.73}$$

Here we see two terms varying sinusoidally with $\phi$, one where both photons bunch in the output arms and one where they split. As expected, this fringe oscillates at a frequency twice that which would be expected classically. In this analysis we have assumed that both photons from the source are degenerate but in general the derivation is also valid when the signal and idler are spectrally distinguishable. In Ref. [90] it is shown that for sources with joint spectra $F_1(\omega_s, \omega_i)$ and $F_2(\omega_s, \omega_i)$ the visibility of a fringe on $\phi$ is given by

$$\mathcal{V} = \left|\int F_2^*(\omega_s, \omega_i)\, F_1(\omega_s, \omega_i)\, d\omega_s d\omega_i\right|^2. \tag{2.74}$$

This shows us that the visibility depends on the overlaps of the joint spectra but is independent of $\omega_s$ and $\omega_i$. Thus, it tells us nothing about the purity but captures the full distinguishability of the sources.

## 2.7 QUANTUM COMPUTING WITH LINEAR OPTICS

We have thus far seen how quantum photonics differs from classical photonics, in both representation and phenomena with no classical analogue. In this section we will briefly look at how a universal quantum computer can be built up with only the components we have discussed in this chapter

### 2.7.1 *State representation and single qubit gates*

The basic unit of quantum information is a qubit (quantum bit). In analogy with the classical information, the qubit requires two well-defined states and information is processed by gates acting on these qubits. However, unlike classical bits, which can only ever be in the state 0 *or* 1, a qubit can be in a linear superposition of it's two basis states

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{2.75}$$

where $\alpha, \beta \in \mathbb{C}$. While we require 4 matrices to fully span the Hilbert space of a single qubit (see sec. 6.6.2), arbitrary qubit operations can be decomposed into rotations around two axes

$$\begin{aligned} \hat{R}_x(\phi) &= \exp\frac{i\phi\sigma_x}{2} \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \hat{R}_z(\phi) &= \exp\frac{i\phi\sigma_z}{2} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.76}$$

Any qubit rotation can now be written as

$$|\psi\rangle = R_z(\phi_1)R_x(\phi_2)R_z(\phi_3)\,|\psi_0\rangle\,. \tag{2.77}$$

One benefit of using photons as quantum information carriers is one we have already seen - the ability to code information on many degrees of freedom. Here we describe a specific encoding known as the 'dual-rail' encoding, where both qubit basis states are defined by a single photonic excitation in one of two orthogonal modes. As mentioned before, many photonic degrees of freedom can be used to define an orthogonal set of modes. For the rest of this section we will restrict our conversation to qubits encoded in the spatial degree of freedom, although schemes exist for other encodings [28, 92, 93]. Given two waveguides described by their creation operators $\hat{a}_1^\dagger$ and $\hat{a}_2^\dagger$, we can define our qubit states to be

$$\begin{aligned} |0\rangle &= \hat{a}_1^\dagger\,|00\rangle = |10\rangle \\ |1\rangle &= \hat{a}_2^\dagger\,|00\rangle = |01\rangle \end{aligned} \tag{2.78}$$

Quantum gates now become operations acting on these modes. In fact, the two necessary rotations can be implemented by components we have already seen. $\hat{R}_x$ changes the amplitude of the state between $|0\rangle$ and $|1\rangle$, which can be mapped the variable reflectivity

beam-splitter. $\hat{R}_z$ changes the relative phase between $|0\rangle$ and $|1\rangle$, an action that can be performed by a phase shifter. Up to a global phase, our arbitrary qubit rotations now become

$$|\psi\rangle = U_{\mathrm{PS}}(\phi_1)U_{\mathrm{MZI}}(\theta)U_{\mathrm{PS}}(\phi_2)\,|\psi_0\rangle\,. \tag{2.79}$$

Single qubit rotations alone are not sufficient for a universal quantum computer, all universal gate sets always include at least one two qubit gate. This is where photons as qubits encounter some problems. Deterministic generation of entanglement is practically impossible with only linear optics. The problem is the lack of photon-photon interaction. Photons are *non-interacting bosons* and while on one hand this means that they interact very weakly with their environment, leading to long coherence times, it does mean that two photon gates cannot be directly implemented. One objection to this one might have is haven't we seen a photon-photon interaction in the HOM experiment described above? While this may look like an interaction, in fact it is just an interference effect, albeit now of the quantum rather than classical field amplitudes. Here, the transition probabilities for each photon are independent on whether the second photon is there or not. For qubit-qubit interaction we need to be able to make a state dependent change on another qubit. Another thought might be that, well aren't quantum computations described by unitary transformations? Surely, we can use a sufficiently large Reck or Clements scheme to implement whatever computation we like? The problem here stems from the fact that the basis of a given quantum computation unitary is the qubit basis states, i.e. for a two qubit gate the basis states are $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$. The basis states for the interferometer would be $|0\rangle_1$, $|1\rangle_1$, $|0\rangle_2$, $|1\rangle_2$. In this basis the same transformation becomes non-unitary [94].

### 2.7.2 *Generating entanglement*

The key to overcoming this inherent draw back of photonic qubits is to use measurement and a technique known as post-selection. With only single photons and linear optics we are limited to probabilistic gates. This is because there are always some possible outcome patterns that take us out of the computational space, for example multiple photons leaving in a single mode. In post-selection, only successful events are counted, and all others are ignored. While simple, this trick is quite powerful in-fact this alone is enough to generate entanglement with linear optics. To see how post-selection can

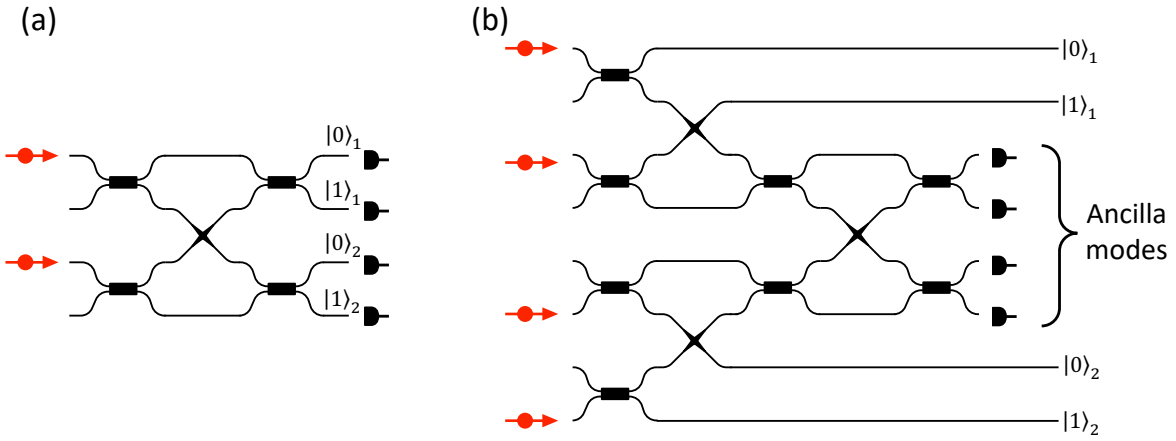


Figure 2.7: Circuits for generating two qubit entangled states. (a) Post-selected Bell state generator using single photons and a Fourier transform interferometer. (b) Heralded Bell state generator

be used to generate entanglement, we start with two qubits initialised in the state $|\psi\rangle_{\text{in}} = |0\rangle_1 \otimes |0\rangle_2$. Using the dual-rail encoding we can map our two qubit states to the creation operators of 4 waveguides

$$\begin{pmatrix} |0\rangle_1 \\ |1\rangle_1 \\ |0\rangle_2 \\ |1\rangle_2 \end{pmatrix} = \begin{pmatrix} \hat{a}_1^\dagger \\ \hat{a}_2^\dagger \\ \hat{a}_3^\dagger \\ \hat{a}_4^\dagger \end{pmatrix}. \tag{2.80}$$

Our input state is therefore $|\psi\rangle_{\text{in}} = \hat{a}_1^\dagger \hat{a}_3^\dagger |00\rangle$. These photons are then passed through an interferometer implementing a four-mode discrete Fourier transform, a generalisation of the balanced beam-splitter to four modes

$$F = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}, \tag{2.81}$$

which can be implemented with the photonic circuit shown in Fig. 2.7(a). Applying this transformation to our input state we are left with

$$|\psi\rangle_{\text{out}} = \frac{1}{4}\left(2\hat{a}_1^\dagger\hat{a}_3^\dagger - 2\hat{a}_2^\dagger\hat{a}_4^\dagger + \hat{a}_1^\dagger\hat{a}_1^\dagger - \hat{a}_2^\dagger\hat{a}_2^\dagger + \hat{a}_3^\dagger\hat{a}_3^\dagger - \hat{a}_4^\dagger\hat{a}_4^\dagger\right)|00\rangle\,, \tag{2.82}$$

which is not an entangled state. However, if we post-select and restrict ourselves to only considering events where both photons leave in separate waveguides the above state reduces to

$$|\psi\rangle_{\text{out}} = \frac{1}{2}\left(\hat{a}_1^\dagger\hat{a}_3^\dagger - \hat{a}_2^\dagger\hat{a}_4^\dagger\right) = \frac{1}{\sqrt{2}}\left(|0\rangle_1|0\rangle_2 - |1\rangle_1|1\rangle_2\right), \tag{2.83}$$

where in the final step we have renormalised and applied the relationship in Eq. 2.7.2. Here we can see that with this caveat we can produce a Bell state using only linear optics. The key issue with this scheme is that in order to determine if the gate has been successful we need to measure, and therefore destroy, the photons that we wish to entangle. This is a large setback if we wish to perform future gates on these entangled qubits. To combat this, we can introduce extra photons, known as ancillas. Measuring a subset of the total photon number can then be used to *herald* the success of a gate. The standard circuit which can generate heralded Bell states is shown in Fig. 2.7(b). This circuit requires 4 total photons in 8 optical modes. These modes can be divided up into 4 signal modes which are used to define the two output qubits, and 4 ancilla modes. The 4 photons are input into the signal modes and balanced beam-splitters are used to couple light from each signal mode into one of the ancilla modes, before a Fourier interferometer is used to erase the which-path information. It can be shown that when exactly two photons are detected in two separate ancilla modes the remaining state is locally equivalent to a Bell state [95]. There are 6 possible heralding patterns, each of which has a probability of $1/32$, and each results in a different maximally entangled state of two photons across 4 modes, which are all equivalent up to mode swaps. This results in a total gate success probability of $3/16$. This can be boosted in a number of ways. Initially, one notices that there are 4 output patterns where the two heralding photons leave in a single ancilla mode. Each of these patterns has a probability of $3/64$ and leaves the output state in a non-maximally entangled state. It is possible to use entanglement distillation techniques and additional linear optical circuitry to convert these output patterns into a maximally entangled Bell state with a probability

of $1/3$ [96]. This results in a total gate probability of $3/16 \times (1 + 1/3) = 1/4$. A more in-depth description can be found in Refs. [95, 97]. Ref. [97] show that this probability can be boosted by applying the coupling and path-erasure sections of the gate many times and stopping when a total of 2 photons have been detected. The gate fails if more than 2 photons are subtracted, this probability can be pushed towards 0 by decreasing the coupling coefficients of the initial beam-splitters. This comes at the cost of increasing the number of times the gate needs to be repeated before two photons are subtracted. By increasing the number of repeats we approach unit probability of measuring in the two-photon subspace. The gate probability now becomes $\approx 0.5 \times (1 + 1/3) = 2/3$.

Another important resource for quantum computing [10] are three-qubit entangled states in the form

$$|\psi\rangle_{\text{GHZ}} = \frac{1}{\sqrt{2}} \left( |000\rangle + |111\rangle \right), \tag{2.84}$$

which are named Greenberger-Horne-Zeilinger (GHZ) states. A similar scheme to the Bell state generator exists to generate this class of states. This scheme utilises 6 input photons in 12 optical modes where a 3-photon click in the ancilla modes heralds a state locally equivalent to a GHZ state. With feed-forward to perform the necessary local rotations on each qubit, conditional on the heralding pattern, all 8 viable heralding patterns can give rise to the desired state, resulting in a total success probability of $1/32$ [98]. Note that a more optimal circuit, if feed-forward is not possible, does exist and has a success probability of $1/54$ [99].

### 2.7.3 *Multiplexing*

So far we have described a variety of building blocks for a quantum computer which are probabilistic. As we wish to scale these devices up, the success probability of the whole device drops exponentially. To combat this unfavourable scaling, we can turn to a technique known as multiplexing. This involves performing many probabilistic processes in parallel and then selecting out the first successful event. These can be repeated either spatially, or temporally, as shown in Fig. 2.8[8]. Spatial multiplexers have many copies of a probabilistic source or gate which are tried simultaneously. Heralding clicks are used to signify when one of these processes was successful. The result of this process is held in a delay line until classical logic can process the click pattern and

8 Other photonic degrees of freedom can also be used, for example frequency [100]

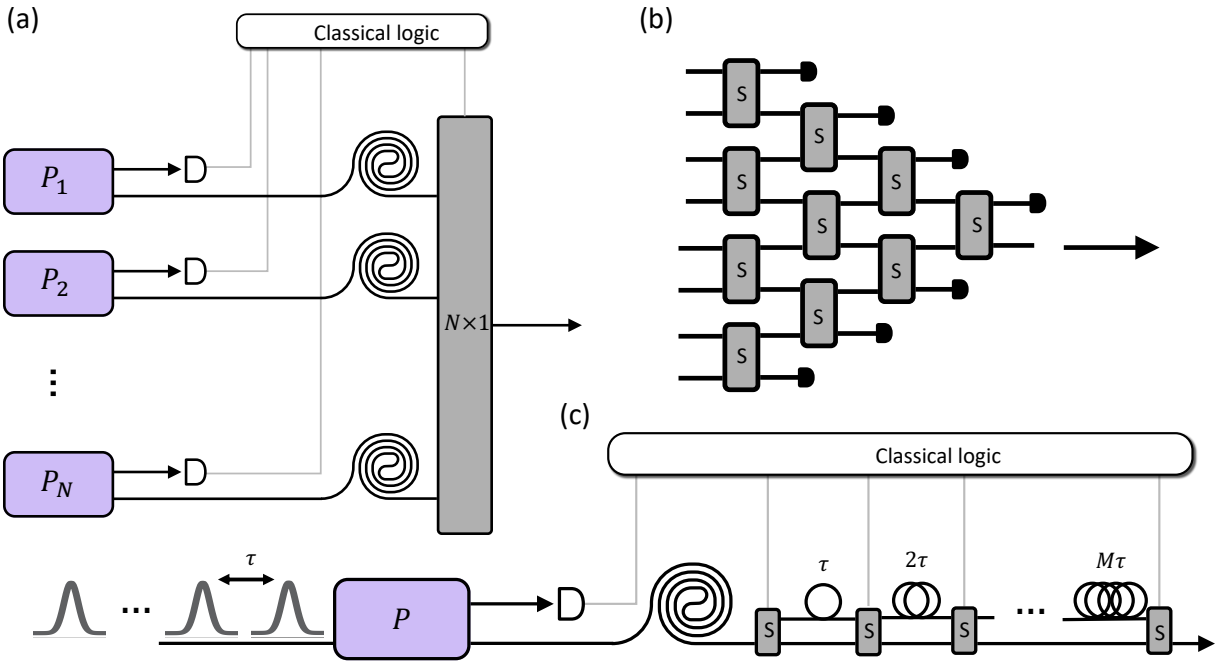


Figure 2.8: (a) Scheme for a spatial multiplexer. $N$ probabilistic devices, $P_i$ are operated in parallel, these could be photon sources or gates. On heralding an event happening, an $N \times 1$ switching network is used to route this event to the rest of the device. Delay lines are required to give enough time for the switching network to be correctly set. (b) Example switching network where each photon only sees $\log_2(N) + 1$ switches. (c) Temporal multiplexer. Here, a single probabilistic device is tried many times, when an event is successful a switching network and delays are used to ensure the event leaves at the same time as any events from other multiplexers. The maximum necessary delay is given by $M = 2^{\lfloor \log_2(N-1) \rfloor}$

send signals to a switching network which is then used to direct the successful event to the desired output. It is possible to design a switching mesh of $N$ switches but where each event only sees $\log_2(N) + 1$ switches. We see here the importance of being able to herald the generation of entanglement, with a post-selected gate it is impossible to multiplex. The probability of a successful event is given by

$$p_f = 1 - (1 - p_0)^N, \tag{2.85}$$

where $p_0$ is the success probability of a single process and $N$ is the total number of repeated processes. We can see that the success probability can be driven arbitrarily high but adding more simultaneous attempts. Another method is to utilise time instead

of space. Here a single copy of the probabilistic process is repeated many times. When a successful event is heralded, any temporal distinguishability between different multiplexers must be removed. To do this the successful event is guided into a set of delay lines which delay the $i^{\text{th}}$ event through enough delays that it exits the same time as the $N^{\text{th}}$ event would have. The maximum delay required is therefore $(N-1)\tau$. This can also be achieved with each mode seeing only $\log_2(N)+1$, number of switches. Delays between switches that double in length each time up to a maximum length of

$$M = 2^{\lfloor \log_2(N-1) \rfloor}. \tag{2.86}$$

Other examples of multiplexing schemes have been found including one for sources that moves the switching to be on the pump rather than the photons [101]. Time multiplexing schemes with less stringent requirements on the time distinguishability have also been invented [94]. Here they note that only two events that are going to be interfered together need to be emitted at the same time. A key component for all of these schemes are fast and low loss switches. We will see in chapter 5 one possible implementation of this kind of switch.

### 2.7.4 *Measurement based quantum computing*

A first pass at designing a scheme for universal quantum computing using linear optics was presented in 2000 by Knill, Laflamme and Milburn (KLM) [9]. This scheme required feed forward and ancilla states. While this scheme was technically efficient, that is has polynomial scaling in all overheads, the scaling factors put this scheme outside the scope of physical implementations - requiring, for example, over 6 million ancilla Bell pairs for a single two qubit operation [94]. Over the following years, however, many improvements to the KLM scheme surfaced [102–104] many which used a different form of quantum computing from the circuit model style of KLM, known as measurement based quantum computation (MBQC). Rather than having an array of qubits on which gates are acted in order to carry out the computation, in MBQC we have a 2D entangled state, known as a cluster state where the rows represent the qubit register and gates are performed by sequentially measuring the qubits along a row. If done correctly, the measurements teleport the information from qubit to qubit. Thus, any quantum algorithm now becomes a set of measurement bases in which the

qubits are sequentially measured. The culmination of this work was Gimeno-Segovia et al. [10], who showed that a large entangled cluster state, a sufficient resource for universal quantum computation [31], could be generated from small entangled states. This scheme relied on the ability to fuse the smaller entangled states together. Two different gates exist for this, one which destroys both the fusing qubits but entangles the rest of the state, known as type-II fusion, and one which only destroys one of the qubits, entangling the remaining qubit with the rest of the state of the destroyed qubit. As with other linear optical gates, these fusion gates are probabilistic, although with ancilla states, these probabilities can be as high as 3/4 [105]. In previous schemes that used these gates, the probabilistic nature was overcome by performing a similar technique to multiplexing, where the gates were repeated until success. In Ref [11] a protocol was demonstrated where small entangled resource states were fused together with probabilistic entangling gates. So long as the probability of these gates is above a threshold, known as the percolation threshold, the incomplete lattice is universal for quantum computation. If the percolation threshold is surpassed then the lattice will be long range connected, in fact, for an infinite lattice there is always a connected path across the lattice. By identifying certain qubits as *central* qubits and others as *path* qubits it is possible to choose the central qubits such that they are all connected by strings of path qubits and can therefore act as the logical nodes of a new *renormalised* lattice. This was further extended in Ref. [10] where it was shown that the resource states could be just three-qubit GHZ states. Resource estimations have shown that even a few hundred logical qubits would be sufficient to simulate small Fermi-Hubbard systems, moving up to between $10^3$ and $10^4$ logical qubits to break RSA encryption, depending on the encryption size, and up to $\approx 10^5$ logical qubits for simulating the chemistry in Li-ion batteries [13] [9].

9 Note here that these are logical qubits which would likely be encoded in a surface code requiring $\approx 1000$ physical qubits. Beyond this Ref. [106] estimates that each physical qubit would require $10^5 - 10^6$ single photon detectors and component loss rates below $10^{-3}$.

# 3

# TIME-RESOLVED TWO-PHOTON INTERFERENCE

"Time is an illusion. Lunchtime doubly so."

Douglas Adams, *Hitchhikers Guide to the Galaxy*

## 3.1 INTRODUCTION AND MOTIVATION

Quantum interference between indistinguishable photons forms the backbone of almost all quantum photonic technologies. High quality quantum interference requires photons that can be reliably generated in a pure state and the ability for many sources to emit photons that are indistinguishable in all degrees of freedom. There are two main classes of photon sources: parametric sources, based on nonlinear optics, and solid-state emitters. Currently, the best solid-state sources are quantum dots, which promise near-deterministic generation of a true single-photon Fock state. While there has been considerable work done to increase the purity and efficiency of these sources [107, 108], making even two dots generate spectrally indistinguishable photons has proven harder. Only recently was quantum interference with a visibility higher than 90% demonstrated between two separate quantum dots [109]. Integrated waveguide parametric sources, on the other hand, can generate spectrally pure photons [110, 111] and the repeatability of the fabrication process allows the fabrication of many indistinguishable sources on one device [15, 112]. Using resonant structures over traditional spiralled waveguides results in sources with a much smaller footprint and lower power consumption due to the field enhancement inside the resonator. It has been shown that high purity photons can be generated from resonant sources [111, 113–115]; however due to random internal phases arising from wavelength-scale path length differences, the emission wavelength of multiple sources will, in general, be different. This internal phase, and therefore the emission wavelength, can be adjusted using on-chip phase shifters. A common method for phase shifting involves thermal control of the refractive index in the ring. Scaling this method to larger numbers of sources becomes impractical due to cross talk between different phase shifters. Many other phase shifting options exist, but none of them are material agnostic, low-loss, and have a proven ability to scale.

In the following work, we experimentally test a different method to solve the problem of spectral distinguishability. Rather than forcing all sources to be intrinsically indistinguishable, we allow a given amount of spectral distinguishability and measure the photons after interference with a high timing resolution. Due to the conjugate nature of time and frequency, projecting the photon state onto a narrow time mode effectively broadens the spectrum of the photon, allowing non-identical photons to interfere. Beyond this, we show that our detectors are fast enough for us to probe interference effects within the wave packet. The chapter is laid out as follows: we introduce the relevant background theory before generalising the model to include realistic experimental imperfections. We then describe and characterise the experimental setup before moving on to the results. We perform HOM interference, first with indistinguishable photons and then with spectrally distinguishable photons to show the efficacy of our model in describing the results. Finally, we investigate interference in a complex Hadamard interferometer.

## 3.2 THE FOURIER TRANSFORM

This chapter describes experiments where we exploit the conjugate nature of time and frequency. To properly understand the later results we first need to understand this relationship — namely the Fourier transform [1]. The Fourier transform can be thought of as a generalisation of the Fourier series to non-periodic functions. The Fourier series defines a periodic function as a sum of sinusoidal components

$$f(t) = \sum_{r=-\infty}^{\infty} c_r e^{i\omega_r t}, \tag{3.1}$$

where $\omega_r = \frac{2\pi r}{T}$ are the characteristic frequencies, $T$ is the function period, and $c_r$ are the Fourier coefficients given by the integral

$$c_r = \frac{\Delta\omega}{2\pi} \int_{-T/2}^{T/2} f(t)\, e^{-i\omega_r t}\, dt. \tag{3.2}$$

1 We note that the following analysis is taken mainly from Ref [116]

This can be substituted into Eq. 3.1 to give

$$f(t) = \sum_{r=-\infty}^{\infty} \frac{\Delta\omega}{2\pi} \int_{-T/2}^{T/2} f(u)\, e^{-i\omega_r u}\, du\, e^{i\omega_r t}. \tag{3.3}$$

As the period, T, tends to infinity the frequency separation between terms, $\Delta\omega$, tends to 0 and the sum over the discrete set of frequencies $\omega_r$ becomes an integral over the continuous variable $\omega$. This allows us to write $f(t)$ as

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\omega)\, e^{i\omega t}\, d\omega = \mathcal{F}^{-1}\left[f(\omega)\right], \tag{3.4}$$

where

$$f(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t}\, dt = \mathcal{F}\left[f(t)\right], \tag{3.5}$$

where the prefactors of $\frac{1}{\sqrt{2\pi}}$ have been chosen arbitrarily, the only condition is that their product is $\frac{1}{2\pi}$. These two transformations, called the inverse Fourier transform and Fourier transform, respectively, represent a bijective map between a parameter space $t$ and the characteristic frequency space $\omega$

$$f(t) \overset{\mathcal{F}}{\longleftrightarrow} f(\omega). \tag{3.6}$$

For the rest of this thesis, $t$ represents the temporal degree of freedom and the Fourier transform takes us to the frequency domain. Now that we have a way to move between the temporal and spectral profile of a signal, we will look at some properties of this transformation that will be important later on. The first situation we will look at are shifts in the frequency domain

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\omega \pm \Delta\omega)\, e^{i\omega t}\, d\omega. \tag{3.7}$$

We define a new variable $\omega' = \omega \pm \Delta\omega$ which gives us

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\omega')\, e^{i\omega' \mp \Delta\omega t}\, d\omega' = \frac{e^{i\Delta\omega t}}{\sqrt{2}} \int_{-\infty}^{\infty} f(\omega')\, e^{i\omega' t} d\omega' = e^{\mp i\Delta\omega t} f(t). \tag{3.8}$$

From this we see that a frequency shift leads to phase shift in the time domain. This will be important later where we use frequency shifts to see interference fringes in the

time domain. The next feature we will examine is how the widths of temporal and spectral functions relate to each other. To see this we will define a Gaussian temporal function, centred at 0 with a standard deviation, $\sigma_t$

$$f(t) = \frac{1}{\sigma_t\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma_t^2}}. \tag{3.9}$$

Taking the Fourier transform of this we are left with

$$f(\omega) = \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} e^{-\frac{\sigma_t^2\omega^2}{2}}. \tag{3.10}$$

This illustrates another feature of the Fourier transform: the Fourier transform of a Gaussian distribution is another Gaussian distribution. This resultant distribution is again a Gaussian centered at 0, this time with a standard deviation $\sigma_\omega = \frac{1}{\sigma_t}$. From this we can see that

$$\sigma_t\sigma_\omega = 1. \tag{3.11}$$

This equation is minimised for Gaussian functions but in general $\sigma_t\sigma_\omega \geq 1$. From this we see that the narrower in time a signal is, the wider the bandwidth of frequencies that make it up.

## 3.3 TIME-RESOLVED INTERFERENCE THEORY

In many quantum optics experiments, the coherence time of the photons is much less than the timing resolution of the detection system. The timing resolution of a system is bounded by the detection jitter, that being the total variance in the measured arrival time of a photon. Jitter can be introduced by the photon detectors themselves and also the amplification electronics and time tagging logic. Even in experiments where the system jitter results in a high enough intrinsic experimental timing resolution, experiments are generally performed with a coincidence window much larger than required. This is because in these experiments any interference effects that could be seen by reducing the coincidence window are less important than the loss this would introduce. In the following chapters we discuss quantum interference experiments where the timing resolution of the experiment is much higher than the coherence time of the interfering

photons. We will see that not only are new interference effects visible but also that, due to the conjugate relationship between time and frequency, it is possible to observe quantum interference with photons that are distinguishable in frequency. The initial work for this was performed by Legero et al. in Ref. [38], we recap the important points here. The general schematic is shown in Fig. 3.1. Here, we have two photons with arbitrary temporal profiles impinging on a balanced beam-splitter. We then detect photons at times $t_0$ and $t_0 + \tau$ on output modes 3 and 4, respectively. We label the top and bottom input modes as modes 1 and 2, respectively and the top and bottom output modes as modes 3 and 4. To perform this analysis we introduce the time-dependent

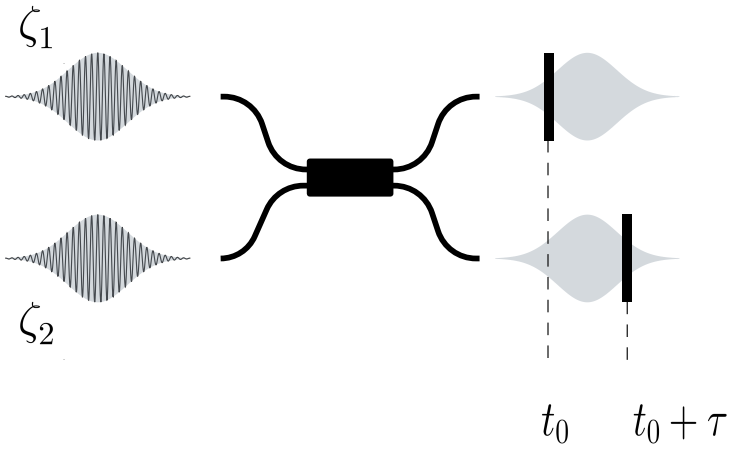


Figure 3.1: Two photons with arbitrary temporal shapes $\zeta_1$ and $\zeta_2$ impinge on a balanced beam-splitter. We then look for coincidence events separated by a time $\tau$. The precision of the measurement of $\tau$ must be much less than the photon's coherence time, defined as the FWHM of the photon wave packet. We integrate $t_0$ as the first coincidence event can happen anywhere in the wave packet.

mode operators $\hat{A}_i^\dagger(t) = \zeta_i(t)\hat{a}_i^\dagger$. Our input state is given by single photons input to the modes $i \in 1, 2$. The output modes, $i \in 3, 4$ are defined by the operation of the beam-splitter on the input modes

$$\begin{aligned} \hat{A}_3^\dagger(t) &= \frac{1}{\sqrt{2}}\left(\hat{A}_1^\dagger(t) + \hat{A}_2^\dagger(t)\right), \\ \hat{A}_4^\dagger(t) &= \frac{1}{\sqrt{2}}\left(\hat{A}_1^\dagger(t) - \hat{A}_2^\dagger(t)\right). \end{aligned} \tag{3.12}$$

We now look at the case where one photon is detected in mode 3 at $t_0$ and the second in mode 4 at $t_0 + \tau$. This leaves the output state $|\Psi\rangle = \hat{A}_3^\dagger(t_0)\,\hat{A}_4^\dagger(t_0 + \tau)\,|00\rangle$. By

substituting in Eq. 3.12 and restricting ourselves to only single photons per input mode, we arrive at the coincidence probability

$$\mathrm{P}\left(t_0, t_0+\tau\right) = \frac{1}{4}\left|\zeta_2\left(t_0\right)\zeta_1\left(t_0+\tau\right) - \zeta_1\left(t_0\right)\zeta_2\left(t_0+\tau\right)\right|^2. \tag{3.13}$$

It is interesting to note here that regardless of the profiles $\zeta_1$ and $\zeta_2$ the probability of a coincidence is always 0 when $\tau = 0$.

### 3.3.1 *Generalised two-photon interference*

Here we generalise the results Sec. 3.3 to include interference via an arbitrary transfer matrix and with mixedness of the photon states. The transformation on the mode operators now becomes

$$\begin{pmatrix}\hat{A}_3^\dagger(t)\\ \hat{A}_4^\dagger(t)\end{pmatrix} = \begin{pmatrix}\mathrm{M}_{11} & \mathrm{M}_{12}\\ \mathrm{M}_{21} & \mathrm{M}_{22}\end{pmatrix}\begin{pmatrix}\hat{A}_1^\dagger(t)\\ \hat{A}_2^\dagger(t)\end{pmatrix}. \tag{3.14}$$

We are looking for the output state where a photon is detected at time $t_0$ in mode 1 and time $t_0+\tau$ in mode 2. This is given by the state $|\psi_{\mathrm{out}}\rangle = \hat{A}_3^\dagger(t_0)\hat{A}_4^\dagger(t_0+\tau)\,|00\rangle$. Applying the transformation in Eq.3.14 and, again, restricting to one photon per input mode, leaves us with

$$|\psi_{\mathrm{out}}\rangle = \left[\mathrm{M}_{12}\mathrm{M}_{21}\hat{A}_2^\dagger(t_0)\hat{A}_1^\dagger(t_0+\tau) + \mathrm{M}_{11}\mathrm{M}_{22}\hat{A}_1^\dagger(t_0)\hat{A}_2^\dagger(t_0+\tau)\right]|00\rangle\,. \tag{3.15}$$

This leads to the probability

$$\mathrm{P}(t_0, t_0+\tau) = |\mathrm{M}_{12}\mathrm{M}_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + \mathrm{M}_{11}\mathrm{M}_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2, \tag{3.16}$$

where we see we have recovered the result from Ref. [117]. In order to extend this further, we model mixedness by adding a non-interfering term

$$\mathrm{P}(t_0, t_0+\tau) = \cos^2\theta\,\mathrm{P}_{\mathrm{int}} + \sin^2\theta\,\mathrm{P}_{\mathrm{mix}}, \tag{3.17}$$

where we have relabelled the probability in Eq.3.16 as $P_{int}$ and introduced the non-interfering probability

$$P_{mix}(t_0, t_0+\tau) = |M_{12}M_{21}\zeta_2(t_0)\zeta_1(t_0+\tau)|^2 + |M_{11}M_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2. \quad (3.18)$$

The parameter $\theta$ in Eq.3.17 quantifies the amount of mixedness in the photons. It can be extracted from the visibility of a Hong-Ou-Mandel fringe through

$$\theta = 2\cos^{-1}\left(\frac{3\mathcal{V}-1}{1+\mathcal{V}}\right). \quad (3.19)$$

To complete this step we need to include the probability that this state is generated and reaches the detectors. The probability of generating $n$ photon pairs from one two-mode squeezed vacuum (TMSV) is given by

$$P_{TMSV}(n) = \text{sech}^2|\xi| \tanh^{2n}|\xi|, \quad (3.20)$$

where $\xi$ is the complex squeeze parameter. We define the transmission of the $i^{th}$ spatial mode as $\eta_i$. This includes the waveguide loss, coupling loss and detector efficiency. The total probability is therefore given by

$$P_{total}(t_0, t_0+\tau) = P_{gen} \times P_{det} \times P(t_0, t_0+\tau), \quad (3.21)$$

where $P_{det} = \eta_1\eta_2$ and $P_{gen} = \text{sech}^2|\xi_1| \tanh^2|\xi_1| \times \text{sech}^2|\xi_2| \tanh^2|\xi_2|$, which allows us to include differences in the squeezing parameter of the two photon source.

### 3.3.2 *Modelling double emissions*

In the later experiments we employ spontaneous sources which generate two-mode squeezed vacuum where, in the regime of low squeezing, we approximate one photon in each of the modes. A predominant source of noise with these sources is the presence of higher photon number terms. Double emissions could affect an experiment in two ways. Firstly, one source could produce two pairs of photons and the second source produce none. Resulting in both photons detected at the output of the beam-splitter arising from one source. However, as in the later experiments we herald the input photon state and post-select on a photon being heralded from each source, this situation

is eliminated. The second, and more relevant, case involves one source producing two pairs and the second source producing one. A full derivation of both cases is provided in Appendix C, however here we only consider the second case. Figure 3.2(a) shows all the possible ways that three photons entering a balanced beam-splitter can lead to a coincidence measurement. We have coloured the photons to indicate that the sources can have, in principle, different central frequencies. Detectors with a finite dead-time,

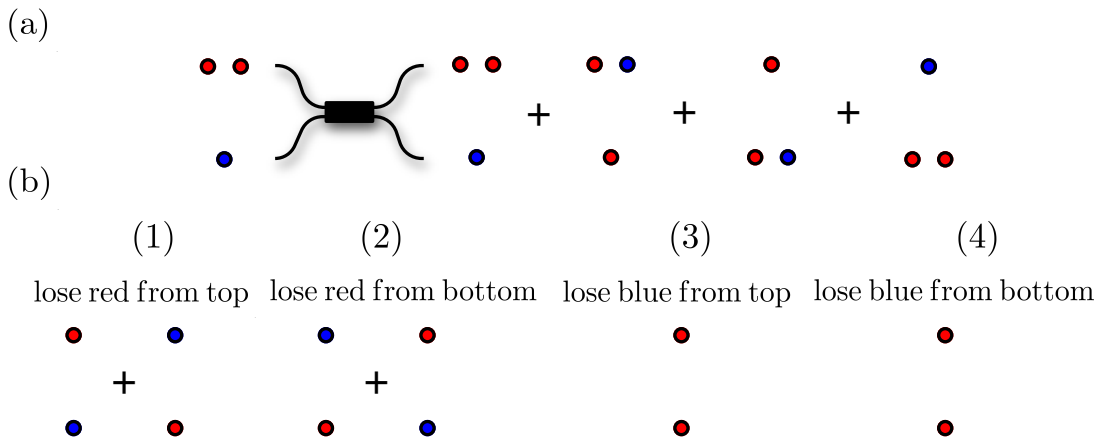


Figure 3.2: (a) illustrating all possible combinations of three photons after a beam-splitter that result in a coincidence click. (b) showing the four different loss possibilities and the resulting state in the coincidence subspace

much longer than the photon's coherence time, will only ever measure two of these photons, therefore, we need only consider the cases where two of the three photons make it to the detectors. In Fig. 3.2(b) we illustrate the possible patterns that arise if one of the three photons is lost. We note that the identical situation is also true for the blue source producing two photons. The total probability including both of these cases as well as the desired interference signal we derived in Eq. 3.16 is given by

$$P_f = A|M_{12}M_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + M_{11}M_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2 + B|\zeta_1(t_0)\zeta_1(t_0+\tau)|^2 + C|\zeta_2(t_0)\zeta_2(t_0+\tau)|^2, \quad (3.22)$$

where

$$\begin{aligned} A = \operatorname{sech}^2\xi_1 \tanh^2\xi_1 \operatorname{sech}^2\xi_2 \tanh^2\xi_2 \eta_1\eta_2 \Big[ 1 + 4\tanh^2\xi_1 \left(|M_{11}|^2(1-\eta_1) + |M_{21}|^2(1-\eta_2)\right) \\ + 4\tanh^2\xi_2 \left(|M_{22}|^2(1-\eta_2) + |M_{12}|^2(1-\eta_1)\right) \Big], \\ B = 4\operatorname{sech}^2\xi_1 \tanh^4\xi_1 \operatorname{sech}^2\xi_2 \tanh^2\xi_2 \eta_1\eta_2 |M_{11}M_{21}|^2 [|M_{12}|^2(1-\eta_1) + |M_{22}|^2(1-\eta_2)], \\ C = 4\operatorname{sech}^2\xi_1 \tanh^2\xi_1 \operatorname{sech}^2\xi_2 \tanh^4\xi_2 |M_{12}M_{22}|^2 [|M_{11}|^2(1-\eta_1) + |M_{21}|^2(1-\eta_2)]. \end{aligned} \tag{3.23}$$

Here, we see that the four possible patterns in Fig. 3.2(b) only contribute in two different ways. If the photon from the single firing source is lost then the remaining two photons produce no fringes and contribute a background with the shape of the envelope of the photon wave packet. This is encompassed in term B for one source and C for the second source. If, however, one of the double emission photons is lost we see that the remaining photons add to the interference signal that we wish to measure. We will see in later sections that it will be possible for us to measure and subtract the terms B and C, leaving us just with the interference term.

## 3.4 EXPERIMENTAL SETUP

In the following section we introduce the experimental components and techniques used in the later experiments. Here we will give a high level overview of the setup and the chip before going into the details about the characterisation of all the components. Figure 3.3 shows the full experimental setup. The results in this chapter are taken using a PriTel wavelength tuneable femtosecond pulsed laser with a repetition rate of 50 MHz emitting pulses centred at 1541.25 nm with a 2 nm bandwidth. As we are using resonant sources with a line-width of around 30 pm we filter pulses from the laser before amplification. This ensures that the majority of the power we are inputting into the chip is in the correct frequency mode. We use a Yenista tuneable filter to reduce the pulse bandwidth to 100 pm. These narrower pulses are amplified using a PriTel Erbium doped fibre amplifier (EDFA). We then use a standard fibre coupled dense wavelength division multiplexing (DWDM) filter to remove any spurious side-bands generated in the amplification process. We also use a polarisation controller to maximise fibre to chip coupling and a 99/1 tap to monitor the input power. To allow us to lock the ring resonances to separate wavelengths we use two Tunics low-power continuous wave

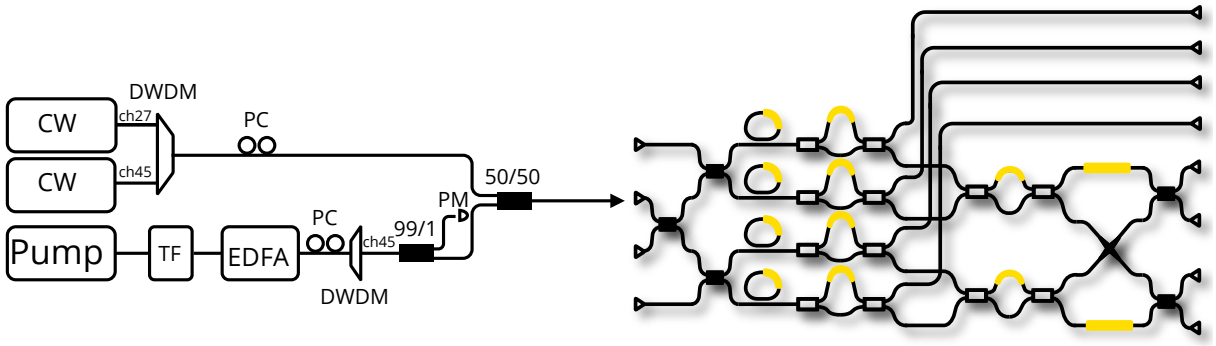


Figure 3.3: Full setup for time-resolved interference experiment. Wavelength tuneable laser from PriTel with a 50 MHz repetition rate and 2 nm pulse width is filtered with a tuneable filter (TF) to 100 pm and then amplified by an erbium doped fibre amplifier (EDFA). This is combined with two CW lasers, used for locking and characterisation, on a balanced beam-splitter. This is then input into a silicon chip where it is evenly split and used to pump four resonant sources. Heralds are directly coupled off chip with grating couplers, signals are input to a reconfigurable interferometer before being filtered and detected off chip.

(CW) lasers at ITU channels 27 (1555.75 nm) and 45 (1541.35 nm), respectively. These two lasers are combined using a 16-channel DWDM. Again, polarisation control on the output of this filter is used to maximise coupling to chip. We then combine both the CW lasers and the pulsed pump laser on a 50/50 beam-splitter which allows us to input all 3 lasers into one grating coupler. On the chip, all light is guided in single mode waveguides with dimensions 500 nm$\times$220 nm. The inputted lasers are split evenly into 4 by balanced MMI couplers and then arrive at 4 microring resonator sources. Here the strong pump pulse is converted to two-mode squeezed vacuum through spontaneous four-wave mixing. We collect the signal and idler photons two resonances away from the pump, 1536.6 nm and 1546.1 nm, respectively. The signal and idlers are spatially separated on chip using AMZI filters (see Sec.2.5.1). The signals are coupled off-chip directly whereas the idlers enter an integrated interferometer before leaving the chip via grating couplers. The pump laser is rejected off-chip using DWDM filters with $\geq$100 dB extinction [14]. Photons are then detected by high efficiency, and most importantly for this experiment, low jitter superconducting nanowire detectors (SNSPD) from PhotonSpot. The voltage pulses output from the SNSPDs are time correlated with a Swabian Ultra time tagger. Powermeters are used at the reflect port of the pump

rejection filters. These powermeters can then be used, along with the CW lasers, to characterise heaters and lock the ring resonances.

### 3.4.1 *Integrated ring resonator sources*

In this work, we use resonant photon sources based on microring resonators. Microring resonators are travelling wave resonators which are made by looping a waveguide back on itself. These offer several benefits over spiralled waveguide sources, namely the small footprint and the low power consumption due to the field enhancement of the resonator. For our purpose they also have the added benefit of generating narrowband photons with a central frequency that can be easily altered by adding a phase in the ring. Light is coupled in, usually by means of an evanescent coupler, only when the round trip of the ring imparts a $2\pi n$ phase shift on the light, where $n \in \mathbb{Z}^+$ - i.e. the round trip length of the ring is exactly an integer number of wavelengths. The frequency spacing between the supported resonances is known as the free spectral range (FSR). In theory, then, the spectral response of a ring resonator should be a series of delta functions at the supported resonances. In practise the inescapable presence of loss in the ring results in resonances with a Lorentzian shape, due to the finite lifetime of the resonator. The ring resonator geometry of interest to us is illustrated in Fig. 3.4(a). Having only a single bus waveguide, these are known as all-pass resonators. The amplitude of the electric field at the output of the bus waveguide is given by

$$\mathrm{E_{out}} = \mathrm{E_{in}} \frac{-\alpha + te^{-i\phi}}{-\alpha t^* + e^{i\phi}}, \tag{3.24}$$

where $E_{in}$ is the input electric field amplitude, $\alpha$ is the round trip transmission and $\phi$ is the round trip phase. As stated above, on resonance the round trip phase, $\phi$, must be equal to an integer multiple of $2\pi$. Inserting this into Eq. 3.24, the numerator becomes $t - \alpha$. From this, we can see that the loss in the ring allows us to define 3 operating regimes for the cavity, depending on the relationship between the round trip loss of the ring and the coupling coefficient into the ring. The main observation is that when the coupling into the ring is equal to the round trip loss, equivalently $t = \alpha$, the amplitude of the transmitted light goes to 0. This is a result of interference between the light transmitted by the coupler and the light is in the ring. As the phase relationship is given by the round trip phase of the ring and the phase of the coupler, this condition

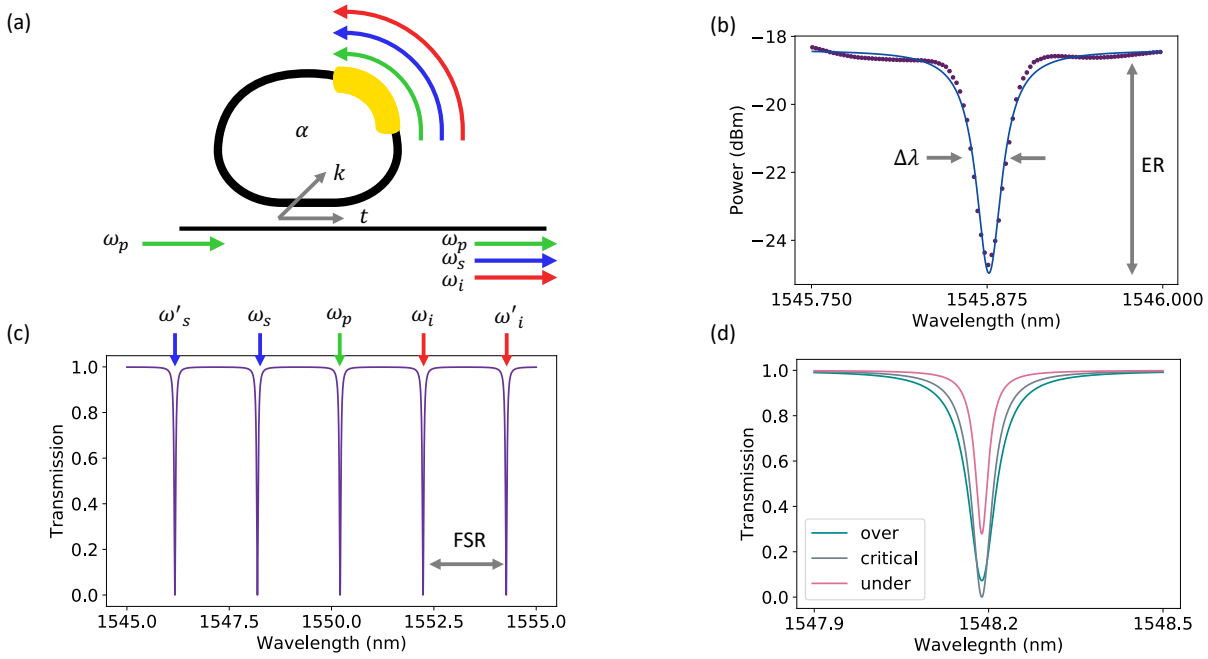


Figure 3.4: Operational principles of ring resonator photon sources. a) Schematic of an all-pass ring resonator as used in the following work. A looped back waveguide with a bus waveguide coupled evanescently. Resonance position is tuned by a thermo-optic phase shifter. b) Measured frequency response of a ring resonator fitted with a Lorentzian line-width ($\Delta\lambda$) and extinction ratio (ER) illustrated. c) Simulated frequency response over many resonances using Eq. 3.24, free spectral range (FSR) shown. A strong pump beam resonant at $\omega_p$ generates pairs of photons on symmetric resonances around the pump. d) Single simulated resonance showing three different coupling regimes. ER maximised when critically coupled but resonance width continues to decrease as coupling into the ring decreases.

matches the amplitude of the transmitted light to that of light in the ring to give perfect constructive interference in the ring and destructive interference in the bus waveguide. This regime, known as critical coupling, maximises the field inside the ring. While advantageous for some applications [118, 119], for photon sources it is not optimal. This is because as the coupling is equal to the ring transmission, any photon generated in the ring has an equal probability to be coupled out or lost, fundamentally limiting the heralding efficiency of a source to 50%. The other two operation regimes, where the either $t < \alpha$ or $t > \alpha$, are named over coupled and under coupled, respectively. For a photon source, the desired regime is slightly over-coupled. The degree of over-coupling is a balancing act between increasing the heralding efficiency of the source, while

maintaining a high enough resonant enhancement of light in the ring. Fig 3.4(d) shows example resonances for these three regimes. We see that the highest extinction ratio (ER), defined as the ratio between on and off resonance powers, is achieved in the critical coupling regime. The ER decreases the coupling coefficient moves away from the round trip loss. The resonance width, however, continues to decrease as the coupling into the ring decreases. This is because the lower the input/output coupling the longer the lifetime of the resonator.

### Classical characterisation

Initially, we classically characterise the response of each of the four rings. This is achieved by sweeping the wavelength of a tuneable CW laser and monitoring the output power with a powermeter. An example ring resonance is shown in Fig 3.4(b), fitted with a Lorentzian. We perform similar analysis on the other rings and have collated the results in table 3.1. We see for all four rings that the FSR is the same to three significant

| Ring | FSR (nm) | $\Delta\lambda$ (pm) | ER (dB) |
|---|---|---|---|
| R1 | 2.27±0.01 | 29.14±1.13 | 3.45±0.02 |
| R2 | 2.27±0.01 | 29.12±1.09 | 3.73±0.02 |
| R3 | 2.27±0.01 | 32.88±1.05 | 7.46±0.03 |
| R4 | 2.27±0.01 | 34.45±1.04 | 12.14±0.05 |

Table 3.1: Table showing the measured ring parameters for the 4 sources employed in this work. FSR is the free spectral range of the resonator. $\Delta\lambda$ is the resonance line-width and ER is the extinction ratio.

figures, this is expected as the FSR depends on the group velocity and path length, both of which are relatively robust to fabrication tolerances. The extinction ratio and line-width vary more as they are dependent on the coupling between the bus waveguide and the ring. Here we use evanescent couplers, which are sensitive to variation in the coupling gap arising from imperfect fabrication. We note here that the resonance width measurements are sensitive to instability in the laser wavelength.

### Quantum characterisation

Now that we have performed the classical characterisation of the resonators and confirmed that they are sufficiently similar, we now turn our attention to how they

perform as photon sources. The roles of energy and momentum in four-wave mixing processes in a resonant source are flipped compared to waveguide sources. A waveguide supports a continuum of wavelengths and the possible four-wave mixing processes are defined by the phase-matching of the waveguide. For ring resonators, all resonances are in phase with each other so the phase-matching condition is always fulfilled and the allowed processes are defined by the supported wavelengths, given by the ring's resonant structure. As shown in Fig. 3.4(c), photons are generated, in pairs, in resonances that are symmetrically spaced in frequency around the pump. In general, the FSR of a ring is wavelength dependent but for small wavelength ranges where the dispersion is roughly flat, the difference in neighbouring FSR's is much less than the resonance width. To achieve proper rejection of the pump, we aim to collect photons at least two DWDM channels away from the pump. In this case, with an FSR of around 2.5 nm, we collect three channels away, which corresponds to two FSR's. We choose the signal and idler to be centered on channels 39 and 51 of the ITU grid (1546.1 nm and 1536.6 nm).

To initially characterise the sources, we will first look at the second order correlation function. Denoted as $g^{(2)}(t)$, this is a measure of intensity (or photon number) correlations in a state of light. Correlation functions have a rich history and have many interesting and useful properties which can be read about in Refs. [120, 121]. For us, the important feature of second order correlation functions is their ability to discriminate between states of light that can be described classically and those which require a quantum description. Classically, it can be shown that the second order correlation should be bounded between 1 and 2, therefore any states outside of this bound cannot be described solely with a classical theory of waves. From this bound, we can define three interesting regimes. Those with $g^{(2)}(t) = 1$ have photon number statistics that obey a Poissonian distribution, that is they arrive randomly. Those with $g^{(2)}(0) > 1$ have super-Poissonian or bunched photon statistics - measuring one photon increases the probability of measuring a second photon, simultaneously. The coherent state and thermal state are examples of states with $g^{(2)}(t) = 1$ and $g^{(2)}(0) = 2$, respectively. The third regime are states with $g^{(2)}(0) < 1$ which have sub-Poissonian statistics, that is detecting one photon decreases the chance of detecting a second photon simultaneously. This is minimised for single Fock state where $g^{(2)}(0) = 0$[2]. The second order correlation

2 This should make intuitive sense as with a true single photon state once the photon is measured, there are no more photons to detect.

function, at 0 time delay, is easily accessible through experiment. The state that we wish to measure is sent to a balanced beam-splitter and coincidence counts between the arms are recorded. We are interested in the squeezing of our source, which is itself directly related to how well our source can approximate a single photon. We therefore measure the $g^{(2)}(0)$ of the signal arm in a heralded manner, using a heralding click on the idler mode to gate the coincidence measurement at the beam-splitter. The $g^{(2)}(0)$ can be calculated as [56]

$$g^{(2)}(0) = C_h \frac{C_{h,1,2}}{C_{h,1}C_{h,2}}, \tag{3.25}$$

where $C_{i,..,k}$ represents a coincidence count between the outputs in the subscript and subscripts $h$, 1 and 2 represent the herald and both beam-splitter outputs. In the end, the quantity we wish to calculate is the squeezing of the source. This should be related to the value of $g^{(2)}(0)$. Heralding on one arm of the source projects out the vacuum term of the two-mode squeezed vacuum state, the higher order Fock terms have the effect of increasing $g^{(2)}(0)$. The amplitude of these terms is directly related to the squeezing parameter. In fact, the squeezing of the source can be calculated from [56]

$$\lambda^2 = \frac{\eta_h}{2\left(1-(1-\eta_h)^2\right)} g^{(2)}(0), \tag{3.26}$$

where $\lambda = \tanh|\xi|$ and $\eta_h$ is the heralding efficiency, calculated as the ratio of coincidences between signal and heralding arms and the single counts in the signal arm. Due to the beam-splitter we are required to include coincidences and singles from both beam-splitter outputs

$$\eta_h = \frac{C_{h,1}+C_{h,2}}{C_1+C_2}. \tag{3.27}$$

This quantity gives the probability of detecting a signal photon given a herald photon is detected. Figure 3.5(a) shows $\lambda^2$ vs power input to the chip. We note here that the power plotted on the x-axis is the average power measured before being coupled into the chip. We see that the squeezing for all four sources are similar. In Fig. 3.5(b) we have plotted the pair generation probability which for low squeezing is $\approx \lambda^2$ as we move the resonance position of the rings across the pump bandwidth. The bold lines show the measured position of the resonance, the shaded lines show the position we set the locking laser to. This discrepancy is probably due to instabilities of the CW laser we used. Another consequence of the laser instabilities is the laser wavelength calibration

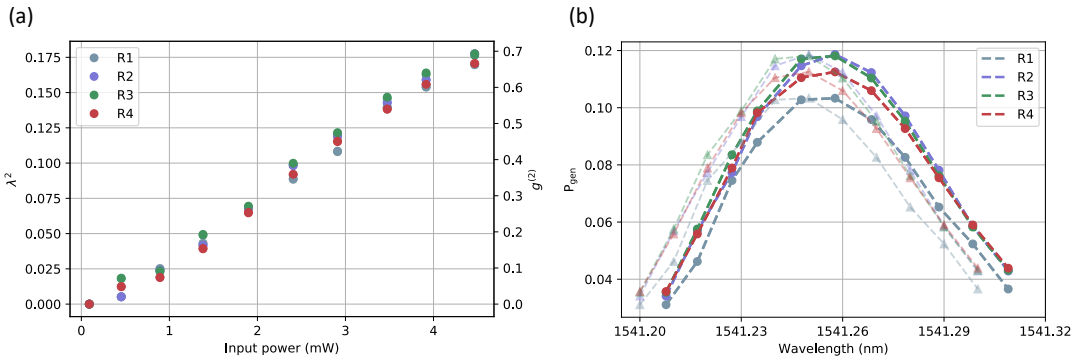


Figure 3.5: Characterisation of ring resonator photon source. a) Squeezing vs input power. Here we plot $\lambda^2$, where $\lambda = \tanh|\xi|$, and $g^{(2)}$ on the same axis. We also see that the squeezing is approximately the same for the different sources. b) Bandwidth of the pump pulse measured by the squeezing from the source. Resonant wavelength of rings is swept via the internal phase shifter and the pair generation probability at the output is measured. We plot the pair generation probability from each source as a function of the resonant wavelength of the rings. The bold plots show the actual measured wavelength of the rings, the lighter lines show the wavelengths that were set in the locking protocol. We see a slight discrepancy based on the stability of the laser. For both plots Poissonian error bars are smaller than the symbol size.

drifted over time, therefore we would perform a sweep similar to Fig 3.5(b) before each measurement. This would allow us to correctly select which wavelengths we should lock the rings to in order to balance the squeezing. We note here that in this chapter we only use the first three sources, we have included the fourth for completeness.

### 3.4.2 *Locking ring resonances*

The ability to tune the resonance position thermally also means that thermal fluctuations in the lab can cause the ring resonances to drift. To combat this the chip rig is attached to a Peltier and a PID temperature controller to keep the temperature of the chip stable. We also enclose the chip rig in a box in order to minimise changes in air temperature from affecting this chip. This works to keep the chip stable on short timescales (this is evident from the smoothness of the resonance dip in Fig. 3.4(b)), however, for timescales longer than a few hours some drift can still be noticed. Also, crosstalk on the chip makes it hard to fix many rings to desired resonance wavelengths. To combat both of these problems we introduced a method of actively locking the ring resonances. By fixing the CW laser at the frequency we wish to lock to, we can define a cost function

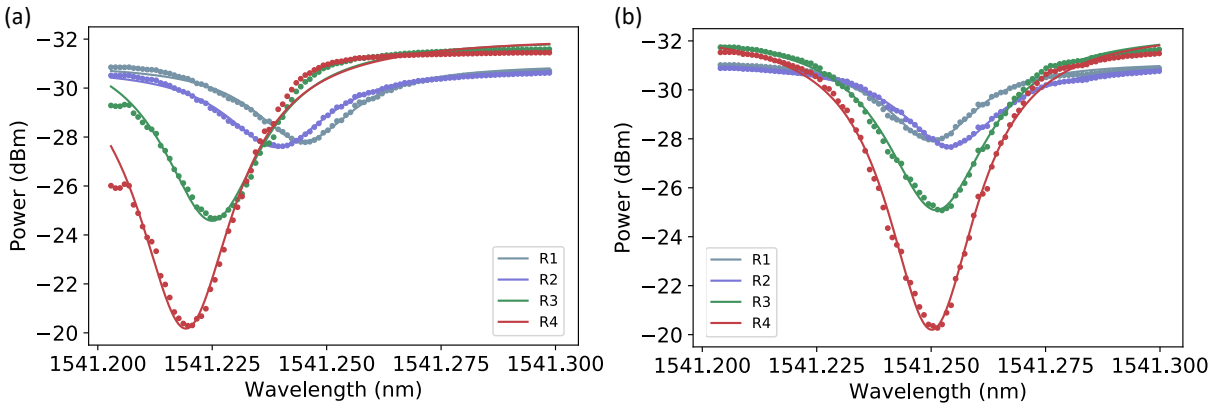


Figure 3.6: Effect of ring locking. Locking wavelength set to 1541.25 nm(a) ring resonances without locking. Here each ring is set separately by setting the voltage that minimises the transmitted power. Standard deviation of ring positions is 10.5 pm (b) rings locked with Nelder-Mead optimisation algorithm. Starting voltages selected from Gaussian distributions centred at the voltages used in (a). Standard deviation of ring positions is 1.6 pm

to be the sum of all powers transmitted by the rings we wish to lock. Using the ring heater voltages as the system parameters we can use the built-in optimisation functions in the python package Numpy to minimise our cost function. When we wish to lock many rings to one wavelength then we only use one CW laser and we lock the pump resonances together. We use the Nelder-Mead optimisation algorithm, as it is a simple algorithm and does not require calculation of the derivative of the cost function. The input settings were chosen to maximise accuracy while still giving the algorithm a good chance to converge in the given number of steps. The starting voltages were sampled from a Gaussian distribution centred at the voltage found by minimising the transmitted power for each ring individually. The standard deviation of the distribution was again chosen to be wide enough to allow restarts to escape any local minima but narrow enough to make sure the algorithm would converge. If the algorithm didn't converge it was restarted with new initial voltages. Once the algorithm had converged, the position of the ring's resonances was measured and checked against some user defined criteria, based on the location and detuning of the rings. If these failed the optimisation was restarted. The effectiveness of this strategy is shown in Fig. 3.6, here we compare the situation where we align each ring to the CW laser independently, by simply choosing the voltage value that minimises the light transmitted through each ring individually, and the case where we use the locking mechanism described above.

### 3.4.3 *Chip characterisation*

In this section we look at the characterisation of the components used in this experiment. We first look at the on-chip linear optics, before turning to the off-chip filters and detectors.

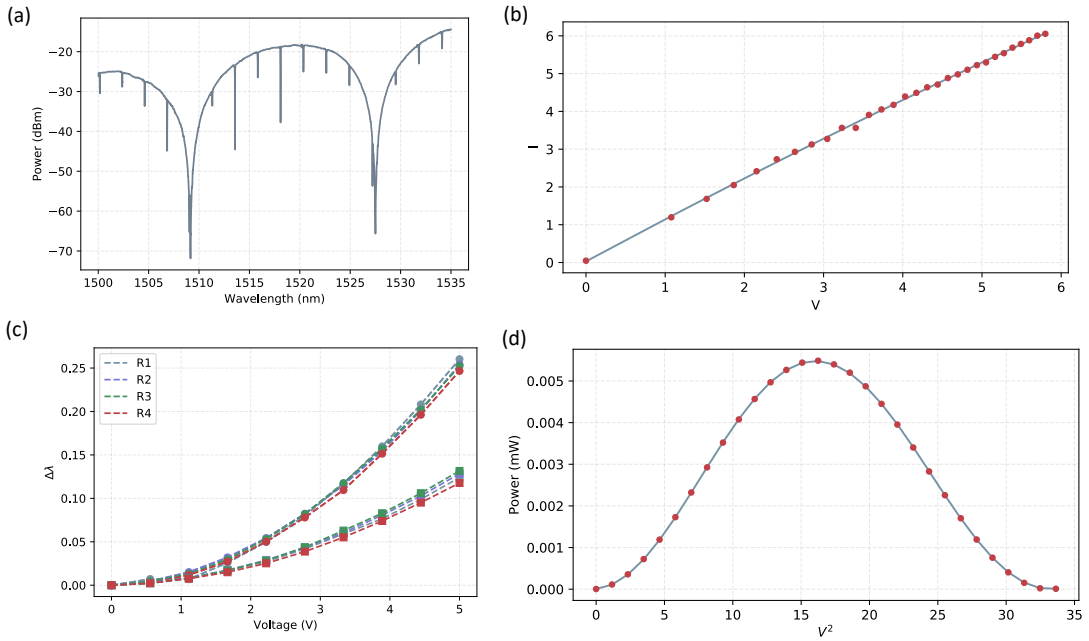


Figure 3.7: Characterisation of on chip linear optics. (a) Frequency scan showing the AMZI and ring response. AMZI FSR is twice the signal and idler separation. Slope is due to the finite grating coupler bandwidth. (b) and (d) Characterisation of a thermo-optic phase shifter. (b) IV curve used to convert set voltage to a power, up to quadratic terms are included to account for non-Ohmic behaviour. (d) Power fringe allows a voltage to be mapped to a transmission. (c) Measure of heater crosstalk on the ring positions. $\Delta\lambda$ is the difference between the ring resonance wavelength compared to its 0 voltage position. The voltage indicated on the x-axis is applied to other sets of heaters, with squares only setting the ring heaters and circles showing the effect of the filter and interferometer heaters.

#### 3.4.3.1 *On-chip components*

Beyond the resonant sources, which we looked at in the previous section, the chip uses linear optical components to filter and interfere the light. For all control of the photons we use thermo-optic phase shifters. These need to be appropriately calibrated to allow us to set the phase we desire. To do this we first measure the current-voltage curve of

the heater, fitting with a second order polynomial to account for non-Ohmic effects, shown in Fig 3.7(b). This allows us to accurately map from the voltage to a power that is dissipated. We then scan an interference fringe using the heater, as shown in Fig. 3.7(d). For heaters inside MZIs this is trivial, for external phase shifters we use other already calibrated MZI phases to implement larger MZIs where the external phase now becomes internal. The fit of this fringe, along with the IV fit, allows us to convert the voltage set into a phase. By inverting this function we can convert from a desired phase into the correct voltage to set. To split the signal and idler photon we use AMZI filters. In order to split the photons correctly, the FSR of the AMZI needs to be twice the frequency separation between the signal and idler photons. For our setup this means the FSR of the AMZI should be 20 nm. Figure 3.7(a) shows the AMZI and ring responses and we can see that the FSR matches correctly. As we are using thermo-optic phase

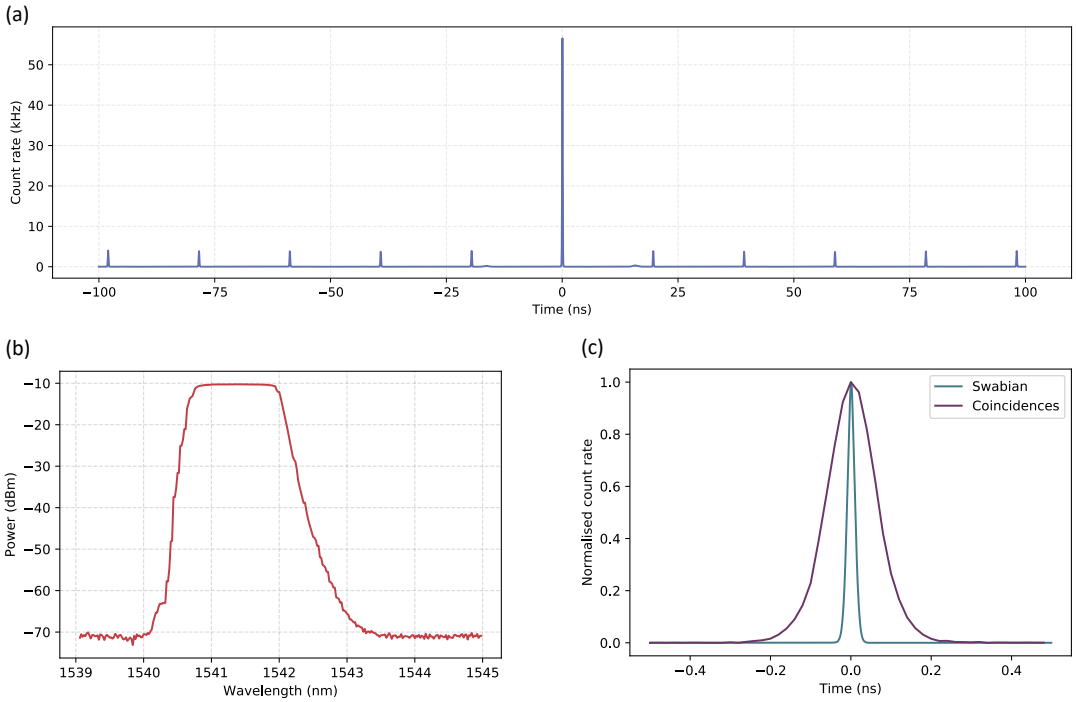


Figure 3.8: Characterisation of off chip components. (a) shows a typical coincidence histogram trace. Central peak corresponds to photons generated in the same pulse. Smaller pulses are coincidences from different pulses. (b) Spectral response of off-chip DWDM. Scanned with a tuneable laser. (c) Histogram using the test signal function on the Swabian to measure the time tagger jitter. Compared to a full coincidence histogram which includes time tagger and detector jitters along with the temporal shape of the photons

shifters, which are positioned 2 µm above the waveguide and radiate power through the cladding equally in all directions, setting phases in one area of the chip can affect components located somewhere else. This is especially relevant for components that are very phase sensitive, such as ring resonators. To quantify this, we look at the position of the ring resonances as we set voltages to the other heaters, shown in Fig 3.7(c). We separate the effect of the other heaters in the other ring resonators (squares) and those in the filters and interferometer (circles). We plot the position of the resonances with a voltage set to the heaters on relative to the position while no voltage is set. For each set of heaters we set all heaters to be the same voltage. From this, we see that the crosstalk affects each ring similarly and that the effect of the filter and interferometer heaters is larger than the ring resonator heaters. The ring resonator contribution shows the need for the locking described above. The filter and interferometer contribution illustrates that we need to re-lock the ring locations after changing phase settings on the chip.

#### 3.4.3.2 *Off-chip components*

The following section describes the more important off chip components used in the rest of this chapter. We use fibre-coupled dense division wavelength multiplexing (DWDM) filters at many points in the setup. These are used to filter any spurious side-bands from the amplification and also to provide the necessary pump rejection before sending the single photons to the detectors. A typical response, measured with a tuneable laser and a powermeter is shown in Fig. 3.8(b). The full extinction ratio can not be extracted from this plot as we are below the noise floor of the powermeter. Once the photons have been filtered to reject the pump light they are sent to high efficiency, low jitter, superconducting nanowire single photon detectors. Counting logic is handled by a Swabian time tagger. A typical coincidence histogram is shown in Fig. 3.8(a). This shows a large central peak, corresponding to coincidences between photons generated in the same laser pulse, and smaller side peaks corresponding to coincidences between photons generated in different pulses. The relative height between these peaks is related to the squeezing in the source [110]. Figure 3.8(c) shows a zoom in of the central peak, the width of this peak combines the detector and time tagger jitter, along with the temporal width of the photons, the FWHM of this peak was measured to be 142 ps. To compare this, we show the histogram peak taken using the test signal, generated by a low-jitter internal clock signal, from the time tagger. The width of this peak corresponds to the combined jitter of two Swabian channels and has a FWHM of 22 ps, giving a

FWHM jitter of a single channel of 15 ps. The detector jitter FWHM are measured by the manufacturer to be between 27 ps and 31 ps for the channels used in these experiments. We define the system jitter to be the combined jitter of two signal channel detectors and time tagger channels. This is because for the later measurements only the jitter on the interfering signal channels is of importance. With the above numbers this is calculated to be, on average, 47 ps.

### 3.4.4 *Data taking and background subtraction*

Swabian provides a Python application programming interface (API) to allow Python code to communicate with the time tagger. This has many useful functions but lacks the ability to access the raw time tag data. For the following experiments, it is the absolute arrival time of the photons that is of importance, and so we have to look beyond the provided Python functionality. We use a custom measurement function in C++, with a Python wrapper, in order to save raw time tags in a binary format. We also employ the ability of the Swabian time tagger to create copies of channels at a hardware level. These virtual channels can be delayed relative to the original. If the channels are delayed by a multiple of the repetition rate of the laser we can measure coincidences between photons generated in different pulses. By changing the pattern of virtual channels we use we can access the double emissions from each ring and also the events with no interference. Figure 3.9 illustrates the different patterns we are interested in. Figure 3.9(a) shows the case where all four detected photons are created in the same pulse. This represents the interference signal that we are interested in. To look at the double emissions of each ring, we shift the other herald channel, as shown in Fig. 3.9(b). This corresponds to two signal photons and one herald photon detected in one pulse, with the second herald in the following pulse. This accounts for both signal photons coming from one source. These counts add to terms B and C in the coincidence probability expression including double emissions, Eq. 3.22, and would have the effect of masking the interference but can be subtracted from the signal counts. Another situation that we will make use off is the pattern where the signal photons are generated in separate pulses and are therefore distinguishable in time and do not interfere. This is shown in Fig. 3.9(c).

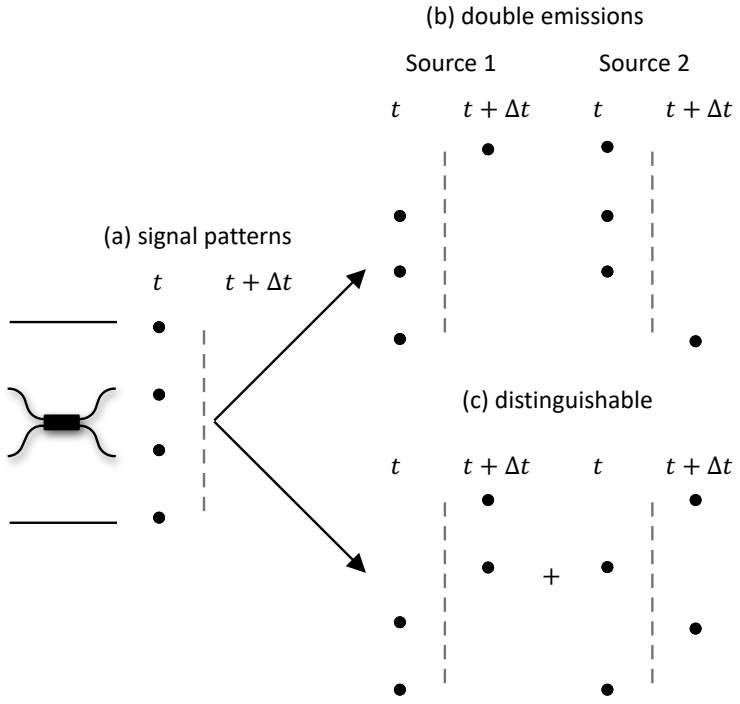


Figure 3.9: Use of virtual channels in order to measure background and distinguishable events. $t$ represents the original channel - i.e. photons generated in one pulse. $t + \Delta t$ represents a copy of a given channel but with all time tags delayed by $\Delta t$, when $\Delta t$ is set to the repetition rate of the laser, we now compare photons generated in the subsequent pulse. (a) All photons generated in the same pulse. These correspond to the interference event we wish to observe. (b) Double emissions. One herald delayed, three photons generated in one pulse - i.e. a double emission from the not delayed source. (c) One pair of photons delayed. Corresponds to interfering photons being generated in subsequent pulses and therefore are distinguishable in time.

### 3.4.5 *Quantum interference characterisation*

In order to characterise the quality of the single photon sources and to demonstrate the effectiveness of our background subtraction technique, we perform a heralded HOM experiment between two of the sources locked to the same emission wavelength. These results, along with the schematic, are shown in Fig. 3.10. We sweep the MZI phase, $\theta$, and measure four-fold coincidences, two herald and two signal photons. To compare fringes with and without background subtraction we normalise each fringe to its maximum counts. For reference, we integrate for 10 minutes per point and the maximum counts per point with and without background subtraction were 527 and

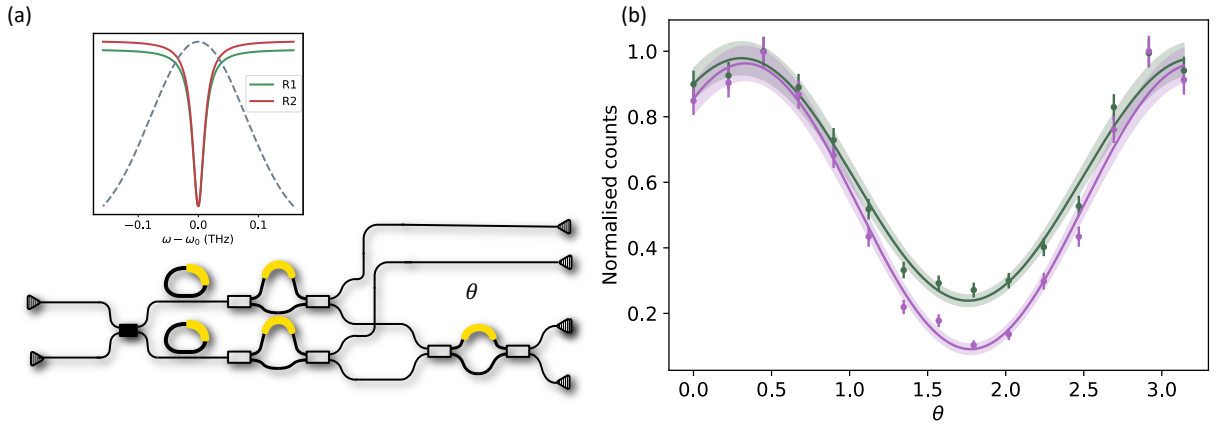


Figure 3.10: Heralded HOM fringe with indistinguishable sources. (a) Schematic for heralded HOM fringe, sweeping angle $\theta$ and measuring four-fold coincidences (two herald and two signal) b) Measured fringe with background (green) and without background (purple) including Poissonian error bars. Shaded area shows the $1\sigma$ interval from the errors on the fitting parameters. Visibility of green curve with no background subtraction $0.608 \pm 0.007$. With background subtraction this increases to $0.828 \pm 0.019$.

426, respectively. The slight shift in the fringe is due to crosstalk from the filters on $\theta$. The fringe visibility without background is $0.828 \pm 0.019$. This maps to a purity of $0.811 \pm 0.023$.

## 3.5 TIME-RESOLVED HOM INTERFERENCE

In this section we will apply the general model, which we derived in sections 3.3.1 and 3.3.2, to the experimental setup described above. We then compare the model to the experimental data.

### 3.5.1 *Simulations*

Previously, we derived a general model for performing two photon interference in a time-resolved manner, and with the inclusion of realistic experimental imperfections. While the full interference effect is described in these equations, it is hard to see what is really going on. In order to visualise the effects we will apply these equations to our experimental setup. Specifically, we will look at a time-resolved HOM fringe with

distinguishable photons from resonant sources. Initially, we need the temporal profiles of the two photons, $\zeta_i$. We obtain this by taking the Fourier transform of the Lorentzian spectral shape of the photons. We require a spectral amplitude function, $l(\omega, \omega_0, \delta\omega)$ such that $|l(\omega, \omega_0, \delta\omega)|^2$ gives us a Lorentzian function peaked at $\omega_0$ with a FWHM $\delta\omega$

$$l(\omega, \omega_0, \delta\omega) = \sqrt{\frac{\delta\omega}{2\pi}} \times \frac{1}{-i\,(\omega - \omega_0) + \frac{\delta\omega}{2}}. \tag{3.28}$$

The general form of Eq. 3.28 matches that in Ref [122], but with an added normalisation

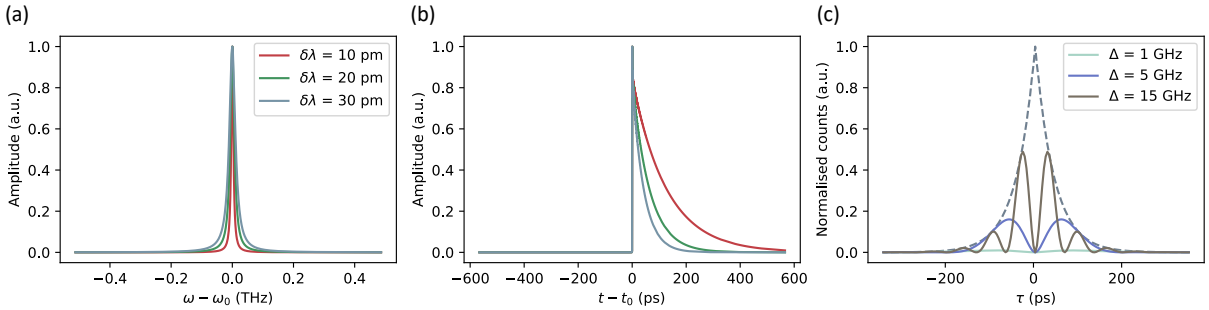


Figure 3.11: (a) Photon spectra for ring resonators with different Q-factors. (b) Corresponding temporal profiles for spectra in (a). (c) Ideal time-resolved fringes for various input photon detuning. Dotted line shows the Lorentzian envelope.

prefactor. The temporal profile, given by the Fourier transform of the spectrum, is a decaying exponential with a decay constant that is proportional to resonance width, $\delta\omega$, as illustrated in Figs 3.11(a) and (b). These temporal profiles can then be input into Eq. 3.13 and, as only relative time differences are important, integrated over $t_0$. The result of this is sinusoidal fringes modulated by a Lorentzian envelope. The frequency of the fringes is equal to the frequency detuning of the photons, shown in Fig.3.11(c). As expected, we see that for $\tau = 0$ the coincidence probability is 0 for all detunings. It is easy to see from this that as the detuning increases, the timing resolution required to capture the 0 probability events increases. This is because the smaller the timing window, the larger the associated uncertainty on the frequency (as per Eq. 3.11). If we are intrinsically uncertain on the frequency of the photon then the possible paths, i.e.

which photon came from which source, are indistinguishable. In general, to interfere photons with a frequency detuning $\Delta$, timing resolution is required such that

$$\delta\tau \ll \frac{1}{|\Delta|}. \tag{3.29}$$

We have seen before how the visibility of a fringe tells us the quality of the quantum interference and Fig. 3.11(c) shows that at the visibility of the fringe can be tuned with the timing window. To map this to the fringe visibility we choose the general transfer

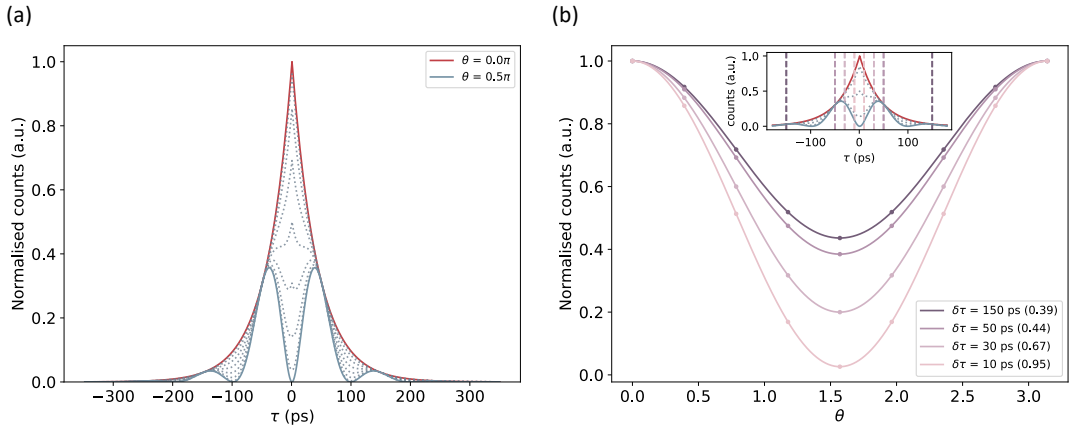


Figure 3.12: Ideal simulation of time-resolved HOM interference. (a) Effect of beam-splitter angle on the time domain fringes for detuning of 80 pm and resonance line-widths of 30 pm. Fringes at $\theta = 0, \frac{\pi}{2}$ have been picked out as these represent the maximum and minimum of the fringe, dotted lines show the transition. (b) HOM fringes for different timing windows, inset shows the intervals that are integrated for each $\delta\tau$. Fringe visibility is shown in brackets in the legend.

matrix, M, to be that of a beam-splitter with tuneable reflectivity, given by[3]

$$\mathrm{M} = \begin{pmatrix} \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \end{pmatrix}. \tag{3.30}$$

We can then vary the angle $\theta$ and see the effect on the time domain fringes, as shown in Fig. 3.12(a). Here we have picked out the fringes representing the maximum and minimum of the HOM fringe. We have illustrated the transition with dotted lines. We see the interference fringe at $\theta = \pi/2$, where $M$ implements a balanced beam-splitter, and the Lorentzian shape representing no interference at $\theta = 0$, where $M$ implements

3 This is equivalent, up to a global phase, to the transfer matrix of the MZI that we use on the chip

a swap. Qualitatively, the transition shows the minimum point rising up eventually becoming the peak of the Lorentzian. In Fig. 3.12(b) we have plotted HOM fringes with differing timing windows. As expected the visibility of the fringes increases as we decrease the timing window. The inset shows the same plot as Fig. 3.12(a) where we have added colour matched lines showing the intervals taken for each fringe. For Fig. 3.12 we have assumed the ideal case with no experimental imperfections. We will now look at the effect of adding in the most common experimental imperfections, as described in Secs 3.3.1 and 3.3.2. In Fig. 3.13 we have separated the effects of double

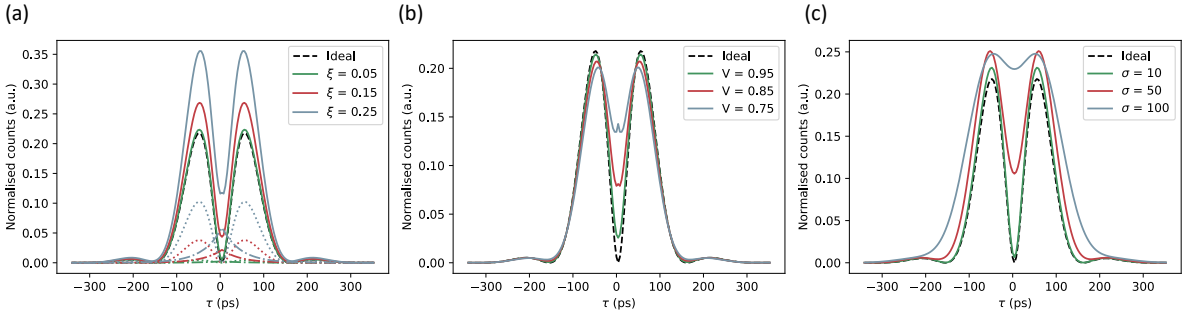


Figure 3.13: Effect of experimental imperfections. In all three plots, a dashed black line shows the ideal fringe with no imperfections. (a) Double emissions. Solid line shows full signal including double emissions. Dotted shows the double emissions that contribute to the interference. Dot-dash line shows the double emissions that do not interfere. (b) Mixedness, lines plotted for different HOM fringe visibilities. (c) Detector and electronics jitter. Signal is convolved with a Gaussian detector response with a FWHM shown in the legend.

emissions, mixedness and timing jitter. Figure 3.13(a) we address double emissions. As described in Sec 3.3.2, the double emissions contributed to the signal in two distinct ways. If the lost photon is one from the double emission then some coherence is retained and the remaining photons do interfere. The second way, when the lost photon is not from the double emission, does not contribute to the interference. These two situations are shown in Fig. 3.13(a) as dotted and dash-dotted line, respectively. We also show the total signal, we can see that the first situation increases the height of the fringe lobes and the second raises the minimum of the interference fringe[4]. Figure 3.13(b) shows the effect of impurity of the interfering photons. We see that increasing the mixedness of the input photon state increases the minimum point of the fringe. This is because the non-interfering term we introduced adds a Lorentzian background.

4 Unsurprisingly, this is a common trend with the experimental imperfections

Finally, and perhaps most importantly, we look at the effect of system timing jitter. Both the detectors and timing electronics used in an experiment have some unavoidable timing uncertainty whether it be from the length of the detectors or from crosstalk in the electronics. To model the effect of this uncertainty we convolve the signal with the temporal response of the detector system, assumed here to be a Gaussian with a standard deviation given by

$$\sigma_f = \sqrt{\sigma_{d,1}^2 + \sigma_{d,2}^2 + \sigma_{t,1}^2 + \sigma_{t,2}^2}, \tag{3.31}$$

where $\sigma_{d,i}$ is the jitter of the $i^{th}$ detector and $\sigma_{t,j}$ is the jitter of the $j^{th}$ time tagger channel. As shown in Fig. 3.13(c), this convolution has the effect of smearing out the interference fringe. In the next section we combine all of these effects to compare to the output of our experiment.

### 3.5.2 *Experimental results*

In order to robustly check the performance of the model we have developed we now perform the experiment. As seen in Fig. 3.10, using a subset of the components on our chip it is possible to perform heralded HOM interference. In the previous HOM experiment, we used a large coincidence window and photons that were indistinguishable in frequency. We now detune the rings by 55 pm (6.8 GHz), symmetrically about the centre of the pump and measure coincidences in a time-resolved manner, as shown in Fig.3.14(b). Unlike in many photon counting experiments, we will be dividing the coincidence window into finely spaced bins. Thus, to record enough coincidence events, long integration times are needed. To combat the effect of drifts in the coupling, that can occur on these timescales, we integrate for a shorter time on each phase setting and repeat this process until we have achieved the desired total integration time. We limit the integration on each of the nine phase points in the fringe to one hour and repeated the fringe 20 times. The ring resonators are re-locked to the desired wavelengths after each one hour integration to avoid crosstalk from setting $\theta$. During the integration time we record all raw time tags for any four-fold coincidences. In post-processing, we combine all the time tags from all 20 repeats and bin them into histograms. The bin size should be chosen to be less than the system jitter but larger than the timescale over which the detector response is roughly flat. We empirically chose this size to be 20 ps. We create

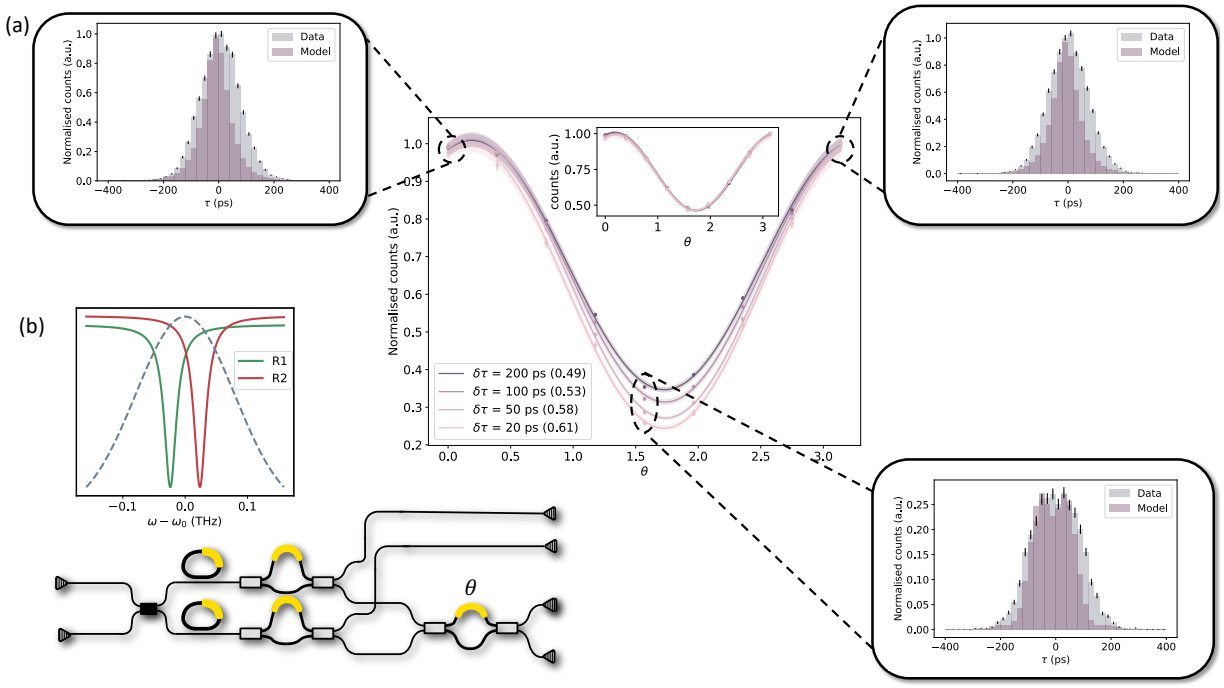


Figure 3.14: Time-resolved HOM fringe. (a) Main figure shows fringe swept on $\theta$ for different timing windows. The visibility for each timing window is shown in the legend and the shaded area around each line shows the $1\sigma$ interval, using the error from the fit. Poissonian error bars are smaller than the symbol size. Inset shows fringe for distinguishable photons (as described in Sec. 3.4.4). We also show some of the histograms and compare to our model. Experimental data includes Poissonian error bars.

histograms for the signal patterns, background for both rings, and distinguishable cases (as described in Sec. 3.4.4). Subtracting the background histograms for both rings from the signal histogram has the effect of removing the double emission Lorentzian that masks the interference. Figure 3.14(a) shows the main results from this experiment. In the centre figure we see the HOM fringes for differing timing windows. We see an increase in the fringe visibility from 0.49 to 0.61 and see visibilities well above the $\frac{1}{3}$ classical limit, illustrating we do have quantum interference of spectrally distinguishable photons. In the inset we show the same fringes but for photons distinguishable in time. Here, we see no change in the visibility of the fringes with timing window and, as expected, the fringes all have a visibility at the classical limit of $\approx \frac{1}{3}$. As with Fig. 3.10, we see a slight offset in the fringe minimum due to crosstalk from the AMZI heaters. We measure this from the fits to be $\Delta\theta = -0.191 \pm 0.013$ rad. and include this in our model. We have also shown a subset of the measured histograms and compared

them to our model. To quantify the distance between the experimental and theoretical histograms we employ the statistical fidelity

$$\mathrm{F}(p,q) = \left(\sum \sqrt{p_i q_i}\right)^2, \tag{3.32}$$

where $p$ and $q$ are the normalised probability distributions for the theory and experimental histograms. We measure an average fidelity of $0.966 \pm 0.005$, showing good agreement between our model and experiment. The deviation from unity fidelity could be explained by several factors. Throughout our modelling we assume that the photons we use are Fourier limited, possessing the narrowest possible temporal profile for a given spectrum. It is possible that the presence of extra Schmidt modes in our photon states, owing to the measured purity being less than one, could temporally broaden the photons. Beyond this, the model is sensitive to changes in the ring parameters. The laser used to measure the ring resonances have limited stability and could have resulted in overestimating the resonance widths.

## 3.6 TIME-RESOLVED COMPLEX HADAMARD PHASE

Having shown the ability of our experimental setup to generate quantum interference between spectrally distinguishable photons, we now turn our attention to controlling the interference within the photon wave packet. To look at this we turn to a specific class of matrices known as the complex Hadamard matrices.

### 3.6.1 *Complex Hadamard matrices*

A Hadamard matrix is defined as an $N \times N$ orthogonal matrix, $H$, where all elements have a modulus of 1, $|H_{ij}| = 1$. We can write the orthogonality condition as

$$HH^\dagger = n\mathbb{I}_N, \tag{3.33}$$

where $n$ is an integer and $\mathbb{I}_N$ represents the $N$ dimension identity matrix [123]. Originally introduced and studied as purely real matrices - i.e $H_{ij} = \{+1, -1\}$ [124], they were later generalised to include complex valued elements in the form $H(q, N)$, where $N$ is the size of the matrix and the elements in the matrix are taken to be the $q^{\text{th}}$ roots of

unity (named Butson-type matrices) [125] and then further to include arbitrary complex elements [124]. The original real valued Hadamard matrices have found applications in classical information processing where they can be used to generate Walsch-Hadamard codes, with uses for transmitting information over noisy channels [126, 127]. The complex generalisation of these matrices are an important class of matrices in quantum information processing, being a core element to teleportation and dense coding protocols [128]. Beyond this, all pairs of mutually unbiased basis (MUB) are linked by a complex Hadamard matrix [129], making the search for MUBs and complex Hadamard matrices in a given dimension, intimately related. We note here that two Hadamard matrices, $H_1$ and $H_2$ are said to be equivalent if they can be transformed into each by

$$H_2 = D_1 P_1 H_1 P_2 D_2, \tag{3.34}$$

where $P_i$ and $D_i$ are permutation and diagonal unitary matrices respectively [5]. Despite their deceptively simple definition, full classification of non-equivalent complex Hadamard matrices has only been completed up to $N = 5$. An important subclass of complex Hadamards are the Fourier matrices, defined as having elements

$$F_{jk}(N) = \frac{1}{\sqrt{N}} e^{\frac{2\pi i j k}{N}}. \tag{3.35}$$

If the Fourier matrix is rescaled by a factor of $\sqrt{N}$ it is a Butson-type matrix of form $H(N, N)$. The Fourier family actually fully defines the complex Hadamards of dimension $N = \{2, 3, 5\}$. For $N = 4$ there are actually an infinite number of non-equivalent complex Hadamard matrices that can be defined by a matrix with a single free parameter, of which the Fourier matrix is one element.

$$F_4(\phi) = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & e^{i\phi} & -1 & -e^{i\phi} \\ 1 & -1 & 1 & -1 \\ 1 & -e^{i\phi} & -1 & e^{i\phi} \end{pmatrix}, \tag{3.36}$$

As we move to dimensions $N > 5$, which non-equivalent families exists is still an open problem in mathematics. However, for every dimension there is a Fourier family,

5 In photonics, this transformation is equivalent to relabeling input and output modes and adding input and output phases

$F_N$, where the Fourier matrix, $F(N)$, either defines the entire family or is an element of the family. In general, any complex Hadamard transformation (up to a scaling factor) can be implemented in a photonic circuit through the decompositions described in Sec. 2.3.2. The Fourier families have a neat mapping to lower depth photonic circuits through a fast Fourier transform algorithm [130]. This consists decomposing, $F_N$ into two smaller Fourier families of size $N/2$, $F^{(1)}_{N/2}$ and $F^{(2)}_{N/2}$. These two transformations can be thought of as acting on the odd and even modes of the larger DFT. These are then recombined by interfering the modes in ascending order - i.e. mode one of $F^{(1)}_{N/2}$ (mode one of $F_N$) is interfered with mode one of $F^{(2)}_{N/2}$ (mode two of $F_N$). Phase control between the top $N/2-1$ is used to tune this interference [6]. Following the decomposition backwards, so long as $N = 2^d$, we will eventually arrive at the $F_2$ matrix - a balanced beam-splitter. This shows that any Fourier matrix with a dimension that is a power of 2 can be decomposed to only balanced beam-splitters, crossings and phase shifters. It is possible to use the same algorithm to decompose Fourier matrices of arbitrary dimension, however this requires beam-splitters with reflectivity other than 50/50 to implement the $F(3)$ transformation [48]. In the following experiments we will make use of an interferometer implementing the transformation in Eq. 3.36, as shown in Fig. 3.15[7]. In Fig. 3.15, we have also illustrated the interesting interference that can occur in Fourier matrices. We can see from the blue and red lines that there are no closed paths within the interferometer. This means that no single photon interference can occur and, therefore, the internal phase has no impact on the single photon statistics. If we now input two photons, the situation changes. If we input photons into different input beam-splitters (for example in Fig. 3.15 we have chosen inputs 1 and 3 but any of the four combinations will have similar result) and look to detect coincidences from different output beam-splitters we now get interference between the red and blue paths. This interference depends on the differences in path lengths between the sum of the red and blue paths. Therefore, by tuning $\phi$ we can tune this interference. This can be seen from the relevant submatrix of Eq.3.36, if we choose inputs 1 and 3 and outputs 1 and 3 then the correct submatrix, accounting for the permutation of inputs and outputs, is

$$F_{\text{sub}} = \frac{1}{2}\begin{pmatrix} e^{i\phi} & 1 \\ 1 & 1 \end{pmatrix}. \tag{3.37}$$

6 These phase shifts correspond to the parameter(s) in the complex Hadamard family
7 This interferometer is equivalent up to permutations of the inputs and outputs

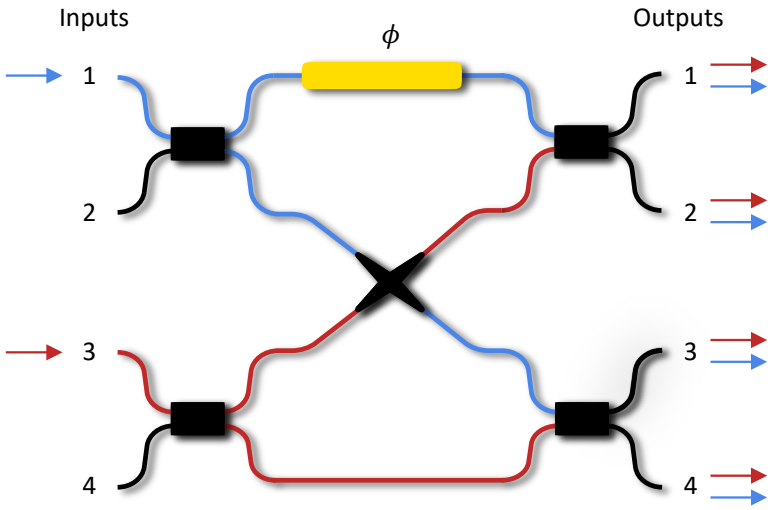


Figure 3.15: Photonic circuit implementing Fourier transform. Angle $\phi$ can be used to continuously tune between all non-equivalent $4 \times 4$ complex Hadamard matrices. Coloured lines show the paths that contribute the internal phase dependent interference.

We then end up with a coincidence probability $P = |\text{Perm}\,(F_{\text{sub}})\,|^2 = \frac{1}{16}|e^{i\phi} + 1|^2$. From this we see that unlike with single photon statistics, the two photon statistics do depend on the internal phase. For three photons we see no dependence on this in internal phase, however for four photons the dependence returns [131]. This can be further generalised to specific $n-$photon suppression laws in $n-$mode Fourier interferometers [132].

### 3.6.2 *Complex Hadamard phase with indistinguishable photons*

For these measurements we employ the same experimental setup as described above, however we now use the first and third rings as our photon sources. A simplified version of the experimental schematic is shown in Fig. 3.16(a). Here we have shown the pump splitting just between rings 1 and 3, however in reality we have static pump splitting. In order to selectively generate photons from specific rings, we use the rings resonant positions and AMZI filters. We sweep a fringe on $\phi$ and integrate for one hour per point. Figure 3.16(b) shows the result of these measurements. We have plotted the four output coincidence patterns that show a dependence on $\phi$. The visibilities of these fringes confirm the results found in Sec. 3.4.5. The inset shows the distinguishable patterns, as expected these show no dependence on $\phi$.

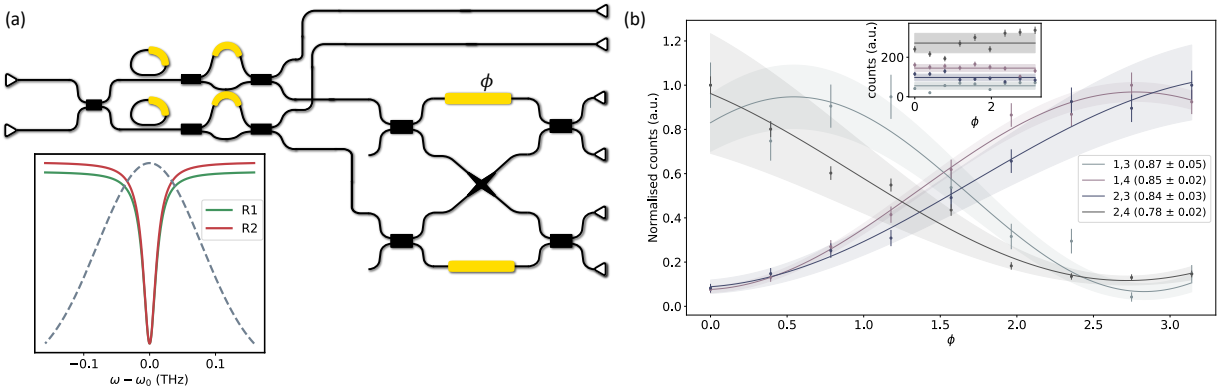


Figure 3.16: Complex Hadamard fringe. (a) Experimental schematic, photons input into inputs one and three. A fringe is swept on $\phi$ and all four output coincidence patterns that show a $\phi$ dependence are measured. (b) Experimental fringes. Error bars assume Poissonian statistics. Shaded area shows $1-\sigma$ interval based on errors on the fit parameters. Fringe visibilities are shown in the caption. The inset shows distinguishable patterns.

### 3.6.3 *Time-resolved simulations*

We now look at this interference effect in a time-resolved manner. In order to find the detection time dependent coincidence probability we replace the transfer matrix in Eq. 3.14 with the relevant submatrix of Eq. 3.36. To illustrate this more explicitly we use Eq. 3.37. This leads to a coincidence probability

$$\mathrm{P}(t_0, t_0+\tau) = \frac{1}{16}|\zeta_2(t_0)\zeta_1(t_0+\tau) + e^{i\phi}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2. \tag{3.38}$$

From this we see the result from Eq.3.13 but with an extra phase factor given by the interferometer internal phase. This has the effect of shifting the minimum of the interference fringe. Figure 3.17(a) shows the result of Eq. 3.38, integrated over all $t_0$. Solid lines show extremes of the interference. Dotted lines show intermediate values of $\phi$. From this we see the effect of $\phi$ moving the interference fringes to the right. Figure 3.17(b) shows the effect of timing window on the visibility of the interference fringes. As expected, increasing the timing window decreases the fringe visibility.

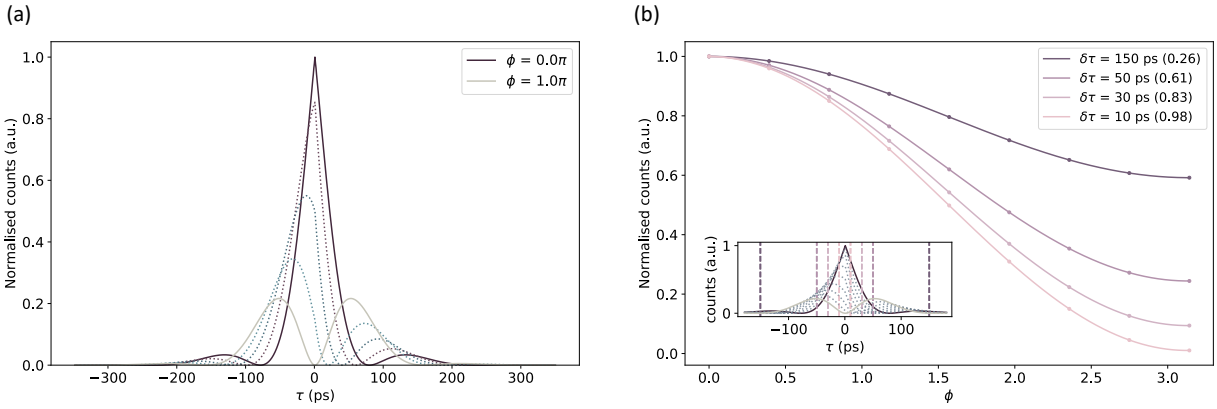


Figure 3.17: Simulated ideal time-resolved complex Hadamard fringe. (a) Time domain fringes, solid lines show $\phi = \{0, \pi\}$. Colour gradient shows change in $\phi$. (b) Complex Hadamard fringes for differing timing windows. Inset shows the intervals used for the different fringes.

### 3.6.4 *Results*

To experimentally probe this interference, we now spectrally detune the two sources by 53 pm (6.7 GHz) and measure in a time-resolved manner. We sweep a five point fringe on $\phi$ and record raw time tags. We integrate for one hour on each point and repeat the sweep 35 times.

These results are shown in Fig. 3.18. Figure 3.18(a) shows the experimental setup. Unlike the heralded HOM measurements where we sum coincidences from all output patterns, here we look at isolated patterns which reduces the count rate by a factor of 4. We choose the pattern with signal photons leaving output modes 1 and 3 (shown with ticks), as this pattern had the lowest output losses. In Fig. 3.18(b) we show the fringe for a variety of coincidence windows and see an increase in visibility from 0.4 to 0.6 The inset shows the fringe for distinguishable patterns, as expected we see no dependence on the phase $\phi$. Figure 3.18(c) compares the experimentally measured histograms with the theoretical model. We see qualitative agreement between the two and we can see the histogram changes as expected with the interference minimum moving off to the left.

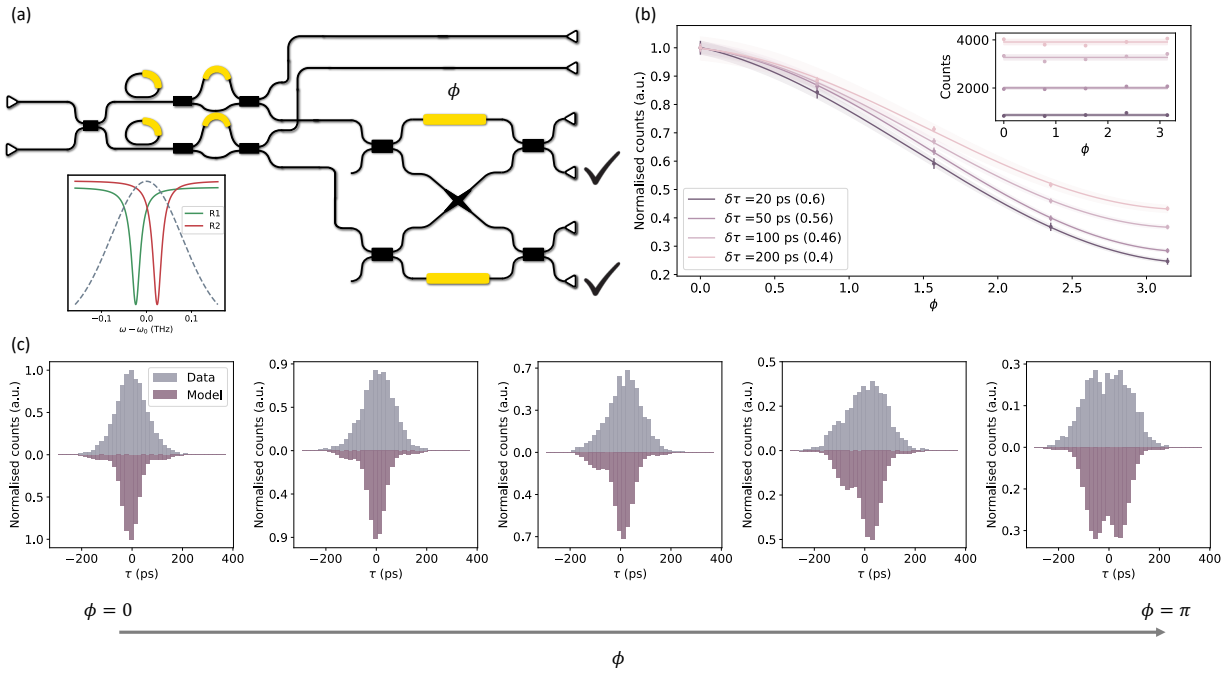


Figure 3.18: Time-resolved complex Hadamard fringe results. (a) Experimental setup, Both input beam-splitters are set to implement balanced beam-splitters and the fringe is swept on $\phi$. We show results for output pattern (1,3) (b) Experimental fringe on $\phi$ for different timing windows. Visibility shown in legend. Inset shows the distinguishable pattern. (c) Histograms for each phase point. We show the experimental histogram and the theory.

## 3.7 DISCUSSION AND OUTLOOK

Most previous experiments on interfering spectrally distinguishable photons used single photon avalanche photodiodes (SPAD) which have intrinsically high jitters ($\approx$ a few hundred picoseconds). The superconducting nanowire single photon detectors used in this work have intrinsic jitters which are an order of magnitude lower than this. In this chapter we have seen how with current, commercially available, photon detectors and timing electronics we can see quantum interference with photons detuned by $\approx$7 GHz. This is a factor of 100 higher than previous experiments [35–37]. We have also shown sufficiently high resolution to allow us to probe interference effects within the coherence times of the photons. All of this has been achieved with standard ring resonator photon sources, generating photons with a bandwidth of several GHz, two orders of magnitude

larger than any previous experiments. This shows the viability of this technique with scalable photon sources.

One of the main factors limiting the visibility of the fringes in the reported experiments is the $\approx 80\%$ spectral purity of the photons. This is likely due to the width of the ring resonances relative to the pump. The maximum photon purity from a standard ring resonator setup, which is bounded at 0.93, is achieved when the pump spectrum is flat across the whole of the ring resonance [122]. It is also possible to push past this limit by manipulating either the pump pulse shape [111, 113], the ring coupling [115, 122] or by adding extra rings to the system [114, 133]. In this chapter we have always modelled impurity in the photon states with a non-interfering term, assuming that the precise timing measurements cannot erase mixedness. This is not strictly true. The coincidence probability for photons with arbitrary spectral purity is given by

$$P_{\text{coinc}}(t_1, t_2) = \sum_{kl} \lambda_k \lambda'_l \left| \zeta_{1k}(t_2)\zeta_{2l}(t_1) - \zeta_{1k}(t_1)\zeta_{2l}(t_2) \right|^2, \tag{3.39}$$

where $\zeta_{ik}$ is the $k^{\text{th}}$ temporal Schmidt mode for the $i^{\text{th}}$ source with $\lambda_k$ ($\lambda'_k$) being the Schmidt coefficients for the first (second) photon state[8]. From this we see that, regardless of the number or shape of the Schmidt modes that for $t_1 = t_2$, $P_{\text{coinc}} = 0$. Full understanding of how the combination or timing jitter and impurity affects the final model would require analysis of at least the first few higher order Schmidt modes. Here we will provide an explanation of why our simplistic description of mixedness describes the data well. Each combination of Schmidt modes will give rise to sinusoidal fringes in the time domain with a frequency set by the difference in central frequencies of the modes. The modulating envelope for each pair will be the convolution of the temporal shapes of the Schmidt modes. The higher order modes have extra lobes that cause broader envelopes but with narrower central peaks. These narrower peaks will be more smeared out by temporal response of the detectors which will add a Lorentzian-like peak to the signal. The size of this peak will be proportional to the weights of the higher order modes and therefore proportional to the mixedness of the photon states.

In terms of our ability to erase frequency information and interfere distinguishable photons, here we are limited by the jitter of our detectors ($\approx$30 ps) and of the timing electronics ($\approx$14 ps). While low for commercially available SNSPDs, this jitter is still an order of magnitude higher than the current state of the art [134]. The timing electronics

8 See Appendix B for a full derivation

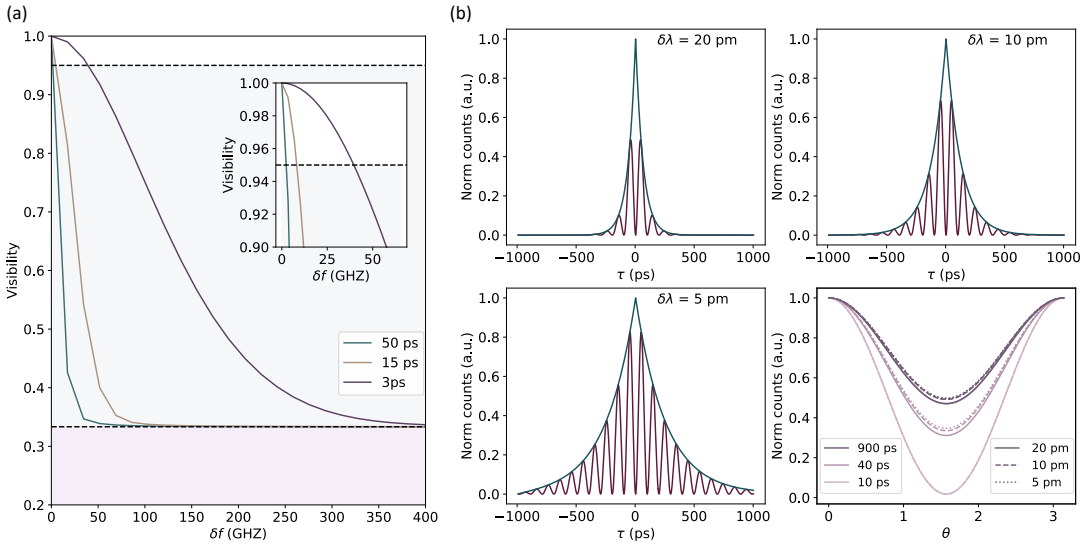


Figure 3.19: (a) HOM fringe visibility as a function of ring detuning for a range of different system jitters. For all fringes we have used a purity of 1 as to isolate only the effect of the timing jitter. We have illustrated three regimes, one comparable to the numbers used in this chapter, an intermediate regime with detector and time tagger assumed to be 10 ps each and the state of the art, assuming all jitter is from the detector and taking the result from [134]. We have added lines for fringe visibility of 0.95 and then at the classical limit of $\frac{1}{3}$. (b) Illustrating the effect of resonance width. Three panels show time domain fringes at $\theta = 0.5$ for different resonance widths. The bottom right panel shows HOM fringes for different timing windows and different resonance widths.

jitter also has room to improve; even new firmware updates can decrease the Swabian jitter below 10 ps.

In order to use this technique in a practical setting, that being a large photonic circuit with many ring resonators sources, we would require that the maximum frequency separation we can erase to be greater than half the FSR of the ring resonators. This would mean that no alignment of the rings would be needed at all. How large the FSR of a ring can be is limited by the resonance width - the FSR needs to be large enough that there is good extinction between neighbouring resonances. Therefore, to increase the FSR we need to decrease the loss in the ring either by engineering the waveguide geometry [135, 136] or changing material [137].

Figure 3.19(a) shows projections for future experiments. Here we have plotted HOM fringe visibility as a function of photon detuning. We have assumed a purity of one, in order to isolate the effect of system timing resolution. We show plots for three scenarios.

The first being currently commercially available technology, similar to what is used in this work. The second scenario we assume that both detector and time tagger jitter are 10 ps. Finally, we show the most optimistic line where we assume the detectors are the only source of jitter and use the current state-of-the-art number from Ref. [134]. We also include a line for the classical limit of $\frac{1}{3}$ and for an example of high quality interference with a visibility of 0.95. We see high visibilities up to ≈40 GHz. This would require ring FSR of around 650 pm. We also see quantum interference up to several hundred GHz. In Fig. 3.19(b) we show the effect of the resonance width on the interference fringe. The resonance line-width only affects the modulating envelope. This has a small effect on the HOM fringes. For very small timing windows, the effect is small as the frequency of the time domain fringes is unaffected. If we integrate the entire fringe we see some difference, this is due to the difference in the overlaps of the photon spectrum.

### ACKNOWLEDGEMENTS AND CONTRIBUTIONS

This work will be reported in Ref. [138]. This part of the work was conceived by Dr. Stefano Paesani and Jacob Bulmer. The simulations were performed by me with significant input from Dr. Alex Jones. The chip was designed by Dr. Stefano Paesani. The experiment was performed and the data analysed by myself. Experimental code for controlling the equipment was based on code written by Alexandre Maïnos. The project was supervised by Dr. Alex Jones and Prof. Anthony Laing.

# 4

# BOSON SAMPLING WITH DISTINGUISHABLE PHOTONS

"'Ford', he said 'there's an infinite number of monkeys outside who want to talk to us about this script for Hamlet they've worked out'"

Douglas Adams, *Hitchhikers Guide to the Galaxy*

## 4.1 INTRODUCTION AND MOTIVATION

Due to the extreme engineering requirements that fault-tolerant quantum computers require, it is useful to ask whether there are any intermediate classes of devices[1] that can exploit quantum mechanical effects for some computational advantage. Research into *noisy intermediate-scale quantum* (NISQ) devices, as named by John Preskill [139, 140], can be roughly divided into two avenues: finding uses for quantum computers made up of some (ranging from a few tens to a few hundred) noisy qubits [141–143], or trying to make a quantum computer that can perform a calculation that is intractable for classical computers, regardless of whether the calculation is "useful" or not [2], known as quantum supremacy [1, 147, 148].

Linear optics remains not only a primary candidate for fault-tolerant quantum computing [101], but is also a key player in the race to show quantum supremacy. Linear optics hopes to show (or in some eyes already has shown) quantum supremacy through a computational problem known as boson sampling [3, 5, 149]. This problem, which will be described in more detail in the next section, consists of interfering photonic states with linear optics and drawing samples from the output probability distribution. It was first introduced by Aaronson and Arkhipov [8] and is linked to calculating matrix permanents of complex matrices, known to be intractable for classical computers. Large-scale implementations of both standard boson sampling [108] and a variant of the original problem, known as Gaussian boson sampling, where the input

1 'Intermediate' here is used to mean between a classical computer and a full fault-tolerant quantum computer

2 Although Scott Aaronson did point out that there may be a use for quantum supremacy experiments in generating certifiable randomness [144]. Others have also suggested that boson sampling could have applications in certain graph problems [145] and simulating molecular vibronic dynamics [146]

state retains the Gaussian nature of the parametric sources [3], have been demonstrated. The probability distribution for the latter case is linked to a matrix function, named the Hafnian, which is in the same complexity class as the permanent [150, 151]. Beyond these two bulk optics implementations, integrated photonics also provides a promising platform for large-scale boson sampling implementations [14, 17].

All current boson sampling implementations either use parametric or solid-state sources. As shown in Ref. [17], ring resonators are a promising candidate for integrated parametric sources due to their low power consumption and small footprint. Both ring resonators and solid state sources suffer from the same problem: it is challenging to make multiple sources that emit at exactly the same wavelength. For ring resonators, the internal phase can be used to tune the emission wavelength. As we saw in the previous chapter, however, experimental errors such as crosstalk make scaling this method impractical. For solid state emitters, such as quantum dots, one solution is demultiplexing many consecutive photons from a single emitter into many spatial modes [108]. This requires complicated and potentially high-loss demultiplexing schemes and is limited by the coherence time of the dot. As we mentioned in the introduction to the previous chapter, quantum interference with a visibility of 93% was only recently shown between two separate quantum dots [109]. Here the emission wavelength was tuned by the Stark effect, however scaling this to the number of photons required for quantum supremacy would be experimentally challenging. In the last chapter, we saw how fast detectors can be used to erase frequency distinguishability between single photon sources; however, we limited ourselves to a fundamental investigation of this phenomenon.

In this chapter, we look to apply this technique to the problem of boson sampling with distinguishable photon sources. To do this, we initially provide some background on the original boson sampling problem and its common variants, then we benchmark our device as a boson sampler using spectrally indistinguishable sources. Next, we perform boson sampling with spectrally distinguishable photons. Initially, we look at the effect of changing the coincidence window on the output probability distribution and then we look at sampling from the arrival times of the photons as well as the output pattern, which is performed for two-photon (four measured photons) and three-photon (six measured photons) interference.

## 4.2 INTRODUCTION TO BOSON SAMPLING

To start we introduce the original boson sampling problem and then the main variants that have appeared in the years following the original paper. All variations of boson sampling follow the same format - a photonic state is prepared, interfered and then detected. We restrict ourselves to only cases with single photon detectors, specifically, the derivations below assume number resolving detectors. In the later experiments, we are limited to using threshold detectors, that is detectors that can only register that 0 photons have been detected or $> 1$ photons have been detected. As we are heralding an input pattern and post selecting on the $n-$fold coincidence events, the following analysis holds, at least approximately, for the samples we measure.

### 4.2.1 *Standard boson sampling*

As mentioned previously, the boson sampling problem, introduced by Aaronson and Arkhipov, concerns itself with drawing coincidence samples from the output of a linear optical interferometer. Initially, we prepare an input state consisting of $n$ photons in the first $n$ modes of an $m$ mode interferometer, $\mathbf{U}$. Here we impose two conditions, namely that there is exactly one photon input in each of the $n$ modes and that the size of the interferometer is sufficiently large that each photon leaves in a separate mode - known as the collision-free case. Generally it is taken that $m = n^2$ should be large enough that collision free events dominate [8]. Taking the first condition, we write the input state as $|\psi\rangle_{\text{in}} = \hat{a}_1^\dagger \hat{a}_2^\dagger ... \hat{a}_n^\dagger \, |0\rangle^{\otimes n}$. The $m \times m$ unitary transformation transforms the input creation operators as

$$\hat{a}_i^\dagger \rightarrow \sum_{j=1}^{m} U_{ij} \hat{b}_j^\dagger, \tag{4.1}$$

Applying this transformation to the input state, $|\psi\rangle_{\text{in}}$, we see our output state is a superposition of the possible $n-$photon output patterns

$$|\psi\rangle_{\text{out}} = \sum_{S}^{\tilde{S}} A_S \left(\hat{b}_1^\dagger\right)^{s_1} \left(\hat{b}_2^\dagger\right)^{s_2} ... \left(\hat{b}_m^\dagger\right)^{s_m} |0\rangle^{\otimes n}, \tag{4.2}$$

where $\tilde{S} = \{S\}$ is the set of all $\binom{m}{n}$ possible output patterns, $A_S$ is the amplitude of pattern $S$, and $s_i$ is the number of photons in the $i^{\text{th}}$ output mode in pattern $S$. We

limit ourselves to the patterns where $s_i \in \{0, 1\}$ - that is one photon in each of $n$ of the possible $m$ modes. Each of the possible patterns in $S$ is associated with a particular transfer matrix, $\mathbf{T}$, taken from a subset of the elements of $\mathbf{U}$. Specifically, to construct $\mathbf{T}$ we take the columns of $\mathbf{U}$ corresponding to the input modes [3] and the rows of $\mathbf{U}$ corresponding to the desired output pattern. The elements of $\mathbf{T}$ are then given by the intersection of the selected rows and columns of $\mathbf{U}$. The probability of detecting a given pattern, $S$, is given by

$$P_S = |A_S|^2 = \left| \frac{\text{Perm}(\mathbf{T})}{\sqrt{s_1! \ldots s_m!}} \right|^2, \tag{4.3}$$

where, in the collision free regime, the normalisation factor is one, and Perm $(\cdot)$ is a matrix function called the permanent. The permanent is defined as

$$\text{Perm}\,(\mathbf{T}) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} T_{i,\sigma(i)}, \tag{4.4}$$

where $T_{ij}$ are the elements of $\mathbf{T}$ and $S_n$ is the symmetric group which contains all the permutations of $(1, 2, \cdots n)$. This matrix function is similar to the determinant but missing the alternating sign. While there exist polynomial time algorithms to calculate the determinant [152] this is not the case for the permanent. In fact, the permanent is in the computational complexity class #P. Unlike the more familiar complexity classes P and NP, which describe decision problems, #P contains the set of problems which count the number of solutions to NP problems and is at least as hard as NP. The permanent can be related to the perfect matchings of a bi-partite graph. A bi-partite graph is a graph whose vertices can be divided into two sets and each edge connects between sets. A perfect matching is then a set of edges of the graph which connect each vertex from one set to one and only one vertex from the other. The total number of perfect matchings is the permanent of the graph's adjacency matrix. This makes the permanent an interesting problem in #P as calculating one perfect matching can be performed by a polynomial time algorithm but counting all the perfect matchings is #P [153]. In boson sampling, the bipartite graph vertices are the input and output modes of the

3 Note, for the standard case this is always the first $n$ columns of $\mathbf{U}$, however we will see later that this is not always the case.

interferometer and the analogue of the adjacency matrix is $\mathbf{T}$ [4]. Each perfect matching now describes the transition amplitude for a possible way in which the input pattern can be mapped to the output pattern. The total transition amplitude is, therefore, the sum of all these possible mappings.

### 4.2.2 *Scattershot boson sampling*

A major draw back with the original imagining of boson sampling is that generally we are limited to photon sources that have less than unity probability of generating a photon, therefore, the probability to generate an $n-$photon input state decreases exponentially with $n$. One revision, named scattershot boson sampling [154], aims to combat this problem when using heralded sources. Here, we make full use of the fact that $m \gg n$ and notice that any input pattern comprises a valid boson sampling experiment. We now pump $m$ parametric sources, one on each input mode, and impose requirements only on the number of input photons, not on their input configuration, allowing a measured herald pattern to inform us of which input pattern a given experiment used. Each parametric source generates a two mode squeezed vacuum state and therefore the probability of a specific input pattern is equivalent to the probability of generating $n$ photons from $m$ two mode squeezers. Thus, the probability of a successful boson sampling input is [154]

$$\mathrm{P}(n,m) = \binom{m}{n} \operatorname{sech}^{2m} \xi \tanh^{2n} \xi, \tag{4.5}$$

where $\xi$ is the squeezing parameter, assumed to be the same for all sources. Here we see that the probability of a successful boson sampling input has been scaled by the combinatorially large number of possible input patterns. To calculate the probability of a sample, we follow the same method as above however now when we construct the transfer matrix, $\mathbf{T}$, we no longer always take the first $n$ columns of $\mathbf{U}$. Now, the required columns of $\mathbf{U}$ are given by the detected heralding pattern. In Ref. [154] they go on to

4 $\mathbf{T}$ is not exactly the adjacency matrix, as described here, as instead of having elements of one if there exists an edge between two vertices and zero if not, the transfer matrix assigns a transition amplitude between vertices (inputs and outputs).

show that for the collision-free regime where $m = n^2$, the success probability at the optimum squeezing level is

$$\mathrm{P}\left(n|\xi_{\max}\right) = \frac{1}{e\sqrt{2\pi(n-1)}}, \tag{4.6}$$

where $\xi_{\max} = 1/\sqrt{n+1}$. Here we see that in the collision free regime, the probability scales approximately inversely proportionally to $\sqrt{n}$[5].

While scattershot boson sampling appears to, at least partially, combat the problem of the decreasing success probability, it is not without its own problems. Scattershot boson sampling in a sense is taking advantage of the naturally Gaussian nature of the parametric sources, however this is destroyed on heralding a given input pattern. By doing this we are restricting ourselves to only events where each detected photon pair emerged from a different source. However, there are several ways that it is possible for us to detect $n$ heralds and $n$ signal photons.

- Each of the $n$ signal photons was generated in a single pair with one of the $n$ heralds
- A subset of the $n$ signal photons were generated from the other $m - n$ photon sources not present in the heralding pattern
- A subset of the $n$ signal photons were generated as higher number Fock states from the heralded sources

In the presence of no loss and number resolving detectors both of the noise processes here can be avoided by heralding on vacuum on all other output modes and ensuring each mode where a photon is detected contains one and only one photon. However, in realistic, experiments which will feature loss and often use threshold detectors, these noise processes are unavoidable. These noise processes mean that if $k > n$ photons are generated but only $n$ are detected, we have no way of knowing which $n$ and therefore no way of knowing what the correct probability for this sample would be. This is discussed more quantitatively in appendix A of Ref. [155]. If the losses in the boson sampler are higher, these noise processes become more prevalent - as higher squeezing

5 We could increase the number of input patterns by increasing the number of modes, but this comes at the cost increasing the experimental overhead.

is required to see the desired number of photons. We will see the impact of this in Sec. 4.7

We have seen that some benefits come by, at least partially, taking the Gaussian nature of parametric sources seriously. This begs the question, what happens if we fully exploit the Gaussian nature of these sources?

### 4.2.3 *Gaussian boson sampling*

The latest variation to the original boson sampling problem, called Gaussian boson sampling (GBS) does away with the notion of single photon inputs entirely, deciding instead to use single mode squeezed states as the input states [6]. In its most general form, GBS involves performing Fock basis measurements on an arbitrary multimode Gaussian state. If a Gaussian state is described by a covariance matrix, $\sigma$, we can define the sampling matrix[7]

$$A = \begin{pmatrix} 0 & \mathbb{1}_m \\ \mathbb{1}_m & 0 \end{pmatrix} \left[\mathbb{1}_{2m} - \sigma_Q^{-1}\right], \tag{4.7}$$

where $\sigma_Q = \sigma + \frac{\mathbb{1}}{2}$. When the covariance matrix $\sigma$ describes a pure state, A reduces to a block diagonal matrix $A = B \bigoplus B^*$. The original setup for GBS involves inputting $n$ single mode squeezed states into and $m-$mode interferometer. At the output of the interferometer we have a multimode squeezed state which can be described by the covariance matrix

$$\sigma = \frac{1}{2} \begin{pmatrix} \mathbf{U} & 0 \\ 0 & \mathbf{U}^* \end{pmatrix} \sigma_0 \begin{pmatrix} \mathbf{U}^\dagger & 0 \\ 0 & \mathbf{U}^T \end{pmatrix}, \tag{4.8}$$

where $\sigma_0$ is the initial squeezing covariance matrix given By

$$\sigma = \frac{1}{2} \begin{pmatrix} \bigoplus_{j=1}^m \cosh \xi_j & \bigoplus_{j=1}^m \sinh \xi_j \\ \bigoplus_{j=1}^m \sinh \xi_j & \bigoplus_{j=1}^m \cosh \xi_j \end{pmatrix} \begin{pmatrix} \bigoplus_{j=1}^m \cosh \xi_j & \bigoplus_{j=1}^m \sinh \xi_j \\ \bigoplus_{j=1}^m \sinh \xi_j & \bigoplus_{j=1}^m \cosh \xi_j \end{pmatrix}^\dagger, \tag{4.9}$$

6 As the results in the rest of this chapter only use standard and scattershot schemes, we include this section only for completeness. For a more thorough description the reader is directed to Refs. [90,150,151]

7 Note, the covariance matrix $\sigma$ must be expressed in the $\hat{a}$, $\hat{a}^\dagger$ basis

where $\xi_j$ represents the squeezing parameter in the $j^{\text{th}}$ input mode. From this we can calculate the sampling matrix to be

$$B = \mathbf{U}\left[\bigoplus_{j=1}^{m} \tanh \xi_j\right] \mathbf{U}^T. \tag{4.10}$$

The probability for a given output pattern is then given by

$$P = \frac{|\text{Haf}(B_S)|}{\sqrt{|\sigma_Q|}}, \tag{4.11}$$

where $B_S$ is a submatrix of $B$ constructed in a similar way to the other variations but, as there is no input pattern, the output pattern defines both the columns and the rows. Haf corresponds to the Hafnian, a matrix function which is in the same complexity class as the permanent and is related to the permanent through

$$\text{Haf}(U) = \text{Perm}\left(\begin{bmatrix} 0 & U \\ U^T & 0 \end{bmatrix}\right). \tag{4.12}$$

## 4.3 VERIFYING A BOSON SAMPLING EXPERIMENT

As with all experiments at the edge of what is possible to simulate classically, verifying that the system is working as expected is a non-trivial problem. This is because it is, by definition, impossible to efficiently, calculate the outcome of a quantum supremacy experiment. In fact, as boson sampling belongs to the #P complexity class, there does not exist an efficient way to perfectly verify that the experiment is working correctly. There are many verification protocols available in the literature for different kinds of quantum supremacy experiments [3, 14, 156]. In the rest of the chapter we choose to use a form of Bayesian updating to verify our experiments. In this we dynamically update our confidence that a given set of samples comes from the desired model vs a given adversarial model. This has benefits for us that a small set of samples can be used to give good evidence towards one model over another, and it can also easily be adapted to compare different models. It however does require all probabilities to be normalised, due to the differing probabilities of bunched events in the different models,

this requires the entire distribution to be calculated, rendering this method ineffective for experiments closer to the quantum supremacy bound. For our proof of principle experiments, with many fewer photons than can be classically simulated, it provides a quick and easy method of confirming our experiment is behaving as expected. The basis of this verification method is Bayes' theorem. This is a compact equation that can be used to calculate the probability of a hypothesis given some evidence, in general it is calculated as

$$P(H|E) = \frac{P(E|H)\,P(H)}{P(E)} \tag{4.13}$$

where $P(E|H)$ is the probability that the evidence arises from the hypothesis, $P(H)$ is the probability of the hypothesis, known as the prior, and $P(E)$ is the total probability of the evidence. As we wish to compare two models, the total probability of the evidence becomes

$$P(E) = P(E|H)\,P(H) + P(E|A)\,P(A), \tag{4.14}$$

where we have introduced an adversarial hypothesis, $A$. In our setting, the evidence is some string of samples from a suspected boson sampler, $D$. We wish to know the probability that this comes from a test model, $M_{test}$, compared to an adversarial model, $M_{adv}$. We assume that the priors of the models $P(E) = P(H) = 0.5$ which cancel, leaving us with

$$P(M_{test}|D) = \frac{P(D|M_{test})}{P(D|M_{test}) + P(D|M_{adv})}. \tag{4.15}$$

If we assume that sequential samples in $D$ are uncorrelated, then to calculate the probability that a given string of samples $D$ comes from a given model, $P(D|M_i)$, we take the product of the probabilities of each sample in $D$. These individual probabilities are calculated according to the specific model. The task of selecting the adversarial model, $M_{adv}$, in a boson sampling experiment is an important one. The aim is to choose a distribution that accurately reflects possible errors in the experiment that could result in samples from a distribution that is easy to simulate [8]. We, therefore, look to find an adversarial model that most closely resembles our experimental setup, but that is easy to sample from.

8 By easy to simulate we refer to distributions that can be sampled from at least approximately using a polynomial time algorithm

## 4.4 EXPERIMENTAL CHARACTERISATION

The experimental setup for the following results is almost identical to the previous chapter. The only change we make is replacing the 50 MHz pulsed laser with a PriTel ultrafast optical clock. This laser has a repetition rate set by an external signal, set here

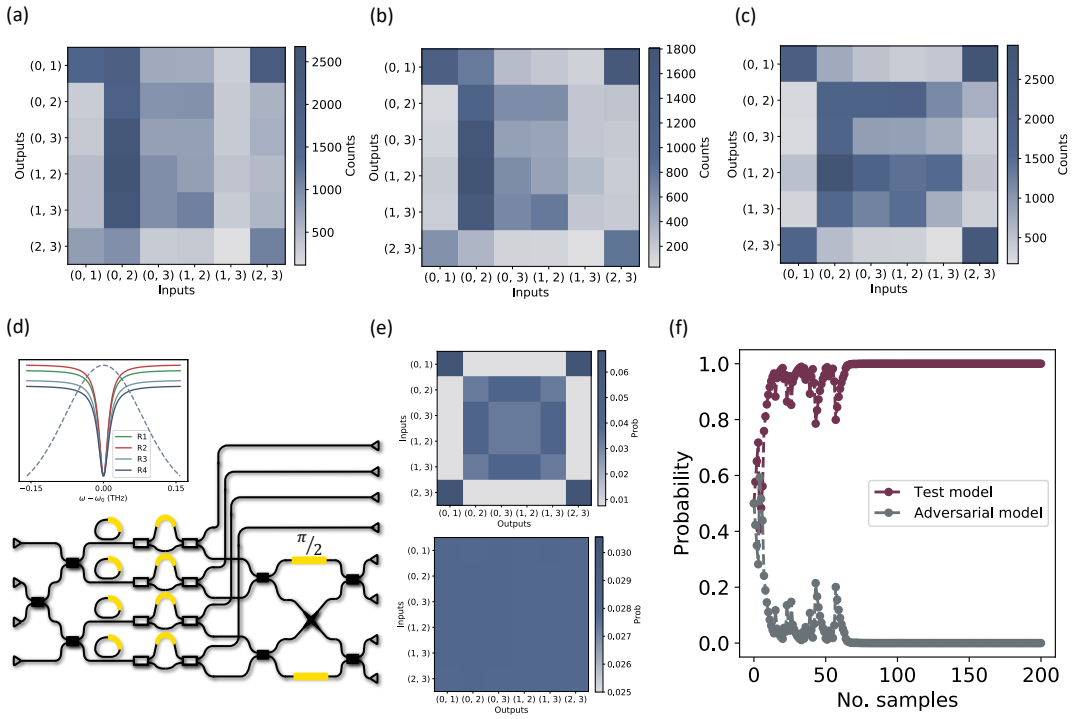


Figure 4.1: (a)-(c) Measured output count distributions. Columns correspond to input patterns, rows correspond to output patterns. (a) Raw uncorrected counts. (b) Background subtracted counts. (c) Background subtraction and normalisation for loss and squeezing differences. (d) Experimental schematic. Unitary implements complex Hadamard matrix with $\phi = \frac{\pi}{2}$. (e) Theoretical normalised probability distributions. Top panel is for indistinguishable photons, bottom for distinguishable photons. (f) Bayesian verification of boson sampling. Test model is that of indistinguishable photons (purple) and the adversarial model is distinguishable photons (grey).

to be 500 MHz. The increase in repetition rate means that we can keep the count rates high while reducing the squeezing and therefore the double emissions. To initially benchmark our chip as a boson sampler, we perform a two-photon scattershot and three-photon standard boson sampling experiments with indistinguishable photons.

### 4.4.1 *Two-photon scattershot*

For the two-photon scattershot experiment all four rings are aligned to the centre of the pump pulse. We set the integrated interferometer to implement the complex Hadamard transformation and set the internal phase $\phi = \frac{\pi}{2}$. This was chosen as it adds some extra structure to the probability distribution. For $\phi = \{0, \pi\}$ all probabilities are either 0 or 1. For our selected $\phi$ value we also have intermediate values, resulting in a qualitatively more interesting distribution. We integrate for a total of 6 hours in one hour chunks. We check the stability of the rings every two hours. For our test and adversarial models we choose indistinguishable and distinguishable photon interference, respectively. This means our test model probabilities are calculated as $P = |\text{Perm}(M_s)|^2$ and the adversarial as $P = \text{Perm}(|M_s|^2)$, where $M_s$ is the submatrix corresponding to the specific sample pattern. We note here that because the suppression laws in the Fourier matrix, the indistinguishable distribution has many input-output combinations that have 0 probability. Any measured counts in these patterns (which will arise experimentally due to noise and impurity of the input photons) will immediately cause the Bayesian validation to fail. To combat this, we include the mixedness in the ideal model. Therefore, ideal model probabilities are calculated as $P = \sin^2(\theta)\,|\text{Perm}(M_s)|^2 + \cos^2(\theta)\,\text{Perm}\left(|M_s|^2\right)$, and we use the measured HOM visibility of 0.82. Figure 4.1(c) shows the experimental setup and (d) shows the theoretical probability distributions for indistinguishable photons (top) and distinguishable photons (bottom). Figure 4.1(d) shows the Bayesian verification confirming that the samples drawn are more likely to have come from quantum interference. We have shown an example 200 samples drawn randomly from the total over $4 \times 10^4$ samples. The total number of samples returns a probability for the test model of 1.0, up to machine precision. In Fig. 4.1(a)-(c) we have shown the experimentally measured distributions. Figure 4.1(a) shows the raw counts we measure from the device. These need to be corrected, first by removing the double emission counts, as described in the previous chapter (Sec. 3.4.4), and then by normalising for differences in output losses and squeezing for each pattern. The final distribution is shown in Fig 4.1(c). To quantify the improvement we look at the statistical fidelity (as introduced in the previous chapter) between the measured and theoretical distributions. Using this, we see an increase in fidelity from 87.4% before correction to 96.7% after correction.

### 4.4.2 *Multi-photon background subtraction*

In the following section, and later in this chapter, we look to generate and interfere three photons. In order to correct for double emissions, here we quickly describe the multi-photon generalisation of our background subtraction method, outlined in the previous chapter. The procedure is generally the same, however, now we sequentially delay each of the $N$ heralds by one repetition rate. With the $i^{\text{th}}$ herald delayed, the resulting $N-$fold coincidence contains all the possible double emissions of the other $N-1$ sources. If we perform this for all $N$ sources and sum the counts, this gives us $N-1$ times the double emission rate. It is $N-1$ fold as each source contributes to every measurement apart from the one where itself is delayed.

### 4.4.3 *Three-photon boson sampling*

As the single photon channel loss of our experiment is high, $\eta \approx 0.03$, the six-fold rate from the chip was low. In order to generate samples at a high enough rate the squeezing in the sources was too large to consider the double emissions negligible. Double emissions from any of the sources that are involved in the interference could be corrected for (see Sec. 4.4.2), however, emissions from the fourth source can not be corrected for in this way. This precludes us from performing scattershot measurements and leaves us without the 4-fold improvement in rates that the scattershot advantage would provide. We therefore pick the input pattern with the lowest coupling losses, those being rings one, two and four. As the three photon distribution in the complex Hadamard interferometer is uniform for both indistinguishable and distinguishable photons, we use the two input MZI's to break this degeneracy. We choose to set both input MZI, phases to 0.9 rad. This number was chosen heuristically in an attempt to roughly minimise the overlap between the distinguishable and indistinguishable distributions, while maximising the total non-bunched probability. These distributions are shown in Figs. 4.2(b) and (c), respectively. Due to the low squeezing rate, we integrate for 93.5 hours. As with all previous measurements, we lock every six hours but check for misalignment every two hours. In this time we recorded a total of 35 six-fold events. After normalisation and background subtraction, we were left with 17 legitimate samples. As can be seen from Fig. 4.2(d) this was enough to pass the verification protocol, confirming that we are likely sampling from the correct distribution.

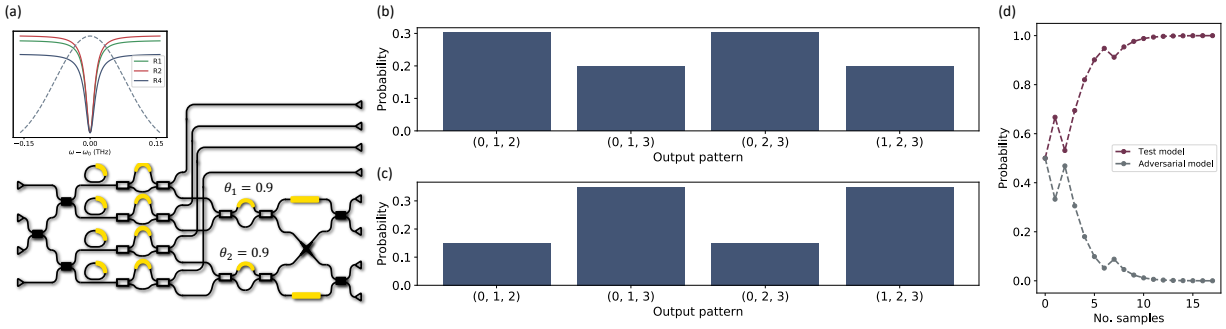


Figure 4.2: Characterisation of three photon (six measured photons) boson sampling with indistinguishable photons. Standard boson sampling approach used with one input pattern, chosen to be rings 1, 2, and 4 due to these having the lowest loss. (a) Experimental schematic, all three sources aligned to the central frequency of the pump. Input MZI phases both set to 0.9, this breaks the degeneracy between indistinguishable and distinguishable distributions and aims to maximise the probability of non-bunched events. (b) Test model distribution, indistinguishable photons but including mixedness. (c) Adversarial model distribution, distinguishable photons. (d) Bayesian verification

## 4.5 TIME-RESOLVED MULTI-PHOTON INTERFERENCE

So far, the arrival time of the photons has been ignored and the photons are assumed to be identical. In the following analysis we look to derive the probability of detecting $N$

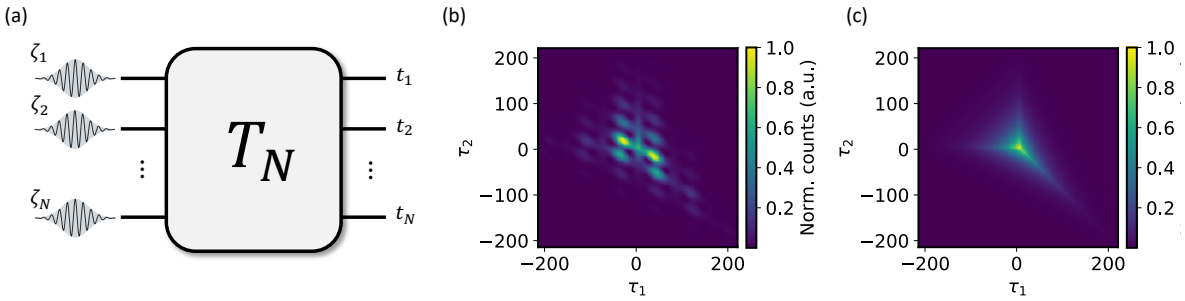


Figure 4.3: (a) Schematic for time-resolved multi-photon interference. $N$ photons with temporal shapes $\zeta_i$ interfere via a transfer matrix, $T_N$. Photon in each mode detected at a separate arrival time $t_i$. (b) Example three photon interference fringe with spectrally distinguishable photons at three different central frequencies. (c) Same fringe but for fully distinguishable photons, no interference.

photons at $N$ different times at the output of a boson sampler, given input photons with

arbitrary temporal spectra, $\zeta_i$, as shown in Fig. 4.3(a). To do this we follow a similar procedure to Sec. 3.3.1. Our desired output state is given by

$$|\psi_{\text{out}}\rangle = \prod_{i=1}^{N} \hat{b}_i^\dagger\left(t_i\right)|0\rangle^{\otimes N}, \tag{4.16}$$

which represents an $N$ fold coincidence with the photon in the $i^{\text{th}}$ mode detected at time $t_i$. The output mode operators are given by

$$\hat{b}_i^\dagger\left(t_i\right) = \sum_{i=1}^{M} T_{ij}\zeta_j\left(t_i\right)\hat{a}_j^\dagger, \tag{4.17}$$

where the $T_{ij}$ are the elements of the submatrix of the boson sampling unitary and by construction $M = N$. Substituting this into Eq. 4.16 we arrive at

$$|\psi_{\text{out}}\rangle = \prod_{i=1}^{N}\sum_{i=1}^{N} T_{ij}\zeta_j\left(t_i\right)\hat{a}_j^\dagger|0\rangle^{\otimes N}. \tag{4.18}$$

This state now contains all possible ways that the $N$ photon output state could arise, including all multi-emission terms. As we are performing scattershot experiments where we herald the input pattern, we only keep terms where the $N$ photons arrive from different sources. As shown in Ref. [157], this can be rewritten as

$$|\psi_{\text{out}}\rangle = \sum_{\sigma\in S_N}\prod_{i=1}^{N} T_{i,\sigma(i)}\zeta_{\sigma(i)}\left(t_i\right)\hat{a}_{\sigma(i)}^\dagger|0\rangle^{\otimes N}, \tag{4.19}$$

where $S_N$ is the group of $N-$dimensional permutations. We can see that the amplitude for this state is the permanent of the transfer matrix, weighted by the temporal shape of the photons. This results in a final probability

$$P_{\text{coinc}}\left(\mathbf{t}\right) = |\text{Perm}\left(\Lambda\right)|^2, \tag{4.20}$$

where the matrix $\Lambda$ has elements given by $\Lambda_{ij} = T_{ij}\zeta_j(t_i)$. Here, we note that we have recovered the result of Ref. [158]. In general, we will integrate over $t_1$, leaving us with interference fringes that depend on $N-1$ different delays. We expect to see separate interference fringes in each of the $N-1$ different dimensions (based on the $N-1$ different time delays), depending on which photons interfere in these specific outputs.

To visualise the multi-photon interference we use our model to simulate three photon interference with spectrally distinguishable photons. We use ring resonator sources with properties similar to those in the experiment but with central wavelengths set to 1541.1 nm, 1541.2 nm, and 1541.3 nm, respectively. For our transfer matrix we choose the submatrix of the complex Hadamard matrix that corresponds to the first three inputs and first three outputs of our chip with $\phi = 0$

$$T = \frac{1}{2}\begin{pmatrix} 1 & i & i \\ i & -1 & 1 \\ i & 1 & -1 \end{pmatrix}. \tag{4.21}$$

We show the fringes from this in Fig. 4.3(b). As expected, we see two-dimensional fringes corresponding to beatings between the spectrally distinguishable photons. At $\tau_1 = \tau_2 = 0$ we recover the indistinguishable case, which for this transfer matrix is $\frac{1}{16}$. For comparison, we also show the case with no interference, calculated as $P(\mathbf{t}) = \text{Perm}\left(|\Lambda|^2\right)$, in Fig. 4.3(c). As expected, we see no fringes.

## 4.6 SCATTERSHOT BOSON SAMPLING WITH VARYING TIMING WINDOWS

As we have shown in Sec. 3.5.2, by decreasing the coincidence window used we can increase the visibility of a quantum interference fringe. In this section, we look to generalise this phenomenon to scattershot boson sampling. Here, we will look at how the measured output distribution changes as a function of the timing window. As to properly show this effect we require enough counts to build up reliable statistics, we restrict ourselves to two photon scattershot experiments. As we only have access to two CW locking lasers, we choose to lock rings in pairs. We use a detuning of 54 pm and lock rings 1 and 4 to one wavelength and 2 and 3 to the other. We integrate for 25 hours in one hour integration chunks. The rings are re-locked to their respective wavelengths every six hours, however every two hours we measure the resonance locations to check they remained aligned.

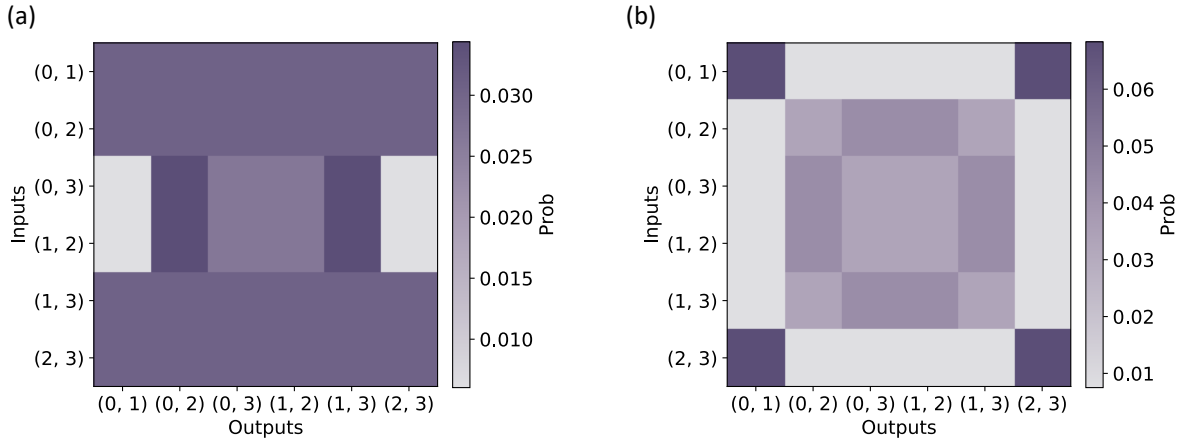


Figure 4.4: (a) Adversarial model for the case where for some patterns the input photons are indistinguishable. (b) Test model calculated for indistinguishable photons with mixedness included.

### 4.6.1 *Ideal and adversarial models*

We now define our test and adversarial models for this experiment. The test model is straight forward, we use the same test model as our characterisation measurements, namely the indistinguishable distribution with mixedness included. The adversarial model is slightly more subtle. Due to our ability to only lock our rings to two different emission wavelengths it would not be a fair comparison to use the fully distinguishable case, as in our characterisation measurements. This is because there are some pairs of photons that are spectrally indistinguishable and would therefore interfere regardless of the quality of our detectors. To account for this, our adversarial model calculates the probabilities from the distinguishable model for all input combinations with spectrally distinguishable sources and uses the indistinguishable model (with mixedness) for the spectrally indistinguishable input patterns. These two adversarial and test models are shown in Figs.4.4(a) and (b), respectively.

### 4.6.2 *Experimental results*

The results of this experiment are shown in Fig 4.5. Figure 4.5(a)-(d) show the experimental distributions for different timing windows. These are taken by only keeping measured samples with the absolute signal-signal time difference being less than the de-

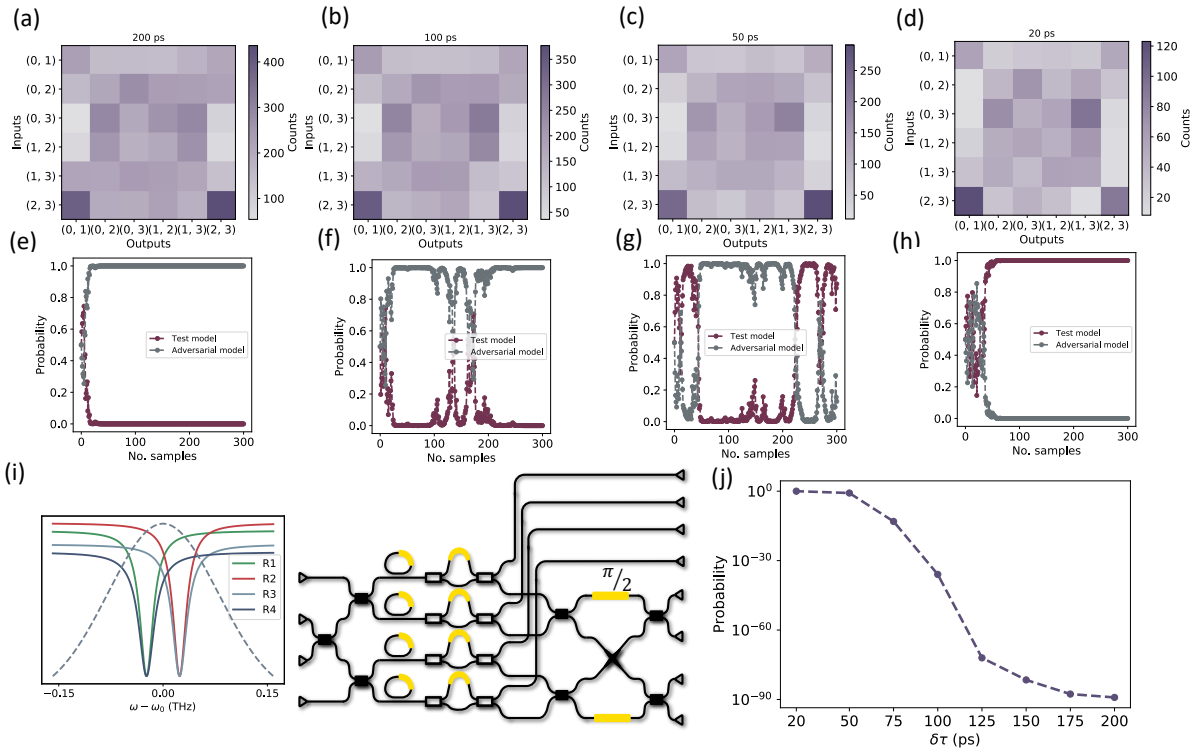


Figure 4.5: (a) - (d) Experimental probability distributions as a function of the timing window. All distributions have been adjusted for loss and squeezing differences. (e)-(h) Corresponding Bayesian verification of 300 samples drawn from the distributions (a) - (d). (i) Chip and source setup. (j) Bayesian probability for large sample sets. Sample sizes of 2000 used and averaged 200 times. If the number of samples within a given timing window are less we take the entire sample list. Error bars, calculated from averaging, are smaller than the symbol size

sired window. We see a qualitative change from a distribution that looks like Fig. 4.4(c) to Fig. 4.4(d). To quantify this we use the Bayesian verification. We illustrate this using a sample list of 300 samples. This is shown in Fig. 4.5(e)-(h). Here we see that, for $\delta\tau$ =200 ps, the adversarial model best describes the sample list. This becomes noisier for $\delta\tau$ =100 ps and $\delta\tau$ =50 ps before flipping and being best described by the test model for $\delta\tau$ =20 ps. To quantify this further, we take larger sample lists and average several times. We use a sample size of 2000 and average 200 times. If the number of samples within the timing window is less than 2000, we take the entire sample list. We plot the final probability on a logarithmic scale in Fig. 4.5. As we have shown above that our test model has more quantum interference in it than the adversarial model, Fig. 4.5

shows clearly that as we decrease the timing window we can increase the quantum interference.

## 4.7 SAMPLING FROM THE TEMPORAL DEGREE OF FREEDOM

The result in the previous section shows that we can, in fact, use fast detectors to increase the quantum interference, and therefore the complexity, of a boson sampler with spectrally distinguishable sources. One key drawback of this scheme is that decreasing the coincidence window necessarily decreases the sample rate of the experiment. In larger experiments where we are operating much lower rates, this would be an unacceptable cost. One thing to notice is that the time-resolved fringes are calculated by integrating over permanents, which is a computationally hard process. This means that drawing samples from these distributions is also hard but does not have the intrinsic reduction in sample rate introduced by the temporal filter used above. This idea, first introduced by Tamma et al. in Ref. [158] and has been further analysed and has been shown to be at least as hard as standard boson sampling [159–162].

In the previous sections the sample was drawn only from the input/output pattern. That is each input/output pattern had a probability assigned to it which was calculated from the required distribution. For the following results we will bin the time-resolved fringes and therefore assign a probability for not only the input/output pattern but also the binned arrival time. This means that every sample detected is a legitimate sample, and we require no temporal filtering.

### 4.7.1 *Two-photon scattershot boson sampling*

In this section we re-analyse the data from Sec. 4.6.2, but now explicitly taking the arrival times of the photons into account.

#### 4.7.1.1 *Ideal and adversarial models*

The test model for this section is acquired from the fringes calculated with Eq. 4.20. For two interfering photons this reduces to the model derived in the previous chapter. Here, we do not use all of the previously describe errors, focusing only on mixedness. For the adversarial model we choose the non-interfering distribution arising from distinguish-

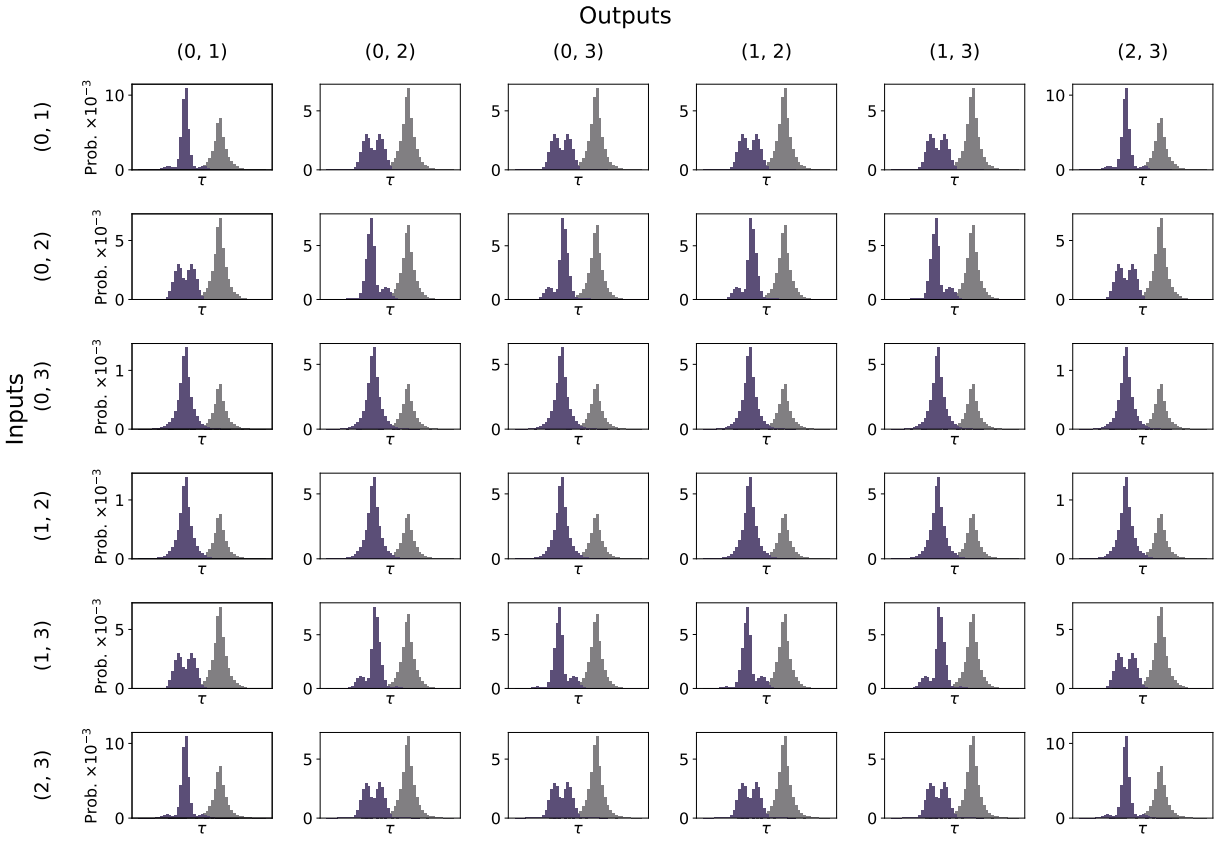


Figure 4.6: Binned theory fringes corresponding to the test and adversarial distributions. Test model in purple and adversarial model in grey. Bin size of 20 ps used. Time axis is arbitrary, and histograms have been shifted in order to make distinguishing the two histograms either. Rows and columns refer to specific input and output patterns, respectively. All probabilities are normalised over all times and all input/output patterns.

able photons, with the same caveat from above that we take the indistinguishable model for the input patterns where the photons are spectrally indistinguishable. Once the theory fringes are plotted from the respective models, we bin them by integrating the curves over each bin. These are then normalised over all time bins and all input/output patterns. We show the entire probability distributions for both models in Fig. 4.7, where we have artificially shifted the histograms with respect to one another to make them easier to see.

#### 4.7.1.2 *Experimental results*

To perform the background subtraction and normalisation, rather than correcting each output pattern we need to correct each individual time bin separately. Each time bin in a given input/output pattern will have the same loss and squeezing normalisation factor, however the background subtraction will be different. To perform the background subtraction, we also bin the double emission counts and then subtract them bin by bin. This means that the final sample count we get depends on the choice of bins. For our choice of bin of 20 ps, we obtain 7316 legitimate samples after correction. We then use the Bayesian verification protocol to compare the two models defined above. In Fig. 4.7(a) we show an example Bayesian verification for a set of 500 samples drawn randomly from the 7316 total samples. In Fig. 4.7(b) we show the final Bayesian probability for samples of different sizes. For each point we randomly draw the set of samples 100 times allowing us to plot a mean with corresponding error bars. We see that it requires a large number of samples in order for the verification to converge appropriately. This is likely due to a combination of the system jitter and scattershot noise, mentioned previously.

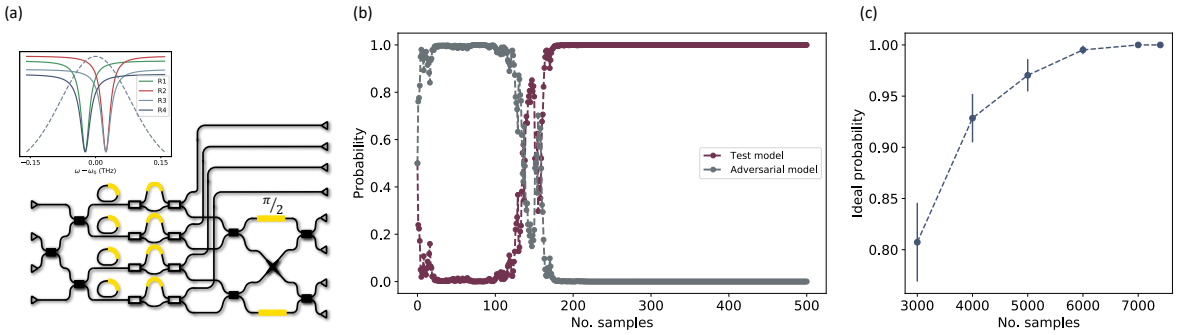


Figure 4.7: (a) Experimental setup two photon scattershot. Sources 2 and 3 are detuned from sources 1 and 4. Interferometer set to implement a Fourier interferometer. (b) Example Bayesian verification for 500 samples. (c) Final Bayesian probabilities for different numbers of sample sizes.

### 4.7.2 *Three-photon standard boson sampling*

As seen in the six-fold characterisation, we are limited to standard boson sampling protocols by the extra noise in the system. For this measurement, we choose the same

input pattern as Sec. 4.4.3, however we detune ring two by 47 pm (5.9 GHz). This detuning is lower than for the two photon experiments in order to keep the squeezing high enough. As we will be also sampling from the photons temporal degree of freedom, the degeneracy between distinguishable and indistinguishable distributions will be broken. Despite this, we choose the same interferometer setup as Sec. 4.4.3. This gives the verification the best chance of converging with a relatively small number of samples. For these measurements, the output losses were optimised further. However, with all of this, we operate at an average six-fold rate of $\approx 2$ per hour. We, therefore, integrate for 212 hours giving us a total of 501 six-folds. As before, we lock the rings to their respective emission wavelengths every six hours and check the resonance positions every two hours.

#### 4.7.2.1 *Ideal and adversarial models*

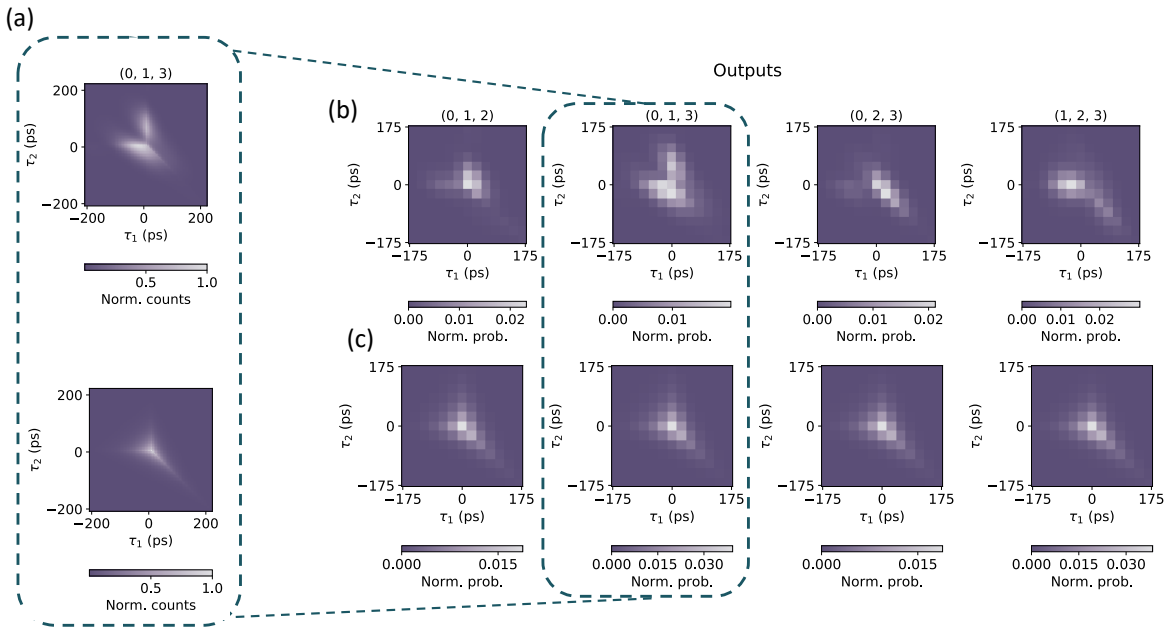


Figure 4.8: Test and adversarial distributions for sampling from the photon arrival times. Title above each panel denotes the output pattern and top and bottom panels correspond to the same pattern. (a) Example unbinned fringes. Top panel is the ideal fringe and bottom panel is the partially distinguishable case. (b) Test model fringe binned with 20 ps bins. (c) Adversarial model fringe binned with 20 ps bins.

The test and adversarial probability distributions are now formed by binning the 2D-theory fringes, similar to those shown in Fig. 4.3(b) and (c). For the test model we use Eq. 4.20 with an adjustment for the mixedness of the photons. As in Sec. 4.6, there is a subtlety here with the adversarial model. The fully distinguishable case is given by $P(\mathbf{t}) = \text{Perm}\left(|\Lambda|^2\right)$, where there is no interference between any of the terms. However, as we only detune one of the three photons, this does not accurately reflect an accurate worst case scenario. We would expect interference between the two indistinguishable photons. For the specific case where it is the second photon that doesn't contribute to the interference, the final probability is given by

$$\begin{aligned} P(\mathbf{t}) = |T_{22}\zeta_{22}|^2 |T_{11}T_{33}\zeta_{11}\zeta_{33} + T_{13}T_{31}\zeta_{13}\zeta_{31}|^2 \\ + |T_{32}\zeta_{23}|^2 |T_{11}T_{23}\zeta_{11}\zeta_{32} + T_{13}T_{21}\zeta_{12}\zeta_{31}|^2 \\ + |T_{12}\zeta_{21}|^2 |T_{21}T_{33}\zeta_{12}\zeta_{33} + T_{23}T_{31}\zeta_{13}\zeta_{32}|^2, \end{aligned} \tag{4.22}$$

where $T_{ij}$ are the elements of the transfer matrix $\mathbf{T}$ and $\zeta_{ij} = \zeta_i\left(t_j\right)$. Once the fringes are calculated, as shown in Fig. 4.8(a), we divide the fringe into bins, integrate across each bin, and normalise total probability over all times and all output patterns. The results of this process are shown in Fig. 4.8(b) and (c) for a bin size of 20 ps.

#### 4.7.2.2 *Experimental results*

With a bin size of 20 ps, after background subtraction and normalisation we are left with 353 legitimate six-fold samples. We perform the Bayesian verification on all of these samples and show the results in Fig. 4.9 with a final probability of 0.9999995, confirming that we are indeed sampling from the correct distribution.

## 4.8 DISCUSSION AND OUTLOOK

In this chapter we have seen two different ways in which our ability to resolve sufficiently small differences in time can be used to retain complexity in boson sampling experiments where the input photons are distinguishable in frequency. The first method, similar to the results presented in the previous chapter, we saw that decreasing the coincidence window moves us from a distribution with less interference to one with more. The limit of this technique and one which makes it impractical for larger scale

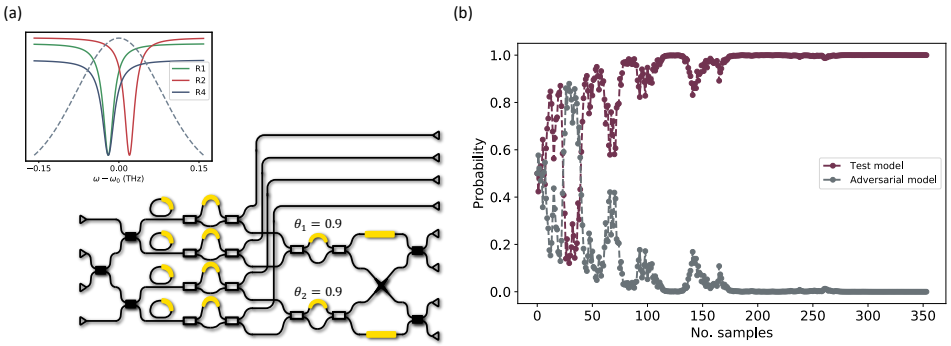


Figure 4.9: (a) Three photon boson sampling with distinguishable photons. Ring 2 is detuned from the other two and the input beam-splitter angles are both set to 0.9 to break the degeneracy between indistinguishable and distinguishable distributions. (b) Bayesian verification including photon arrival time. Protocol assigns a probability to the combination of output pattern and arrival time. Performed for all 353 legitimate six-fold events. Final probability is 0.9999995.

experiments is that there is a fundamental trade off between the number of accepted samples and the amount of recovered interference. In order to combat this, and using the approach introduced in Ref. [158], we note that the complexity remains if we sample not only from the output pattern but also from the probability distribution over the arrival times. We implement this for two-photon scattershot and three photon standard boson sampling. Here we also noticed the effect of one of the main restrictions of scattershot boson sampling, that being the effect of spurious emissions from the unheralded sources. This becomes especially prevalent in systems with high output loss, as the squeezing must be increased to see the $n-$fold processes at an acceptable rate. This problem helps motivate the idea of Gaussian boson sampling, as mentioned in the introduction of this chapter, where we no longer use parametric photon sources to approximate single photon Fock states but rather embrace the Gaussian nature of these photon sources fully. The same principles used in this work should map nicely to Gaussian boson sampling although, to the best of my knowledge, this is missing from the literature. However, recent work on performing Gaussian optics with spectrally multimode sources [163] signify that it should be reasonably straight forward to include the photon arrival time into the GBS framework. A further limitation of the work in this chapter is that we have only qualitatively analysed the increase in quantum interference. In general, throughout this work, we have chosen two distributions with differing amounts of interference and shown that our experiment is more likely to be generating

samples from the distribution with higher interference. It would be interesting to see, quantitatively, how quantum interference in an experiment changes as a function of system jitter. Ref. [164] introduced a way of simulating partial distinguishability by representing $n$ partially distinguishable photons as $k$ interfering indistinguishable photons and $n-k$ non-interfering fully distinguishable photons. It would be interesting to apply similar analysis as a function of timing resolution.

### ACKNOWLEDGEMENTS AND CONTRIBUTIONS

This work will be reported in Ref. [138]. This part of the work was conceived by myself with significant input from Dr. Alex Jones and Dr. Stefano Paesani. The simulations were performed by me with help from Dr. Alex Jones. The experiment was performed and the data analysed by myself. The project was supervised by Prof. Anthony Laing.

# 5

# ALL-OPTICAL SWITCH: INTERMODAL CROSS-PHASE MODULATION

"For a moment, nothing happened. Then, after a second or so, nothing continued to happen"

Douglas Adams, *Hitchhikers Guide to the Galaxy*

## 5.1 INTRODUCTION AND MOTIVATION

Fast and low-loss switching of light is a necessary resource for many quantum photonic technologies. Measurement-based quantum computers need a reconfigurable measurement basis so future qubits can be measured based on previous results [30, 101, 103, 165]. This is in order to account for the random measurement outcomes from previous measurements. Beyond this, in Sec. 2.7.3, we have seen how linear optical quantum computing will require large amounts of multiplexing in order to build the large resource cluster state [94, 106, 166, 167]. All multiplexing schemes require switching networks that must be reconfigurable on the scale of the lifetime of the photons. The faster the switch, therefore, the smaller the delay lines and the lower the subsequent probability of the photon being lost. In an integrated platform, the switching network will need to be integrated on chip with the superconducting nanowire single photon detectors and will, therefore, need to operate at cryogenic temperatures. Other quantum technologies, such as quantum key distribution (QKD) [168, 169] and quantum metrology [170] also use fast reconfigurability.

Despite the widespread need for high-quality fast phase shifting, the perfect solution that combines high bandwidth, low loss, the ability to operate at cryogenic temperatures, and is material agnostic, has not yet been found. A leading candidate for phase shifters that maximise speed and minimise loss are those based on an electro-optic effect called the Pockels effect. Where applying a DC electric field across a waveguide subsequently changes the refractive index of the waveguide proportional to the strength of the field. This is attractive for many reasons: the switching speed is effectively limited by how fast the electric field can be switched, the loss is limited by overlap of the waveguide

mode and the electrodes, which can be minimised by the design, and these devices can be operated at cryogenic temperatures. The natural drawback is that, generally, strong electro-optic effects require a $\chi^{(2)}$ nonlinearity, which is, as we have seen in Sec. 2.4.1, only found in materials with a crystal structure with certain symmetry properties. While these materials have been shown to be promising candidates for integrated photonic technologies, they often demand more sophisticated fabrication processes that limit their scalability [19–27].

Many alternatives to $\chi^{(2)}$ electro-optic phase shifting methods exist: plasmonic dispersion modulators, micro-electro mechanical systems (MEMS), and thermo-optic phase shifters are just some. While each has attractive properties and has found uses in different areas of photonics, none ticks all the desired boxes for scalable integrated quantum applications [171–176]. One possible solution, cross-phase modulation (XPM), uses a strong laser pulse to change the waveguide refractive index. As with the Pockels effect, this is also a nonlinear-optical process, but is controlled by the strength of the material's $\chi^{(3)}$ nonlinearity [1]. This nonlinearity has no special requirements on the material crystal structure and has even been demonstrated in amorphous materials [85, 177, 178], liquids [179–181] and gases [182–185]. Demonstrations of XPM-based all-optical switches have been shown in optical fibre setups [186], in integrated ring resonators [187] and in optical waveguides [188]. In most cases, frequency serves as the degree of freedom to separate pump and signal beams. This requires low-loss and broadband frequency multiplexers, which remain challenging to implement in an integrated platform. In Ref. [189], transverse mode is used as the separating degree of freedom, creating a transverse mode MZI. The optical pump, therefore, co-propagates with both arms of the MZI and the difference in XPM efficiencies between TE0-TE0 and TE0-TE1 is used to implement a relative phase shift between the arms. This difference is much less than either of the processes alone and, therefore, requires a large amount of power to generate a large enough relative phase shift.

In this chapter we present, model, and test an intermodal XPM modulator in the form of an MZI with the pump located on only one arm. The rest of the chapter proceeds as follows. First, we introduce the general model for cross phase modulation, including the transverse profile of the modes. Following this, we address the specifics of designing

1 It is worth noting here that Pockels effect is a DC electro-optic effect, requiring a static electric field. Cross-phase modulation is an AC electro-optic effect and depends on the strength of an oscillating optical field.

the experimental circuit by first determining the necessary parameters for the MZI, before describing the process of designing the photonic components. Finally, we move to the experimental setup and characterisation of the devices.

## 5.2 CROSS-PHASE MODULATION

Cross-phase modulation is a third-order nonlinear optical process whereby a strong pump beam can impart a phase shift on a signal beam. In the following section we introduce the necessary mathematical model to understand this process. Initially, we introduce the simplest model, with plane waves propagating through a material. We then look at the more realistic case of pulses propagating in waveguides.

### 5.2.1 *Intensity dependent refractive index*

In this section we will look at how waves propagating through a nonlinear material can perturb the material refractive index. We first model this with a monochromatic wave $\mathbf{E}(\omega) = E(z,t)\, e^{i(kz-\omega t)} + E^*(z,t)\, e^{-i(kz-\omega t)}$ propagating through a material with a nonlinear polarisation

$$\mathbf{P}(\omega) = \chi^{(1)}\epsilon_0\mathbf{E}(\omega) + \chi^{(3)}\epsilon_0\mathbf{E}(\omega)\mathbf{E}(\omega)\mathbf{E}(\omega). \tag{5.1}$$

Expanding out the nonlinear part and keeping only terms without frequency generation, we arrive at

$$\mathbf{P}(\omega) = \chi^{(1)}\epsilon_0\mathbf{E}(\omega) + 3\chi^{(3)}\epsilon_0|E(z,t)|^2\mathbf{E}(\omega) = \chi_{\text{eff}}\epsilon_0\mathbf{E}(\omega), \tag{5.2}$$

where $\chi^{eff} = \chi^{(1)} + 3\chi^{(3)}|E(z,t)|^2$. To see how this new effective susceptibility affects the material refractive index, we note that in the linear case the refractive index is related to the susceptibility by

$$n_0 = \sqrt{1+\chi^{(1)}}. \tag{5.3}$$

The nonlinear susceptibility adds a perturbation to the refractive index

$$n_0 + \delta n = \sqrt{1+\chi_{\text{eff}}}. \tag{5.4}$$

We now rearrange and expand in order to solve for $\delta n$

$$\delta n = \sqrt{1 + \chi^{(1)} + 3\chi^{(3)} |E(z,t)|^2} - \sqrt{1 + \chi^{(1)}}. \tag{5.5}$$

To simplify this further, we rearrange it into

$$\delta n = \sqrt{1 + \chi^{(1)}} \left( \sqrt{1 + \frac{3\chi^{(3)} |E(z,t)|^2}{1 + \chi^{(1)}}} - 1 \right). \tag{5.6}$$

If we let $x = \frac{3\chi^{(3)}|E(z,t)|^2}{1+\chi^{(1)}}$ we can then Taylor expand $y = \sqrt{1+x}$. Keeping only the first two terms we see that

$$\delta n \simeq \frac{3\chi^{(3)} |E(z,t)|^2}{2n_0}. \tag{5.7}$$

Taking the intensity of the plane wave to be $I = 2n_0\epsilon_0 c |E(z,t)|^2$ we arrive at

$$\delta n \simeq \frac{3\chi^{(3)} I}{4n_0^2 \epsilon_0 c}. \tag{5.8}$$

It is convention to define a new nonlinear parameter, $n_2$ [2], such that the total refractive index is now $n = n_0 + n_2 I$ and $n_2$ is related to the nonlinear susceptibility by

$$n_2 = \frac{3\chi^{(3)}}{4n_0^2 \epsilon_0 c}, \tag{5.9}$$

where we have now shown that a pump beam travelling through a nonlinear material changes the refractive index of the material by an amount proportional to the intensity of the beam. This effect is known as self-phase modulation. It is natural to now ask what happens when two waves propagate together. How does the second wave see this change in refractive index? We now define a bi-chromatic input field $\mathbf{E}(\omega_1, \omega_2) = \mathbf{E}(\omega_1) + \mathbf{E}(\omega_2)$ where $\mathbf{E}(\omega_i) = E_i(z,t)\, e^{i(k_i z - \omega_i t)}$. The new nonlinear polarisation becomes

$$\mathbf{P}(\omega_2) = \chi^{(1)}\epsilon_0 \mathbf{E}(\omega_2) + 6\chi^{(3)}\epsilon_0 |E_1(z,t)|^2 \mathbf{E}(\omega_2) = \chi_{\text{eff}}\epsilon_0 \mathbf{E}(\omega_2), \tag{5.10}$$

2 There are many conventions on how exactly to define this parameter, a full breakdown of the different definitions and their relationships are given in Refs. [57, 190]

where we have only kept the terms affecting the propagation of $\mathbf{E}(\omega_2)$ and have omitted the self phase term. Performing a similar analysis to above we arrive at $n = n_0 + n_2^{\text{cross}} I_1$ with $I_1$ being the intensity of the light at $\omega_1$ and

$$n_2^{\text{cross}} = \frac{3\chi^{(3)}}{2n_0^2\epsilon_0 c} = 2n_2, \tag{5.11}$$

where we now see that the second beam also sees a change in refractive index proportional to the intensity of the first, but the constant of proportionality is twice that we found before. This effect is known as cross-phase modulation. Over a given length, this change in refractive index results in a phase shift given by

$$\phi(L) = \frac{2\pi L}{\lambda}\tilde{n}_2 I, \tag{5.12}$$

where $\tilde{n}_2$ is $n_2$ and $n_2^{\text{cross}}$ for self- and cross-phase modulation respectively.

### 5.2.2 *Including transverse mode profile*

The above analysis demonstrates the phenomenon of self- and cross-phase modulation in an intuitive way however there are many assumptions underlying the derivation which miss some subtleties that will be important in understanding the results presented later in this chapter. To rectify this, we will derive the same effects more rigorously. To start we, once more, define a bi-chromatic input field

$$\mathbf{E}\left(\omega_p, \omega_s\right) = \mathrm{E_p}\left(x, y, z, t\right) e^{i\left(\beta_\mathrm{p} z - \omega_\mathrm{p} t\right)} + \mathrm{E_s}\left(x, y, z, t\right) e^{i(\beta_\mathrm{s} z - \omega_\mathrm{s} t)} + \text{c.c.}, \tag{5.13}$$

where $\mathrm{E_i}\left(x, y, z, t\right) = \mathrm{F_i}\left(x, y\right) \mathrm{A_i}\left(z, t\right)$. Here we have also included the transverse profile of the field, $\mathrm{F_i}\left(x, y\right)$ and the spatio-temporal amplitude $\mathrm{A_i}\left(z, t\right)$. We have replaced the labels 1 and 2 for the different frequencies with the subscripts p and s, for pump and signal. This change of convention will make it easier to map this derivation to the later results. The polarisation of the material, in the presence of a $\chi^{(3)}$ nonlinearity is then

$$\mathbf{P}_{\text{tot}} = \chi^{(1)}\epsilon_0\mathbf{E} + \chi^{(3)}\epsilon_0\mathbf{EEE} = \mathbf{P}_\mathrm{L} + \mathbf{P}_\mathrm{NL}. \tag{5.14}$$

Previously, we expanded the nonlinear polarisation, however, we only focused on one term at a time. Now, we consider all the ways in which the original frequencies can evolve. To do this, we ignore any frequency generating terms, but keep all terms where $\omega_{ijk} = \omega_i - \omega_j + \omega_k$ and $\omega_i = \omega_j$. This leaves us with

$$\begin{aligned} \mathrm{P_{k,NL}} = &\underbrace{3\epsilon_0\chi^{(3)}_{\mathrm{ksss}}\mathrm{E_s}\mathrm{E_s^*}\mathrm{E_s}e^{i(\beta_\mathrm{s}z-\omega_\mathrm{s}t)} + 3\epsilon_0\chi^{(3)}_{\mathrm{kppp}}\mathrm{E_p}\mathrm{E_p^*}\mathrm{E_p}e^{i\left(\beta_\mathrm{p}z-\omega_\mathrm{p}t\right)}}_{\text{self-phase modulation}} \\ &+ \underbrace{6\epsilon_0\chi^{(3)}_{\mathrm{kssp}}\mathrm{E_s}\mathrm{E_s^*}\mathrm{E_p}e^{i\left(\beta_\mathrm{p}z-\omega_\mathrm{p}t\right)} + 6\epsilon_0\chi^{(3)}_{\mathrm{kpps}}\mathrm{E_p}\mathrm{E_p^*}\mathrm{E_s}e^{i(\beta_\mathrm{s}z-\omega_\mathrm{s}t)}}_{\text{cross-phase modulation}}, \end{aligned} \tag{5.15}$$

where $\mathrm{P_{k,NL}}$ is the $k^{\text{th}}$ component of the nonlinear polarisation, $k, s, p \in \{x, y, z\}$ are the polarisation direction of the material polarisation, signal and pump fields and $\chi^{(3)}_{\mathrm{ijkl}}$ is the corresponding element of the $\chi^{(3)}$ tensor. Generally, we will consider light that is all polarised in the same direction and therefore will only need one element of this tensor. To determine the dynamics of these waves we look for solutions of the non-linear wave equation

$$\nabla^2\mathbf{E_i} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{E_i} = \mu_0\frac{\partial^2}{\partial t^2}\mathbf{P}_{\mathrm{tot}}, \tag{5.16}$$

where $\mathbf{E_i}$ represents the $i^{\text{th}}$ frequency component of the field $\mathbf{E}$. Without the perturbation to the polarisation, $\mathbf{P}_{\mathrm{NL}}$, this would result in $N$ independent uncoupled differential equations for an input field with $N$ frequency components. Specifically, for a waveguide system, where we have constraints on the transverse profile, the equations describe the evolution of the amplitude $\mathrm{A_i}(z,t)$. As we have seen, the presence of this perturbation couples the propagating fields together and therefore also couples these differential equations. For our case, where we have ignored frequency generating processes these coupled equations take the form

$$\begin{aligned} \frac{\partial \mathrm{A_p}}{\partial z} + \sum_{m\geq 1}\frac{i^{k-1}\beta_{k,p}}{k!}\frac{\partial^k \mathrm{A_p}}{\partial t^k} &= i\left(\gamma_{pp}P_p|\mathrm{A_p}|^2 + 2\gamma_{ps}P_s|\mathrm{A_s}|^2\right)\mathrm{A_p} \\ \frac{\partial \mathrm{A_s}}{\partial z} + \sum_{m\geq 1}\frac{i^{k-1}\beta_{k,s}}{k!}\frac{\partial^k \mathrm{A_s}}{\partial t^k} &= i\left(\gamma_{ss}P_s|\mathrm{A_s}|^2 + 2\gamma_{sp}P_p|\mathrm{A_p}|^2\right)\mathrm{A_s}, \end{aligned} \tag{5.17}$$

where we have ignored nonlinear losses and have employed the slowly varying envelope approximation, which assumes $\frac{\partial^2 \mathrm{A_i}}{\partial z^2} \ll \frac{\partial \mathrm{A_i}}{\partial z}$ and we can, therefore, ignore the second

derivative terms. We have introduced the initial power of each field such that the intensity of the field is given by $I_i = P_i|A_i|^2/A_{\text{eff}}$ where $A_{\text{eff}}$ is the effective mode area of the field. The nonlinear coupling coefficient $\gamma_{ij}$ is given by

$$\gamma_{ij} = \frac{3\omega_i n_{g,i} n_{g,j}}{4\epsilon_0 c^2} \frac{\iint_{\text{wg}} \mathrm{F}_\mathrm{j}^*(x,y)\, \chi^{(3)}_{ijji} \mathrm{F}_\mathrm{i}(x,y)\, \mathrm{F}_\mathrm{i}^*(x,y)\, \mathrm{F}_\mathrm{j}(x,y)\, dxdy}{\left(\iint_\infty n^2(x,y)|\mathrm{F}_\mathrm{i}(x,y)|^2 dxdy\right)\left(\iint_\infty n^2(x,y)|\mathrm{F}_\mathrm{j}(x,y)|^2 dxdy\right)}, \quad (5.18)$$

where $F_i(x,y)$ is the transverse mode field profile of the $i^{\text{th}}$ field. Full derivations of these equations are reported in Refs. [63, 191]. So far we have only made two assumptions which require justification. Firstly, we ignored the frequency generating terms. This is a reasonable assumption to make as the frequency generating terms are very sensitive to the wavelengths of the fields and waveguide parameters. As the created and destroyed fields are always the same for self- and cross- phase modulation, these terms do not have these same requirements. This means that it is easy to select a pump wavelength and waveguide geometry in a way to suppress the generation terms. Secondly, we have assumed that nonlinear loss is negligible. This is only valid at either low powers or when the sum of the energies of the two pump photons is less than the band gap of the nonlinear material, $2\hbar\omega_p \leq E_{\text{bandgap}}$. This is not true, for example, for devices made in silicon using telecom wavelength photons [87, 192]. However, as we plan to fabricate devices from silicon nitride, a material which has a much larger bandgap than silicon (and more than three times the combined energy of two telecom photons), this is a reasonable assumption for us to make. Even with these two assumptions we see that solving the coupled equations in Eq. 5.17 is non-trivial and, in general, can only be done numerically. There are, however, two further assumptions we can make that will make solving these equations simpler. Firstly we define a dispersion length, $L_D = T_0/\beta_2$ which quantifies the second order dispersion effects in the waveguide. If this length is longer than the length of the waveguide then we can safely ignore the time dependence of $\mathrm{A_i}$. For the waveguide geometries and pulse widths we consider this number is 2 to 3 orders of magnitude larger than the waveguide length, making this a valid assumption. The second assumption we make is that $\mathrm{A_s} \ll \mathrm{A_p}$. As we are designing this modulator

with eyes towards quantum applications with single photons, this is also a reasonable assumption. With all of these employed the coupled equations become

$$\begin{aligned}\frac{dA_p}{dz} &= i\gamma_{pp}P_p|A_p|^2A_p,\\ \frac{dA_s}{dz} &= i2\gamma_{sp}P_p|A_p|^2A_s.\end{aligned} \tag{5.19}$$

These simplified equations have analytic solutions

$$\begin{aligned}A_p &= |A_p|e^{i\gamma_{pp}P_p|A_p|^2z},\\ A_s &= |A_s|e^{i2\gamma_{sp}P_p|A_p|^2z}.\end{aligned} \tag{5.20}$$

From these we can calculate the additional phase accumulated for a wave propagating for a length $L$ to be $\phi_{SPM}(L) = \gamma_{pp}P_p|A_p|^2L$ and $\phi_{XPM}(L) = 2\gamma_{sp}P_p|A_p|^2L$. Writing the XPM phase explicitly we have

$$\phi_{XPM} = 2P_p|A_p|^2L\frac{3\omega_s n_{g,s}n_{g,p}}{4\epsilon_0c^2}\frac{\iint_{wg}F_p^*\chi^{(3)}_{sppp}F_sF_s^*F_p dxdy}{\left(\iint_\infty n^2(x,y)|F_s|^2dxdy\right)\left(\iint_\infty n^2|F_p|^2dxdy\right)}. \tag{5.21}$$

Later we will look at the case where the pump and signal are both in transverse electric (TE) modes in both silicon and silicon nitride waveguides. For TE modes most of the field propagates within the waveguide and for both silicon and silicon nitride the majority of the nonlinear interaction takes place in the core. Therefore, $\chi^{(3)}$ and $n^4$ can be moved outside the integrals. We can then use the definition of the non-linear refractive index, Eq. 5.9 to re-write the non-linear phase shift as

$$\phi_{XPM} = 2P_p|A_p|^2Ln_2\frac{2\pi}{\lambda_s}\frac{n_{g,s}n_{g,p}}{n^2}\frac{\iint_{wg}F_p^*F_sF_s^*F_p dxdy}{\left(\iint_\infty |F_s|^2dxdy\right)\left(\iint_\infty |F_p|^2dxdy\right)}, \tag{5.22}$$

The overlap integral now defines the inverse of the effective interaction area. Therefore, in the situation with no dispersion, and therefore $n_{g,s} = n_{g,p} = n$, we recover the result from Sec. 5.2.1.

## 5.3 CIRCUIT DESIGN

We have seen that the presence of a strong laser pulse can impart a phase shift on a weaker pulse propagating in the same waveguide. This opens the door to creating photonic switches that use light as the switching mechanism. In the following section we will describe the process of designing an integrated all-optical switch based on an MZI design with an XPM phase shifter in one of the arms. Specifically, we will aim to use a pump and signal which propagate in different transverse waveguide modes. In principle the same analysis could apply to any pair of transverse modes [3], however for our device we wish to have the signal propagating in the fundamental transverse electric mode (TE0) and the pump in the first order transverse electric mode (TE1). The reasoning behind this decision is two-fold. Firstly, for rectangular strip waveguides, the transverse electric modes have the majority of the field is confined to the waveguide core. For most integrated photonic platforms, this is where the highest nonlinearity resides. This implies that for a given length we should expect higher phase shifts for transverse electric modes, as opposed to the transverse magnetic modes where the majority of the mode propagates in the cladding. Secondly, the choice to have the pump propagate in the higher order mode means we can employ low loss and broadband mode converters to multiplex and demultiplex the pump from the phase shifted arm of the MZI. We are predominantly designing devices for silicon nitride (SiN). While the nonlinear refractive index for SiN is an order of magnitude lower than silicon (Si), the lack of two-photon absorption at telecom wavelengths allows for higher peak powers. In order to determine the optimal switch geometry there are several factors that need to be considered - the nonlinear mode overlap, effective interaction length, the expected XPM phase shift, and how to excite specific transverse modes. In the following sections we take a closer look at each of these factors.

### 5.3.1 *Nonlinear mode overlap*

In this section we look at the behaviour of different transverse modes and the overlap integrals between them as a function of waveguide geometry. Using higher order modes makes it likely that unwanted mode mixing in the device as could be a large problem. In a straight unperturbed waveguide each supported mode propagates with

3 In fact later on we will discuss the possibility of using higher order pairs of modes

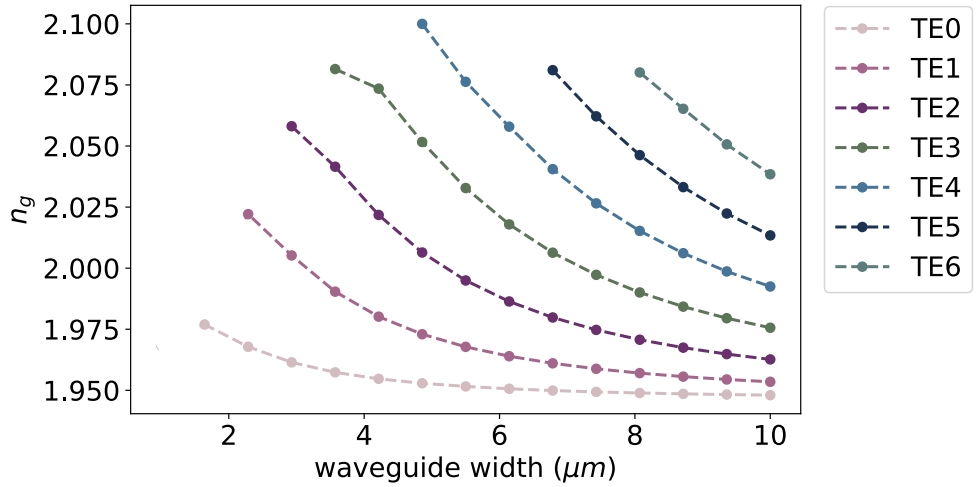


Figure 5.1: Plot of group index vs waveguide width for higher order transverse modes in silicon nitride waveguides.

its own distinct propagation constant and there is zero power transfer between the modes. Perturbations to the waveguide, such as changes in refractive index, sidewall roughness and waveguide bending, modify the propagation constants of the guided modes resulting in a non-zero overlap between modes. As a consequence, the different modes are now coupled with a coupling coefficient that depends on two factors, the degree of perturbation - i.e. how much the propagation constants are changed, and the original difference between the propagation constants. Multimode waveguides must be wide in order to accommodate the higher order modes, this reduces the overlap between the mode and the sidewalls. Therefore, it is the waveguide bends that will provide the largest perturbation to the mode propagation constants. We can minimise this by using larger bend radii. As mentioned previously, the degree of waveguide perturbation is not the only contributing factor; the initial difference in propagation constants between the modes is also important. To the first order, this is given by $\beta_1 = 1/v_g = n_g/c$. Figure. 5.1 shows how the group index, $n_g$ varies as a function of waveguide width for higher order modes. We see as the waveguide width increases the difference between the group indices of higher order modes decreases. From this we can see that larger waveguide widths require larger bend radii. To minimise mode coupling while also minimising bend radii, we, initially, restrict ourselves to only the fundamental and first order transverse mode. We look at the overlap integral in Eq. 5.22,

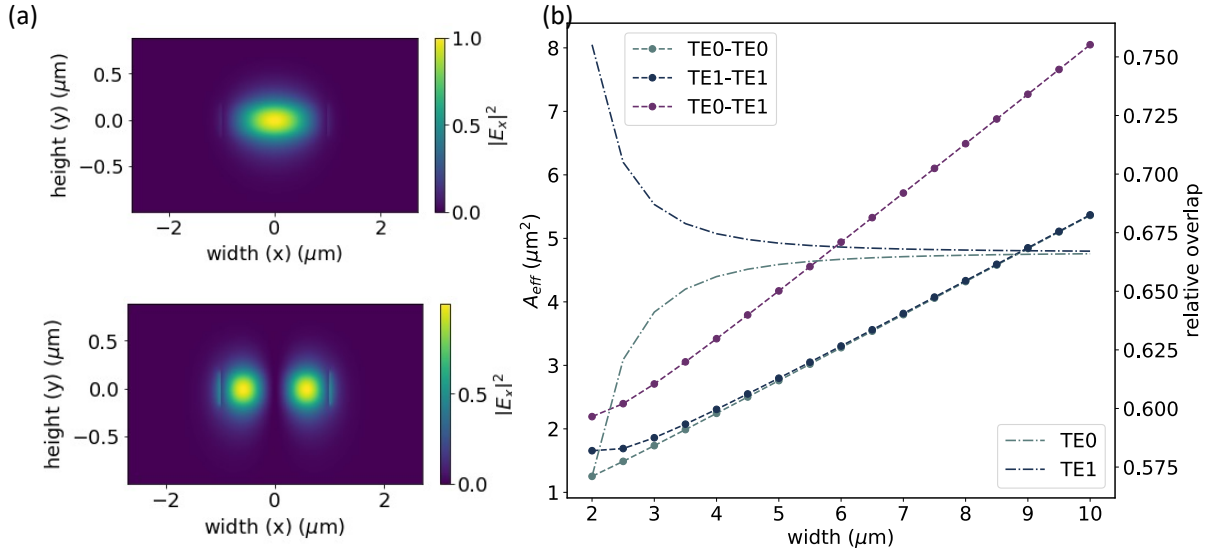


Figure 5.2: (a) Intensity profiles for the fundamental and first order transverse modes. (b) Effective area comparison. Circles and dashed line show the effective areas of the two individual modes and of the combined modes. The values for these are shown on the left-hand axis. The right-hand axis, and the dot-dashed lines correspond, to the relative overlap between the nonlinear area and the individual modes

specifically we plot the inverse of this overlap which gives us the effective mode area of the nonlinear interaction

$$A_{eff} = \frac{\left(\iint_{\infty} |\mathrm{F_s}|^2 dxdy\right)\left(\iint_{\infty} |\mathrm{F_p}|^2 dxdy\right)}{\iint_{\mathrm{wg}} \mathrm{F_p^* F_s F_s^* F_p} dxdy}. \tag{5.23}$$

In Fig. 5.2(a) we show the intensity profile of the two modes in question. Fig. 5.2(b) we plot both $A_{eff}$ (dots and dashed lines) for both the individual modes and for the combined nonlinear interaction area. We see that the nonlinear interaction area is always higher than the modal area of the individual modes, corresponding to the lower efficiency of these intermodal processes. We also quantify this difference by plotting the relative overlap, that is the ratio between the nonlinear interaction area and the effective mode area of the individual modes. In general, as expected, narrower waveguides result in lower effective area and therefore a higher phase shift. Although, as we mentioned previously, there are some technical problems with moving to higher order transverse modes, perhaps there is some gain that would make them worthwhile? To this end, we plot all the pairwise overlaps for the first 7 transverse modes in Fig. 5.3 Here we

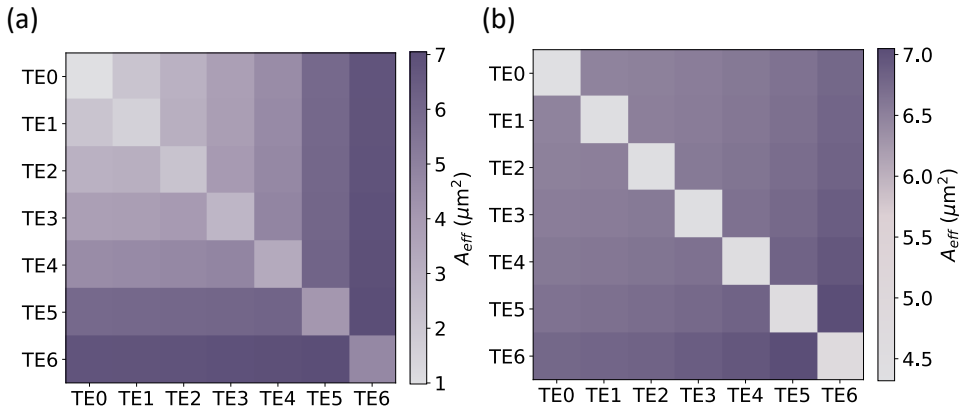


Figure 5.3: Plots showing the pairwise nonlinear effective areas between the first seven higher order transverse modes. (a) Here the waveguide width for each pair is taken to be the lowest width that supports both modes (b) Here the waveguide width for all pairs is the same (8 µm).

have chosen two different ways to interpret this question. In Fig. 5.3(a), the waveguide width for a given square is taken to be the lowest value that would support the two modes. From this we see that increasing the mode order increases the effective area and therefore decreases the phase shift. This, however, could simply be an artefact of the fact that the larger waveguide widths required for the higher order modes, necessarily result in a lower confinement of the light and therefore a lower intensity. To combat this and isolate effects purely of the differences in spatial distribution of the fields, we fix the waveguide width for all calculations to be 8 µm. These results are shown in Fig. 5.3(b). The trend that lower mode orders have lower effective areas remains and we are, therefore, justified in our decision to use only the lowest two mode orders.

### 5.3.2 *Characteristic lengths*

We have already talked about a characteristic length that is important to consider for these simulations, that being the dispersive length, which allowed us drop the time dependence of the pulse amplitude. This is, however, not the only length scale we need to consider. As we are talking about two pulses propagating, in general, at different wavelengths and in different transverse modes, there is no guarantee that these pulses will propagate at the same velocity. Any difference in the group index of the two pulses

will result in the pulses walking off from each other, reducing the effective interaction. We can then define an effective interaction length, that being the length that two pulses with identical group velocities would need to propagate in order to result in the same interaction strength. The maximum interaction length is achieved when the pulse are delayed such that the pulses overlap maximally in the centre of the waveguide. With this caveat, the general form of this effective length is given by [193]

$$L_{eff} = \frac{\int_0^L \int_0^\infty I_1(z,t)\, I_2(z,t)\, dZdt}{\int_0^\infty I_1(z = L/2, t)\, I_2(z = L/2, t)\, dt}, \tag{5.24}$$

where $I_1(z,t)$ and $I_2(z,t)$ are the intensities of the two pulses. This equation can be simplified for the extreme walk-off regimes. To quantify these regimes, we introduce a dimensionless quantity [193]

$$\Gamma = \sigma_{eff}\tau = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}|\Delta\beta_1|L, \tag{5.25}$$

where $\Delta\beta_1 = \beta_1(\omega_1, TE0) - \beta_1(\omega_2, TE1)$, and $\sigma_i$ is the standard deviation of the Gaussian pulse. To convert this to be in terms of the FWHM of the pulse in the time domain we note that, for a Gaussian pulse the FWHM in time is related to the standard deviation in angular frequency as

$$\sigma_i = \frac{0.44 \times 2\pi}{2\sqrt{2\log 2}\delta T_i}. \tag{5.26}$$

Inserting this into Eq. 5.25, we arrive at

$$\Gamma = \frac{0.44 \times 2\pi}{2\sqrt{2\log 2}} \times \frac{1}{\sqrt{\delta T_1^2 + \delta T_1^2}}|\Delta\beta_1|L. \tag{5.27}$$

This final form helps us to intuitively understand the two regimes. For $\Gamma \ll 1$, we have minimal walk off. This occurs for small $\Delta\beta_1$, long pulses, or short waveguides. Here, the interaction length can be considered to be the waveguide length, $L$. The second regime, $\Gamma \gg 1$ is where we have maximal walk off. Here, the fast pulse is delayed so that there is initially minimal interaction and as the pulses propagate the fast pulse walks through the slow pulse. The delay is chosen so maximum interaction is achieved

at $L/2$ and, again, there is minimal interaction at $L$. In this case, which occurs for large $\Delta\beta_1$, short pulses, or long waveguides, the effective length becomes [193]

$$L_{eff} = \frac{\sqrt{2}}{\Gamma} L. \tag{5.28}$$

To see how the geometry of our waveguide and pulse choices affects the effective length we first plot $\Gamma$ to see which regime we will be in. Figures 5.4(a) and (b) show $\Gamma$ parameter

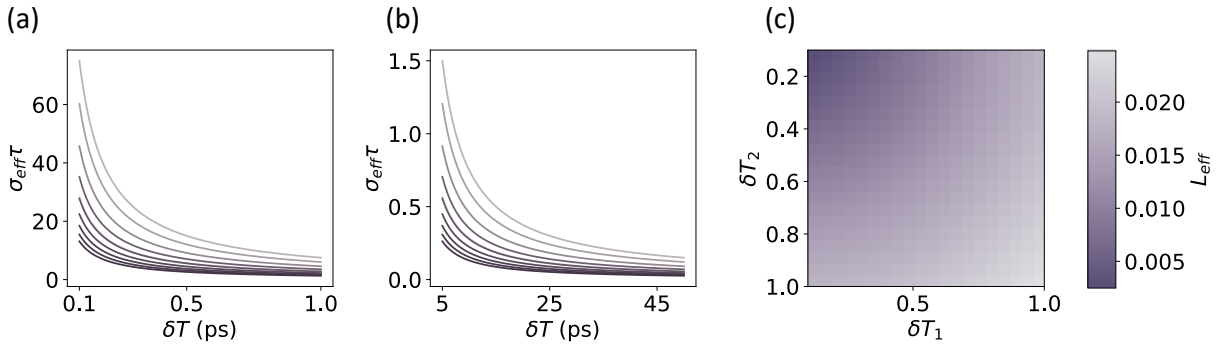


Figure 5.4: (a) and (b) $\Gamma$ as a function of pulse width where $\delta T_1 = \delta T_2$. Different colours correspond to different waveguide widths. Lightest (darkest) shade corresponds to 8.2 µm (2.3 µm). (c) $L_{eff}$ for differing pulse widths. Waveguide width fixed to 2 µm

as a function of pulse width. Here we have set $\delta T_1 = \delta T_2$ and a waveguide length of 3 cm, this reflects the experimental setup we use later. We have also included differing waveguide widths. We can see that for small pulse widths, we are in the second regime however there is a cut-off at approximately a few ps where we transition to $\Gamma < 1$. As for our setup we wish the pump intensity to be as high as possible, short pump pulses are desirable. However, this has the added effect of reducing the interaction length. One way to combat this may be to use shorter pump pulses but longer signal pulses. This could increase the interaction length but allow high pump intensities. To illustrate this we plot $L_{eff}$ as a function of both $\delta T_1$ and $\delta T_2$. This is shown in Fig. 5.4(c) where we have fixed the waveguide width and length to be 2 µm and 3 cm, respectively. From this, we see that the interaction length can be increased by lengthening one pulse with respect to the other.

### 5.3.3 *Simulating phase shift*

In this section, we look to combine all these factors to simulate the phase shift we can expect for different setups. The free parameters we have are the waveguide width and waveguide length, before fabrication and then the pulse parameters after fabrication. For the following simulations we fix the laser repetition rate to 50 MHz and the average power to 20 mW. We set the power in Eq. 5.22 to be the peak power of the pulse, this is calculated as

$$P_{\text{peak}} = \frac{P_{\text{avg}}}{RR \times \delta T}, \tag{5.29}$$

where $RR$ is the laser repetition rate. We fix the wavelengths of the signal and pump to be 1555 nm and 1560 nm, respectively. As we have seen in the previous sections, with our waveguide geometries and pulse widths we are around the transitional point $\Gamma$. To account for this, we take the phase shift length to be the minimum of the waveguide length and $L_{eff}$. We initially sweep the waveguide width, shown in Fig. 5.5(a). For this,

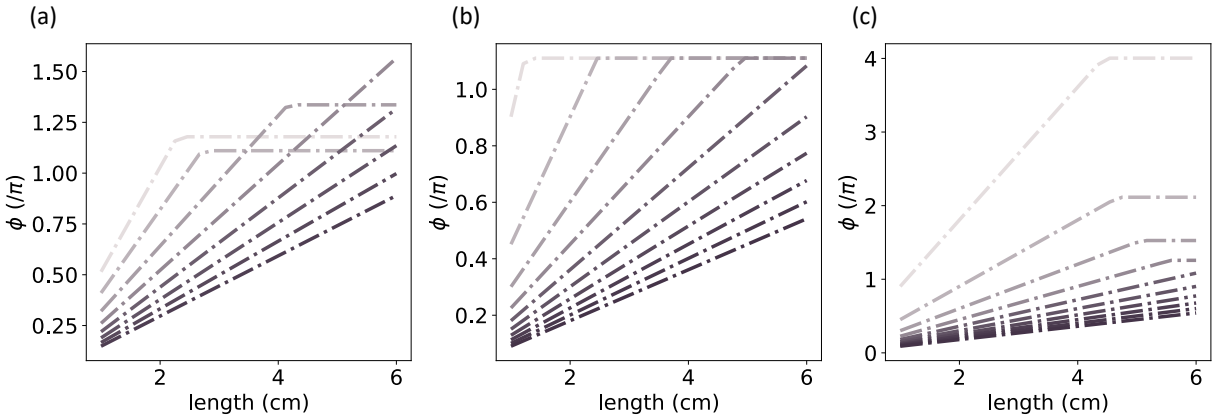


Figure 5.5: Simulation of phase shift from cross-phase modulation. All plots show phase shift as a function of waveguide length. Phase shift is shown as a fraction of $\pi$. (a) Different lines show different waveguide widths. Darker colour indicates a wider waveguide. We show 8 widths evenly spaced from 2 µm to 9 µm. Pulse width of both pulses set to 2.2 ps. (b) and (c) Waveguide width fixed to 3 µm. Each line corresponds to a different pulse width with darker shades corresponding to longer pulses. We show 10 lines evenly spaced from 1 ps to 10 ps (b) Both pulse widths the same (c) Signal width fixed to 5 ps

we fix the pulse width of both signal and pump pulses to be 2.2 ps. This value is chosen as it is the corresponding temporal FWHM of a Gaussian with a FWHM of 200 GHz, the

width of the standard DWDM filters we use [4]. We see that for the narrower waveguide widths we see a saturation length. For shorter waveguide lengths $L_{eff} > L$ and we take the interaction length to be the waveguide length. After the saturation point $L_{eff} < L$ and there is no gain in phase shift by increasing the waveguide length. This saturation point is lower for smaller waveguide widths. This is understandable as for narrower wavelengths the difference in propagation constants, $\Delta\beta_1$ is larger and the resulting effective length is shorter. We see that we are able to achieve phase shifts above $\pi$ for realistic waveguide lengths and widths. In fact, with this pulsing setup, even waveguide widths up to 7 µm show phase shifts above $\pi$ for a few cm of waveguide. In Figs. 5.5(b) and (c) we look at the effect of pulse widths. To this end, we fix the waveguide width to be 3 µm. We look at pulse widths between 1 ps to 10 ps as this is a range that is experimentally accessible. In Fig. 5.5(b) we keep the pump and signal pulse widths the same, we see that the shorter pulses saturate the maximum phase shift quickly as the waveguide length overtakes the effective length. Regardless of pulse width, we are limited to phase shifts of around $\pi$. In general, the reason we desire short temporal pulses is for the increase in peak power. However, there is no reason that the signal pulse also needs to be narrow. As we have seen before the effective length can be increased by lengthening only one pulse. In Fig. 5.5(c) we scan the pump pulse width again, however now we fix the signal pulse at 5 ps. We see now that the effective length is increased, moving the saturation point and allowing phase shifts well in excess of $\pi$.

### 5.3.4 *Mode converters*

In order to multiplex and demultiplex the pump field from one arm of the MZI, we require a method to selectively exciting the desired transverse waveguide mode. Many options exist for this, such as off-centre edge coupling, designed to couple directly into a specific mode, [194], asymmetric trenches [195] or tailored grating structures [196, 197]. The method we choose to use are asymmetric directional couplers. Our main reasoning is that these are low loss, broadband and allow the pump pulse to be multiplexed into and out of the MZI arm. These couplers work on the same principle as standard directional couplers (see Sec. 2.5.1). However, in order to achieve power transfer into higher order modes we need to match the propagation constants of the fundamental

4 We appreciate that the spectrum after filtering will be closer to a top hat than a Gaussian but this value will do for a starting point

mode of one waveguide with the desired higher order mode in the other. This condition, known as phase-matching, consists of matching the effective refractive indices of the two modes

$$n_{\text{eff},1}\,(TE0) = n_{\text{eff},2}\,(TE1)\,. \tag{5.30}$$

With this condition fulfilled, the power transfer between the modes can be tailored by changing the gap and the length of the coupler. The power transfer for a given coupler can be calculated analytically

$$P_{\text{t}}(L) = P_0 \sin^2\left(\frac{\pi L \Delta n}{\lambda_0}\right), \tag{5.31}$$

where $\Delta n$ is the difference in effective indices of the two supermodes of the coupled waveguides, and is a function of the gap between the waveguides. For mode converters, full power transfer is desired, that is $P_{\text{t}}(L) = P_0$. The minimum length that achieves this can be calculated as

$$l_{\text{tot}} = \frac{\lambda_0}{2\Delta n}. \tag{5.32}$$

The difference in the effective indices of the supermodes decreases as the distance between the waveguides increases, leading to an increase in the total coupling length, as expected. We show the results of these simulations in Fig. 5.6. In Fig. 5.6(a) we show the phase matching conditions for a TE0-TE1 coupler. We fix the initial width of the TE0 waveguide to be the single mode waveguide width of 1.3 µm. With this we arrive at a phase-matched TE1 waveguide width of 2.9 µm. In Fig. 5.6(b) we sweep the separation between the directional coupler waveguides and plot the corresponding power crossover length. As expected, increasing the gap increases the crossover length. The selection of which gap to use is a balancing act - increasing the gap increases the tolerance of the splitting ratio to fabrication defects but it also increases the size of the component. As a middle ground, we select the separation to be 0.3 µm which results in a coupling length of 70.6 µm.

### 5.3.5 *Final design*

Figure 5.7 shows a simplified schematic and the final design which was fabricated. To measure the phase shift we will look at interference in an MZI and will include mode converters in both arms to balance loss. We also include a slight asymmetry in the MZI

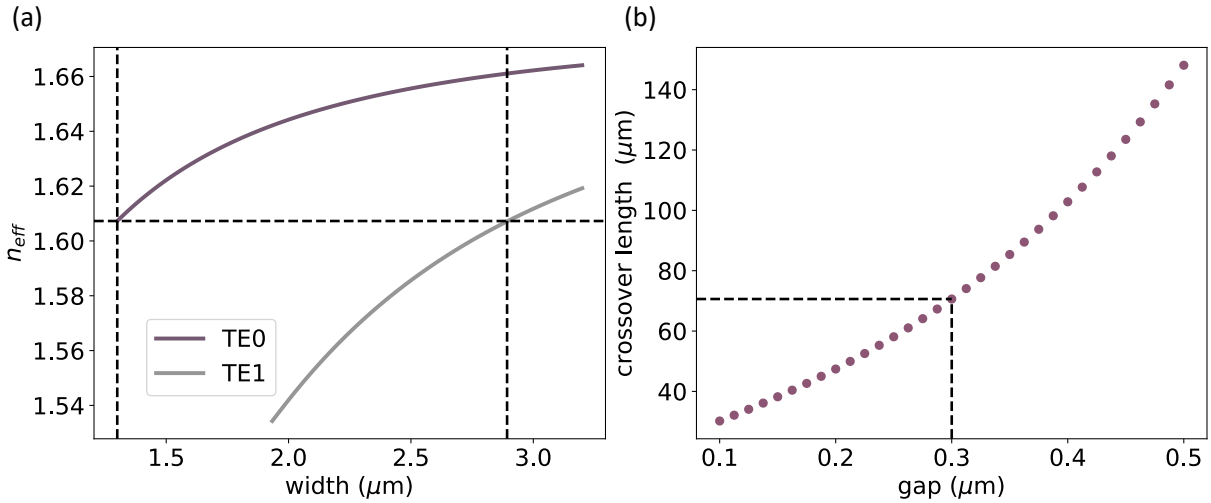


Figure 5.6: Simulation of mode converters. (a) Phase matching condition for TE0 and TE1 modes. We fix the width of the TE0 waveguide to be the single mode width 1.3 µm the corresponding TE1 waveguide width is 2.9 µm. (b) Power crossover length as a function of coupler separation with waveguide widths determined in (a). We will use a gap of 0.3 µm which gives a coupling length of 70.6 µm

paths. This allows us to frequency tune the signal pulse to be at either identity or swap, meaning we only need to aim to reach a $\pi$ phase shift. In this section we briefly describe the decisions leading to the final device layout. The starting point is the waveguide width. This is a trade-off - decreasing the width will decrease $A_{eff}$, however it also

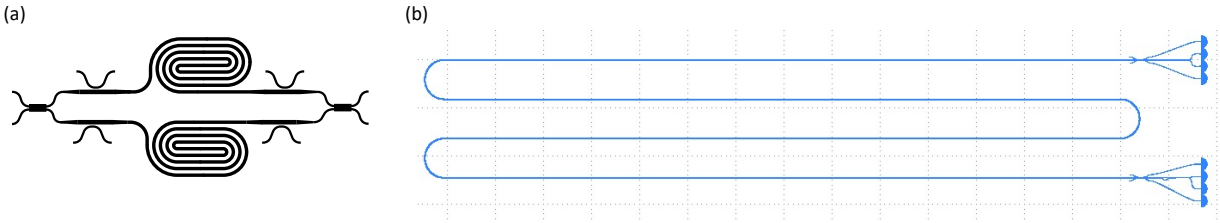


Figure 5.7: (a) Simplified schematic of the final devices. MZI with mode multiplexing and demultiplexing in each arm. Only one arm will be used, the other mode converters are included to balance loss. (b) Final design which was fabricated. Full etched silicon nitride device with grating couplers.

increases the mode overlap with the sidewalls, and therefore the loss of the device, as well as increasing $\Delta\beta_1$ and decreasing $L_{eff}$. For these initial tests, the primary aim is to demonstrate a phase shift. We therefore aim to maximise the mode confinement, choosing the waveguide width to be 2 µm, the narrowest that supports both modes.

After the device is fabricated, we have some control over the exact pumping scheme we use and, as we have seen, the combination of pulse widths is the primary decider of the interaction length. With this in mind, we select a waveguide length that is somewhere in the middle of the range we investigated, that being 3 cm. The final component to place are the mode converters. As we have seen, for our chosen gap of 0.3 µm we should set the coupling length to be 70.6 µm. This, however, assumes the coupling turns on and off instantaneously. In reality, we have to bend the waveguides into the coupling region and the coupling starts adiabatically. To account for this we need to perform full 3D FDTD simulations of the bending regions. This allows us to estimate the contribution of the bent regions and we can adjust the coupling region appropriately. We find that the bending regions contribute $\approx 10\%$ coupling. The corrected coupling length therefore becomes 55 µm.

## 5.4 EXPERIMENTAL RESULTS

### 5.4.1 *Experimental setup*

To experimentally probe the resulting XPM phase shift we look to generate two, spectrally distinguishable, laser pulses [5]. One will be used as the signal to be switched, the other will be amplified before being used as the pump pulse. To ensure that each pair of pulses overlap properly in the waveguide, we will carve both pulses from one broadband laser pulse. We use the same technique as shown in Ref. [14]. A seed pulse is generated by a PriTel wavelength tuneable laser. The laser pulse is centred at 1558 nm and emits pulses with a 2.3 nm bandwidth. We then use a PriTel temporal compressor which uses self phase modulation to spectrally broaden the pulse. We can then use a standard multi-channel DWDM to select two 200 GHz slices from the compressor output. One pulse is sent to two EDFAs - the first to ensure that the input power to the second is above its amplification threshold. The signal passes through a tuneable optical delay. Each pulse has its own polarisation controller and the pump path includes a 99/1 tap to allow us to monitor the input power. These pulses then are coupled via grating couplers onto the chip. The full setup is illustrated in Fig. 5.8. With the pulses

5 While in theory the pulses need only be distinguishable in one degree of freedom, in practice the filtering given by the mode converters is unlikely to be enough to reduce the pump level below the noise floor of either powermeters or single photon detectors

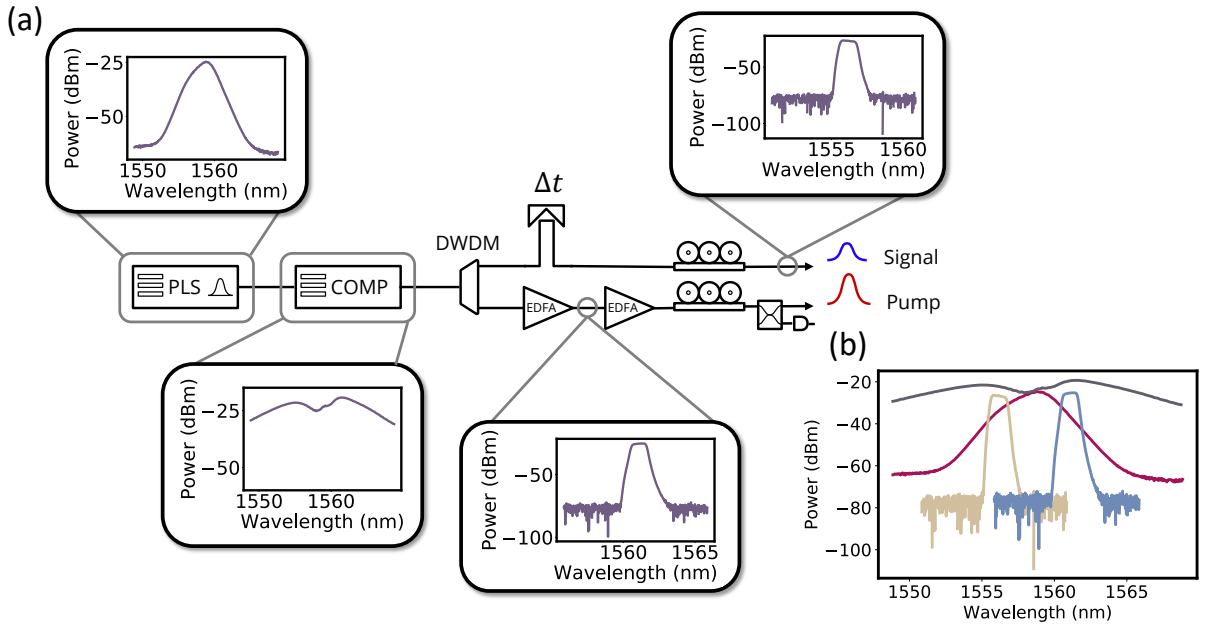


Figure 5.8: Here we show how pairs of spectrally distinguishable and temporally correlated laser pulses are generated. (a) Full schematic. A temporal compressor is used to spectrally broaden pulses from a seed laser. We then carve 200 GHz pulses symmetrically around the centre of the seed pulse using DWDM. One of these is then amplified by a chain of EDFAs. A tuneable optical delay is included in the signal arm in order for the pulses to be temporally matched. A 99/1 tap is used in the pump arm to monitor input power and polarisation controllers are used on both arms to maximise fibre-to-chip coupling. (b) Overlay of all pulses. Seed laser (red), broadened pulse (purple), signal pulse (beige) and pump pulse (blue).

generated we must now tackle the problem of how we ensure that pulses are properly temporally aligned in the waveguide. As mentioned previously, pulses travelling with different group velocities will walk off from each other. To maximise the interaction length we, therefore, need the pulses to be maximally overlapped in the centre of the waveguide. This delay can be calculated as [110]

$$\tau_{\text{opt}} = -\frac{L}{2}\left(\frac{1}{v_{\text{g,TE1}}} - \frac{1}{v_{\text{g,TE0}}}\right). \tag{5.33}$$

For our chosen waveguide length of 3 cm we find $\tau_{\text{opt}} \approx 4$ ps. In order to align the pulse with sub-picosecond resolution, we use an Oz Optics optical delay line. This delay line has the necessary resolution but only has maximum delay of 160 ps. We, therefore, need

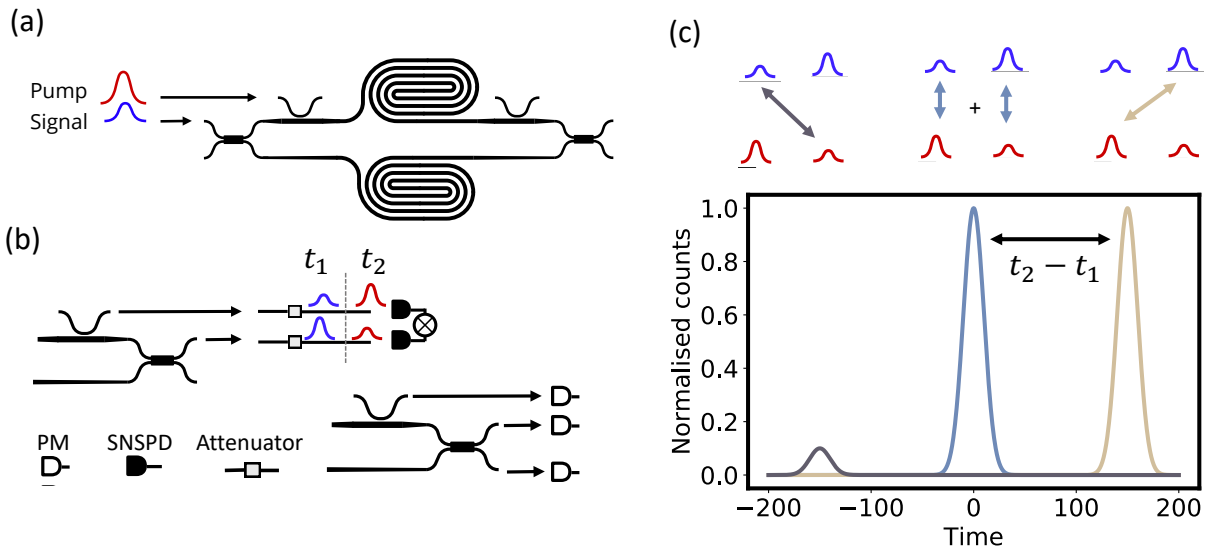


Figure 5.9: Measurement schematics for calibration and phase measurements. (a) Chip schematic showing pump inputs. (b) Setups for the two required measurements. Top shows how we use SNSPDs to coarsely overlap the pulses in the waveguide. We use the fact that the mode converters aren't perfect and therefore some of each pulse can be found in both the mode converter and MZI output. By attenuating these signals we can use SNSPDs and time tagging logic to measure a coincidence histogram. (c) Example histogram showing two dominant peaks. The first arises from pairs of photons generated in the same pulse. The second from pairs of photons generated in different pulses.

to balance the delay differences to within around 2 cm [6]. To coarsely align the pulses we use SNSPDS and a Swabian Ultra time tagger to record coincidence histograms. This lets us align the pulses to within the combined detector and time tagger jitter. As we saw in chapter 3, this is around 50 ps. The general protocol for this coarse alignment is shown in Fig. 5.9. The pump pulse is input to the mode converter input and the signal into the MZI input, Fig. 5.9(a). We then use the fact that the mode converters do not perfectly convert the mode, effectively acting as a highly reflective (transmissive) beam-splitter for the pump (signal) pulse. We can then attenuate the outputs and send the light to the SNSPDS, Fig. 5.9(b). A diagram of the resulting coincidence histogram is shown in Fig. 5.9(c). We can see from this that there are two dominant peaks. The central peak, at $t = 0$ corresponds to coincidence events where both photons came from the same pulse i.e. $t_1 \leftrightarrow t_1$ or $t_2 \leftrightarrow t_2$ events. The second peak corresponds to events

6 This is no mean feat due to our lab being generally only stocked with a few different lengths of fibre (0.33 m, 1 m, 2 m etc)

where the photons came from different pulses. Specifically, it is events between the higher pulse amplitudes - that is the part of pump (signal) that is reflected (transmitted). In the case we have illustrated in Figs. 5.9(b) and (c), this is the case in which the bottom mode clicks at $t_1$ and the top at $t_2$. There is a third peak that corresponds to the opposite events however this is usually lower than the histogram background. In Fig. 5.9(c) we have drawn these peaks to be the same height, however this is not always the case. The central peak depends on the sum of the total counts going through and can therefore be increased by increasing either of the pump or signal power. The second peak depends on the power in both pulses and therefore can be best tuned by changing the on-chip coupling. We note that the attenuation on each arm was also independently tuneable. By overlapping the two histogram peaks, the temporal difference between the pulses was within the range of the optical delay line. From here we use the splitting ratio of the MZI as our signal to optimise. Whenever the pulses are optimally delayed with respect to one another we should see the maximum change in the MZI splitting ratio. As both the signal and pump are pulsed, it is enough for us to look at the average power over time, rather than trying to resolve the laser pulse width. We therefore, remove the attenuation and swap the SNSPDs for power meters, Fig. 5.9(b). We also monitor the output pump power, which along with the off-chip power before the chip provides enough information to estimate the on chip power.

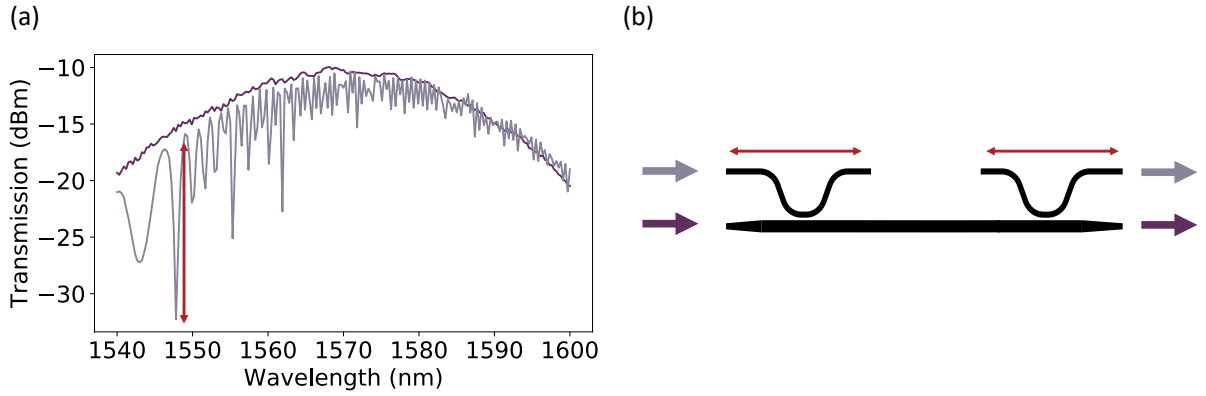


Figure 5.10: Testing the mode converter performance. (a) Measured transmission spectrum for the TE0-TE0 and TE1-TE1 powers. (b) Measurement schematic. Colours correspond to the plots in (a). Periodic oscillations in TE1-TE1 power comes from resonant effects due to unterminated waveguides.

### 5.4.2 *Results*

Initially, we characterise the performance of the mode converters. To do this we compare the power transmitted from the TE0 input to TE0 output to the power transmitted from the TE1 input to TE1 output. If the input and output coupling for both these paths are the same then any difference in the transmitted powers can be attributed to imperfect mode conversion. We show this plot in Fig. 5.10(a). We see oscillatory behaviour in the TE1-TE1 transmission. This is attributed to resonant behaviour in the mode converters due to unterminated waveguides. To determine the mode converter efficiency we therefore only use the peaks of these oscillations. We define the powers input to the mode converter (grey arrow) and multimode waveguide (purple arrow) as $P_{\mathrm{i,TE1}}$ and $P_{\mathrm{i,TE0}}$, respectively. The output powers are similarly named $P_{\mathrm{o,TE1}}$ and $P_{\mathrm{o,TE0}}$. The mode converter maps the TE0 field to the TE1 as $E_{\mathrm{TE1}} = \sqrt{\eta} E_{\mathrm{TE0}}$, where $\eta$ is the converter transmission. The transmitted power at the TE0 and TE1 outputs are therefore

$$\begin{aligned} P_{\mathrm{o,TE0}} &= \epsilon_{i,TE0}\epsilon_{0,TE0} P_{\mathrm{i,TE0}} \\ P_{\mathrm{o,TE1}} &= \epsilon_{i,TE1}\epsilon_{0,TE1}\eta^2 P_{\mathrm{i,TE1}}, \end{aligned} \tag{5.34}$$

where the $\epsilon_{j,k}$, with $j \in \{i, o\}$, represents the input/output coupling efficiency for the $k^{\mathrm{th}}$ mode, and the $\eta^2$ represents that there are two mode converters. If we assume the input and output couplings are the same for each mode, $\epsilon_{j,TE0} = \epsilon_{j,TE1}$, and that $P_{\mathrm{i,TE1}} = P_{\mathrm{i,TE0}}$, then we can write the transmission as

$$\eta = \sqrt{\frac{P_{\mathrm{o,TE1}}}{P_{\mathrm{o,TE0}}}}. \tag{5.35}$$

With this method we find a mode converter transmission of 0.92. It is hard, via this method to determine if the drop in transmission comes from the assumptions we have made or from problems in the simulation or fabrication process. Now, we move on to determine the temporal profiles of the pump pulses. As our pulses may not be transform limited, spectral intensity measurements are not sufficient to determine the temporal shape of the pulse. To properly reconstruct the spectral and temporal shape of the pulses we use a frequency resolved optical grating (FROG). A FROG is an autocorrelator that uses nonlinear optics (in our case second harmonic generation) in order to determine both the spectrum and spectral phase of an input pulse. This

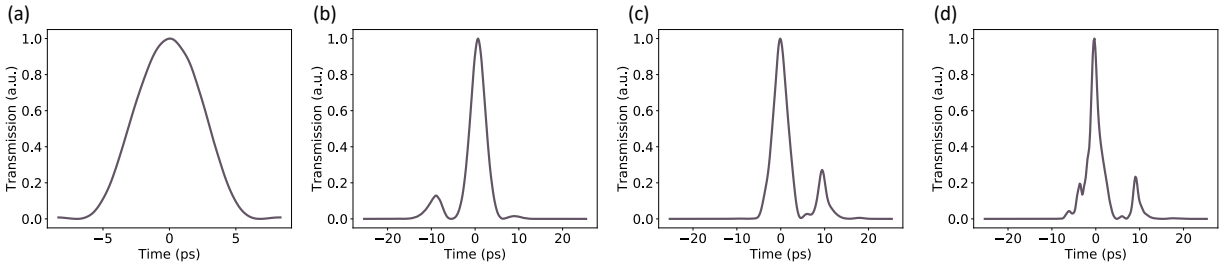


Figure 5.11: Temporal intensity profiles of the pulses after the waveguide. (a) Signal pulse (b) - (d) Pump pulses at different EDFA currents (b) EDFA current 760 mA (output power approximately 19 mW) (c) EDFA current 1040 mA (output power approximately 54 mW) (b) EDFA current 1250 mA (output power approximately 76 mW)

information allows the full temporal profile of the pump to be reconstructed. We measure the pump and signal pulses after they have travelled through the chip. We perform this measurement for a variety of different EDFA currents. As the signal is unaffected by the EDFA we only show one of these measurements. These results are shown in Fig. 5.11. Figure. 5.11(a) shows the reconstructed temporal intensity profile of the signal pump. We find a FWHM of $\approx 6$ ps. Note, as expected, this is much larger than the 2.2 ps expected for a Gaussian pulse with the same spectral FWHM as the filter. Figures. 5.11(b) - (d) show the signal pulses for three different EDFA currents, corresponding to average output powers between approximately 19 mW to 76 mW. We aim to maximise the peak power of the pulses, therefore we opted to use an EDFA with a larger amplification range (30 dB). This EDFA is not optimised for pulses with high peak powers, and therefore we expect to see a large amount of self-phase modulation due to the longer fibre inside the EDFA [7]. This SPM is evident in Figs. 5.11(b) - (d). We see as we increase the EDFA current the pulse width reduces. At the lower powers the temporal width is $\approx 4$ ps. The extra lobes are due to the Fourier transform of the spectral profile of the pump which includes the resonant effects of the mode converters. We now turn our attention to measuring the phase shift. To do this, we need to find the optimal delay between the pulses. This delay is given by $\tau/2$, where $\tau$ is defined in Eq.5.25. For our setup with $L = 3$ cm and waveguide width of 2 µm, we find this delay to be 4.14 ps. To find this, we use the tuneable optical delay using the splitting ratio of the switch, defined as $P_1/(P_1 + P_2)$, with $P_i$ being the power at the $i^{\text{th}}$ MZI output. By scanning the delay and measuring the splitting ratio, the optimal delay is

7 This will have the effect of narrowing the pulse in time and increasing the intensity - a benefit in our case

then given by the delay that maximises the change in splitting ratio (with our choice of $P_1$ this corresponds to minimising the splitting ratio). We perform this sweep for a range of EDFA currents. These dips are shown in Fig. 5.12(b). In Fig. 5.12(a) we plot the

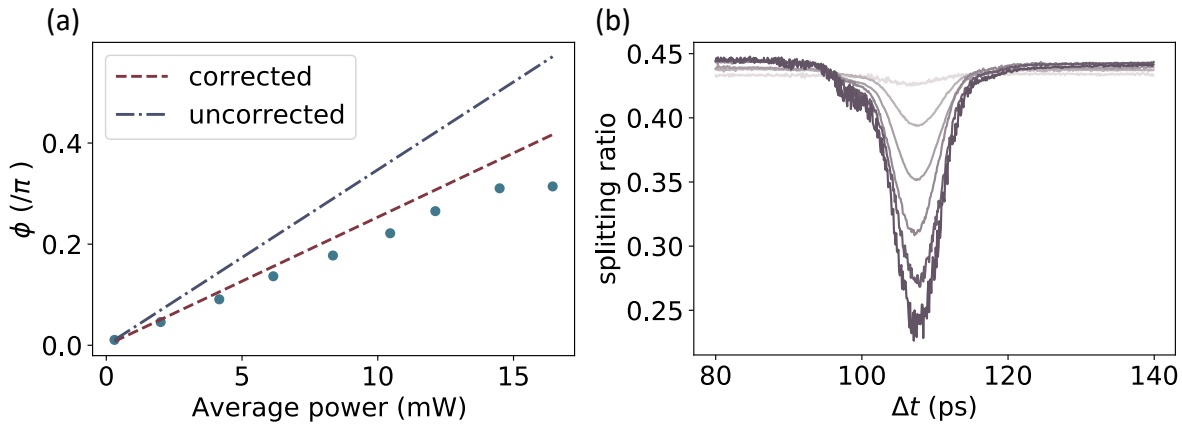


Figure 5.12: (a)splitting ratio at the dip minimum plotted against estimated waveguide power. Theory lines included with and without correcting for the extra lobes on the pulse. (b) Splitting ratio as a function of delay. Darker colours correspond to higher EDFA currents.

change in splitting ratio of each dip versus the estimated power in the waveguide. As expected for small values of $\phi$ where $\sin\phi \approx \phi$, we see the phase increases linearly with power. We also include a comparison to the theory predictions. The blue line shows the theory line as we defined previously, where the peak power is calculated using the average power and linewidth. We see some discrepancy between the theory and data lines. This could be attributed to the extra lobes in the temporal intensity profile, as seen in Fig. 5.11. These extra lobes result in a measured average power that is not all contributing to the switching pulse. A rough correction for this can be acquired by scaling the average power based on the relative heights of the two peaks. With this the peak power becomes

$$P_{\text{peak}} = (1-p)\frac{P_{\text{avg}}}{RR \times \delta T}, \tag{5.36}$$

where $p$ is the height of the extra lobe, relative to the main peak. The height of the second lobe changes with the EDFA power. We therefore choose a value in the middle of the amplification range we used. We measured the relative height of the lobe to be 0.27. The theory line, with this correction in place, is shown in red. We see that, with this correction included, our model is now in better agreement with the data. There is still

some discrepancy remaining, this could be related to the simplicity of our correction. A more likely reason, however, is our choice of interaction length. With our pulse parameters and waveguide width, we calculate $\Gamma$ to be 0.68. This puts us in the regime where the effective length, given complete walk-off of the pulses is larger than the waveguide length. We therefore take the interaction length to be the waveguide length. However, as we mentioned this is only strictly true when $\Delta\beta_1 = 0$, and therefore $\Gamma = 0$. As we still have a finite but not complete walk-off of the pulses, we overestimate the interaction length. Another reason we may have overestimated the interaction length is due to losses in the waveguide. In nonlinear optical processes, where the strength of the interaction depends on the intensity of light in the waveguide, we can define another length scale due to losses [63]

$$L_{NL} = \frac{1 - e^{-\alpha L}}{\alpha}, \tag{5.37}$$

where $\alpha$ is the attenuation coefficient of the waveguide. This length can be thought of the equivalent length of a lossless waveguide that would result in the same interaction strength. This length is shorter than the waveguide with loss and therefore as we haven't taken this into account, we have again overestimated the interaction length. To include this properly we would need to properly characterise the waveguide propagation loss and, unfortunately, these preliminary fabrication runs did not have structures capable of performing these measurements. A keen eye may have noticed that, as the laser power is increased, the dips we have plotted start to become noisy and less smooth. This trend continues as the power increases. We plot the temporal dip for the highest EDFA current, corresponding to an average power at the EDFA output of ≈85 mW, in Fig. 5.13(a), where we see that the fluctuations in the splitting ratio almost obscure the dip completely. Initially we need to determine whether this is a process happening in the waveguide, i.e. it is dependent on the absolute power in the waveguide, or whether it is happening before the chip, likely in the EDFA. To this end we repeat the above measurements but with some added loss before the waveguide. We see that the instability starts at the same EDFA current, which now corresponds to a lower power in the waveguide. This shows that the fluctuations are likely coming from the EDFA. In order to find the cause of these fluctuations we also plot the power in each of the MZI outputs. This allows one to tell whether the fluctuations in the power are correlated or anti-correlated. If the powers are correlated then this is likely a result from power leaking from the pump into the signal mode. If they are anti-correlated then this is likely

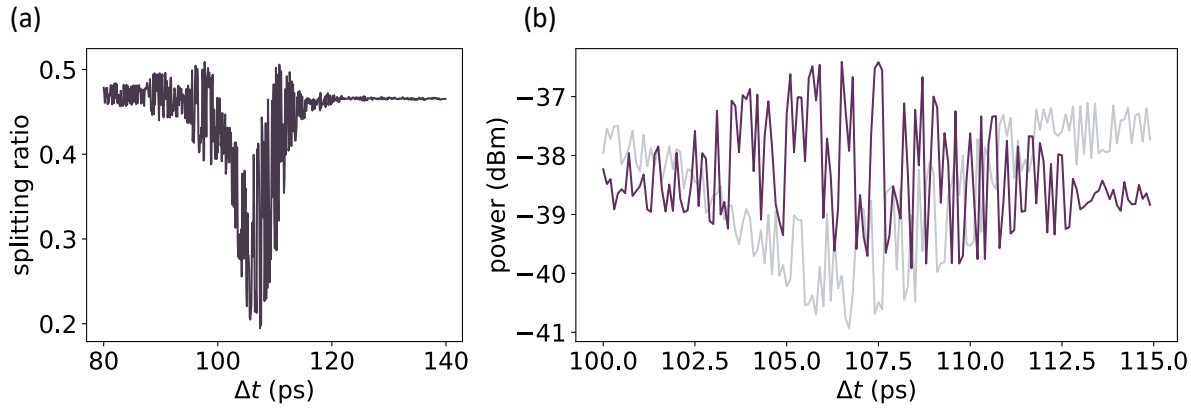


Figure 5.13: (a) Splitting ratio versus temporal delay for the highest EDFA current (b) Measured power at each of the MZI outputs as a function of the temporal delay. We have shifted the purple line up by 2 dB so it overlaps with the grey line

a phase error due to changes in the pump power. These measured powers are shown in Fig. 5.13(b), where we have zoomed in on the bottom of the dip and have shifted one of the powers, so they overlap. We see in places the powers are potentially anti-correlated but in many places they are uncorrelated completely. It is likely that the fluctuations are a combination of both effects. At high amplification, modulation instability inside the EDFA which leads to the generation of spectral side-bands. These side-bands, which appear in the time domain as amplitude modulation, cause fluctuations in the gain of the EDFA. This in turn causes fluctuations in the SPM which, due to the proximity of the pump and signal wavelengths, can leak into the signal mode. Once averaged over many pulses this would result in the powers being uncorrelated.

### 5.4.3 *Dual lasers*

One limiting factor with the experimental setup described above is that we are limited in the pulse widths we can use. As we carve the two pulses from a single spectrally broad pulse we are effectively limited by the bandwidth of the DWDM used. We saw that some help was afforded through SPM in the EDFA however this only allowed pulse widths of 4 ps. In principle, we could use a second compressor on the pump pulse after filtering, but we are limited here by the number of compressors in the laboratory. One thing we do have access to in the laboratory is a pair of PriTel ultrafast optical clocks[8].

8 One of these was used for the measurements in chapter 4

These lasers have a repetition rate that is governed by an external clock. If both lasers are driven by the same clock then they will be locked to generate pulses at the same time. Now that we have two separate lasers generating the pump and signal pulses we are free to manipulate the spectral properties of each laser separately. In this section we will look at whether this extra controllability gives enough flexibility to reach a $\pi$ phase shift with the current waveguide geometry. At first glance it would seem that this is an easy question to answer. With the current setup we managed to see a phase shift of $\phi > 0.25/\pi$ so if we can increase the spectral bandwidth by a factor of 4 then we should be able to reach $\pi$. However, the internal clock that is shipped with the UOCs, which we could use to trigger the second laser, is set to 500 MHz, an order of magnitude higher than the laser used in the previous measurements. This means for a given average power that can be supplied by the EDFA, the peak power is a factor of 10 lower for the UOC. Initial tests show that we should be able to get pulse widths of at least 1 ps and potentially lower. We can use the 200 GHz DWDM to filter the signal pulse. This is necessary to prevent any light from the pump that scatters into the TE0 mode from reaching the powermeters. Beyond this, having a longer signal pulse will increase the interaction length of the pulses. These simulations are shown in Fig. 5.14,

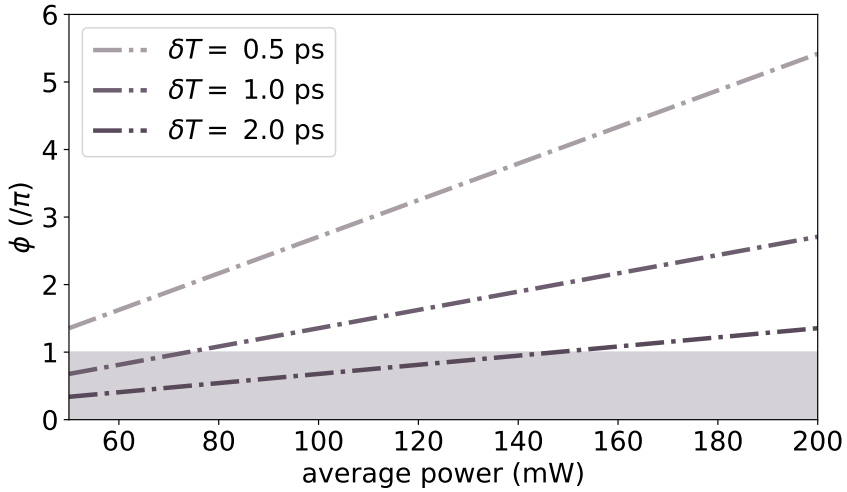


Figure 5.14: Simulations for double laser setup. Signal pulse with set to 5 ps. Shaded line shows $\pi$ phase shift. $\Gamma$ calculated to be 0.9, 0.95, 0.97 for the pulse widths of 2 ps,1 ps, and 0.5 ps, respectively.

where we have set the signal pulse width to 5 ps. Preliminary tests have shown that the standard EDFAs in our laboratory are capable of generating up to at least 120 mW. We

would therefore expect with the higher gain EDFA, used in the previous measurements, that the powers we have plotted to be realistic. We see from this plot that phase shifts of above $\pi$ would be reachable. This should be caveated with the fact that the narrower pulses bring us even closer to the $\Gamma$ transitional point ($\Gamma > 0.9$ for all of the pulse widths plotted). This means the effective length is likely less than we have estimated. To determine this fully, the integral in Eq. 5.24 would need to be evaluated.

## 5.5 DISCUSSION AND OUTLOOK

In this chapter, we have seen how cross-phase modulation can be used effectively to make an all-optical switch. We used a pump and signal pulses in different transverse modes. This gives us an extra degree of freedom to filter the pump but, more importantly, it allows us to easily make low loss and broadband multiplexers. We designed and tested a silicon nitride device. Comparing to simulations we see good agreement between theory and experiment. This will allow future devices to be designed with more confidence. With the first generation of device we managed to implement a phase shift of $0.3\pi$, short of the $\pi$ shift required for quantum photonic applications. We have discussed one possible option for increasing the phase shift by using two lasers locked together. We are currently in the process of testing this possibility. Currently, we measure a fibre to fibre transmission through the chip of $\approx -13$ dB, mainly due to the angle of the fibre array not being that which the grating couplers were designed for. It is likely we can improve this by at least 3 dB and in future devices we could consider using edge couplers for higher coupling efficiencies. Another limiting factor was the extra lobes in the temporal profile, coming from resonant effects due to unterminated waveguides in the mode converters, this can be completely eradicated in future devices. It would also be interesting to look at the possibility of using different waveguide geometries such as rib waveguides. So far we have considered only devices in silicon nitride. The main reasoning here was the lower linear and nonlinear loss, however, the nonlinear parameter in silicon is a factor of 10 larger. Beyond this, due to the increased index contrast, much smaller waveguides can be made with much higher mode confinement. Ignoring the losses, this results in much higher phase shifts for silicon devices. The linear losses in silicon can be reduced, although often at the cost of increasing the mode area, however at telecom wavelengths two photon absorption will limit the maximal pulse powers that can be reached. A key feature of cross-phase modulation is that there

is no requirement on phase matching meaning that in principle we can use any pump wavelength. There are two relevant non-linear loss processes: two-photon absorption where two pump photons excite over the bandgap, depleting power in the pump, and cross two-photon absorption where one signal and one pump photon excite across the bandgap, depleting both the pump and the signal. The silicon bandgap is $\approx 1.1$ eV, if the signal photon is at 1550 nm then a pump pulse above 4160 nm will not provide enough energy to excite electrons across the bandgap resulting in cross-TPA. Cross-TPA becomes less probable with low photon number signal pulses. If we wish to combat two-photon absorption then pump pulses around 2200 nm are sufficient [70, 87]. It would have to be seen how the dispersion at these higher wavelengths would affect the spectrally broad pulses that would be required, however, this could be a route to fast switching in silicon devices. Using silicon would reduce the requirements of peak power for a $\pi$ phase shift, which is a large problem with extending this switching method to higher speeds - that being managing to amplify to high enough average powers.

### ACKNOWLEDGEMENTS AND CONTRIBUTIONS

This work was conceived of and designed by myself. I implemented the simulations and designed the device. The device was fabricated by Dr. Francesco Lenzini, Emma Lomonte and their colleagues at the university of Münster. The experiment was performed by myself and Dr. Alex Jones with input from Alexandre Maïnos and Jacob Bulmer. The project was supervised by Prof. Anthony Laing and Prof. Wolfram Pernice.

# 6

# COHERENCE OPTIMISATION THROUGH $\mathcal{PT}$-SYMMETRIC HAMILTONIANS

"Zaphod scribbled a few sums, crossed them out and threw the pencil away."

Douglas Adams, *Hitchhikers Guide to the Galaxy*

## 6.1 INTRODUCTION AND MOTIVATION

One of the main results of quantum mechanics, the Schrödinger equation, tells us that the time dynamics of a closed quantum system are entirely decided by the *Hamiltonian*, $H$, acting on the Hilbert space of the system. One of the central postulates of quantum mechanics asserts that $H$ must be *Hermitian*. That is, for a finite dimensional system, $H = H^\dagger$. Hermiticity enforces a real set of eigenvalues on $H$. Given that the eigenvalues of $H$ correspond to the possible results of energy measurements on a system, this seems like a sensible postulate. Beyond this, as the time dynamics of a system are given by exponentiating $H$, a Hermitian $H$ will result in dynamics that are unitary. Unitary dynamics preserve energy, as expected for a closed system.

The above argument seems sound, and you would be forgiven for thinking that was the end of the story and that we need not worry about systems described by non-Hermitian Hamiltonians. This is, however, not the case. In reality, most quantum systems interact too strongly with their environments to be described as closed systems. In an open quantum system, that is a desired quantum system coupled to an environment that we do not, or more often, cannot fully describe, the dynamics can be described, at least in part, by a non-Hermitian Hamiltonian. Here, the complex eigenvalues and non-unitary evolution correspond to energy flowing either from or to the environment.

So, it seems that non-Hermitian systems can be comfortably studied within the traditional framework of quantum mechanics. However, the Hermiticity postulate is just that, a postulate that cannot be derived from first principles. This still leaves the question open of whether we can formulate quantum mechanics and do away with the Hermiticity. This question was answered in Ref. [198] by Bender et al. who

showed that there are other classes of Hamiltonian that also possess real eigenspectra. Specifically, Bender et al. look at a class of Hamiltonians that are symmetric under combined parity and time reversal, known as $\mathcal{PT}$-symmetry. This was later expanded to show that it is possible to use this class of Hamiltonians to construct quantum theories with real spectra and unitary evolution [199]. Whether or not these quantum theories represent physical closed quantum systems remains a topic of debate [200, 201]. However, they most definitely can be used to describe sub-systems of larger closed Hermitian systems. In this context, these systems have been shown to possess many desirable qualities [202–210].

In this work, we look at the ability of quantum $\mathcal{PT}$-systems to increase the coherence of a qubit by driving it towards a more coherent target state. We simulate a parameterised $\mathcal{PT}$-system on a linear optical chip capable of implementing arbitrary unitary transformations on 6 modes. We use an adaptive feedback loop to tune the system parameter and optimise the qubit coherence. In the coming chapter, we initially introduce the description of open quantum systems before moving on the specific example of $\mathcal{PT}$-symmetric Hamiltonians. Next, we turn to the experimental setup, before describing how we can use the chip's unitary operation to simulate the required non-unitary dynamics. We then look at the experimental results and, finally, at a more comprehensive numerical study of the algorithm.

## 6.2 OPEN QUANTUM SYSTEMS AND THEIR DESCRIPTIONS

Open quantum systems permit several different, but equivalent, descriptions [211]. Here, we reproduce the Lindblad master equation analysis and show how these dynamics can be modelled, in part, by an effective non-Hermitian Hamiltonian. The Lindblad equation describes the time evolution of a general density matrix subject to Markovian noise, that is noise that is independent of previous times. The general form of the Lindblad equation is

$$\mathcal{L}(\rho) = \frac{d\rho}{dt} = \underbrace{-i[H,\rho]}_{\text{coherent dynamics}} + \underbrace{\sum_k 2L_k\rho L_k^\dagger - \sum_k \rho L_k^\dagger L_k + L_k^\dagger L_k \rho}_{\text{dissipative dynamics}}, \tag{6.1}$$

where $H$ is the Hermitian Hamiltonian describing the system and $L_k$ are called the Lindblad operators, which describe the coupling between the system and the environment. The above form of the Lindblad equation ensures that the $\rho$ remains physical, that is $\mathrm{Tr}\,(\rho) = 1$ and $\rho$ is positive.

In general, to simplify the master equation derivation for a specific system some approximations are required. Two common approximations are Markovianity which, as mentioned above, assumes no temporal correlations, and the Born approximation. This second approximation assumes that if the states are in a product state at $t = 0$ then they remain in a product state at time $t$. This states that quantum correlations between the system and environment are negligible. Both of these are realistic assumptions given that the environment size is much larger than the system size or the coupling between the system and environment is sufficiently small [212]. Eq. 6.1 can be thought of as the Hermitian (coherent) dynamics, given by term 1, with the two extra terms acting as a dissipative channel. To see how these same dynamics can be described by a non-Hermitian Hamiltonian we introduce the effective Hamiltonian

$$H_{\text{eff}} = H - i\sum_k L_k^\dagger L_k. \tag{6.2}$$

We can then redefine the Lindbladian dynamics as

$$\mathcal{L}(\rho) = -i\left(H_{\text{eff}}\rho - \rho H_{\text{eff}}^\dagger\right) + \sum_k 2L_k\rho L_k^\dagger. \tag{6.3}$$

From this we can see that the Lindbladian dynamics can be reproduced by a non-Hermitian effective Hamiltonian plus some extra terms, commonly known as the recycling or jump terms. The evolution by a non-Hermitian Hamiltonian alone does not preserve the trace of $\rho$. If subject to non-Markovian noise the norm of $\rho$ always decreases, and with non-Markovian noise it can also increase. The extra recycling terms, therefore, act to renormalise the state $\rho$.

## 6.3 INTRODUCTION TO PARITY-TIME SYMMETRIC HAMILTONIANS

In this section we introduce a specific class of non-Hermitian Hamiltonians, symmetric under combined parity and time reversal, known as $\mathcal{PT}$-symmetric Hamiltonians.

To understand these Hamiltonians we need to first understand the effects of parity reflection, $\mathcal{P}$ and time reversal, $\mathcal{T}$, operators

$$\begin{aligned} \mathcal{P} &: \{p \to -p,\, x \to -x,\, i \to i\}, \\ \mathcal{T} &: \{p \to -p,\, x \to x,\, i \to -i\}, \end{aligned} \tag{6.4}$$

where the complex conjugation that occurs with applying $\mathcal{T}$ is required to preserve the canonical commutation relation $[x, p] = i\hbar$. This also means that $\mathcal{T}$, and therefore also $\mathcal{PT}$, is conjugate-linear [213]. Some other properties of these operators are $\mathcal{P}^2 = \mathcal{T}^2 = (\mathcal{PT})^2 = \mathbb{1}$ and $[\mathcal{P}, \mathcal{T}] = 0$. A $\mathcal{PT}$-symmetric Hamiltonian is thus a Hamiltonian that is invariant under the combined action of these two operations

$$\mathcal{PT}H\mathcal{PT} = H. \tag{6.5}$$

We can see that if a Hamiltonian obeys Eq.6.5 then

$$\begin{aligned} (\mathcal{PT})(\mathcal{PT})H(\mathcal{PT}) &= (\mathcal{PT})H, \\ H(\mathcal{PT}) &= (\mathcal{PT})H, \end{aligned} \tag{6.6}$$

which therefore means $[\mathcal{PT}, H] = 0$. Despite this, due to the conjugate-linear nature of $\mathcal{PT}$, in general eigenstates of $H$ will not be eigenstates of $\mathcal{PT}$. If, however, the eigenstates of $H$ are also eigenstates of $\mathcal{PT}$ then it can be shown that the Hamiltonian eigenvalues are real and positive [199, 214]. For Hamiltonians where some of the eigenstates are not simultaneously eigenstates of $\mathcal{PT}$, the eigenvalues become complex. These two regimes are known as $\mathcal{PT}-$unbroken and $\mathcal{PT}-$broken, respectively. The transitional point is an exceptional point where at least two of the eigenvectors coalesce. This is different to other degeneracy points where two or more eigenvalues may be the same, but the corresponding eigenvectors remain distinct. In the following work we consider a particular set of $\mathcal{PT}$-symmetric Hamiltonians, parameterised by a single parameter, $\gamma$. Their general form is given by

$$H_N(\gamma) = -\sum_{k=1}^{N-1} \left[|k\rangle\langle k+1| + |k+1\rangle\langle k|\right] + i\gamma\,|1\rangle\langle 1| - i\gamma\,|N\rangle\langle N|\,. \tag{6.7}$$

This class of Hamiltonian can be thought of as describing a linear chain of $N$ modes with symmetric nearest neighbour coupling, which we have set to $-1$, and the first

and last modes having symmetric gain and loss, respectively. Specifically, due to the limitations of our experimental setup we will only look at the cases $N = \{2, 3\}$. For these system sizes we write the Hamiltonians explicitly

$$H_2(\gamma) = \begin{pmatrix} i\gamma & -1 \\ -1 & -i\gamma \end{pmatrix} \qquad H_3(\gamma) = \begin{pmatrix} i\gamma & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & -i\gamma \end{pmatrix}. \tag{6.8}$$

These have eigenvalues of $\lambda_2 = \pm\sqrt{1-\gamma^2}$ and $\lambda_3 = 0, \pm\sqrt{2-\gamma^2}$. We can see that for both of these there is a transitional point, at $\gamma_c = 1$ for the 2 mode case and $\gamma_c = \sqrt{2}$ for the 3 mode case, where the eigenvalues transition from purely real to imaginary. For the 2D case these eigenvalues lead to the eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ i\gamma - \sqrt{1-\gamma^2} \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 1 \\ i\gamma + \sqrt{1-\gamma^2} \end{pmatrix}, \tag{6.9}$$

where we see at the $\gamma_c$ these both coalesce to $\begin{pmatrix} 1 \\ i \end{pmatrix}$. A similar analysis of the 3 mode case shows that at the $\gamma_c$ all three eigenvectors coalesce. The parity operator for these Hamiltonians is given by the anti-diagonal permutation matrix

$$\mathcal{P} = \sum_{j=1}^{N} |j\rangle \langle N+1-j| . \tag{6.10}$$

## 6.4 EXPERIMENTAL SETUP

In the following section we describe the experimental setup used to perform the experimental tests in later sections.

### 6.4.1 *Generating photons with parametric down conversion*

We generate single photons in a bulk optics single photon source capable of generating up to four single photons from two independent two-mode squeezed vacuum states. A mode-locked Titanium:Sapphire laser generates 140 fs wide pulses at a rate of approximately 80 MHz. We tune the laser to emit at around 808 nm. To alter the

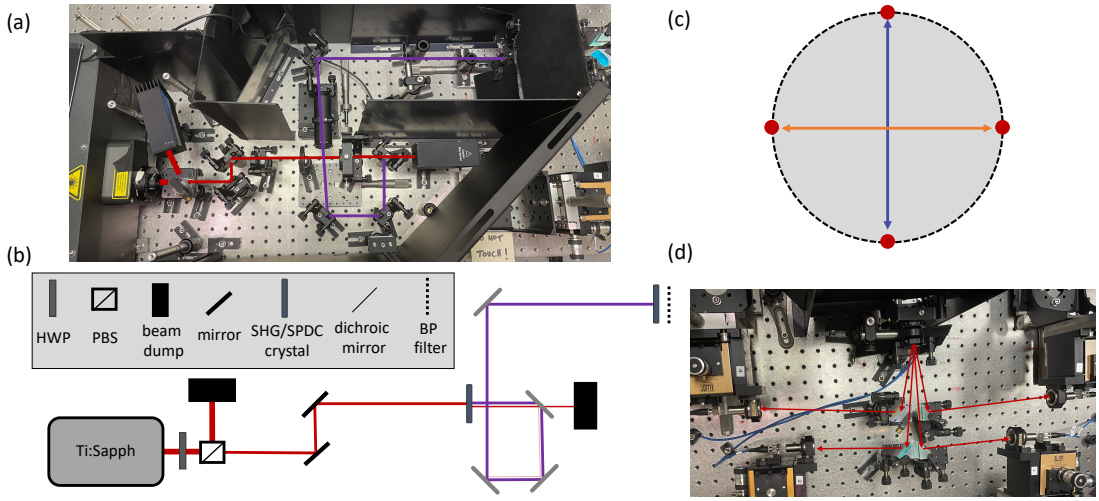


Figure 6.1: Experimental setup for generating single photons. (a) Photograph showing the pump preparation. (b) Schematic for the pump preparation. Titanium:Sapphire laser emitting femtosecond pulses at 808 nm. Half-wave plate (HWP) and polarising beam-splitter (PBS) are used to tune the laser power. Beam is directed to barium borate (BBO) crystal that performs second harmonic generation (SHG) resulting in a beam of 404 nm light. Any remaining 808 nm light is filtered by a series of dichroic mirrors before the pump is guided to bismuth borate (BiBO) crystal generating pairs of single photons through spontaneous parametric down conversion. (c) Two-mode squeezed states generated at symmetric points on a cone. We call the orange line source 1 and the blue line source 2. (d) Photograph showing the collection stages.

emitted power, and therefore give us fine control of the squeezing level in the sources, the laser first hits a half-wave plate and then a polarising beam-splitter. The light is then guided to a barium borate (BBO) crystal where it performs type-I second harmonic generation. A series of dichroic mirrors are used to filter any remaining 808 nm light while simultaneously guiding the 404 nm pulses to a second nonlinear crystal. This time a bismuth borate crystal (BiBO) which is used to generate two-mode squeezed states through type-I spontaneous parametric down conversion. A simplified schematic of this process is shown in Fig. 6.1(b) and a photograph of the full setup in Fig. 6.1(a) In both we have illustrated 808 nm in red and 404 nm light in purple. The generated squeezed states, which are degenerate in wavelength, are generated in spatially symmetric points around the SPDC emission cone. This allows us collect photons from two independent photon sources from separate squeezed states, as shown in Fig. 6.1(c). We refer to the horizontal source (orange) as source 1 and the vertical source (blue) as source 2. A photograph of the collection stage shows how angled mirrors are used

to isolate photons from the different sources. The collection stages have six degrees of freedom (three translational and three rotational) that can be used to optimise the fibre coupling. One collection stage on each source has a motorised translational stage on the longitudinal axis. This allows the temporal distinguishability of the photons to be tuned, an important feature for performing HOM interference. For a detailed description of the construction of this source the reader is directed towards Ref. [215]

### 6.4.2 *Reprogrammable interferometer*

We use a fully reprogrammable integrated photonic interferometer capable of implementing arbitrary unitary operations on up to six modes. This chip was fabricated and packaged by Nippon Telegraph and Telephone (NTT). The device consists of germanium doped silica waveguides with undoped silica cladding. Due to the low index contrast between the waveguide and cladding, the waveguides have dimensions of 3.5 µm × 3.5 µm. While this low index contrast means the device footprint is on the centimeter scale (10 cm × 4 cm), it also results in a large overlap between the waveguide mode in the edge coupler and the fibre mode and, therefore, a high input/output coupling. This, along with low propagation loss in silica, gives an end to end transmission of the device of ≈ 58%. The device is based on the Reck style decomposition (see Sec. 2.3.2) and consists of a triangular mesh of 15 MZIs and 15 external phases. All the thermo-optic phase shifters are constructed from 1.5 mm × 50 µm of tantalum nitride ($Ta_2N$) deposited on top of the waveguides. To improve the performance, and reduce crosstalk, trenches are etched either side of the heaters. As air has a lower thermal conductivity than silica this helps confine the heat to the waveguide. The MZIs consist of two directional couplers with a single thermo-optic phase shifter between. Figure 6.2 shows a simplified schematic of the chip. This image is simplified in a few ways. For clarity, we have omitted the path length matching that is used to make sure that photons from different inputs can interfere on any of the beam-splitters. Also, due to the size of the components, in order to fit the device onto a single wafer the circuit has been folded on itself so the inputs and outputs are on the same side. A large bend radius is used to minimise excess loss and mode mixing in the bends. The chip packaging includes glued fibres at the input and output, electrical packaging for the 30 heaters, and a Peltier heater to control the temperature of the chip. This Peltier heater is controlled by an

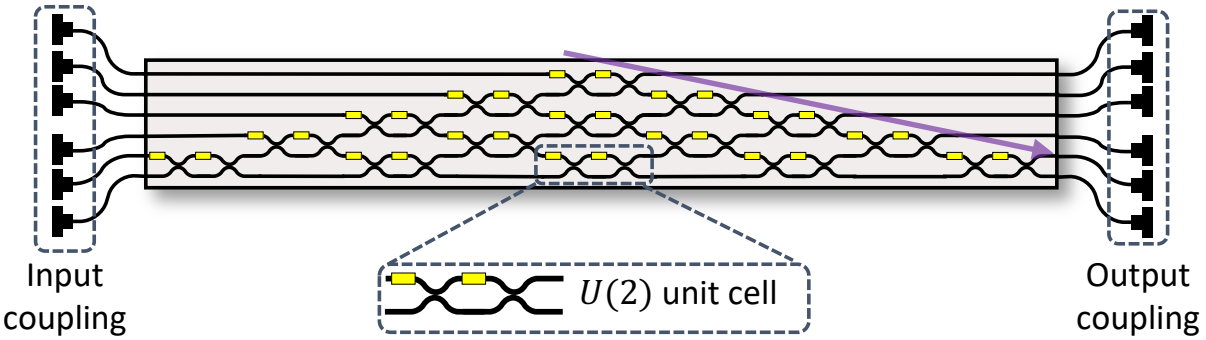


Figure 6.2: Schematic of the silica-on-insulator arbitrary linear optical interferometer used in this chapter. Triangular mesh of beam-splitters and phase shifters used to implement arbitrary unitary transformation over 6 modes. Inset shows an example of the individual $U(2)$ elements used to construct the device. MZI consisting of two directional couplers used to adjust amplitude between two modes and external phase shifter to adjust relative phase between the arms. We have also included an arrow showing the lead diagonal, the first set of beam-splitters that need to be characterised.

Arroyo PID temperature controller which is set to a temperature slightly above the ambient room temperature.

### 6.4.3 *Photon detection*

As we are operating at 808 nm, superconducting detectors are not the only option for high efficiency detectors. We employ silicon single photon avalanche photodiodes (SPAD) which can operate at room temperatures and have nominal efficiencies between 50% and 60%. Due to the symmetric waveguide dimensions, both TE and TM modes are guided equally. We employ polarisation maintaining fibre throughout, however chip bends and waveguide sidewall roughness can result in mode conversion. TE and TM modes do not interfere and so any mode conversion will appear as noise in the data. To combat this, we use a polarising beam-splitter (PBS) at each output only sending the desired polarisation to the detectors. In order to determine when photons are coincident on the detectors, we use a DPC-230 time correlated single photon counter (TCSPC) unit made by Becker and Hickle. This specific module has 16 low voltage TTL (LVTTL) inputs. The outputs of the SPADs are attenuated in order to bring the voltage level below the LVTTL threshold. This card returns 32 bit binary tags containing information

such as the channel and absolute time of arrival. This card was originally designed for fluorescence measurement however, through code written by Pete Shadbolt, it has been adapted for quantum photonic experiments. Due to large volume of time tags we receive for a relatively small number of desired coincidence events, saving all the data and post-processing is an undesirable option. The time tags are recorded and analysed in one second chunks. While the $n^{th}$ chunk is recorded, the $(n-1)^{th}$ chunk is analysed and only the coincidence events are saved. To allow for the relatively large detector and time tagger jitter ($\approx$ few hundred ps) the coincidence window is set to 2 ns.

### 6.4.4 *Characterisation and calibration*

In this section we look at benchmarking the performance of the experimental setup. First, we look at the linear optical characterisation of the chip and then at the quantum performance of the source. For the initial characterisation of the linear optics, a relationship between the desired phase and required voltage is necessary. Taking into account possible non-Ohmic response of the heaters and noting that the phase shift should be proportional to the dissipated power (voltage squared) we arrive at

$$\phi(V) = a + bV^2 + cV^3 \tag{6.11}$$

The coefficients can be extracted by fitting an interference fringe on the desired heater. For the MZI phases this is simply achieved by inputting single photons into one input and measuring photons at one output as the internal heater voltage is changed. For the external phases MZIs before and after are set to 50/50 constructing a larger MZI. In order to properly characterise a given heater within a larger optical network, it is important to make sure that there are no closed loops, apart from the interference path. This is equivalent to saying there should only be light incident on input of the MZI and that the light on the detector can only have left from one MZI output. Due to this, to fully characterise a universal interferometer, the order of characterisation matters. In a Reck-style interferometer, such as the one used here, the heaters are characterised along the diagonals. Starting with the longest diagonal, as shown in Fig. 6.2, all the heaters can be characterised without any assumptions on the starting phase of the uncharacterised heaters. For a more detailed discussion of the heater characterisation, the reader is directed to Refs. [214, 215]. With the characterisation complete, we can

now map a desired unitary matrix to a set of phases, thanks to the Reck decomposition (see Sec. 2.3.2), and then map these phases to voltages to be set on the on-chip heaters.

In an ideal world, this would be the end of the story. However, in the real world each phase shifter cannot be treated independently. In general, there are many effects that can lead to correlation between different phase shifters in an integrated interferometer. The two main process in a silica device are thermal crosstalk and electrical crosstalk. Thermal crosstalk was mitigated in this design by the inclusion of etched trenches either side of the heaters. Electrical crosstalk, however, remains an issue that needs to be accounted for. Electrical crosstalk stems from two main areas - parasitic resistances of the electrical connections and shared grounds between heaters. The parasitic resistances mean that the resistance from the heater to ground is non-zero, reducing the current flowing through a given heater. As well as this, the parasitic resistances also dissipate heat in unwanted places, changing the waveguide phase. The second problem, many heaters sharing a single ground connection, means that setting voltages on other heaters changes the effective potential of the ground. Initial crosstalk characterisation was performed by Jacques Carolan et al. where they measured the shared resistances and built a voltage model which included all these parasitic resistances [215]. This, therefore, accounts for the decrease in current flow through each heater, for a given voltage drop, and allows the potential shifts to be calculated, but does not account for the extra heat being dissipated. A further, more thorough, experimental analysis of the crosstalk was performed by Nicola Maraviglia which aimed to quantify these processes as well [214].

Once the individual heaters are characterised, it is important to benchmark the performance of the whole device. In order to check the ability of the device to implement a given unitary we must correct for imbalances in the output losses and detector efficiencies. To do this we implement a collection of random unitaries and use single photons injected into one of the modes to retrieve the absolute squared of one column of the dialled unitary. To compare the data, we then define a cost function as $C(d, p, q) = 1 - F(d, p, q)$ with $F(d, p, q)$ being the statistical fidelity between the measured counts and the corresponding unitary column, defined as

$$F(d, p, q) = \sum_{i=1}^{6} \sqrt{\frac{p_i d_i}{N} q_i}, \tag{6.12}$$

where $p$ is the vector of measured counts, $q$ is the absolute value squared of the corresponding unitary column and $N = \sum_{i=1}^{6} p_i d_i$ with $d_i$ as a correction factor included

to account for the $i^{th}$ detector efficiency. We can then use $d$ as our free parameters to minimise $C$. In general, we average the fidelity over $M$ randomly generated unitary matrices. The cost function then becomes

$$C(d) = 1 - \frac{1}{M}\sum_{i=1}^{M} F(d, p^{(i)}, q^{(i)}), \tag{6.13}$$

where $p^{(i)}$ ($q^{(i)}$) is the $i^{\text{th}}$ measurement vector (unitary column). Using this approach, and setting $M = 500$, we retrieve fidelities in excess of 0.995.

With the linear optics characterised and benchmarked, we now look at the performance of the single photon source. To determine the quality of the photons being generated, we use the standard HOM technique. The motorised collection stage allows us to change the temporal distinguishability between the input photons. We can then either use an on chip MZI, set to 50/50, or an off chip fibre beam-splitter. In an ideal case, that is with no entanglement between the photons, the shape of the HOM dip is given by the convolution of the Fourier transforms of the spectral shapes of the individual photons. For our setup, this includes the Gaussian spectral shape of the down conversion and the top-hat shape of the interference filter. In general, the convolution of these functions can be very complicated and in practise it is enough to use a function which captures the important features of the dip. We therefore, use the fitting function

$$f(x) = \left(a_1 + a_2 x + a_3 x^2\right) \times \left(1 - V \exp\left(-\frac{(x - x_0)^2}{2\Delta x^2}\right) \operatorname{sinc}\left(s_1 \left(x - x_0\right)\right)\right), \tag{6.14}$$

where the fitting parameters $a_i$ describe possible drifts in the coupling of the photons into fibre. The important parameters of the dip, i.e the visibility, width and central position, are given by fitting parameters $V$, $\Delta x$, and $x_0$, respectively. The effects of the spectral shape of the spectral filter is given by the sinc function, parameterised by $s_1$, and the down conversion is described by the Gaussian prefactor. An example dip, using two photons from the same (horizontal) source and an on-chip beam-splitter, is shown in Fig. 6.3 with a visibility calculated from the fit to be $0.98 \pm 0.01$. It is also possible to perform a HOM dip with photons from different sources, where we herald on the non-interfering mode to project into an approximate Fock state. These have a lower fidelity of $\approx 0.92$, which is attributed to mixedness of the single photon states due to spectral correlations between the signal and idler. In the following work we

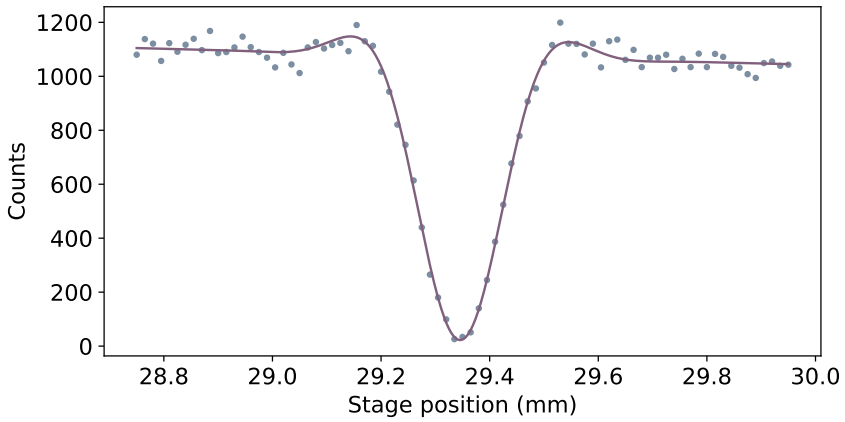


Figure 6.3: Example HOM dip using an on-chip beam-splitter and two photons from one source. One second integration per point. Visibility of the fringe measured from the fit to be $0.98 \pm 0.01$.

only require two photons from a single source, however, information on higher photon number experiments with this source are detailed in Refs. [214–216]

## 6.5 UNITARY DILATION APPROACH TO SIMULATING PT SYSTEMS

In the previous section we have described in detail how our optical chip can be used to implement *unitary* transformations. However, as the Hamiltonians we wish to study are non-Hermitian, their respective evolution operators are non-unitary. In this section we describe the method we use to circumvent this problem.

### 6.5.1 *Unitary dilation*

In the following experiments we use a technique known Halmos unitary dilation [217]. Here, we embed the $N$-dimensional evolution operator in a $2N$-dimensional unitary. This process requires that we first renormalise the operator, where the operator norm, $\| \cdot \|$, is given by the largest singular value. This gives us

$$\tilde{G}_N(\gamma, t) = \frac{G_N(\gamma, t)}{\| G_N(\gamma, t) \|}, \tag{6.15}$$

where $G_N(\gamma,t) = e^{-iH_N(\gamma)t}$. We note here that the normalisation factor is also time dependent, and therefore the normalisation procedure needs to be repeated for each time step. We can then embed this into a larger unitary matrix

$$U_{2N}(\gamma,t) = \begin{pmatrix} \tilde{G}_N(\gamma,t) & iD_N(\gamma,t) \\ iD_N^*(\gamma,t) & \tilde{G}_N^\dagger(\gamma,t) \end{pmatrix}, \tag{6.16}$$

where $D_N(\gamma,t) = \sqrt{1-\tilde{G}_N(\gamma,t)\,\tilde{G}_N^\dagger(\gamma,t)}$. This unitary matrix can now be implemented on our photonic chip. To simulate the time dependence of $U_{2N}(\gamma,t)$, we discretise the time steps and implement $U_{2N}(\gamma,t_i)$ for each time step. By inputting photons in only the top $N$ modes, and post-selecting on these photons leaving the top $N$ modes, we can then recover the dynamics of $\tilde{G}_N(\gamma,t)$. An interesting point to note here is that, $U_{2N}(\gamma,t)$ is unitary for any value of $\gamma$ and therefore changing $t$ and $\gamma$ simply correspond to changing the set of phases we dial on the chip. This renders changing between Hermitian, $\mathcal{PT}$-unbroken, and $\mathcal{PT}$-broken regimes easy, and allows us to investigate the entire parameter space of $\tilde{G}_N(\gamma,t)$. The unitary, $U_{2N}(\gamma,t)$, can be thought of as coupling the system $\tilde{G}_N(\gamma,t)$ with its Hermitian conjugate (time reversed), with a coupling strength dependent on the non-Hermitcity of the Hamiltonian. In the following work we always restrict ourselves to the post-selected case where we restrict ourselves to the dynamics of the $\tilde{G}_N(\gamma,t)$, however, in Ref. [218], the full nature of the coupled time reversed systems is analysed. To aid clarity, here we differentiate between the interferometer modes, into which we inject photons, and the modes of the $\mathcal{PT}$-system. In general the input state is

$$|\psi\rangle_{\mathrm{F}} \bigotimes |\psi\rangle_{\mathrm{R}} = \left(\hat{a}_1^\dagger\right)^{s_1}\left(\hat{a}_2^\dagger\right)^{s_2}\ldots\left(\hat{a}_N^\dagger\right)^{s_N}|0\rangle^{\otimes N} \bigotimes \left(\hat{a}_{N+1}^\dagger\right)^{s_{N+1}}\ldots\left(\hat{a}_{2N}^\dagger\right)^{s_{2N}}|0\rangle^{\otimes N}, \tag{6.17}$$

where $\hat{a}_i^\dagger$ represents the creation operator of the $i^{\text{th}}$ mode of the interferometer and $s_i$ is the occupation of this mode. We have also introduced the subscripts F and R to denote states initialised in the 'forward' system (capturing the dynamics of $\tilde{G}_N(\gamma,t)$) and 'reverse' system (for $\tilde{G}_N^\dagger(\gamma,t)$) In all of the following we restrict ourselves to the case where $s_i \in \{0,1\}$. We also restrict ourselves to only single input photons, in which case the input states can be written as

$$\{|k\rangle = |0_p\rangle^{\otimes k-1}|1_p\rangle\,|0_p\rangle^{\otimes 2N-k}\}_{k=1}^{2N}, \tag{6.18}$$

where the subscript $p$ corresponds to a photonic mode, rather than to a mode label of $\tilde{G}_N(\gamma, t)$. The modes of $\tilde{G}_N(\gamma, t)$ are labelled as qudit modes, i.e. $|0\rangle$, $|1\rangle$, ... $|N-1\rangle$. We input single photons but the linear optics allows us to also create superposition states.

### 6.5.2 *Experimental results for two and three mode systems*

To initially characterise the performance of the chip and the unitary dilation procedure we look at the dynamics of a single excitation in a two and a three mode system. To match the convention used later on, we label the modes of the N-mode system starting at $|0\rangle$. For the two mode case, we input photons into the top mode of the interferometer. This initialises the state

$$|0\rangle_{\mathrm{F}} = |1_p 0_p\rangle_F \bigotimes |0_p 0_p\rangle_R , \tag{6.19}$$

where the subscript F denotes the forward system. We then implement the $4 \times 4$ unitary matrix $U_4(\gamma, t)$ on the top 4 modes of the interferometer. This leaves us with a $4-$dimensional output density matrix

$$\rho = U_4(\gamma, t)\, \rho_0 U_4^{\dagger}(\gamma, t) , \tag{6.20}$$

where $\rho_0 = |0\rangle_{\mathrm{F}} \langle 0|_{\mathrm{F}}$. We can then trace out the reverse system to leave us with the dynamics of only the forward system. This is normalised to also include the probability of no photons being detected in the forward system, $|\mathrm{vac}\rangle = |0_p 0_p\rangle_F$, that is the case where the photon is detected in either of the modes corresponding to the reverse subsystem. We repeat this for different values of $\gamma$, allowing us to investigate all regions of the parameter space. The results of this measurement are shown in Fig. 6.4. In Fig. 6.4(a), with solid lines for the theory lines and triangular points for the data, we show the time dynamics for the unbroken symmetry regime, $\gamma = 0.25$. We see an oscillating pattern with a non-zero probability for the photon to leak into the reverse system. The period of this oscillation, $\tau$, is related to $\gamma$ by $\tau = 2\pi/\sqrt{1-\gamma^2}$ [214]. We also see a singularity in the first derivative, this is an artefact of how we normalise the state. If we normalise probabilities all in the forward system, this disappears. This stems from the fact that there are discontinuities in the evolution of $D_N(\gamma, t)$. Probabilities of excitations tunnelling from one system to another, those which we include as $|\mathrm{vac}\rangle$, are dependent on $D_N(\gamma, t)$ and the discontinuities appear in the state evolution. If

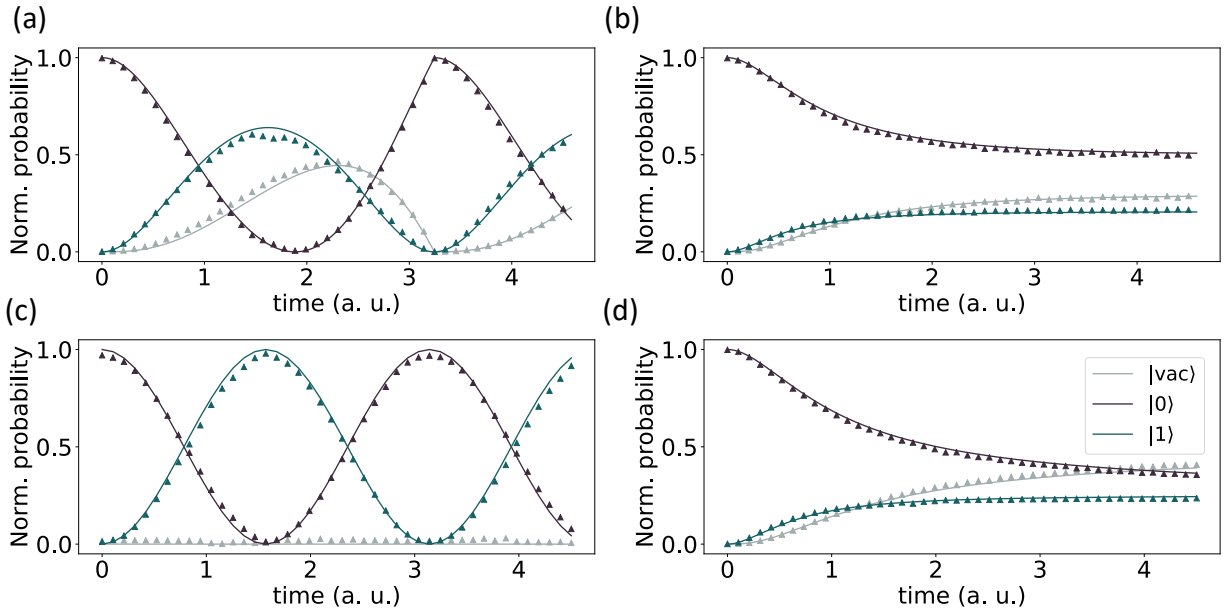


Figure 6.4: Simulated dynamics for two-mode $\mathcal{PT}$ system. All panel's colours match the legend in (d). $|\text{vac}\rangle$ represents the probability of the photon leaking into the reverse system. Triangular points show the experimental data and solid line represents the theoretical prediction (a) The dynamics in the unbroken symmetry region with $\gamma = 0.25$. (b) Broken symmetry region with $\gamma = 1.1$ (c) The dynamics for the Hermitian case, with $\gamma = 0$. (d) Evolution at the exceptional point, $\gamma = 1$

we instead post-select only on excitations which remain in the desired subsystems the probabilities are independent of $D_N(\gamma, t)$ and the discontinuities disappear. Figure 6.4(c) shows the Hermitian case, $\gamma = 0$, as expected here we see oscillating between the two modes of the system with the photon always contained in the top two modes. Figure 6.4(b) shows the dynamics in the broken symmetry regime, $\gamma = 1.1$. We no longer see the oscillating pattern, seen in the unbroken region, instead the probabilities settle to a steady state. Figure 6.4(c) show the dynamics at the exceptional point, which also decay to a steady state but over a longer timescale than the broken region.

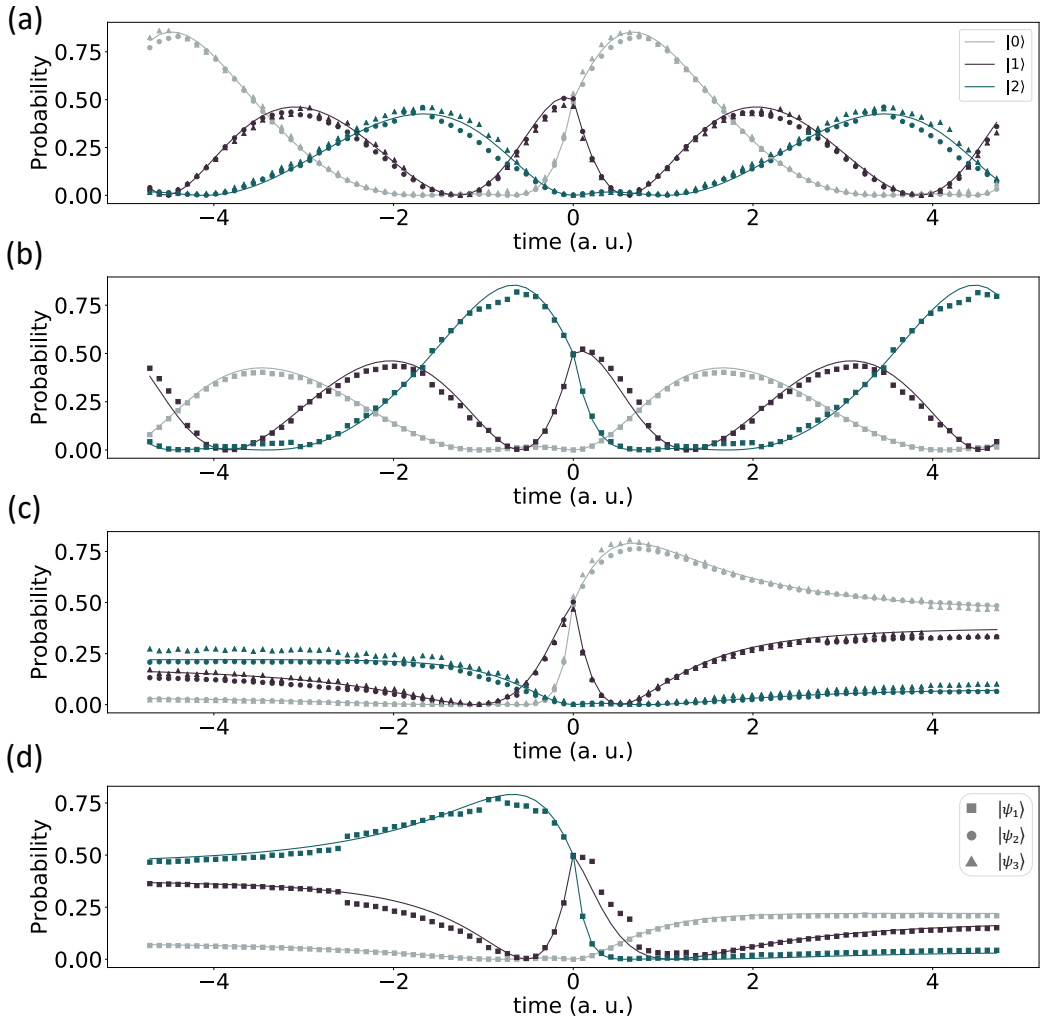


Figure 6.5: Simulated dynamics for a three-mode $\mathcal{PT}$ system. The colours in all four panels map to the qutrit levels shown in the legend on (a) and the symbols map to the states shown in the legend on (d). The top two panels show the unbroken symmetry regime, $\gamma = 0.5\gamma_c$ and the bottom two panels show the broken symmetry regime, $\gamma = 1.1\gamma_c$. The odd panels show the plots for states $|\psi\rangle_1$ and $|\psi\rangle_3$. These are shown together as they should demonstrate the same dynamics. The even panels show $|\psi\rangle_2$.

For the three mode case we have more flexibility in our choice of input states. We therefore choose to look at three different input states.

$$
\begin{aligned}
|\psi\rangle_1 &= \frac{|0\rangle_{\mathrm{F}} - i\,|1\rangle_{\mathrm{F}}}{\sqrt{2}} = \frac{|1_p 0_p 0_p\rangle - i\,|0_p 1_p 0_p\rangle}{\sqrt{2}} \bigotimes |0_p 0_p 0_p\rangle \\
|\psi\rangle_2 &= \frac{i\,|1\rangle_{\mathrm{F}} + |2\rangle_{\mathrm{F}}}{\sqrt{2}} = \frac{i\,|0_p 1_p 0_p\rangle + |0_p 0_p 1_p\rangle}{\sqrt{2}} \bigotimes |0_p 0_p 0_p\rangle \\
|\psi\rangle_3 &= \frac{|0\rangle_{\mathrm{R}} + i\,|1\rangle_{\mathrm{R}}}{\sqrt{2}} = |0_p 0_p 0_p\rangle \bigotimes \frac{|1_p 0_p 0_p\rangle + i\,|0_p 1_p 0_p\rangle}{\sqrt{2}}.
\end{aligned}
\tag{6.21}
$$

These states are related by the following symmetry transformations

$$\begin{aligned} |\psi\rangle_2 &= \mathcal{PT}\,|\psi\rangle_1\,, \\ |\psi\rangle_3 &= \mathcal{T}\,|\psi\rangle_{1,R}\,, \end{aligned} \tag{6.22}$$

where $|\psi\rangle_{1,R}$ represents the same superposition as $|\psi\rangle_1$ but initialised in the reverse subsystem. The superposition states are created with the on chip beam-splitters. For the three-mode system, the entire device is required to implement the evolution matrix, therefore we are required to compile the state generation with the evolution matrix

$$U_{\text{chip}} = U_6\,(\gamma, t) \cdot U_{\text{prep}}. \tag{6.23}$$

Similar to the two mode case, we normalise to include the probability of the photon leaving in the reverse system, however to avoid clutter we only show the qutrit levels. These results are shown in Fig. 6.5, where we have organised the plots in a way to illustrate the state symmetries. We also show plots for both unbroken and broken symmetry regimes. The state symmetries are visible in both regimes. In Figs. 6.5(a) and (c) we show both $|\psi\rangle_1$ and $|\psi\rangle_3$. As expected from the symmetry transformations above, these two states follow the same dynamics. Figures. 6.5(b) and (d) show the qutrit levels for $|\psi\rangle_2$. From this we can see that if we reverse the time direction, then the dynamics of $|\psi\rangle_2$ match those of $|\psi\rangle_1$, if we swapped the levels $|0\rangle$ and $|2\rangle$, equivalent to applying $\mathcal{P}$.

## 6.6 COHERENCE OPTIMISATION THROUGH ADAPTIVE FEEDBACK LOOP

State engineering using open quantum systems is a rich and fruitful area of research. Here, the system Hamiltonian and system-environment coupling can be tuned to transform an arbitrary initial state into a desired, often entangled, target state [219–221]. In fact, in Ref. [222] it was shown that for any quantum circuit it is possible to derive a Lindblad master equation which evolves to a steady state that encodes the output of the quantum circuit. This evolution is then independent of the initial state and therefore more robust to errors. Similarly, in the following section we will examine how $\mathcal{PT}$-symmetric systems can be used to drive an arbitrary qubit state towards a given, more coherent, target state.

### 6.6.1 *Algorithm*

In order to optimise the coherence of a qubit we simulate a Hamiltonian with a time-varying value of $\gamma$ using an adaptive feedback algorithm. The schematic for the

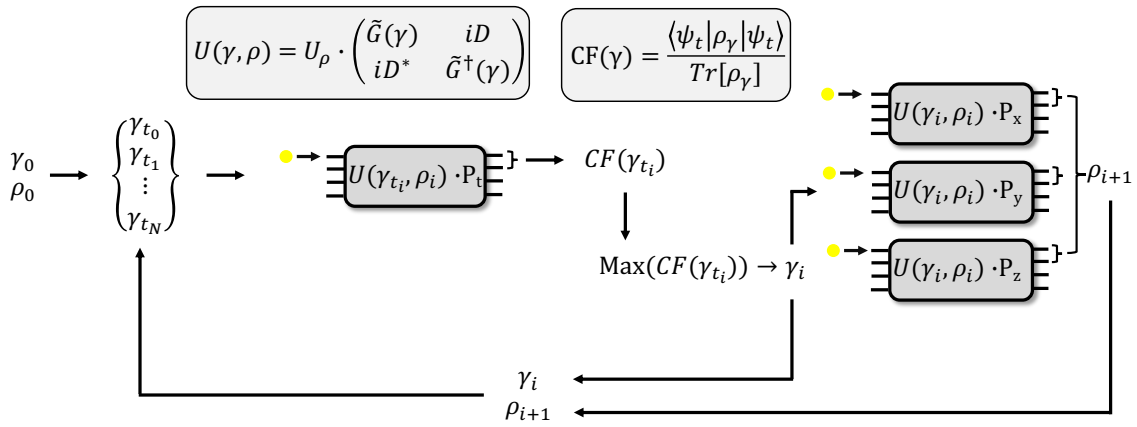


Figure 6.6: Schematic of coherence optimisation algorithm. $P_t$ represents the projector onto the target state and $P_x$, $P_y$, $P_z$ are the projectors onto the qubit bases for tomography.

algorithm is shown in Fig. 6.6. In general, we aim to maximise a cost function

$$CF(\rho) = \frac{\langle\psi_t|\,\rho\,|\psi_t\rangle}{\mathrm{Tr}\,[\rho]}, \tag{6.24}$$

which is the projection of the state onto a desired, in principle, more coherent, target state. To simulate a time dependent $\gamma(t)$ we discretise the time evolution into steps of $\delta t$, and then aim to find a set of $\gamma$ values, $\{\gamma_i\}$ where at each time step $\gamma_i$ maximises the cost function. At each step of the optimisation, we select a range of $\gamma$ values to test,$\{\gamma_{t_i}\}$, centred at the optimum $\gamma_{i-1}$ from the previous step. Each of these $\gamma_{t_i}$ values are tested - the state is evolved subject to $U_{2N}(\gamma_{t_i}, \Delta t)$ and then projected on the target state in order to evaluate $CF$. Note that this testing time, $\Delta t$, can, in principle, be different from the discritised evolution time step. The $\gamma_{t_i}$ which maximises the cost function is then chosen to be current $\gamma_i$ value. The state is then evolved by $U_{2N}(\gamma_i, \delta t)$, and we perform full state tomography on the resulting state. To simulate the effect of decoherence on the qubit we classically transform the reconstructed density matrix as

$$\rho \rightarrow \rho + p\,(\sigma_d \rho \sigma_d - \rho)\,\delta t, \tag{6.25}$$

where $p$ is the dephasing rate and $\sigma_d$ is the Pauli matrix in the direction of the dephasing. This state, after decoherence, becomes the input state for the next step of the algorithm and $\gamma_i$ becomes the central value for the new $\gamma$ range.

### 6.6.2 *Quantum state tomography*

As the algorithm requires us to reconstruct the qubit state at each step, it will be useful here to recap the important points of quantum state tomography. Quantum state tomography aims to reconstruct the state of a quantum system by performing measurements along different bases. We are concerned with reconstructing the state of a single particle, the process however can be generalised to multi-particle systems [223, 224]. To perform quantum state tomography on a qudit of dimension $d$ we require the generators of the $SU(d)$ group as this, along with the $d-$dimensional identity, form a complete qudit basis [224]. For a qubit, with $d = 2$, these generators are the well known Pauli matrices

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.26}$$

We can now use these matrices to represent an arbitrary qubit density matrix as

$$\rho = \frac{1}{2} \sum_{k=0}^{3} S_k \sigma_k. \tag{6.27}$$

Quantum state tomography, therefore, reduces to determining the parameters $\{S_k\}_{k=0}^{3}$. These parameters are given by $S_k = \mathrm{Tr}\left[\sigma_k \rho\right]$ and can be linked to pairs of projective measurements on the qubit

$$\begin{aligned} S_0 &= \mathrm{P}_{|0\rangle} + \mathrm{P}_{|1\rangle} \\ S_1 &= \mathrm{P}_{|+\rangle} - \mathrm{P}_{|-\rangle} \\ S_2 &= \mathrm{P}_{|+i\rangle} - \mathrm{P}_{|-i\rangle} \\ S_3 &= \mathrm{P}_{|0\rangle} - \mathrm{P}_{|1\rangle}. \end{aligned} \tag{6.28}$$

From this we can see that, by normalisation, $S_0 = 1$. The unitary that corresponds to the projector onto a given basis are constructed from the eigenstates of the basis, where the $i^{th}$ row of the unitary matrix is the $i^{th}$ eigenvector of the $\sigma_k$. This construction method

also holds for dimensions $d \neq 2$. The density matrix, $\rho$, can, in principle, describe a mixed state made up of an arbitrary ensemble of pure states. However, it is always possible to diagonalise a density matrix so that it is described by a mixture of just $d$ orthogonal pure states. This is how we can implement arbitrary mixed qubit states on our chip. We diagonalise the $2 \times 2$ density matrix, equivalent to taking its singular value decomposition, and evolve each pure state separately. Once we have performed tomography on each state individually they are combined to give the final density matrix

$$\rho = \lambda_1 \rho_1 + \lambda_2 \rho_2, \tag{6.29}$$

where $\lambda_i$ is the relative weight (singular value) associated with the $i^{\text{th}}$ basis state, $\rho_i$ is the reconstructed density matrix. On our chip, we also have to take into account that the fact that the non-unitary evolution of our chip means that there is always some probability of the photon being detected in the modes corresponding to the reverse system. To combat this we also include an extra normalisation step, which is taken by recording the proportion of counts detected in the forward and reverse systems, for each of the evolved pure states. This is then used to normalise the final calculated $S_k$ parameters.

### 6.6.3 *Experimental results*

To perform the coherence optimisation algorithm experimentally we use the 2 mode Hamiltonian used in sec. 6.5.2 and choose the target state to be $|i\rangle$. In Fig. 6.7 we show the results of the algorithm for an input state of $|-i\rangle$. We see that the cost function converges to close to 1 for 0 dephasing but, as expected, is lower as we add dephasing. To see the effect of the dephasing rate without the algorithm, we simulate the Hermitian regime i.e. $\gamma = 0$, where we see decaying oscillations. By fitting these oscillations we can extract the real dephasing rates. The set dephasing rates were $\{0, 0.18, 0.36\}$ and the measured rates were $\{0.03, 0.2, 0.35\} \pm 0.01$. We see that there is some intrinsic dephasing introduced by the simulator itself. This input state was chosen as it has the minimal overlap with the target state. However, it is already a maximally coherent state.

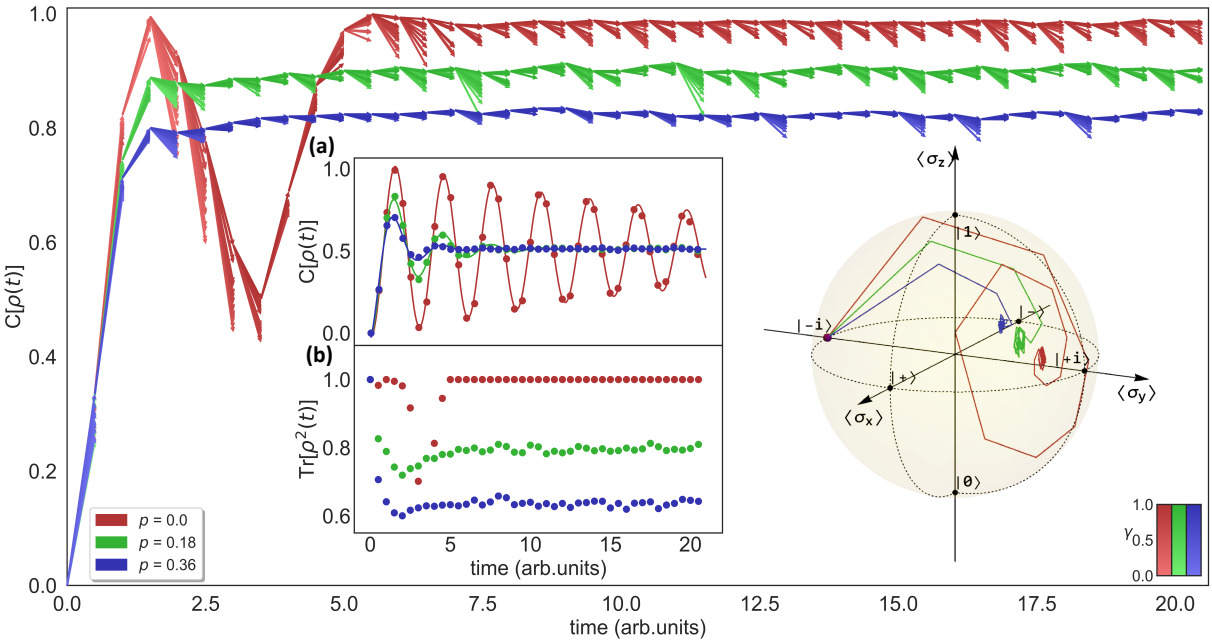


Figure 6.7: Coherence optimisation algorithm with input state of $|-i\rangle$ for three dephasing rates, $p$. At each step we use arrows to denote the value of the cost function for each of the $\gamma_{t_i}$ values tested. The saturation of the colour of the arrows corresponds to the magnitude of $\gamma_{t_i}$; The inset saturation colour bar (bottom right) applies only to the main figure. Inset (a) shows the state evolution with $\gamma = 0$, subject to the same dephasing channel. Dots represent experimental data and a numerical fit is shown with a solid line. This allows the measured dephasing rates to be extracted. Set dephasing of $\{0, 0.18, 0.36\}$ and the measured rates were $\{0.03, 0.2, 0.35\} \pm 0.01$ . Inset (b) shows the evolution of the state purity. We also show the state evolution, through the algorithm, on the Bloch sphere.

In order to confirm that the algorithm also performs with mixed states, we perform the same measurements but with two mixed density matrices

$$\rho_1 = \begin{pmatrix} 0.6 & 0.2 + 0.3i \\ 0.2 - 0.3i & 0.4 \end{pmatrix} \qquad \rho_2 = \begin{pmatrix} 0.5933 & 0.187 - 0.1149i \\ 0.187 + 0.1149i & 0.4067 \end{pmatrix} \tag{6.30}$$

The results for these two states are shown in Fig. 6.8(a) and (b), respectively. From this we see that algorithm does work for states inside and on the surface of the Bloch sphere. In the following sections, we will use a numerical simulator of the experiment to more thoroughly understand the scope and range of the algorithm.

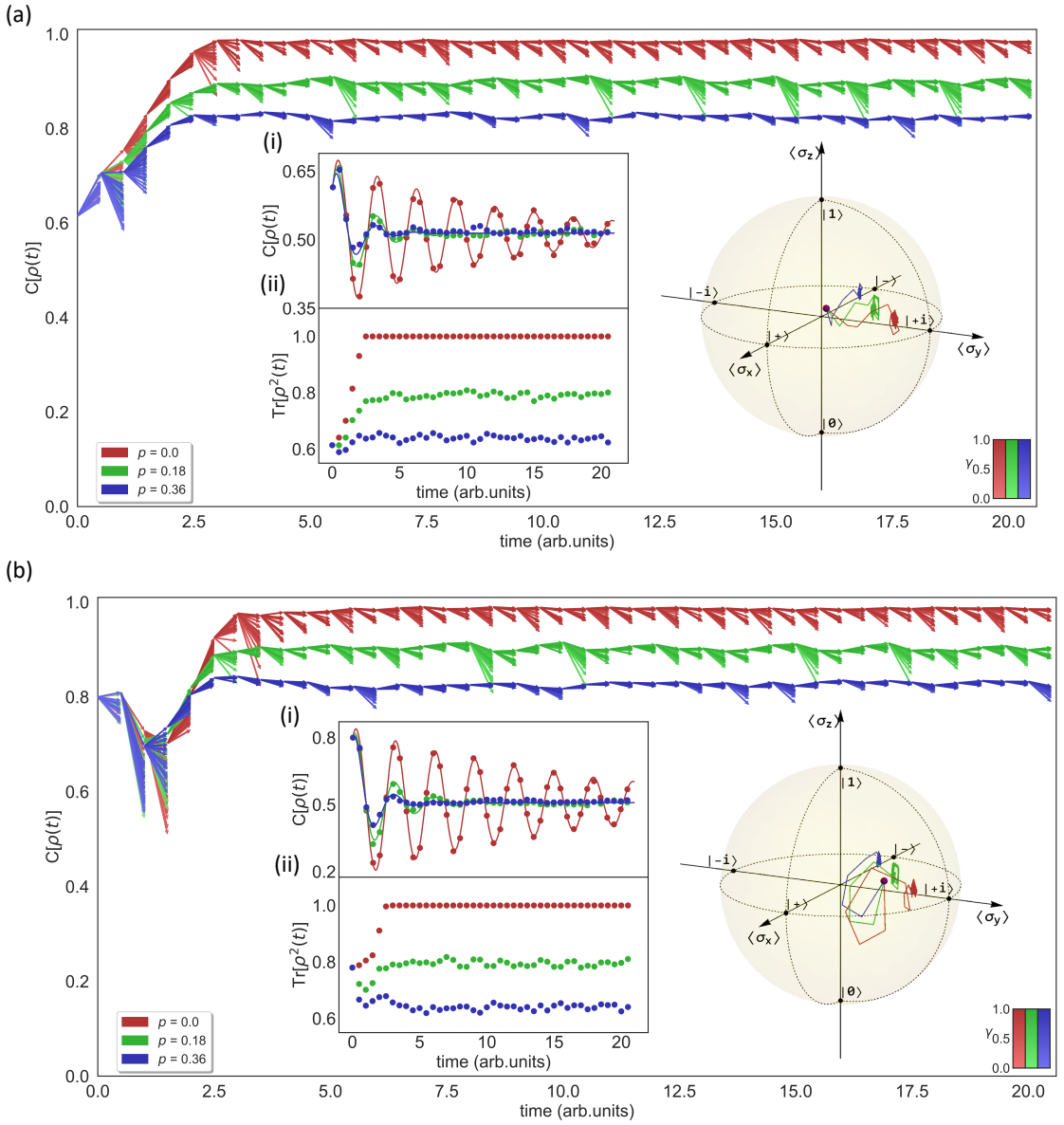


Figure 6.8: Coherence optimisation algorithm for mixed input states and a range of different dephasing rates. Plot is the same format as Fig. 6.7. (a) Results for state $\rho_1$ (b) Results for state $\rho_2$.

### 6.6.4 *Algorithm optimisation*

The optimisation algorithm has a variety of free parameters. In the above experiment, we are limited by the time needed for each iteration step. However, in simulation we

have more scope to vary these parameters. The most important parameters that we have control over are the evolution discretisation step size, $\delta t$, and the cost function evolution time, $\Delta t$. How small we discretise the evolution affects how closely we simulate the dynamics of $H_2(\gamma(t))$. In order find the optimum value to trade of accuracy and

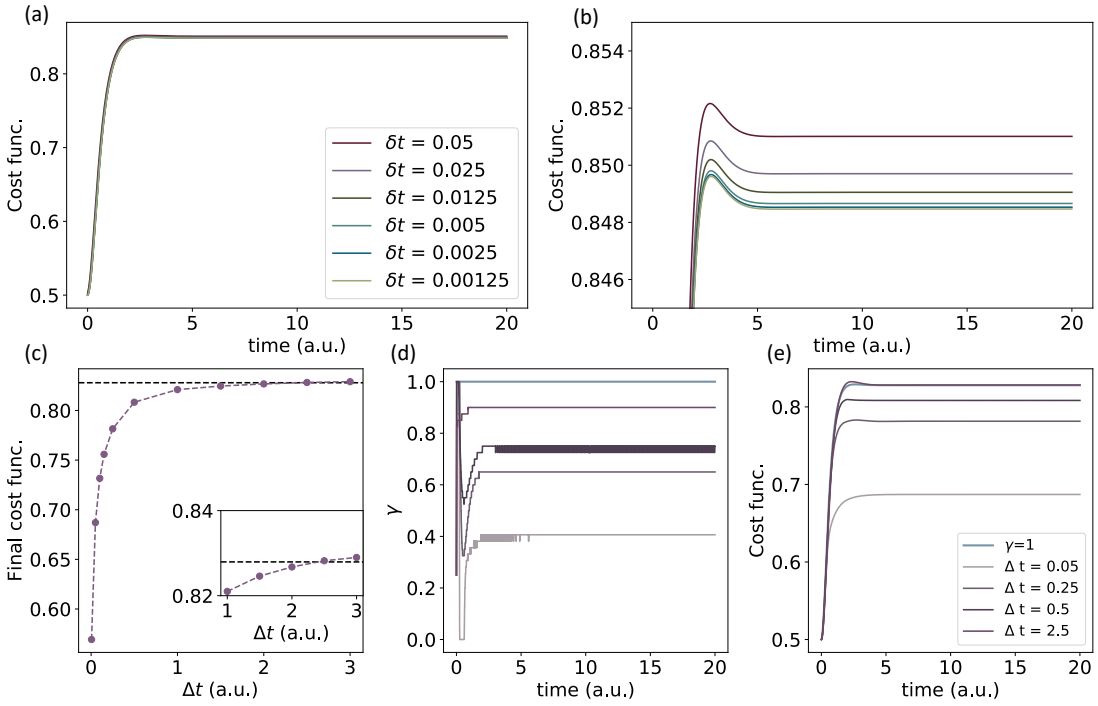


Figure 6.9: Effect of changing algorithm parameters. For all plots initial state taken to be the maximally mixed state $\rho = 0.5\,|0\rangle\langle 0| + 0.5\,|1\rangle\langle 1|$, target state chosen to be $|+i\rangle$, with dephasing along $\sigma_x$ of 0.2. (a) Cost function evolution for fixed $\gamma = 1$ for different evolution step sizes. (b) Same as (a) but with tighter scale to aid visualisation. (c) Final cost function value for algorithm as a function of $\Delta t$. (d) $\gamma$ value through the algorithm. Colours match legend in (d). (d) Cost function evolution for different values of $\Delta t$

run time, we simulate a Hamiltonian with a fixed $\gamma = 1$, plotting the cost function value over the evolution. We choose the initial state to be a maximally mixed state $\rho = 0.5\,|0\rangle\langle 0| + 0.5\,|1\rangle\langle 1|$ and the target state to be $|+i\rangle$. These results are shown in Fig. 6.9(a) and (b), we see that the final cost function value converges to some value as we decrease the step size, $\delta t$. From this, we choose a value of $\delta t = 0.005$ for all following simulations. Another important parameter is the time we evolve by before checking the cost function at each step, $\Delta t$. Figure. 6.9(e) shows the results of the optimisation algorithm for a range of $\Delta t$ values. We keep the same target and initial

states as before, and we keep the method for selecting the $\gamma$ range at each step the same as the experimental setup. As a reference, we also include the evolution by a fixed Hamiltonian $H_2(\gamma = 1)$. We see that increasing the time, increases the end value of the cost function. Figure. 6.9(c) shows the final cost function value, we see that it approaches the value attained with $\gamma = 1$ and for large values it overtakes it by a very small amount. From this we choose a value of $\Delta t = 3$ for the following simulations.

### 6.6.5 *Effect of varying initial and target states*

We now use our simulation to more thoroughly look at the effect of changing the initial states as well as look at changing the target states. Initially, we generate 20 random mixed states. The coherence optimisation algorithm is then simulated for each of these states with the algorithm parameters determined in the previous section and a target state of $|+i\rangle$. We find that for the 20 states we have a standard deviation of the final

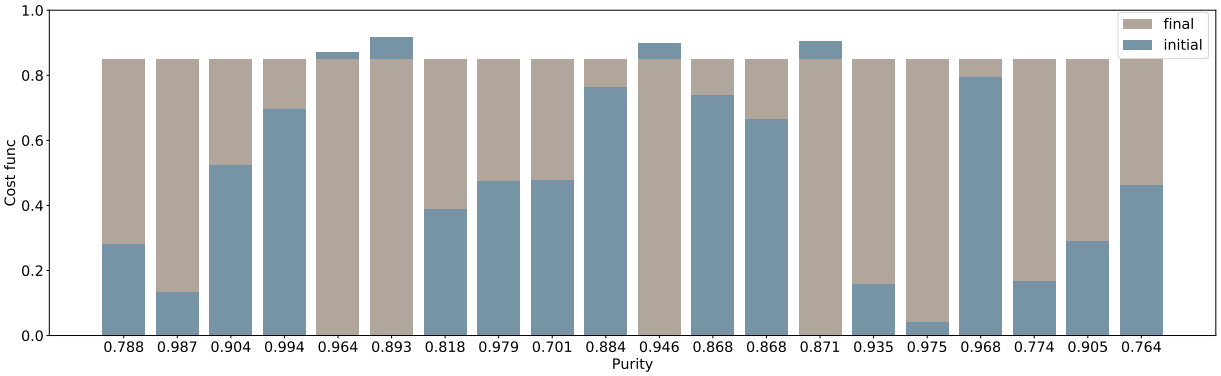


Figure 6.10: Changing the initial states. 20 randomly generated mixed states. State purity is shown on the x axis. We show the initial (blue) and final (brown) cost function values.

cost function of $2.6 \times 10^{-6}$. These results are shown in Fig. 6.10, where we show the purity of the initial density matrix, initial and final cost function values for each state. We see that the initial overlap and purity have no effect on the algorithm. To look at the effect of the target state, we define a general qubit state as

$$|\psi_t\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle\,. \tag{6.31}$$

We can then sweep $\theta$ between 0 and $\pi$ and $\phi$ from $-\pi$ to $\pi$. We fix the dephasing to be 0.2 in the $\sigma_x$ direction. The initial state is taken to be the maximally mixed state $\rho = 0.5\,|0\rangle\langle 0| + 0.5\,|1\rangle\langle 1|$. Figure 6.11 shows a colour map of the results. We see that

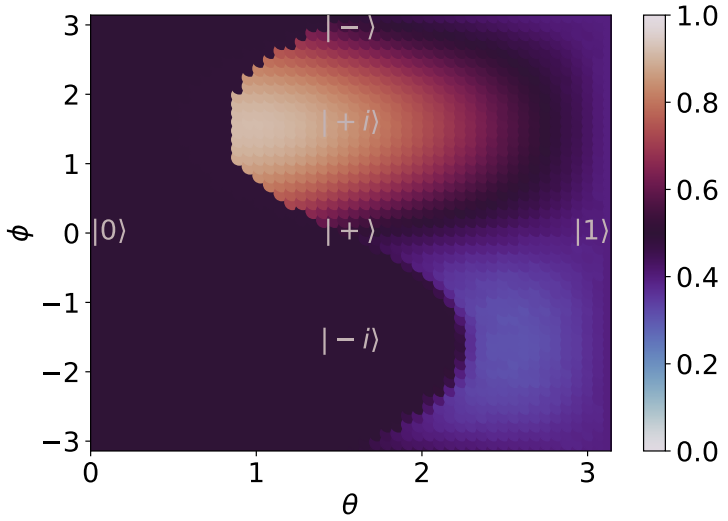


Figure 6.11: Changing the target states. Testing target states from the whole Bloch sphere surface. For clarity, we've included labels on the pairs of canonical qubit basis states.

the algorithm works best around $|+i\rangle$. Many regimes of the Bloch sphere remain in the maximally mixed input state and interestingly, some states around $|1\rangle$ end with lower cost function values than in initial state. Although we note that these end states are no longer maximally mixed, and have been driven closer to $|+i\rangle$.

### 6.6.6 *Comparison with fixed $\gamma$*

Initially, when optimising the algorithm parameters, we compared to evolving the state with a Hamiltonian with a fixed $\gamma$ value. Specifically, we compare to fixing $\gamma = 1$. This Hamiltonian is the exceptional point and therefore has the state $|+i\rangle$ as its only remaining eigenvalue. There, we saw that by increasing the evolution time before measuring the cost function, $\Delta t$, we could bring the performance algorithm close to, and even slightly exceed, the fixed $\gamma$ evolution. Although the degree that we exceeded by was very small and could be attributed to the discretisation of the evolution. In this section we will look at this comparison more thoroughly. Initially, we look to change the initial state, similar to Fig. 6.10. We use $|+i\rangle$ as the target state and add a dephasing

of 0.2 in the $\sigma_x$ direction. Figure 6.12 shows these results for 20 randomly generated

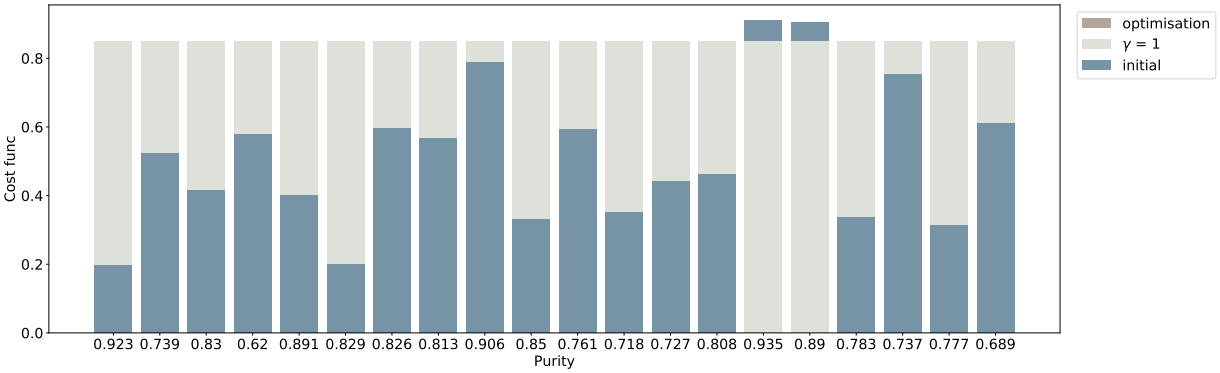


Figure 6.12: Comparison of our optimisation algorithm with evolution with a fixed $\gamma = 1$. Optimisation performed for 20 randomly generated mixed states. State purity shown on the x axis. Initial cost function value shown on y axis along with optimisation and $\gamma = 1$ final cost function values.

initial states. As with the optimisation algorithm we see that fixed $\gamma$ evolution always converges to the same value, regardless of initial state. In fact, for the 20 states we find a standard deviation of the final value of $\approx 10^{-16}$. We also see that this final value is slightly lower than the optimisation algorithm. The final cost function value is found to be 0.850 (0.849) for the optimisation algorithm ($\gamma = 1$). This difference is very small and can likely be attributed to errors introduced by the discrete evolution. Next we look at the effect of dephasing rate and axis on the final cost function value. We choose our initial state to be the maximally mixed state and $|+i\rangle$ to be our target state. We then sweep different dephasing values across all three dephasing axes. The results for this are shown in Fig. 6.13. Here we see that both perform similarly across the different dephasing rates and directions. We also see that the different dephasing axes have largely different effects on both the algorithm and fixed gamma evolution. Adding a dephasing along a certain axis has the effect of moving states from the surface of the Bloch sphere towards the centre, around that axis. We have illustrated this, for $\sigma_x$, on a 2D projection of the Bloch sphere in Fig. 6.13(b). This helps make sense of the plot in Fig. 6.13(a). We can see that states on the poles of the dephasing axis are unaffected by the dephasing. This matches with Fig. 6.13(a) where the dephasing along $\sigma_y$ has little effect on both the optimisation and fixed $\gamma$ evolution. We would expect dephasing along $\sigma_x$ and $\sigma_z$ to have equal effect when the target state is $|+i\rangle$. We see that this is not the case and also that there is a small drop in the cost function with dephasing in

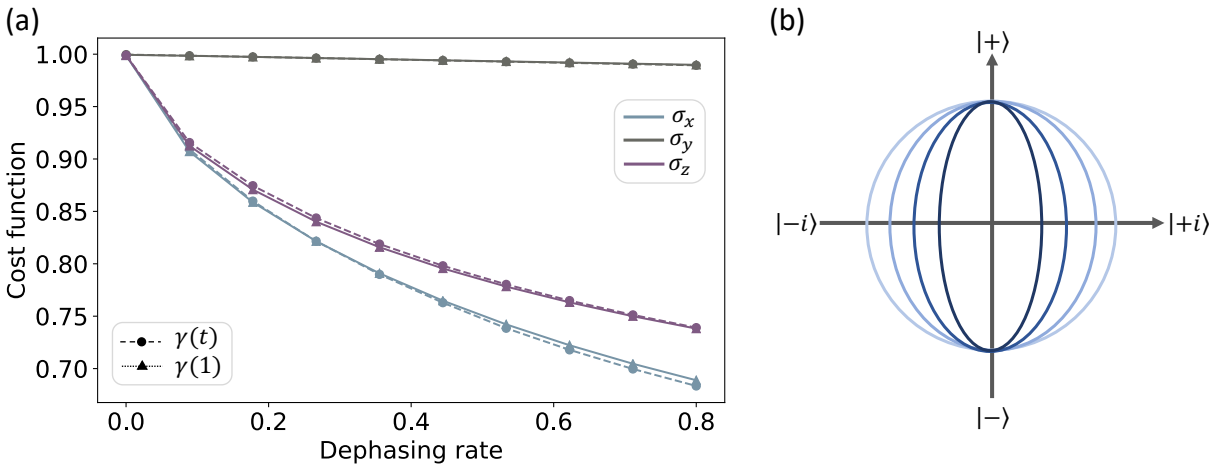


Figure 6.13: Comparison of our optimisation algorithm with evolution with a fixed $\gamma = 1$ for different dephasing rates and axes. Initial state is the maximally mixed state, target state is $|i\rangle$. Line colour illustrates the corresponding dephasing axis. Circular points with dashed lines are data for the optimisation algorithm, triangular points with dotted lines show $\gamma = 1$. (b) Visualisation of the effect of dephasing along the $\sigma_x$ direction. The darker the shade of blue, the larger the value of dephasing.

the $\sigma_y$ direction. This possibly hints that errors are being introduced due to the finite size of the evolution steps. Finally, we look at changing the target state. Again, we fix the initial state to be the maximally mixed state. Due to run time constraints, instead of sweeping the entire space, we plot the time evolutions for the six canonical qubit basis states. We perform this for three dephasing rates with a fixed dephasing axes, chosen to be $\sigma_x$. We also fix the dephasing rate to be 0.3 and then plot the sweeps for all three dephasing axes. The results of this are shown in Figure. 6.14. Figs. 6.14(a) to (c) varies the dephasing axis while fixing the dephasing rate at 0.3 and Figs. 6.14(d) to (f) vary the dephasing rate but fixing the dephasing axis to $\sigma_x$. We show three dephasing rates - 0, 0.3 and 0.7. With no dephasing (or dephasing angle $\sigma_y$) the $\gamma = 1$ the final cost functions represent the overlap between the target state and the exceptional point eigenstate. We see for certain states the optimisation algorithm results in a final cost function value is higher than for fixed $\gamma$. This is, however, due to the fact that $\gamma = 1$ is driving the state towards $|i\rangle$ which has a low overlap with certain target states. For the optimisation algorithm the state is driven towards a mixed state, with an overlap of 0.5, while this has a higher cost function value it is less coherent. Overall, again, we see that the algorithm performs similarly to the fixed $\gamma$ evolution.

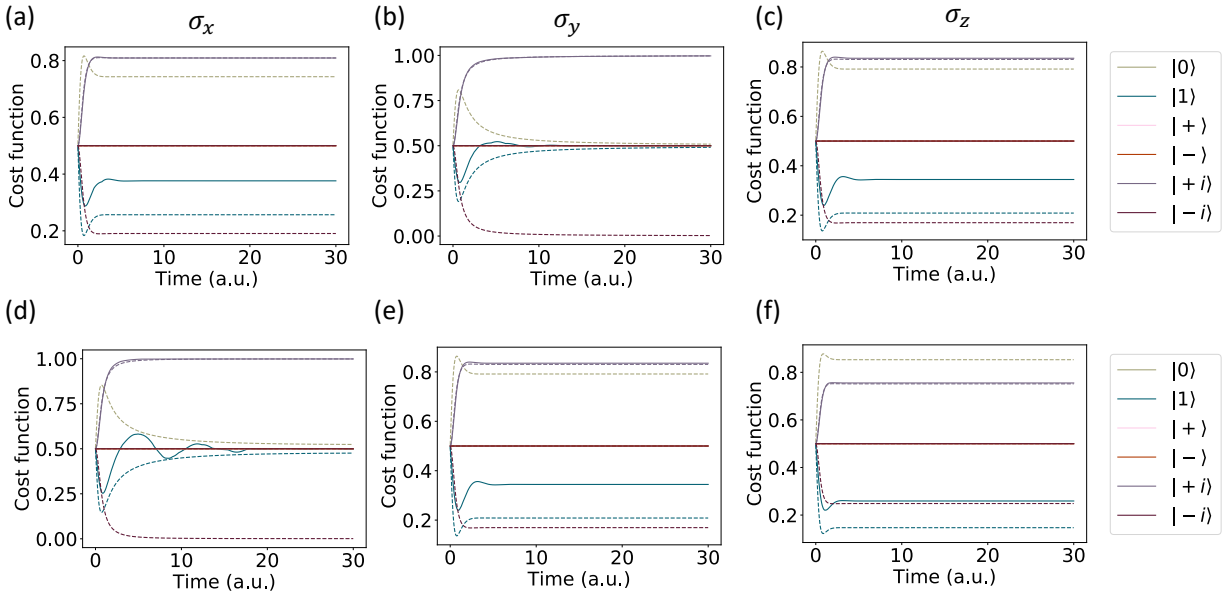


Figure 6.14: Comparing target state dependence of $\gamma(t)$ and $\gamma = 1$ for a different dephasing. For each subplot, solid lines represent the optimisation algorithm and dashed lines are fixed $\gamma$ evolution. Colours on all subplots correspond to the legend at the end of the row. (a) to (c) Varying the dephase axis with dephasing rate set to 0.3. Plot title shows the corresponding dephase axis. (d) to (f) Varying the dephasing rate. Dephase axis fixed to $\sigma_x$. Dephasing rates are 0,0.3,0.7 respectively.

## 6.7 DISCUSSION AND CONCLUSIONS

In this chapter we have looked at an interesting class of non-Hermitian Hamiltonians, those that exhibit $\mathcal{PT}$-symmetry. We have then shown how arbitrary linear optical unitary transformations can be used to simulate the non-Unitary dynamics of these systems using unitary dilation. We experimentally benchmarked this simulation protocol and showed good agreement between theoretical curves and experimental data. The unitary dilation method used here suffers from drawbacks which limit its potential scalability. A key problem is that the post-selection probability, that is the probability that an excitation remains in the desired subsystem, is inversely proportional to the degree of non-Hermiticity. Other schemes exist where rather than rescaling by the largest singular value, squeezing is used to implement positive singular values [225]. In Ref. [226], they also show how these protocols can be implemented using Trotterisation rather than a singular value decomposition. This removes another problem with unitary dilation, namely that one must exponentiate the desired Hamiltonian, a computationally

expensive task. Finally, we demonstrated an interesting property of $\mathcal{PT}$-symmetric Hamiltonians, namely their ability to drive a state towards a more coherent final state. We experimentally tested this property by implementing a Hamiltonian whose non-Hermiticity was time dependent. This was achieved using an adaptive feedback algorithm. We successfully showed, for a variety of input states, that this algorithm does in fact increase the final coherence of an initial state. A thorough numerical study of the algorithm showed that it works independently of initial state but that the target state is important. We then compared the algorithm to evolving a state with a fixed $\gamma$ value, specifically we choose to operate with $\gamma = \gamma_c$. We saw that both the fixed $\gamma$ and the optimisation algorithm performed similarly. To understand why this is the case we need to understand exactly what causes the non-Hermitian system to increase coherence in this way. To do this we derive the general evolution operator

$$G(\gamma, t) = e^{-iH_2(\gamma)t} = e^{\begin{pmatrix} \gamma t & it \\ it & -\gamma t \end{pmatrix}}. \tag{6.32}$$

This can be calculated easily so long as the matrix to be exponentiated can be diagonalised. This is because if a matrix $M$ can be written as $M = VDV^{-1}$, where $D$ is a diagonal matrix, then $e^M = Ve^DV^{-1}$. Following this through we arrive at

$$\begin{aligned} &G(\gamma, t) \\ &= \begin{pmatrix} \cosh\left(t\sqrt{\gamma^2-1}\right) + \frac{\gamma}{\sqrt{\gamma^2-1}}\sinh\left(t\sqrt{\gamma^2-1}\right) & \frac{i}{\sqrt{\gamma^2-1}}\sinh t\sqrt{\gamma^2-1} \\ \frac{i}{\sqrt{\gamma^2-1}}\sinh t\sqrt{\gamma^2-1} & \cosh\left(t\sqrt{\gamma^2-1}\right) - \frac{\gamma}{\sqrt{\gamma^2-1}}\sinh\left(t\sqrt{\gamma^2-1}\right) \end{pmatrix}. \end{aligned} \tag{6.33}$$

For $H_2(\gamma)$ the exceptional point is at $\gamma_c = 1$. We can, therefore, see here how if $\gamma < \gamma_c$ the evolution is periodic and sinusoidal, however for $\gamma > \gamma_c$ the evolution operator is hyperbolic. Calculating the evolution operator at the exceptional point using Eq. 6.33 results in division by 0. We can take the limit as $\gamma \to 1$, however a clearer method is using the power series definition of a matrix exponential, $e^M = \mathbb{I} + M + \frac{M \cdot M}{2} \cdots$. We

first calculate $M \cdot M = \left(\gamma^2 t^2 - t^2\right) \mathbb{I}$, which, along with the higher powers, vanishes at $\gamma = 1$. This leaves us with the EP evolution operator

$$G_{\mathrm{EP}}(t) = \mathbb{I} + M = \begin{pmatrix} 1+t & it \\ it & 1-t \end{pmatrix}. \tag{6.34}$$

To see how this affects a given state, we evolve an initial density matrix, $\rho_0$. For illustrative purposes we choose a maximally mixed state, however, from above we can see that the process is independent of initial state. We calculate

$$\rho(t) = G_{\mathrm{EP}}(t) \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} G_{\mathrm{EP}}^{\dagger}(t), \tag{6.35}$$

This leaves us with the density matrix

$$\rho(t) = \frac{1}{2\mathrm{Tr}\left[\rho(t)\right]} \begin{pmatrix} (1+t)^2 + t^2 & -2it^2 \\ 2it^2 & (1-t)^2 + t^2 \end{pmatrix}. \tag{6.36}$$

where the normalisation factor $\mathrm{Tr}\left[\rho(t)\right] = 1 + 2t^2$. If we now take the limit as $t$ goes to infinity we see

$$\rho \xrightarrow{t\to\infty} \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} = |i\rangle \langle i| \tag{6.37}$$

which is exactly the pure density matrix of the only EP eigenstate. From this we can see that as we are changing the target state but keeping $\gamma$ fixed at the EP, we are effectively measuring $|\langle \psi_t| + i\rangle|^2$. The optimisation algorithm on the other hand tries to find a specific time dependent $\gamma(t)$ to maximise the cost function. As the algorithm also relies on the fact that the state to evolve has some overlap with the EP eigenstate, it is most effective and struggles in the same areas of the target state space as the fixed $\gamma$ evolution. Experimentally, forcing a system to operate at its exceptional point requires careful balancing of the system parameters. The optimisation algorithm shows that similar performance can be achieved without needing to operate constantly at this delicate point.

ACKNOWLEDGEMENTS

The results from these chapters will be published in Refs. [218, 227]. The project was conceived by Prof. Anthony Laing and Prof. Yogesh Joglekar. Many people contributed to the experimental setup and characterisation [48, 214, 216]. The two and three site data was taken and analysed by Dr. Nicola Maraviglia. The coherence optimisation data was taken and analysed by myself and Nicola. The numerical analysis of the optimisation algorithm was performed by me, with input from Nicola and Yogesh.

# 7

# CONCLUSION

"'Well', said Zaphod airily, 'it's partly the curiosity, partly a sense of adventure but mostly i think it's the fame and the money....'"

Douglas Adams, *Hitchhikers Guide to the Galaxy*

This thesis describes a collection of experiments that aim to explore the role of time in integrated photonic experiments. Here we will recap the main results. Initially, we have seen how the high timing precision offered by modern superconducting nanowire single photon detectors can be used to erase frequency information, allowing photons which are intrinsically spectrally distinguishable to be interfered. Using integrated microring resonators on a silicon chip, we demonstrated an increase of interference visibility from 0.49 to 0.61 by decreasing the coincidence window used. We also saw good agreement between our model and the experimental data with a fidelity of $0.966 \pm 0.005$. Beyond this we used the internal phase in a $4 \times 4$ complex Hadamard interferometer to tune the time-domain interference fringes. Again, we saw good agreement with our model and also an increase in interference visibility from 0.4 to 0.6. All of these experiments were performed with standard microring resonator sources emitting telecom photons with a bandwidth of several GHz and with a spectral detuning of $\approx$7 GHz. Both the frequency detuning and photon bandwidth are two orders of magnitude higher than existing experiments and prove that this technique is compatible with standard silicon photonics components. Projective simulations have shown that with current state-of-the-art detectors we could expect quantum interference with photons detuned by several hundred GHz and high visibility up to $\approx$50 GHz. Future theoretical work would properly incorporate higher order Schmidt modes into the model.

In the following chapter, we extended this principle and applied it to the boson sampling problem. Initially, we employ a similar technique whereby the timing window is used to change the quality of the quantum interference. We show that, for a two photon scattershot experiment, changing the timing window can move use away from a distribution with minimal quantum interference to one with more. This has an intrinsic drawback that the temporal filtering associated with a smaller coincidence window necessarily reduces the number of useable samples. To combat this we performed

two-photon scattershot and three-photon standard boson sampling where the arrival time of the photons was used explicitly. This retains the complexity of boson sampling however now every generated sample is a valid one. Further experimental work would increase the number of sources, ideally with FSRs small enough that no tuning of the emission wavelength is required. This would allow fully passive boson sampling experiments but with the small footprint, low power consumption and high squeezing of ring resonator sources. Theoretical extensions would be to quantify the amount of interference that accompanies any remaining partial distinguishability of the photons as in Ref. [164]. It would also be interesting to see how time resolved measurement can be used in Gaussian boson sampling experiments.

We designed and tested an all optical switch based on cross-phase modulation where the pump and signal fields are in different transverse waveguide modes. Preliminary results, with a silicon nitride device, showed a phase shift of $0.3\pi$ at a speed of 50 MHz. While not measured directly, the insertion loss of this device is potentially $<$1 dB. This is based of measured propagation loss of single mode waveguides of 0.3 dB cm$^{-1}$. This phase shift can be altered with the pumping scheme: narrower pulses lead to higher intensity but also to a reduced interaction length. In silicon nitride, as the propagation loss is low, increasing the length is a viable route to a higher phase shift. Moving to silicon, where the nonlinear coefficient is an order of magnitude higher and the field confinement is also higher, would need proper modelling of the effect of nonlinear losses. This effect could also be mitigated by moving to higher wavelengths For example both fields around 2.1 µm or the signal at telecom with a pump wavelength $>$4.2 µm.

Finally, we used linear optics to simulate the dynamics of systems described by non-Hermitian but $\mathcal{PT}$-symmetric Hamiltonians. Initially, we showed the ability of our simulator to recreate the dynamics of two and three mode systems. We then showed how these systems can be used to drive a qubit towards a desired, more coherent, target state even in the presence of a dephasing channel. To simulate a time dependent Hamiltonian we used an adaptive feedback loop. We experimentally verified the algorithm for both pure and mixed initial states. Following this, we used a numerical simulator of the algorithm to compare it to evolution at the systems exceptional point. We saw that the algorithm can be brought arbitrarily close to this evolution, even surpassing by a small amount. We showed that both the fixed evolution and algorithm are independent of input state, however both work best with target states close to the exceptional point eigenstate.

## OUTLOOK

The results in this thesis augment the existing integrated photonics literature and add evidence to support the feasibility and scalability of integrated quantum photonics. The requirements for fault-tolerant linear-optical quantum computing are quite daunting, specifically the millions of pure indistinguishable deterministic single photons and component loss rates below $10^{-3}$. While new architectures could possibly reduce overheads, these numbers are currently well beyond state-of-the art capabilities and may well require more than incremental improvements. That being said, the limiting factors in, for example, component loss in integrated platforms is the fabrication techniques rather than intrinsic material absorption. Giving room for the component loss to be reduced significantly. Our demonstration of interference of distinguishable photons with fast detectors relaxes slightly the indistinguishably criteria and shows that even commercially available single photon detectors, which have not been optimised for jitter, can be used to increase interference visibility of spectrally detuned photons.

Despite the component level engineering challenges which need to be overcome, there are higher level building blocks that can, and should, be demonstrated with current technology. Higher photon number experiments are inevitable and even with only current state-of-the art components we can expect integrated photonic experiments with tens of photons to become more common. This will open the door to heralded Bell and GHZ state generation, a vital step towards the deterministic generation of small entangled resource states. Another necessary building block for integrated quantum photonics is feed-forward, both multiplexing and adaptive measurement. To this end our demonstration of a high speed, and in principle low loss all-optical photonic switch in a CMOS compatible material is a step along this route.

Bespoke quantum computers designed to solve a specific problem are a potential intermediate step on the road to a universal quantum computer. The controllability and manufacturability of integrated photonics makes it a leading candidate for these devices. The results in chapter 6 show the feasibility of general integrated photonic devices to also simulate open quantum systems. The flexibility of photonics also allows specific devices to simulate open-quantum systems such as that outlined in Ref. [214].

# APPENDICES

"'These things will become clear to you', said the old man gently, 'at least', he added with slight doubt in his voice, 'clearer than they are at the moment.'"

Douglas Adams, *Hitchhikers Guide to the Galaxy*

# A

# TIME-RESOLVED REVERSE HOM

In this appendix we look to model time-reversed Hong-Ou-Mandel interference in a time-resolved manner. We use the analysis method shown in Appendix B of Ref. [90], adapted for our needs. We look to calculate the probability of a coincidence measurement in a reverse HOM experiment as a function of the arrival times of the two photons.

## A.1 TIME-RESOLVED REVERSE HOM

In the time domain we can write the rev-HOM input state as

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{2}} \iint dt_s dt_i \underbrace{T_1(t_s,t_i)\hat{a}_1^\dagger(t_s)\hat{a}_1^\dagger(t_i)\,|00\rangle}_{\text{source 1}} + e^{i2\phi} \underbrace{T_2(t_s,t_i)\hat{a}_2^\dagger(t_s)\hat{a}_2^\dagger(t_i)\,|00\rangle}_{\text{source 2}}, \tag{A.1}$$

where $T_i(t_s,t_i)$ is the joint temporal amplitude, the Fourier transform of the joint spectral amplitude. The beam-splitter relationships for a balanced beam-splitter are

$$\begin{aligned} \hat{a}_1^\dagger &\to \frac{1}{\sqrt{2}}\left(\hat{a}_3^\dagger + \hat{a}_4^\dagger\right), \\ \hat{a}_2^\dagger &\to \frac{1}{\sqrt{2}}\left(\hat{a}_3^\dagger - \hat{a}_4^\dagger\right). \end{aligned} \tag{A.2}$$

Applying these we arrive at

$$\begin{aligned} |\psi_{\text{out}}\rangle = \frac{1}{2\sqrt{2}} \iint dt_s dt_i T_1(t_s,t_i) &\left(\hat{a}_3^\dagger(t_s) + \hat{a}_4^\dagger(t_s)\right)\left(\hat{a}_3^\dagger(t_i) + \hat{a}_4^\dagger(t_i)\right)|00\rangle \\ &+ e^{i2\phi} T_2(t_s,t_i)\left(\hat{a}_3^\dagger(t_s) - \hat{a}_4^\dagger(t_s)\right)\left(\hat{a}_3^\dagger(t_i) - \hat{a}_4^\dagger(t_i)\right)|00\rangle . \end{aligned} \tag{A.3}$$

Expanding the brackets we can group terms where both photons leave in one waveguide and those where they leave in different waveguides. We then arrive at

$$|\psi_{\text{out}}\rangle = \frac{1}{2\sqrt{2}} \iint dt_s dt_i \underbrace{\left[T_1(t_s,t_i) + e^{i2\phi}T_2(t_s,t_i)\right]\left[\hat{a}_3^\dagger(t_s)\hat{a}_3^\dagger(t_i) + \hat{a}_4^\dagger(t_s)\hat{a}_4^\dagger(t_i)\right]}_{|\text{bunch}\rangle} + \underbrace{\left[T_1(t_s,t_i) - e^{i2\phi}T_2(t_s,t_i)\right]\left[\hat{a}_4^\dagger(t_s)\hat{a}_3^\dagger(t_i) + \hat{a}_3^\dagger(t_s)\hat{a}_4^\dagger(t_i)\right]}_{|\text{split}\rangle}. \tag{A.4}$$

We now introduce the time-dependent measurement projector. On a single mode this is given as $\hat{P}_i(t) = \hat{a}_i^\dagger(t)\,|0\rangle\,\langle 0|\,\hat{a}_i(t)$. The coincidence projector is therefore

$$\hat{P}(t_1,t_2) = \hat{P}_3(t_1) \otimes \hat{P}_4(t) = \hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle\,\langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1). \tag{A.5}$$

We can then see that $\langle\text{bunch}|\,\hat{P}(t_1,t_2)\,|\text{bunch}\rangle = 0\ \forall t_1,t_2$, and we can therefore continue with just $|\text{split}\rangle$. The coincidence probability is therefore given by

$$\begin{aligned}
P_{\text{coinc}}(t_1,t_2) &= \langle\text{split}|\,\hat{P}(t_1,t_2)\,|\text{split}\rangle \\
&= \frac{1}{8}\iiiint dt_s dt_i dt_s' dt_i' \left[T_1^*(t_s',t_i') - e^{-i2\phi}T_2^*(t_s',t_i')\right]\left[T_1(t_s,t_i) - e^{i2\phi}T_2(t_s,t_i)\right] \\
&\times \langle 0|\left[\hat{a}_4(t_s')\hat{a}_3(t_i') + \hat{a}_3(t_s')\hat{a}_4(t_i')\right]\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle \\
&\times \langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1)\left[\hat{a}_4^\dagger(t_s)\hat{a}_3^\dagger(t_i) + \hat{a}_3^\dagger(t_s)\hat{a}_4^\dagger(t_i)\right]|0\rangle \\
&= \iiiint dt_s dt_i dt_s' dt_i' \left[T_1^*(t_s',t_i') - e^{-i2\phi}T_2^*(t_s',t_i')\right]\left[T_1(t_s,t_i) - e^{i2\phi}T_2(t_s,t_i)\right] \\
&\times \Big[\,\langle 0|\,\hat{a}_4(t_s')\hat{a}_1(t_i')\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle\,\langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_3^\dagger(t_i)\hat{a}_4^\dagger(t_s)\,|0\rangle \\
&+ \langle 0|\,\hat{a}_4(t_s')\hat{a}_3(t_i')\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle\,\langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_3^\dagger(t_s)\hat{a}_4^\dagger(t_i)\,|0\rangle \\
&+ \langle 0|\,\hat{a}_4(t_i')\hat{a}_3(t_s')\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle\,\langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_3^\dagger(t_i)\hat{a}_4^\dagger(t_s)\,|0\rangle \\
&+ \langle 0|\,\hat{a}_4(t_i')\hat{a}_3(t_s')\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle\,\langle 0|\,\hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_3^\dagger(t_s)\hat{a}_4^\dagger(t_i)\,|0\rangle\,\Big].
\end{aligned} \tag{A.6}$$

We can simplify this further by noting that $\langle 0|\, \hat{a}_4(t_2')\hat{a}_3(t_1')\hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\,|0\rangle = \delta_{t_1 t_1'}\delta_{t_2 t_2'}$. This mean that from the integral only four terms are not 0. The probability therefore becomes

$$\begin{aligned} P_{\text{coinc}}(t_1,t_2) = \frac{1}{8}\Big( & \left[T_1^*(t_2,t_1) - e^{-i2\phi}T_2^*(t_2,t_1)\right]\left[T_1(t_2,t_1) - e^{i2\phi}T_2(t_2,t_1)\right] \\ & + \left[T_1^*(t_2,t_1) - e^{-i2\phi}T_2^*(t_2,t_1)\right]\left[T_1(t_1,t_2) - e^{i2\phi}T_2(t_1,t_2)\right] \\ & + \left[T_1^*(t_1,t_2) - e^{-i2\phi}T_2^*(t_1,t_2)\right]\left[T_1(t_2,t_1) - e^{i2\phi}T_2(t_2,t_1)\right] \\ & + \left[T_1^*(t_1,t_2) - e^{-i2\phi}T_2^*(t_1,t_2)\right]\left[T_1(t_1,t_2) - e^{i2\phi}T_2(t_1,t_2)\right]\Big) \\ = \frac{1}{2} & \left[T_1^*(t_1,t_2) - e^{-i2\phi}T_2^*(t_1,t_2)\right]\left[T_1(t_1,t_2) - e^{i2\phi}T_2(t_1,t_2)\right], \end{aligned} \tag{A.7}$$

where we have used the symmetry of the JTA $T(t_1,t_2) = T(t_2,t_1)$. Expanding this last step we arrive at

$$\begin{aligned} P_{\text{coinc}}(t_1,t_2) &= \frac{1}{4}\left(|T_1(t_1,t_2)|^2 + |T_2(t_1,t_2)|^2\right) - \frac{1}{2}\text{Re}\left(e^{i2\phi}T_1^*(t_1,t_2)T_2(t_1,t_2)\right) \\ &= \frac{1}{4}\left(|T_1(t_1,t_2)|^2 + |T_2(t_1,t_2)|^2\right) - \frac{1}{2}|T_1^*(t_1,t_2)T_2(t_1,t_2)|\cos(\phi+\theta_0), \end{aligned} \tag{A.8}$$

where $\theta_0 = \arg\left(T_1^*(t_1,t_2)T_2(t_1,t_2)\right)$ and we have also applied the complex identity $z + z^* = 2\text{Re}(z)$.

## A.2 SIMULATING RING JOINT TEMPORAL AMPLITUDE

To simulate the time resolved reverse HOM we first need to calculate the joint temporal amplitude of the photon sources. We will use photons generated with ring resonator sources with the same parameters as measured in chapter 3. First, we calculate the joint spectral amplitude. As shown in Sec. 2.4.3, the JSA for a SFWM source can be calculated as

$$F(\omega_s,\omega_i) = \frac{1}{P}\int \alpha(\omega_p)\alpha(\omega_s+\omega_i-\omega_p)\phi(\omega_p,\omega_s,\omega_i)d\omega_p. \tag{A.9}$$

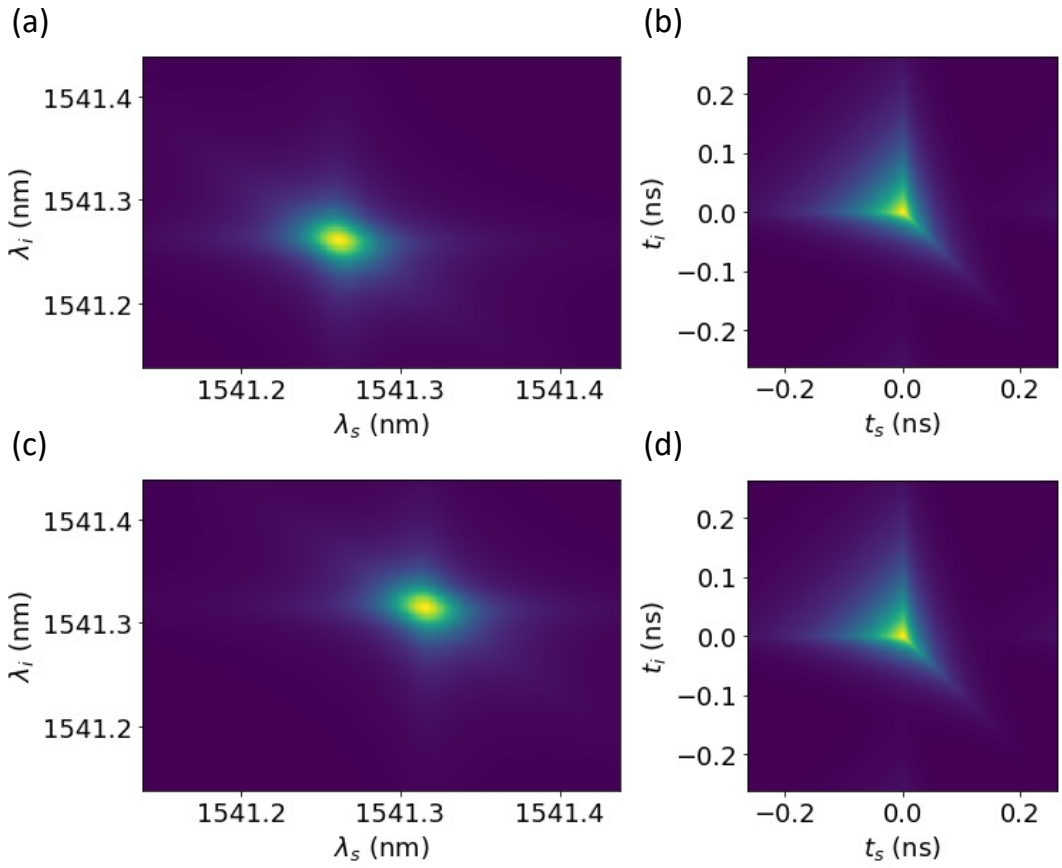


Figure A.1: Absolute values of JSA and JTA for two resonator photon sources. (a) JSA and (b) JTA for ring resonator source 1 (c) JSA and (d) JTA for ring resonator source 2, spectrally detuned by 54 pm from source 1.

For the simulation, we can set $P = 1$ and the geometry of ring resonator photon sources means that $\phi(\omega_p, \omega_s, \omega_i) \approx 1 \ \forall \omega_p, \omega_s, \omega_i$. With this in hand, we can retrieve the joint temporal amplitude with an inverse Fourier transform

$$T(t_s, t_i) = \frac{1}{2\pi} \iint F(\omega_s, \omega_i) e^{i(\omega_s t_s + \omega_i t_i)} d\omega_s d\omega_i. \tag{A.10}$$

In Fig. A.1 we show examples of the absolute value of the joint spectral and temporal amplitudes for the two sources. We have detuned them by 54 pm.

## A.3 TIME-RESOLVED REVERSE HOM FRINGES

With this we can use Eq. A.8 to simulate the reverse HOM fringe. Notice, in chapter 3 we set $t_2 = t_1 + \tau$ and looked at relative delays between the output modes. For this case that is not an option, the interference fringes are found in the absolute arrival times of the photons. For each setting of the phase shifter phase, we can therefore

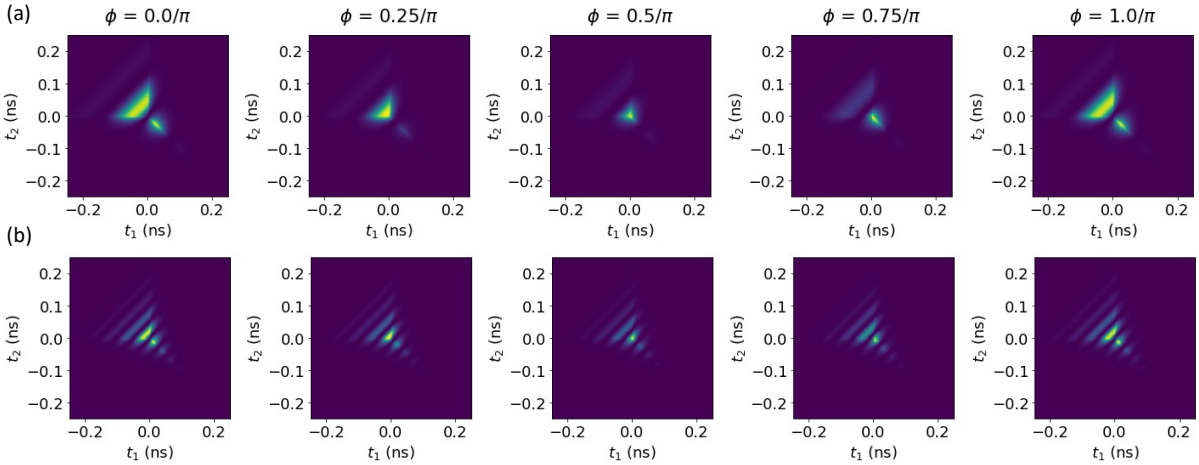


Figure A.2: Time resolved reverse HOM fringes. Phase $\phi$ is shown as a fraction of $\pi$ above each panel. (a) Sources detuned by 54 pm (b) Sources detuned by 100 pm

plot a 2D interference fringe in $t_s$ and $t_i$. In Fig. A.2 we have plotted these fringes for different phase values. As expected we see fringes in the time domain that oscillate with a frequency dependent on twice the applied phase. To see how the visibility of a rev-HOM fringe is affected by our timing window we would need to integrate these plots within a square area. We could then see how changing the size of this integration window changes our visibility.

# B

# TIME-RESOLVED HOM WITH MIXED PHOTON STATES

## B.1 TIME-RESOLVED HOM WITH MIXED INPUTS

While the results of Ref. [38] considers interference of photons with arbitrary spectral/temporal profiles the photons are assumed to be in pure states. In the following we detail the derivation of the coincidence probability of time-resolved HOM interference with photons in spectrally mixed state. To start we define an arbitrary photon density operator to be

$$\rho_i = \sum_k \lambda_k \left|\phi_{ik}\right\rangle \left\langle\phi_{ik}\right| , \tag{B.1}$$

where

$$\left|\phi_{ik}\right\rangle = \int dt\ \zeta_{ik}(t)\hat{a}_i^\dagger(t)\left|0\right\rangle_i . \tag{B.2}$$

We can then write the two photon input state $|11\rangle$ as

$$\rho_{\text{in}} = \rho_1 \otimes \rho_2 = \sum_{kl} \lambda_k \lambda_l' \left|\psi_{kl}\right\rangle \left\langle\psi_{kl}\right| , \tag{B.3}$$

with

$$\left|\psi_{kl}\right\rangle = \iint dtdt'\ \zeta_{1k}(t)\zeta_{2l}(t')\hat{a}_1^\dagger(t)\hat{a}_2^\dagger(t')\left|00\right\rangle . \tag{B.4}$$

Applying the beam-splitter relations to $\rho_{in}$ reduces to applying them to each of basis states $|\psi_{kl}\rangle$

$$\begin{aligned}\left|\psi_{kl}\right\rangle \rightarrow \left|\psi_{kl}\right\rangle_{\text{out}} &= \iint dtdt'\ \zeta_{1k}(t)\zeta_{2l}(t')\left[\hat{a}_3^\dagger(t)+\hat{a}_4^\dagger(t)\right]\left[\hat{a}_3^\dagger(t')-\hat{a}_4^\dagger(t')\right]\left|00\right\rangle \\ &= \iint dtdt'\ \zeta_{1k}(t)\zeta_{2l}(t')\left[\hat{a}_3^\dagger(t)\hat{a}_3^\dagger(t')+\hat{a}_4^\dagger(t)\hat{a}_3^\dagger(t')-\hat{a}_3^\dagger(t)\hat{a}_4^\dagger(t')-\hat{a}_4^\dagger(t)\hat{a}_4^\dagger(t')\right]\left|00\right\rangle\end{aligned} \tag{B.5}$$

to give a final density matrix

$$\rho_{\text{out}} = \sum_{kl} \lambda_k \lambda_l' \left|\psi_{kl}\right\rangle_{\text{out}} \left\langle\psi_{kl}\right|_{\text{out}} . \tag{B.6}$$

Using the time-resolved coincidence projector, $\hat{P}(t_1, t_2) = \hat{a}_3^\dagger(t_1)\hat{a}_4^\dagger(t_2)\left|0\right\rangle\left\langle 0\right|\hat{a}_4(t_2)\hat{a}_3(t_1)$, we arrive at

$$\begin{aligned}
P_{\text{coinc}}(t_1, t_2) &= \text{Tr}\left[\rho_{\text{out}} \times \hat{P}(t_1, t_2)\right] \\
&= \sum_{kl} \lambda_k \lambda_l' \,{}_{\text{out}}\left\langle \psi_{kl}\right| \hat{P}(t_1, t_2) \left|\psi_{kl}\right\rangle_{\text{out}} \\
&= \sum_{kl} \lambda_k \lambda_l' \left|\left\langle 0\right| \hat{a}_4(t_2)\hat{a}_3(t_1) \left|\psi_{kl}\right\rangle_{\text{out}}\right|^2 \\
&= \sum_{kl} \lambda_k \lambda_l' \Bigg| \iint dt dt' \zeta_{1k}(t)\zeta_{2l}(t') \\
&\qquad \times \left[\left\langle 0\right| \hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_4^\dagger(t)\hat{a}_3^\dagger(t') \left|0\right\rangle - \left\langle 0\right| \hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_3^\dagger(t)\hat{a}_4^\dagger(t') \left|0\right\rangle\right]\Bigg|^2
\end{aligned} \tag{B.7}$$

In the last step we have only kept the terms from Eq.B.5 where both photons leave from different beam-splitter outputs. To simplify this further, similar to Appendix A, we note that $\left\langle 0\right| \hat{a}_4(t_2)\hat{a}_3(t_1)\hat{a}_4^\dagger(t_2')\hat{a}_3^\dagger(t_1') \left|0\right\rangle = \delta_{t_1 t_1'}\delta_{t_2 t_2'}$. This leaves us with only two terms from the double integral leaving a final probability of

$$P_{\text{coinc}}(t_1, t_2) = \sum_{kl} \lambda_k \lambda_l' \left|\zeta_{1k}(t_2)\zeta_{2l}(t_1) - \zeta_{1k}(t_1)\zeta_{2l}(t_2)\right|^2. \tag{B.8}$$

Interestingly here we notice that if $t_1 = t_2 = t$ $P_{\text{coinc}}(t, t) = 0$, regardless of the number or specific shapes of the Schmidt modes which make up the photon states used.

# C

## MODELLING DOUBLE EMISSIONS

Here, we provide the full derivation of the two-photon interference probability including double emissions. Initially, we look at the case where only two photons are generated and both are produced from a single source, then we extend this to look at the case where two photons are detected but three photons are generated - two from one source and one from the other. We begin by looking at the case where either source fires twice, $|\psi\rangle_i = \frac{\hat{A}_i^\dagger \hat{A}_i^\dagger}{2}|00\rangle$, where the subscript $i$ denotes which source fires into the corresponding spatial mode. Applying the transformations in Eq. 3.14 and including the transmission of each mode, $\eta_i$, and the detection time of the photons we arrive at

$$\begin{aligned} |\psi_1\rangle &= 2M_{11}M_{21}\sqrt{\eta_1\eta_2}\zeta_1(t_0)\zeta_1(t_0+\tau)\hat{a}_1^\dagger\hat{a}_2^\dagger|00\rangle\,, \\ |\psi_2\rangle &= 2M_{12}M_{22}\sqrt{\eta_1\eta_2}\zeta_2(t_0)\zeta_2(t_0+\tau)\hat{a}_1^\dagger\hat{a}_2^\dagger|00\rangle\,. \end{aligned} \tag{C.1}$$

The total probability is given by $P_i = P_{\text{TMSV}}(2) \times |\langle 0|\hat{a}_1\hat{a}_2|\psi_i\rangle|^2$ which results in:

$$\begin{aligned} P_1 &= 4\,\text{sech}^2\xi_1\tanh^4\xi_1\eta_1\eta_2|M_{11}M_{21}\zeta_1(t_0)\zeta_1(t_0+\tau)|^2, \\ P_2 &= 4\,\text{sech}^2\xi_2\tanh^4\xi_2\eta_1\eta_2|M_{12}M_{22}\zeta_2(t_0)\zeta_2(t_0+\tau)|^2. \end{aligned} \tag{C.2}$$

This case was included only for completeness, in the experiments in the main text we herald the input photon state and post-select on a photon being heralded from each source. This means that we always have at least one photon pair generated from each source.

Finally, we look at the case where one source fires twice and the second source fires once. This consists of the predominant noise terms that we encounter in the later experiments. To aid clarity but without loss of generality we will assume that the two sources have different central frequencies. The two input states are then

$$\begin{aligned} |\psi\rangle_{\text{in},1} &= \hat{a}_{1,\omega_1}^\dagger\hat{a}_{1,\omega_1}^\dagger\hat{a}_{2,\omega_2}^\dagger|0\rangle = |2_{1,\omega_1}, 1_{2,\omega_2}\rangle\,, \\ |\psi\rangle_{\text{in},2} &= \hat{a}_{1,\omega_1}^\dagger\hat{a}_{2,\omega_2}^\dagger\hat{a}_{2,\omega_2}^\dagger|0\rangle = |1_{1,\omega_1}, 2_{2,\omega_2}\rangle\,, \end{aligned} \tag{C.3}$$

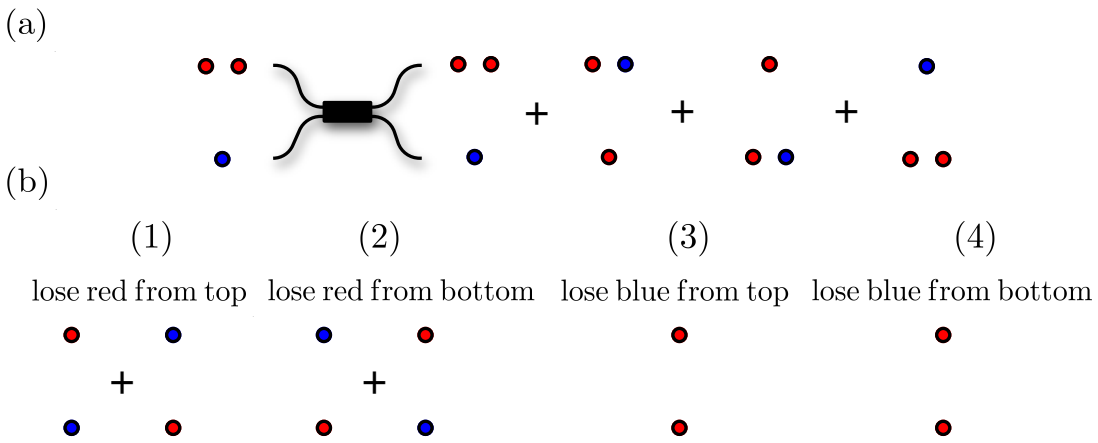


Figure C.1: Reproducing the figure from chapter 3 to help clarity of this derivation. (a) Shows the possible output patterns that arise if one source fires twice and the other source fires once. (b) Shows the remaining patterns if one photon is lost.

where $\hat{a}^{\dagger}_{i,\omega_j}$ denotes a photon generated in spatial mode $i$ and frequency mode $j$. In our scenario we note that source $i$ emits a photon centred at frequency $\omega_i$. Initially, we will focus on $|\psi\rangle_{\text{in},1}$, after interference through $M$. We see 4 distinct output patterns which could lead to a coincidence click

$$|\psi_{\text{out}}\rangle = M_{11}M_{11}M_{22}\,|2_{1,\omega_1}, 1_{2,\omega_2}\rangle + 2M_{11}M_{21}M_{11}\,|1_{1,\omega_1}, 1_{2,\omega_1}, 1_{2,\omega_2}\rangle + 2M_{11}M_{21}M_{12}\,|1_{1,\omega_1}, 1_{2,\omega_1}, 1_{1,\omega_2}\rangle + M_{21}M_{21}M_{12}\,|2_{2,\omega_1}, 1_{1,\omega_2}\rangle\,, \quad \text{(C.4)}$$

where some output states have a multiplicative factor indicating that there is more than one way that this pattern can arise. This is shown illustratively in Fig. C.1(a). Detectors with a finite dead-time, much longer than the photon's coherence time, will only ever measure two of these photons, therefore, we need only consider the cases where two of the three photons make it to the detectors. We now sum over the four ways to lose one photon from this state $\sum_i^2 \sum_j^2 \hat{a}_{i,\omega_j}\ |\psi\rangle_{\text{out}}$, from this we only keep terms in within

the coincidence subspace. Again, this is illustrated in Fig C.1(b). For simplicity, we will take each of the 4 scenarios in Fig C.1(b) separately and then incoherently sum them

$$\begin{aligned}
|\psi\rangle_1 &= \mathrm{M}_{11}\mathrm{M}_{11}\mathrm{M}_{22} \times 2\sqrt{\eta_1\eta_2(1-\eta_1)}\,|1_{1,\omega_1}, 1_{2,\omega_2}\rangle \\
&\quad + 2\mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{12} \times \sqrt{\eta_1\eta_2(1-\eta_1)}\,|1_{1,\omega_2}, 1_{2,\omega_1}\rangle\,, \\
|\psi\rangle_2 &= \mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{22} \times 2\sqrt{\eta_1\eta_2(1-\eta_2)}\,|1_{1,\omega_1}, 1_{2,\omega_2}\rangle \\
&\quad + 2\mathrm{M}_{21}\mathrm{M}_{21}\mathrm{M}_{12} \times \sqrt{\eta_1\eta_2(1-\eta_1)}\,|1_{1,\omega_2}, 1_{2,\omega_1}\rangle\,, \\
|\psi\rangle_3 &= 2\mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{12}\sqrt{\eta_1\eta_2(1-\eta_1)}\,|1_{1,\omega_1}, 1_{2,\omega_1}\rangle \\
|\psi\rangle_4 &= 2\mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{22}\sqrt{\eta_1\eta_2(1-\eta_2)}\,|1_{1,\omega_1}, 1_{2,\omega_1}\rangle\,.
\end{aligned} \tag{C.5}$$

Here, the factor of 2 in front of matrix elements illustrates that there are two ways for this pattern to arise and the factor of 2 in front of the transmissions shows that there are two ways to lose a photon and end up with this pattern. We now map the photon frequency to the time dependent field operators $\hat{a}^\dagger_{i,\omega_j} \to \hat{A}^\dagger_j(t_i)$, with $t_1 = t_0$ and $t_2 = t_0 + \tau$ and calculate the probabilities to be

$$\begin{aligned}
\mathrm{P}_{1,1} &= 4|\mathrm{M}_{11}|^2\eta_1\eta_2(1-\eta_1)|\mathrm{M}_{12}\mathrm{M}_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + \mathrm{M}_{11}\mathrm{M}_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2, \\
\mathrm{P}_{1,2} &= 4|\mathrm{M}_{21}|^2\eta_1\eta_2(1-\eta_2)|\mathrm{M}_{12}\mathrm{M}_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + \mathrm{M}_{11}\mathrm{M}_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2, \\
\mathrm{P}_{1,3} &= 4\eta_1\eta_2(1-\eta_1)|\mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{12}\zeta_1(t_0)\zeta_1(t_0+\tau)|^2, \\
\mathrm{P}_{1,4} &= 4\eta_1\eta_2(1-\eta_2)|\mathrm{M}_{11}\mathrm{M}_{21}\mathrm{M}_{22}\zeta_1(t_0)\zeta_1(t_0+\tau)|^2,
\end{aligned} \tag{C.6}$$

where $\mathrm{P}_{i,j}$ represents the probability of pattern $j$ given a double emission from source $i$. It is interesting to note here that, in two outcomes where one of the double emission photons is lost, there is still interference between the remaining photons. The final probability needs to be scaled by the probability of generating the input state

$$\mathrm{P}_{\mathrm{double},1} = \mathrm{P}_{\mathrm{gen},1}(\mathrm{P}_{1,1} + \mathrm{P}_{1,2} + \mathrm{P}_{1,3} + \mathrm{P}_{1,4}), \tag{C.7}$$

where $P_{gen,1} = \operatorname{sech}^2 \xi_1 \tanh^4 \xi_1 \times \operatorname{sech}^2 \xi_2 \tanh^2 \xi_2$. A similar analysis can be used for the mirror situation, double emission from source two

$$\begin{aligned} P_{2,1} &= 4|M_{22}|^2\eta_1\eta_2(1-\eta_2)|M_{12}M_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + M_{11}M_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2 \\ P_{2,2} &= 4|M_{12}|^2\eta_1\eta_2(1-\eta_1)|M_{12}M_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + M_{11}M_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2 \\ P_{2,3} &= 4\eta_1\eta_2(1-\eta_1)|M_{11}M_{22}M_{12}\zeta_2(t_0)\zeta_2(t_0+\tau)|^2 \\ P_{2,4} &= 4\eta_1\eta_2(1-\eta_2)|M_{12}M_{21}M_{22}\zeta_2(t_0)\zeta_2(t_0+\tau)|^2. \end{aligned} \tag{C.8}$$

Combing this with the generation probability, $P_{gen,2} = \operatorname{sech}^2 \xi_1 \tanh^2 \xi_1 \times \operatorname{sech}^2 \xi_2 \tanh^4 \xi_2$, we arrive at the final probability

$$P_{double,2} = P_{gen,2}(P_{2,1} + P_{2,2} + P_{2,3} + P_{2,4}), \tag{C.9}$$

We can now see that all the probabilities, Eqs. 3.21,C.7,C.9, only contain either interference terms or noise terms, we can thus write the final probability as

$$\begin{aligned} P_f = A|M_{12}M_{21}\zeta_2(t_0)\zeta_1(t_0+\tau) + M_{11}M_{22}\zeta_1(t_0)\zeta_2(t_0+\tau)|^2 + \\ B|\zeta_1(t_0)\zeta_1(t_0+\tau)|^2 + C|\zeta_2(t_0)\zeta_2(t_0+\tau)|^2, \end{aligned} \tag{C.10}$$

where

$$\begin{aligned} A &= \operatorname{sech}^2 \xi_1 \tanh^2 \xi_1 \operatorname{sech}^2 \xi_2 \tanh^2 \xi_2 \eta_1\eta_2 \Big[1 + 4\tanh^2 \xi_1 \left(|M_{11}|^2 (1-\eta_1) + |M_{21}|^2 (1-\eta_2)\right) \\ &\qquad + 4\tanh^2 \xi_2 \left(|M_{22}|^2 (1-\eta_2) + |M_{12}|^2 (1-\eta_1)\right)\Big] \\ B &= 4\operatorname{sech}^2 \xi_1 \tanh^4 \xi_1 \operatorname{sech}^2 \xi_2 \tanh^2 \xi_2 \eta_1\eta_2 |M_{11}M_{21}|^2 [|M_{12}|^2(1-\eta_1) + |M_{22}|^2(1-\eta_2)] \\ C &= 4\operatorname{sech}^2 \xi_1 \tanh^2 \xi_1 \operatorname{sech}^2 \xi_2 \tanh^4 \xi_2 |M_{12}M_{22}|^2 [|M_{11}|^2(1-\eta_1) + |M_{21}|^2(1-\eta_2)]. \end{aligned} \tag{C.11}$$

# BIBLIOGRAPHY


[1] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. ⟨1, 82⟩

[2] Edwin Pednault, John A Gunnels, Giacomo Nannicini, Lior Horesh, and Robert Wisnieff. Leveraging secondary storage to simulate deep 54-qubit sycamore circuits. *arXiv preprint arXiv:1910.09534*, 2019. ⟨1⟩

[3] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, et al. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020. ⟨1, 82, 83, 89⟩

[4] Jacob FF Bulmer, Bryn A Bell, Rachel S Chadwick, Alex E Jones, Diana Moise, Alessandro Rigazzi, Jan Thorbecke, Utz-Uwe Haus, Thomas Van Vaerenbergh, Raj B Patel, et al. The boundary for quantum advantage in Gaussian boson sampling. *arXiv preprint arXiv:2108.01622*, 2021. ⟨1⟩

[5] Han-Sen Zhong, Yu-Hao Deng, Jian Qin, Hui Wang, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Dian Wu, Si-Qiu Gong, Hao Su, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Jelmer Renema, Chao-Yang Lu, and Jian-Wei Pan. Phase-programmable Gaussian boson sampling using stimulated squeezed light, 2021. ⟨1, 82⟩

[6] Yulin Wu, Wan-Su Bao, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, et al. Strong quantum computational advantage using a superconducting quantum processor. *arXiv preprint arXiv:2106.14734*, 2021. ⟨1⟩

[7] John Preskill. Quantum computing and the entanglement frontier. *arXiv preprint arXiv:1203.5813*, 2012. ⟨1⟩

[8] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 333–342, 2011. ⟨1, 82, 84⟩

[9] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001. ⟨1, 42⟩

[10] Mercedes Gimeno-Segovia, Pete Shadbolt, Dan E Browne, and Terry Rudolph. From three-photon Greenberger-Horne-Zeilinger states to ballistic universal quantum computation. *Physical Review Letters*, 115(2):020502, 2015. ⟨1, 40, 43⟩

[11] Konrad Kieling, Terry Rudolph, and Jens Eisert. Percolation, renormalization, and quantum computing with nondeterministic gates. *Physical Review Letters*, 99(13):130501, 2007. ⟨1, 43⟩

[12] Sara Bartolucci, Patrick Birchall, Hector Bombin, Hugo Cable, Chris Dawson, Mercedes Gimeno-Segovia, Eric Johnston, Konrad Kieling, Naomi Nickerson, Mihir Pant, et al. Fusion-based quantum computation. *arXiv preprint arXiv:2101.09310*, 2021. ⟨1⟩

[13] Isaac H Kim, Eunseok Lee, Ye-Hua Liu, Sam Pallister, William Pol, and Sam Roberts. Fault-tolerant resource estimate for quantum chemical simulations: case study on Li-ion battery electrolyte molecules. *arXiv preprint arXiv:2104.10653*, 2021. ⟨1, 43⟩

[14] Stefano Paesani, Yunhong Ding, Raffaele Santagati, Levon Chakhmakhchyan, Caterina Vigliar, Karsten Rottwitt, Leif K Oxenløwe, Jianwei Wang, Mark G Thompson, and Anthony Laing. Generation and sampling of quantum states of light in a silicon chip. *Nature Physics*, 15(9):925–929, 2019. ⟨2, 30, 53, 83, 89, 124⟩

[15] Jianwei Wang, Stefano Paesani, Yunhong Ding, Raffaele Santagati, Paul Skrzypczyk, Alexia Salavrakos, Jordi Tura, Remigiusz Augusiak, Laura Mančinska, Davide Bacco, et al. Multidimensional quantum entanglement with large-scale integrated optics. *Science*, 360(6386):285–291, 2018. ⟨2, 44⟩

[16] Caterina Vigliar, Stefano Paesani, Yunhong Ding, Jeremy C Adcock, Jianwei Wang, Sam Morley-Short, Davide Bacco, Leif K Oxenløwe, Mark G Thompson, John G

Rarity, et al. Error-protected qubits in a silicon photonic chip. *Nature Physics*, 17(10):1137–1143, 2021. ⟨2⟩

[17] JM Arrazola, V Bergholm, K Brádler, TR Bromley, MJ Collins, I Dhand, A Fumagalli, T Gerrits, A Goussev, LG Helt, et al. Quantum circuits with many photons on a programmable nanophotonic chip. *Nature*, 591(7848):54–60, 2021. ⟨2, 83⟩

[18] Caterina Taballione, Tom AW Wolterink, Jasleen Lugani, Andreas Eckstein, Bryn A Bell, Robert Grootjans, Ilka Visscher, Dimitri Geskus, Chris GH Roeloffzen, Jelmer J Renema, et al. 8× 8 reconfigurable quantum photonic processor based on silicon nitride waveguides. *Optics Express*, 27(19):26842–26857, 2019. ⟨2, 14⟩

[19] Mian Zhang, Cheng Wang, Yaowen Hu, Amirhassan Shams-Ansari, Tianhao Ren, Shanhui Fan, and Marko Lončar. Electronically programmable photonic molecule. *Nature Photonics*, 13(1):36–40, 2019. ⟨2, 30, 107⟩

[20] Di Zhu, Linbo Shao, Mengjie Yu, Rebecca Cheng, Boris Desiatov, CJ Xin, Yaowen Hu, Jeffrey Holzgrafe, Soumya Ghosh, Amirhassan Shams-Ansari, et al. Integrated photonics on thin-film lithium niobate. *Advances in Optics and Photonics*, 13(2):242–352, 2021. ⟨2, 30, 107⟩

[21] Zhaohui Ma, Jia-Yang Chen, Yong Meng Sua, Zhan Li, Chao Tang, and Yu-Ping Huang. Efficient photon pair generation in thin-film lithium niobate microring resonators. In *Quantum 2.0*, pages QM6B–2. Optical Society of America, 2020. ⟨2, 30, 107⟩

[22] Mengyue Xu, Mingbo He, Hongguang Zhang, Jian Jian, Ying Pan, Xiaoyue Liu, Lifeng Chen, Xiangyu Meng, Hui Chen, Zhaohui Li, et al. High-performance coherent optical modulators based on thin-film lithium niobate platform. *Nature Communications*, 11(1):1–7, 2020. ⟨2, 107⟩

[23] Mian Zhang, Cheng Wang, Prashanta Kharel, Di Zhu, and Marko Lončar. Integrated lithium niobate electro-optic modulators: when performance meets scalability. *Optica*, 8(5):652–667, 2021. ⟨2, 107⟩

[24] Felix Eltes, Gerardo E Villarreal-Garcia, Daniele Caimi, Heinz Siegwart, Antonio A Gentile, Andy Hart, Pascal Stark, Graham D Marshall, Mark G Thompson,

Jorge Barreto, et al. An integrated optical modulator operating at cryogenic temperatures. *Nature Materials*, 19(11):1164–1168, 2020. ⟨2, 107⟩

[25] Jianwei Wang, Alberto Santamato, Pisu Jiang, Damien Bonneau, Erman Engin, Joshua W Silverstone, Matthias Lermer, Johannes Beetz, Martin Kamp, Sven Höfling, et al. Gallium arsenide (GaAs) quantum photonic waveguide circuits. *Optics Communications*, 327:49–55, 2014. ⟨2, 107⟩

[26] Marlon Placke and Sven Ramelow. Engineering AlGaAs-on-insulator toward quantum optical applications. *Optics Letters*, 45(24):6763–6766, 2020. ⟨2, 107⟩

[27] Trevor J Steiner, Joshua E Castro, Lin Chang, Quynh Dang, Weiqiang Xie, Justin Norman, John E Bowers, and Galan Moody. Ultrabright entangled-photon-pair generation from an AlGaAs-On-Insulator microring resonator. *PRX Quantum*, 2(1):010337, 2021. ⟨2, 107⟩

[28] Peter C Humphreys, Benjamin J Metcalf, Justin B Spring, Merritt Moore, Xian-Min Jin, Marco Barbieri, W Steven Kolthammer, and Ian A Walmsley. Linear optical quantum computing in a single spatial mode. *Physical Review Letters*, 111(15):150501, 2013. ⟨2, 36⟩

[29] Robert Raussendorf, Jim Harrington, and Kovid Goyal. A fault-tolerant one-way quantum computer. *Annals of Physics*, 321(9):2242–2270, 2006. ⟨2⟩

[30] Robert Raussendorf and Hans J Briegel. A one-way quantum computer. *Physical Review Letters*, 86(22):5188, 2001. ⟨2, 106⟩

[31] Robert Raussendorf. Measurement-based quantum computation with cluster states. *International Journal of Quantum Information*, 7(06):1053–1203, 2009. ⟨2, 43⟩

[32] Robert Raussendorf, Daniel E Browne, and Hans J Briegel. Measurement-based quantum computation on cluster states. *Physical Review A*, 68(2):022312, 2003. ⟨2⟩

[33] Naomi Nickerson and Héctor Bombín. Measurement based fault tolerance beyond foliation. *arXiv preprint arXiv:1810.09621*, 2018. ⟨2⟩

[34] Michael Vasmer and Dan E Browne. Three-dimensional surface codes: Transversal gates and fault-tolerant architectures. *Physical Review A*, 100(1):012312, 2019. ⟨2⟩

[35] Thomas Legero, Tatjana Wilk, Markus Hennrich, Gerhard Rempe, and Axel Kuhn. Quantum beat of two single photons. *Physical Review Letters*, 93(7):070503, 2004. ⟨2, 78⟩

[36] Xu-Jie Wang, Bo Jing, Peng-Fei Sun, Chao-Wei Yang, Yong Yu, Vincenzo Tamma, Xiao-Hui Bao, and Jian-Wei Pan. Experimental time-resolved interference with multiple photons of different colors. *Physical Review Letters*, 121(8):080501, 2018. ⟨2, 78⟩

[37] Tian-Ming Zhao, Han Zhang, Jian Yang, Zi-Ru Sang, Xiao Jiang, Xiao-Hui Bao, and Jian-Wei Pan. Entangling different-color photons via time-resolved measurement and active feed forward. *Physical Review Letters*, 112(10):103602, 2014. ⟨2, 78⟩

[38] Thomas Legero, Tatjana Wilk, Axel Kuhn, and Gerhard Rempe. Time-resolved two-photon quantum interference. *Applied Physics B*, 77(8):797–802, 2003. ⟨2, 48, 177⟩

[39] Matthew J Collins, Chunle Xiong, Isabella H Rey, Trung D Vo, Jiakun He, Shayan Shahnia, Christopher Reardon, Thomas F Krauss, MJ Steel, Alex S Clark, et al. Integrated spatial multiplexing of heralded single-photon sources. *Nature Communications*, 4(1):1–7, 2013. ⟨2⟩

[40] Rodney Loudon. *The quantum theory of light*. OUP Oxford, 2000. ⟨4⟩

[41] Daniel F Walls and Gerard J Milburn. *Quantum optics*. Springer Science Business Media, 2007. ⟨4⟩

[42] A Hurwitz. Über die Erzeugung der Invarianten durch Integration. *Akademie Wissenschaften Göttingen*, pages 546–564, 1897. ⟨11⟩

[43] Michael Reck, Anton Zeilinger, Herbert J Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Physical Review Letters*, 73(1):58, 1994. ⟨11⟩

[44] J Nolen. Synthesis of multiple beam networks for arbitrary illuminations. *Bendix Corp*, 1965. ⟨11⟩

[45] Daniel Jost Brod and Michał Oszmaniec. Classical simulation of linear optics subject to nonuniform losses. *Quantum*, 4:267, 2020. ⟨13⟩

[46] Michał Oszmaniec and Daniel J Brod. Classical simulation of photonic linear optics with lost particles. *New Journal of Physics*, 20(9):092002, 2018. ⟨13⟩

[47] William R Clements, Peter C Humphreys, Benjamin J Metcalf, W Steven Kolthammer, and Ian A Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3(12):1460–1465, 2016. ⟨13⟩

[48] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J Russell, Joshua W Silverstone, Peter J Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, et al. Universal linear optics. *Science*, 349(6249):711–716, 2015. ⟨14, 30, 74, 166⟩

[49] Chaitali Joshi, Alessandro Farsi, Avik Dutt, Bok Young Kim, Xingchen Ji, Yun Zhao, Andrew M Bishop, Michal Lipson, and Alexander L Gaeta. Frequency-domain quantum interference with correlated photons from an integrated microresonator. *Physical Review Letters*, 124(14):143601, 2020. ⟨14⟩

[50] Toshiki Kobayashi, Rikizo Ikuta, Shuto Yasui, Shigehito Miki, Taro Yamashita, Hirotaka Terai, Takashi Yamamoto, Masato Koashi, and Nobuyuki Imoto. Frequency-domain Hong-Ou-Mandel interference. *Nature Photonics*, 10(7):441–444, 2016. ⟨14⟩

[51] Yu He, X Ding, Z-E Su, H-L Huang, J Qin, C Wang, S Unsleber, C Chen, H Wang, Y-M He, et al. Time-bin-encoded boson sampling with a single-photon device. *Physical Review Letters*, 118(19):190501, 2017. ⟨14⟩

[52] Aseema Mohanty, Mian Zhang, Avik Dutt, Sven Ramelow, Paulo Nussenzveig, and Michal Lipson. Quantum interference between transverse spatial waveguide modes. *Nature Communications*, 8(1):1–7, 2017. ⟨14⟩

[53] Keith R Motes, Alexei Gilchrist, Jonathan P Dowling, and Peter P Rohde. Scalable boson sampling with time-bin encoding using a loop-based architecture. *Physical Review Letters*, 113(12):120501, 2014. ⟨14⟩

[54] William Clements. *Linear quantum optics: components and applications*. PhD thesis, University of Oxford, 2018. ⟨15, 16⟩

[55] Peter J Mosley. *Generation of heralded single photons in pure quantum states*. PhD thesis, University of Oxford, 2007. ⟨20⟩

[56] Alexander Edward Jones. *Distinguishability in quantum interference*. PhD thesis, Imperial College London, 2019. ⟨20, 58⟩

[57] Robert W Boyd. *Nonlinear optics*. Academic press, 2020. ⟨20, 109⟩

[58] Jesper Bjerge Christensen. *Tailored four-wave-mixing processes for optical quantum information science*. PhD thesis, Technical University of Denmark, 2018. ⟨20⟩

[59] U Fischer, T Zinke, J-R Kropp, F Arndt, and K Petermann. 0.1 db/cm waveguide losses in single-mode SOI rib waveguides. *IEEE Photonics Technology Letters*, 8(5):647–648, 1996. ⟨23⟩

[60] Yurii A Vlasov, Martin O'boyle, Hendrik F Hamann, and Sharee J McNab. Active control of slow light on a chip with photonic crystal waveguides. *Nature*, 438(7064):65–69, 2005. ⟨23⟩

[61] Yohei Hamachi, Shousaku Kubo, and Toshihiko Baba. Slow light with low dispersion and nonlinear enhancement in a lattice-shifted photonic crystal waveguide. *Optics Letters*, 34(7):1072–1074, 2009. ⟨23⟩

[62] Pisu Jiang and Krishna C Balram. Suspended gallium arsenide platform for building large scale photonic integrated circuits: passive devices. *Optics Express*, 28(8):12262–12271, 2020. ⟨23⟩

[63] Govind P Agrawal. *Nonlinear fiber optics*. Academic press, 2001. ⟨24, 112, 131⟩

[64] Zhaoyang Li, Cheng Wang, Shuai Li, Yi Xu, Lei Chen, Yaping Dai, and Yuxin Leng. Fourth-order dispersion compensation for ultra-high power femtosecond lasers. *Optics Communications*, 357:71–77, 2015. ⟨24⟩

[65] Faraz Najafi, Jacob Mower, Nicholas C Harris, Francesco Bellei, Andrew Dane, Catherine Lee, Xiaolong Hu, Prashanta Kharel, Francesco Marsili, Solomon Assefa, et al. On-chip detection of non-classical light by scalable integration of single-photon detectors. *Nature Communications*, 6(1):1–8, 2015. ⟨26⟩

[66] Patrik Rath, Oliver Kahl, Simone Ferrari, Fabian Sproll, Georgia Lewes-Malandrakis, Dietmar Brink, Konstantin Ilin, Michael Siegel, Christoph Nebel, and Wolfram Pernice. Superconducting single-photon detectors integrated with

diamond nanophotonic circuits. *Light: Science & Applications*, 4(10):e338–e338, 2015. ⟨26⟩

[67] Samuel Gyger, Julien Zichi, Lucas Schweickert, Ali W Elshaari, Stephan Steinhauer, Saimon F Covre da Silva, Armando Rastelli, Val Zwiller, Klaus D Jöns, and Carlos Errando-Herranz. Reconfigurable photonics with on-chip single-photon detectors. *Nature Communications*, 12(1):1–8, 2021. ⟨26⟩

[68] Emma Lomonte, Martin A Wolff, Fabian Beutel, Simone Ferrari, Carsten Schuck, Wolfram HP Pernice, and Francesco Lenzini. Single-photon detection and cryogenic reconfigurability in lithium niobate nanophotonic circuits. *arXiv preprint arXiv:2103.10973*, 2021. ⟨26⟩

[69] Yunhong Ding, Christophe Peucheret, Haiyan Ou, and Kresten Yvind. Fully etched apodized grating coupler on the soi platform with -0.58 db coupling efficiency. *Optics Letters*, 39(18):5348–5350, 2014. ⟨27⟩

[70] Lawrence M Rosenfeld. *Mid-infrared quantum optics in silicon: light work with long waves*. PhD thesis, University of Bristol, 2020. ⟨28, 31, 135⟩

[71] Marvin J Weber. *Handbook of optical materials*. CRC press, 2018. ⟨30⟩

[72] Marie Wandel. *Attenuation in silica-based optical fibers*. PhD thesis, Technical University of Denmark, 2005. ⟨30⟩

[73] Ulrich Gubler and Christian Bosshard. Optical third-harmonic generation of fused silica in gas atmosphere: Absolute value of the third-order nonlinear optical susceptibility $\chi^{(3)}$. *Physical Review B*, 61(16):10702, 2000. ⟨30⟩

[74] Lianghong Yin and Govind P Agrawal. Impact of two-photon absorption on self-phase modulation in silicon waveguides. *Optics Letters*, 32(14):2031–2033, 2007. ⟨30⟩

[75] Kazuhiro Ikeda, Robert E Saperstein, Nikola Alic, and Yeshaiahu Fainman. Thermal and Kerr nonlinear properties of plasma-deposited silicon nitride/silicon dioxide waveguides. *Optics Express*, 16(17):12987–12994, 2008. ⟨30⟩

[76] DTH Tan, Kazuhiro Ikeda, PC Sun, and Yeshaiahu Fainman. Group velocity dispersion and self phase modulation in silicon nitride waveguides. *Applied Physics Letters*, 96(6):061101, 2010. ⟨30⟩

[77] HK Tsang and Y Liu. Nonlinear optical properties of silicon waveguides. *Semiconductor Science and Technology*, 23(6):064007, 2008. ⟨30⟩

[78] David J Moss, Roberto Morandotti, Alexander L Gaeta, and Michal Lipson. New CMOS-compatible platforms based on silicon nitride and Hydex for nonlinear optics. *Nature Photonics*, 7(8):597–607, 2013. ⟨30⟩

[79] E Güler, G Uğur, Ş Uğur, and M Güler. A theoretical study for the band gap energies of the most common silica polymorphs. *Chinese Journal of Physics*, 65:472–480, 2020. ⟨30⟩

[80] W Bludau, A Onton, and W Heinke. Temperature dependence of the band gap of silicon. *Journal of Applied Physics*, 45(4):1846–1848, 1974. ⟨30⟩

[81] AH Atabaki, E Shah Hosseini, AA Eftekhar, S Yegnanarayanan, and A Adibi. Optimization of metallic microheaters for high-speed reconfigurable silicon photonics. *Optics Express*, 18(17):18312–18323, 2010. ⟨30⟩

[82] Hwan Soo Dow, Woo Sik Kim, and Jung Woo Lee. Thermal and electrical properties of silicon nitride substrates. *AIP Advances*, 7(9):095022, 2017. ⟨30⟩

[83] Amir Arbabi and Lynford L Goddard. Measurements of the refractive indices and thermo-optic coefficients of $Si_3N_3$ and $SiO_x$ using microring resonances. *Optics Letters*, 38(19):3878–3881, 2013. ⟨30⟩

[84] Ian H Malitson. Interspecimen comparison of the refractive index of fused silica. *Josa*, 55(10):1205–1209, 1965. ⟨30⟩

[85] O Alibart, J Fulconis, GKL Wong, SG Murdoch, WJ Wadsworth, and JG Rarity. Photon pair generation using four-wave mixing in a microstructured fibre: theory versus experiment. *New Journal of Physics*, 8(5):67, 2006. ⟨31, 107⟩

[86] Alex R McMillan, J Fulconis, M Halder, C Xiong, JG Rarity, and William J Wadsworth. Narrowband high-fidelity all-fibre source of heralded single photons at 1570 nm. *Optics Express*, 17(8):6156–6165, 2009. ⟨31⟩

[87] Lawrence M Rosenfeld, Dominic A Sulway, Gary F Sinclair, Vikas Anant, Mark G Thompson, John G Rarity, and Joshua W Silverstone. Mid-infrared quantum optics in silicon. *Optics Express*, 28(25):37092–37102, 2020. ⟨31, 112, 135⟩

[88] Chong-Ki Hong, Zhe-Yu Ou, and Leonard Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59(18):2044, 1987. ⟨31⟩

[89] Agata M Brańczyk. Hong-Ou-Mandel interference. *arXiv preprint arXiv:1711.00080*, 2017. ⟨33⟩

[90] Stefano Paesani. *Large-scale integrated quantum photonics*. PhD thesis, University of Bristol, 2019. ⟨33, 35, 88, 172⟩

[91] Jeremy C Adcock. *Generating optical graph states*. PhD thesis, University of Bristol, 2019. ⟨34⟩

[92] Joseph M Lukens and Pavel Lougovski. Frequency-encoded photonic qubits for scalable quantum information processing. *Optica*, 4(1):8–16, 2017. ⟨36⟩

[93] Demid V Sychev, Alexander E Ulanov, Egor S Tiunov, Anastasia A Pushkina, A Kuzhamuratov, Valery Novikov, and AI Lvovsky. Entanglement and teleportation between polarization and wave-like encodings of an optical qubit. *Nature Communications*, 9(1):1–7, 2018. ⟨36⟩

[94] Mercedes Gimeno-Segovia. *Towards practical linear optical quantum computing*. PhD thesis, Imperial College London, 2015. ⟨37, 42, 106⟩

[95] Qiang Zhang, Xiao-Hui Bao, Chao-Yang Lu, Xiao-Qi Zhou, Tao Yang, Terry Rudolph, and Jian-Wei Pan. Demonstration of a scheme for the generation of "event-ready" entangled photon pairs from a single-photon source. *Physical Review A*, 77(6):062316, 2008. ⟨39, 40⟩

[96] Jaewoo Joo, Peter L Knight, Jeremy L O'Brien, and Terry Rudolph. One-way quantum computation with four-dimensional photonic qudits. *Physical Review A*, 76(5):052326, 2007. ⟨40⟩

[97] Sara Bartolucci, Patrick M Birchall, Mercedes Gimeno-Segovia, Eric Johnston, Konrad Kieling, Mihir Pant, Terry Rudolph, Jake Smith, Chris Sparrow, and Mihai D Vidrighin. Creation of entangled photonic states using linear optics. *arXiv preprint arXiv:2106.13825*, 2021. ⟨40⟩

[98] Michael Varnava, Daniel E Browne, and Terry Rudolph. How good must single photon sources and detectors be for efficient linear optical quantum computation? *Physical Review Letters*, 100(6):060502, 2008. ⟨40⟩

[99] FV Gubarev, IV Dyakonov, M Yu Saygin, GI Struchalin, SS Straupe, and SP Kulik. Improved heralded schemes to generate entangled states from single photons. *Physical Review A*, 102(1):012604, 2020. ⟨40⟩

[100] Chaitali Joshi, Alessandro Farsi, Stéphane Clemmen, Sven Ramelow, and Alexander L Gaeta. Frequency multiplexing for quasi-deterministic heralded single-photon sources. *Nature Communications*, 9(1):1–8, 2018. ⟨40⟩

[101] Terry Rudolph. Why I am optimistic about the silicon-photonic route to quantum computing. *APL Photonics*, 2(3):030901, 2017. ⟨42, 82, 106⟩

[102] Daniel E Browne and Terry Rudolph. Resource-efficient linear optical quantum computation. *Physical Review Letters*, 95(1):010501, 2005. ⟨42⟩

[103] Michael A Nielsen. Optical quantum computation using cluster states. *Physical Review Letters*, 93(4):040503, 2004. ⟨42, 106⟩

[104] Nadav Yoran and Benni Reznik. Deterministic linear optics quantum computation with single photon qubits. *Physical Review Letters*, 91(3):037903, 2003. ⟨42⟩

[105] Fabian Ewert and Peter van Loock. 3/4-efficient bell measurement with passive linear optics and unentangled ancillae. *Physical Review Letters*, 113(14):140403, 2014. ⟨43⟩

[106] Ying Li, Peter C Humphreys, Gabriel J Mendoza, and Simon C Benjamin. Resource costs for fault-tolerant linear optical quantum computing. *Physical Review X*, 5(4):041007, 2015. ⟨43, 106⟩

[107] Hamza Abudayyeh, Boaz Lubotzky, Anastasia Blake, Jun Wang, Somak Majumder, Zhongjian Hu, Younghee Kim, Han Htoon, Riya Bose, Anton V Malko, et al. Single photon sources with near unity collection efficiencies by deterministic placement of quantum dots in nanoantennas. *APL Photonics*, 6(3):036109, 2021. ⟨44⟩

[108] Hui Wang, Jian Qin, Xing Ding, Ming-Cheng Chen, Si Chen, Xiang You, Yu-Ming He, Xiao Jiang, L You, Z Wang, et al. Boson sampling with 20 input photons and

a 60-mode interferometer in a $10^{14}$-dimensional Hilbert space. *Physical Review Letters*, 123(25):250503, 2019. ⟨44, 82, 83⟩

[109] Liang Zhai, Giang N. Nguyen, Clemens Spinnler, Julian Ritzmann, Matthias C. Löbl, Andreas D. Wieck, Arne Ludwig, Alisa Javadi, and Richard J. Warburton. Quantum interference of identical photons from remote quantum dots, 2021. ⟨44, 83⟩

[110] Stefano Paesani, Massimo Borghi, Stefano Signorini, Alexandre Maïnos, Lorenzo Pavesi, and Anthony Laing. Near-ideal spontaneous photon sources in silicon quantum photonics. *Nature Communications*, 11(1):1–6, 2020. ⟨44, 63, 125⟩

[111] Ben M Burridge, Imad I Faruque, John G Rarity, and Jorge Barreto. High spectro-temporal purity single-photons from silicon micro-racetrack resonators using a dual-pulse configuration. *Optics Letters*, 45(14):4048–4051, 2020. ⟨44, 79⟩

[112] Joshua W Silverstone, Damien Bonneau, Kazuya Ohira, Nob Suzuki, Haruhiko Yoshida, Norio Iizuka, Mizunori Ezaki, Chandra M Natarajan, Michael G Tanner, Robert H Hadfield, et al. On-chip quantum interference between silicon photon-pair sources. *Nature Photonics*, 8(2):104–108, 2014. ⟨44⟩

[113] Jesper Bjerge Christensen, Jacob Gade Koefoed, Karsten Rottwitt, and CJ McKinstrie. Engineering spectrally unentangled photon pairs from nonlinear microring resonators by pump manipulation. *Optics Letters*, 43(4):859–862, 2018. ⟨44, 79⟩

[114] IN Chuprina, NS Perminov, D Yu Tarankova, and AA Kalachev. Generating pure single-photon states via spontaneous four-wave mixing in a system of coupled microresonators. *Laser Physics Letters*, 15(10):105104, 2018. ⟨44, 79⟩

[115] Yingwen Liu, Chao Wu, Xiaowen Gu, Yuechan Kong, Xinxin Yu, Renyou Ge, Xinlun Cai, Xiaogang Qiang, Junjie Wu, Xuejun Yang, et al. High-spectral-purity photon generation from a dual-interferometer-coupled silicon microring. *Optics Letters*, 45(1):73–76, 2020. ⟨44, 79⟩

[116] KF Riley, MP Hobson, and SJ Bence. Mathematical methods for physics and engineering, 2006. ⟨45⟩

[117] Benjamin Kambs and Christoph Becher. Limitations on the indistinguishability of photons from remote solid state sources. *New Journal of Physics*, 20(11):115003, 2018. ⟨49⟩

[118] John E Heebner and Robert W Boyd. Enhanced all-optical switching by use of a nonlinear fiber ring resonator. *Optics Letters*, 24(12):847–849, 1999. ⟨55⟩

[119] Amnon Yariv. Critical coupling and its control in optical waveguide-ring resonator systems. *IEEE Photonics Technology Letters*, 14(4):483–485, 2002. ⟨55⟩

[120] Christopher Gerry, Peter Knight, and Peter L Knight. *Introductory quantum optics*. Cambridge University Press, 2005. ⟨57⟩

[121] Roy J Glauber. Optical coherence and photon statistics. *Quantum Optics and Electronics*, pages 63–185, 1965. ⟨57⟩

[122] Z Vernon, M Menotti, CC Tison, JA Steidle, ML Fanto, PM Thomas, SF Preble, AM Smith, PM Alsing, M Liscidini, et al. Truly unentangled photon pairs without spectral filtering. *Optics Letters*, 42(18):3638–3641, 2017. ⟨67, 79⟩

[123] Wojciech Tadej and Karol Życzkowski. A concise guide to complex Hadamard matrices. *Open Systems & Information Dynamics*, 13(2):133–177, 2006. ⟨72⟩

[124] SS Agaian. Application of Hadamard matrices. In *Hadamard matrices and their applications*, pages 134–177. Springer, 1985. ⟨72, 73⟩

[125] AT Butson. Generalized Hadamard matrices. *Proceedings of the American Mathematical Society*, 13(6):894–898, 1962. ⟨73⟩

[126] Kathy J Horadam and Parampalli Udaya. Cocyclic Hadamard codes. *IEEE Transactions on Information Theory*, 46(4):1545–1550, 2000. ⟨73⟩

[127] A Hedayat and Walter Dennis Wallis. Hadamard matrices and their applications. *The Annals of Statistics*, pages 1184–1238, 1978. ⟨73⟩

[128] Reinhard F Werner. All teleportation and dense coding schemes. *Journal of Physics A: Mathematical and General*, 34(35):7081, 2001. ⟨73⟩

[129] Thomas Durt, Berthold-Georg Englert, Ingemar Bengtsson, and Karol Życzkowski. On mutually unbiased bases. *International Journal of Quantum Information*, 8(04):535–640, 2010. ⟨73⟩

[130] Michel E Marhic. Discrete Fourier transforms by single-mode star networks. *Optics letters*, 12(1):63–65, 1987. ⟨74⟩

[131] Alex E Jones, Adrian J Menssen, Helen M Chrzanowski, Tom AW Wolterink, Valery S Shchesnovich, and Ian A Walmsley. Multiparticle interference of pairwise distinguishable photons. *Physical Review Letters*, 125(12):123603, 2020. ⟨75⟩

[132] Malte Christopher Tichy, Markus Tiersch, Fernando de Melo, Florian Mintert, and Andreas Buchleitner. Zero-transmission law for multiport beam splitters. *Physical Review Letters*, 104(22):220405, 2010. ⟨75⟩

[133] Xiaoge Zeng, Cale M Gentry, and Miloš A Popović. Four-wave mixing in silicon coupled-cavity resonators with port-selective, orthogonal supermode excitation. *Optics Letters*, 40(9):2120–2123, 2015. ⟨79⟩

[134] Boris Korzh, Qing-Yuan Zhao, Jason P Allmaras, Simone Frasca, Travis M Autry, Eric A Bersin, Andrew D Beyer, Ryan M Briggs, Bruce Bumble, Marco Colangelo, et al. Demonstration of sub-3 ps temporal resolution with a superconducting nanowire single-photon detector. *Nature Photonics*, 14(4):250–255, 2020. ⟨79, 80, 81⟩

[135] Yuguang Zhang, Xiao Hu, Daigao Chen, Lei Wang, Miaofeng Li, Peng Feng, Xi Xiao, and Shaohua Yu. Design and demonstration of ultra-high-Q silicon microring resonator based on a multi-mode ridge waveguide. *Optics Letters*, 43(7):1586–1589, 2018. ⟨80⟩

[136] Aleksandr Biberman, Michael J Shaw, Erman Timurdogan, Jeremy B Wright, and Michael R Watts. Ultralow-loss silicon ring resonators. *Optics letters*, 37(20):4236–4238, 2012. ⟨80⟩

[137] Matthew W Puckett, Jiawei Wang, Debapam Bose, Grant M Brodnik, Jianfeng Wu, Karl Nelson, and Daniel J Blumenthal. Silicon nitride ring resonators with 0.123 db/m loss and Q-factors of 216 million for nonlinear optical applications. In *The*

*European Conference on Lasers and Electro-Optics*, page ce_11_3. Optical Society of America, 2019. ⟨80⟩

[138] Patrick Yard, Alexander Jones, Stefano Paesani, Alexandre Maïnos, Jacob Bulmer, and Anthony Laing. Time-resolved multi-photon interference in a silicon chip. *In preparation*, 2022. ⟨81, 105⟩

[139] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018. ⟨82⟩

[140] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum (NISQ) algorithms. *arXiv preprint arXiv:2101.08448*, 2021. ⟨82⟩

[141] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):1–7, 2014. ⟨82⟩

[142] A Yu Kitaev. Quantum measurements and the Abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995. ⟨82⟩

[143] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016. ⟨82⟩

[144] Aaronson Scott. Certified randomness from quantum supremacy. Beyond Crypto: A TCS Perspective, 2018. ⟨82⟩

[145] Kamil Brádler, Pierre-Luc Dallaire-Demers, Patrick Rebentrost, Daiqin Su, and Christian Weedbrook. Gaussian boson sampling for perfect matchings of arbitrary graphs. *Physical Review A*, 98(3):032310, 2018. ⟨82⟩

[146] Joonsuk Huh, Gian Giacomo Guerreschi, Borja Peropadre, Jarrod R McClean, and Alán Aspuru-Guzik. Boson sampling for molecular vibronic spectra. *Nature Photonics*, 9(9):615–620, 2015. ⟨82⟩

[147] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, 2019. ⟨82⟩

[148] Aram W Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549(7671):203–209, 2017. ⟨82⟩

[149] Lu Chao-Yang. Gaussian boson sampling for quantum computational advantage. Quantum research seminars toronto 11, 2021. ⟨82⟩

[150] Craig S Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. Gaussian boson sampling. *Physical Review Letters*, 119(17):170501, 2017. ⟨83, 88⟩

[151] Regina Kruse, Craig S Hamilton, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. Detailed study of Gaussian boson sampling. *Physical Review A*, 100(3):032326, 2019. ⟨83, 88⟩

[152] Helmut Alt. *Computational discrete mathematics: advanced lectures*, volume 2122. Springer, 2003. ⟨85⟩

[153] Leslie G Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189–201, 1979. ⟨85⟩

[154] Austin P Lund, Anthony Laing, Saleh Rahimi-Keshari, Terry Rudolph, Jeremy L O'Brien, and Timothy C Ralph. Boson sampling from a Gaussian state. *Physical Review Letters*, 113(10):100502, 2014. ⟨86⟩

[155] Alex Neville. *Classical sampling and optimisation algorithms for quantum photonics*. PhD thesis, University of Bristol, 2017. ⟨87⟩

[156] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, 2018. ⟨89⟩

[157] Stefan Scheel. Permanents in linear optical networks. *arXiv preprint quant-ph/0406127*, 2004. ⟨95⟩

[158] Vincenzo Tamma and Simon Laibacher. Multiboson correlation interferometry with arbitrary single-photon pure states. *Physical Review Letters*, 114(24):243601, 2015. ⟨95, 99, 104⟩

[159] Vincenzo Tamma and Simon Laibacher. Boson sampling with non-identical single photons. *Journal of Modern Optics*, 63(1):41–45, 2016. ⟨99⟩

[160] Vincenzo Tamma and Simon Laibacher. Multi-boson correlation sampling. *Quantum Information Processing*, 15(3):1241–1262, 2016. ⟨99⟩

[161] Simon Laibacher and Vincenzo Tamma. From the physics to the computational complexity of multiboson correlation interference. *Physical Review Letters*, 115(24):243605, 2015. ⟨99⟩

[162] Simon Laibacher and Vincenzo Tamma. Toward quantum computational supremacy of boson sampling with random overlap in the photonic spectra. *arXiv preprint arXiv:1801.03832*, 2018. ⟨99⟩

[163] Oliver F Thomas, Will McCutcheon, and Dara PS McCutcheon. A general framework for multimode Gaussian quantum optics and photo-detection: Application to Hong–Ou–Mandel interference with filtered heralded single photon sources. *APL Photonics*, 6(4):040801, 2021. ⟨104⟩

[164] Jelmer J Renema, Adrian Menssen, William R Clements, Gil Triginer, William S Kolthammer, and Ian A Walmsley. Efficient algorithm for boson sampling with partially distinguishable photons. *arXiv preprint arXiv:1707.02793*, 2017. ⟨105, 168⟩

[165] Robert Prevedel, Philip Walther, Felix Tiefenbacher, Pascal Böhi, Rainer Kaltenbaek, Thomas Jennewein, and Anton Zeilinger. High-speed linear optics quantum computing using active feed-forward. *Nature*, 445(7123):65–69, 2007. ⟨106⟩

[166] Fumihiro Kaneda and Paul G Kwiat. High-efficiency single-photon generation via large-scale active time multiplexing. *Science advances*, 5(10):eaaw8586, 2019. ⟨106⟩

[167] Robert JA Francis-Jones, Rowan A Hoggarth, and Peter J Mosley. All-fiber multiplexed source of high-purity single photons. *Optica*, 3(11):1270–1273, 2016. ⟨106⟩

[168] Henry Semenenko, Philip Sibson, Mark G Thompson, and Chris Erven. Interference between independent photonic integrated devices for quantum key distribution. *Optics Letters*, 44(2):275–278, 2019. ⟨106⟩

[169] Henry Semenenko, Philip Sibson, Andy Hart, Mark G Thompson, John G Rarity, and Chris Erven. Chip-based measurement-device-independent quantum key distribution. *Optica*, 7(3):238–242, 2020. ⟨106⟩

[170] J Sabines-Chesterking, R Whittaker, SK Joshi, PM Birchall, Paul-Antoine Moreau, A McMillan, HV Cable, JL O'Brien, JG Rarity, and JCF Matthews. Sub-shot-noise transmission measurement enabled by active feed-forward of heralded single photons. *Physical Review Applied*, 8(1):014016, 2017. ⟨106⟩

[171] Hamed Sattari, Teodoro Graziosi, Marcell Kiss, Tae Joon Seok, Sangyoon Han, Ming C Wu, and Niels Quack. Silicon photonic MEMS phase-shifter. *Optics Express*, 27(13):18959–18969, 2019. ⟨107⟩

[172] David J Thomson, Frederic Y Gardes, Jean-Marc Fedeli, Sanja Zlatanovic, Youfang Hu, Bill Ping Piu Kuo, Evgeny Myslivets, Nikola Alic, Stojan Radic, Goran Z Mashanovich, et al. 50-Gb/s silicon optical modulator. *IEEE Photonics Technology Letters*, 24(4):234–236, 2011. ⟨107⟩

[173] Graham T Reed, G Mashanovich, F Yand Gardes, and DJ Thomson. Silicon optical modulators. *Nature Photonics*, 4(8):518–526, 2010. ⟨107⟩

[174] Hani Nejadriahi, Steve Pappert, Yeshaiahu Fainman, and Paul Yu. Efficient and compact thermo-optic phase shifter in silicon-rich silicon nitride. *arXiv preprint arXiv:2106.09476*, 2021. ⟨107⟩

[175] Nicholas C Harris, Yangjin Ma, Jacob Mower, Tom Baehr-Jones, Dirk Englund, Michael Hochberg, and Christophe Galland. Efficient, compact and low loss thermo-optic phase shifter in silicon. *Optics Express*, 22(9):10487–10493, 2014. ⟨107⟩

[176] Johannes Henriksson, Tae Joon Seok, Jianheng Luo, Kyungmok Kwon, Niels Quack, and Ming C Wu. Digital silicon photonic MEMS phase-shifter. In *2018 International Conference on Optical MEMS and Nanophotonics (OMN)*, pages 1–2. IEEE, 2018. ⟨107⟩

[177] RH Stolen and A Ashkin. Optical Kerr effect in glass waveguide. *Applied Physics Letters*, 22(6):294–296, 1973. ⟨107⟩

[178] Maksim Zalkovskij, AC Strikwerda, Krzysztof Iwaszczuk, A Popescu, D Savastru, Radu Malureanu, AV Lavrinenko, and Peter Uhd Jepsen. Terahertz-induced Kerr effect in amorphous chalcogenide glasses. *Applied Physics Letters*, 103(22):221102, 2013. ⟨107⟩

[179] Margaux Barbier, Isabelle Zaquine, and Philippe Delaye. Spontaneous four-wave mixing in liquid-core fibers: towards fibered Raman-free correlated photon sources. *New Journal of Physics*, 17(5):053031, 2015. ⟨107⟩

[180] Qin Zhong and John T Fourkas. Optical Kerr effect spectroscopy of simple liquids. *The Journal of Physical Chemistry B*, 112(49):15529–15539, 2008. ⟨107⟩

[181] PP Ho and RR Alfano. Optical Kerr effect in liquids. *Physical Review A*, 20(5):2170, 1979. ⟨107⟩

[182] Claude Watson Bruce. Electro-optical Kerr effect in gases. *Physical Review*, 44(8):682, 1933. ⟨107⟩

[183] AD Buckingham and DA Dunmur. Kerr effect in inert gases and sulphur hexafluoride. *Transactions of the Faraday Society*, 64:1776–1783, 1968. ⟨107⟩

[184] Federico Belli, Amir Abdolvand, John C Travers, and Philip St J Russell. Highly efficient deep UV generation by four-wave mixing in gas-filled hollow-core photonic crystal fiber. *Optics Letters*, 44(22):5509–5512, 2019. ⟨107⟩

[185] M Ghotbi, M Beutler, and F Noack. Generation of 2.5$\mu$J vacuum ultraviolet pulses with sub-50 fs duration by noncollinear four-wave mixing in argon. *Optics Letters*, 35(20):3492–3494, 2010. ⟨107⟩

[186] Matthew A Hall, Joseph B Altepeter, and Prem Kumar. All-optical switching of photonic entanglement. *New Journal of Physics*, 13(10):105004, 2011. ⟨107⟩

[187] Vilson R Almeida, Carlos A Barrios, Roberto R Panepucci, Michal Lipson, Mark A Foster, Dimitre G Ouzounov, and Alexander L Gaeta. All-optical switching on a silicon chip. *Optics Letters*, 29(24):2867–2869, 2004. ⟨107⟩

[188] Shatrughna Kumar and Mrinal Sen. Integrable all-optical switch for photonic integrated circuits. *JOSA B*, 38(2):611–620, 2021. ⟨107⟩

[189] Niklas M Lüpken, Tim Hellwig, Martin Schnack, Jörn P Epping, Klaus-J Boller, and Carsten Fallnich. Low-power broadband all-optical switching via intermodal cross-phase modulation in integrated optical waveguides. *Optics Letters*, 43(8):1631–1634, 2018. ⟨107⟩

[190] Paul N Butcher and David Cotter. *The elements of nonlinear optics*. Cambridge university press, 1990. ⟨109⟩

[191] RM Osgood, NC Panoiu, JI Dadap, Xiaoping Liu, Xiaogang Chen, I-Wei Hsieh, E Dulkeith, WMJ Green, and Yurii A Vlasov. Engineering nonlinearities in nanoscale optical systems: physics and applications in dispersion-engineered silicon nanophotonic wires. *Advances in Optics and Photonics*, 1(1):162–235, 2009. ⟨112⟩

[192] Joshua W Silverstone, J Wang, D Bonneau, P Sibson, R Santagati, C Erven, JL O'Brien, and MG Thompson. Silicon quantum photonics. In *2016 International Conference on Optical MEMS and Nanophotonics (OMN)*, pages 1–2. IEEE, 2016. ⟨112⟩

[193] Bin Fang, Offir Cohen, Jamy B Moreno, and Virginia O Lorenz. State engineering of photon pairs produced through dual-pump spontaneous four-wave mixing. *Optics Express*, 21(3):2707–2717, 2013. ⟨118, 119⟩

[194] Stefano Signorini, Mattia Mancinelli, Massimo Borghi, Martino Bernard, Mher Ghulinyan, Georg Pucker, and Lorenzo Pavesi. Intermodal four-wave mixing in silicon waveguides. *Photonics Research*, 6(8):805–814, 2018. ⟨121⟩

[195] Weimin Ye, Xiaodong Yuan, Yang Gao, and Jianlong Liu. Design of broadband silicon-waveguide mode-order converter and polarization rotator with small footprints. *Optics Express*, 25(26):33176–33183, 2017. ⟨121⟩

[196] Tim Hellwig, Till Walbaum, Martin Schnack, and Carsten Fallnich. Optically induced transverse mode conversion of ultrashort-pulses in optical waveguides. In *Nonlinear Optics*, pages NTh2A–5. Optical Society of America, 2013. ⟨121⟩

[197] Zhuo Cheng, Jun Wang, Zeyuan Yang, Lina Zhu, Yuanqing Yang, Yongqing Huang, and Xiaomin Ren. Sub-wavelength grating assisted mode order converter on the SOI substrate. *Optics express*, 27(23):34434–34441, 2019. ⟨121⟩

[198] Carl M Bender and Stefan Boettcher. Real spectra in non-Hermitian Hamiltonians having $\mathcal{PT}$ symmetry. *Physical Review Letters*, 80(24):5243, 1998. ⟨136⟩

[199] Carl M Bender. Making sense of non-Hermitian Hamiltonians. *Reports on Progress in Physics*, 70(6):947, 2007. ⟨137, 139⟩

[200] Stefan Scheel and Alexander Szameit. $\mathcal{PT}$-symmetric photonic quantum systems with gain and loss do not exist. *EPL (Europhysics Letters)*, 122(3):34001, 2018. ⟨137⟩

[201] Yi-Chan Lee, Min-Hsiu Hsieh, Steven T Flammia, and Ray-Kuang Lee. Local $\mathcal{PT}$-symmetry violates the no-signaling principle. *Physical Review Letters*, 112(13):130404, 2014. ⟨137⟩

[202] Shruti Dogra, Artem A Melnikov, and Gheorghe Sorin Paraoanu. Quantum simulation of parity–time symmetry breaking with a superconducting quantum processor. *Communications Physics*, 4(1):1–8, 2021. ⟨137⟩

[203] Joseph Schindler, Ang Li, Mei C Zheng, Fred M Ellis, and Tsampikos Kottos. Experimental study of active LRC circuits with $\mathcal{PT}$-symmetries. *Physical Review A*, 84(4):040101, 2011. ⟨137⟩

[204] Haitan Xu, David Mason, Luyao Jiang, and JGE Harris. Topological energy transfer in an optomechanical system with exceptional points. *Nature*, 537(7618):80–83, 2016. ⟨137⟩

[205] Peng Peng, Wanxia Cao, Ce Shen, Weizhi Qu, Jianming Wen, Liang Jiang, and Yanhong Xiao. Anti-parity–time symmetry with flying atoms. *Nature Physics*, 12(12):1139–1145, 2016. ⟨137⟩

[206] Pei Miao, Zhifeng Zhang, Jingbo Sun, Wiktor Walasik, Stefano Longhi, Natalia M Litchinitser, and Liang Feng. Orbital angular momentum microlaser. *Science*, 353(6298):464–467, 2016. ⟨137⟩

[207] Liang Feng, Zi Jing Wong, Ren-Min Ma, Yuan Wang, and Xiang Zhang. Single-mode laser by parity-time symmetry breaking. *Science*, 346(6212):972–975, 2014. ⟨137⟩

[208] Bo Peng, Şahin Kaya Özdemir, Fuchuan Lei, Faraz Monifi, Mariagiovanna Gianfreda, Gui Lu Long, Shanhui Fan, Franco Nori, Carl M Bender, and Lan Yang. Parity-time-symmetric whispering-gallery microcavities. *Nature Physics*, 10(5):394–398, 2014. ⟨137⟩

[209] M Brandstetter, M Liertzer, C Deutsch, P Klang, J Schöberl, Hakan E Türeci, G Strasser, K Unterrainer, and S Rotter. Reversing the pump dependence of a laser at an exceptional point. *Nature Communications*, 5(1):1–7, 2014. ⟨137⟩

[210] Jing Zhang, Bo Peng, Şahin Kaya Özdemir, Kevin Pichler, Dmitry O Krimer, Guangming Zhao, Franco Nori, Yu-xi Liu, Stefan Rotter, and Lan Yang. A phonon laser operating at an exceptional point. *Nature Photonics*, 12(8):479–484, 2018. ⟨137⟩

[211] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002. ⟨137⟩

[212] Victor V Albert. Lindbladians with multiple steady states: theory and applications. *arXiv preprint arXiv:1802.00010*, 2018. ⟨138⟩

[213] Alex Krasnok, Nikita Nefedkin, and Andrea Alu. Parity-time symmetry and exceptional points: a tutorial. *arXiv preprint arXiv:2103.08135*, 2021. ⟨139⟩

[214] Nicola Maraviglia. *Simulating open quantum systems in integrated photonics*. PhD thesis, University of Bristol, 2020. ⟨139, 144, 145, 147, 149, 166, 169⟩

[215] Jacques Carolan. *Universal linear optics: characterisation, verification and computation*. PhD thesis, University of Bristol, 2015. ⟨142, 144, 145, 147⟩

[216] Chris Sparrow, Enrique Martín-López, Nicola Maraviglia, Alex Neville, Christopher Harrold, Jacques Carolan, Yogesh N Joglekar, Toshikazu Hashimoto, Nobuyuki Matsuda, Jeremy L O'Brien, et al. Simulating the vibrational quantum dynamics of molecules using photonics. *Nature*, 557(7707):660–667, 2018. ⟨147, 166⟩

[217] Paul R Halmos. Unitary dilations. In *A Hilbert space problem book*, pages 120–127. Springer, 1982. ⟨147⟩

[218] Nicola Maraviglia, Patrick Yard, Ross Wakefield, Jacques Carolan, Chris Sparrow, Levon Chakhmakhchyan, Chris Harrold, Toshikazu Hashimoto, Nobuyuki Matsuda, Andrew K. Harter, Yogesh N. Joglekar, and Anthony Laing. Photonic quantum simulations of coupled $\mathcal{PT}$-symmetric hamiltonians. *Physical Review Research*, 4:013051, 2022. ⟨148, 166⟩

[219] Sebastian Diehl, A Micheli, Adrian Kantian, B Kraus, HP Büchler, and Peter Zoller. Quantum states and phases in driven open quantum systems with cold atoms. *Nature Physics*, 4(11):878–883, 2008. ⟨152⟩

[220] Hanna Krauter, Christine A Muschik, Kasper Jensen, Wojciech Wasilewski, Jonas M Petersen, J Ignacio Cirac, and Eugene S Polzik. Entanglement generated by dissipation and steady state entanglement of two macroscopic objects. *Physical Review Letters*, 107(8):080503, 2011. ⟨152⟩

[221] Jaeyoon Cho, Sougato Bose, and MS Kim. Optical pumping into many-body entanglement. *Physical Review Letters*, 106(2):020504, 2011. ⟨152⟩

[222] Frank Verstraete, Michael M Wolf, and J Ignacio Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Physics*, 5(9):633–636, 2009. ⟨152⟩

[223] Joseph B Altepeter, Daniel FV James, and Paul G Kwiat. 4 qubit quantum state tomography. In *Quantum state estimation*, pages 113–145. Springer, 2004. ⟨154⟩

[224] RT Thew, Kae Nemoto, Andrew G White, and William J Munro. Qudit quantum-state tomography. *Physical Review A*, 66(1):012303, 2002. ⟨154⟩

[225] Nora Tischler, Carsten Rockstuhl, and Karolina Słowik. Quantum optical realization of arbitrary linear transformations allowing for loss and gain. *Physical Review X*, 8(2):021017, 2018. ⟨163⟩

[226] Ross Wakefield, Yogesh N. Joglekar, and Anthony Laing. $\mathcal{PT}$-symmetry in nonlinear quantum optics. *In preparation*, 2022. ⟨163⟩

[227] Patrick Yard, Nicola Maraviglia, Ross Wakefield, Jacques Carolan, Chris Sparrow, Chakhmakhchyan Levon, Chris Harrold, Toshikazu Hashimoto, Nobuyuki Matsuda, Andrew K. Harter, Yogesh N. Joglekar, and Anthony Laing. Coherence optimisation with $\mathcal{PT}$-symmetric Hamiltonians. *In preparation*, 2022. ⟨166⟩