



QUANTUM INTERFERENCE IN UNIVERSAL
LINEAR OPTICAL DEVICES FOR QUANTUM
COMPUTATION AND SIMULATION

CHRISTOPHER SPARROW

Department of Physics
Imperial College London

A thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy

September 2017

Abstract

It is believed that the exotic properties of quantum systems can be harnessed to perform certain computational tasks more efficiently than classical theories allow. The production, manipulation and detection of single photons constitutes a potential platform for performing such non-classical information processing. The development of integrated quantum photonics has provided a miniaturised, monolithic architecture that is promising for the realisation of near-term analog quantum devices as well as full-scale universal quantum computers.

In this thesis we investigate the viability of these photonic quantum computational approaches from an experimental and theoretical perspective. We implement the first universally reconfigurable linear optical network; a key capability for the rapid prototyping of photonic quantum protocols. We propose and demonstrate the use of these devices as a new platform for the programmable quantum simulation of molecular vibrational dynamics.

We then tackle an important outstanding problem in linear optical quantum computing; quantifying how partial-distinguishability amongst photons affects logical error rates. Finally, we propose a series of schemes aimed at counteracting these distinguishability errors in order to achieve practical quantum technologies with imperfect photonic components.

Acknowledgements

I would first like to thank my supervisor Anthony Laing for advice, guidance and scientific insight throughout my PhD. I would also like to thank Jeremy O'Brien and Terry Rudolph for the inspiration to explore the many beautiful and frustrating quirks of photonic quantum computing.

I am hugely grateful to all those with whom the results in this thesis were obtained. Whether in the lab or at the whiteboard; Enrique Martín-López, Nick Russell, Jacques Carolan, Chris Harrold, Patrick Birchall, Hugo Cable, Alex Neville and Nicola Maraviglia, I learned a great deal from all of you.

There are too many people to acknowledge at the Centre for Quantum Photonics (and latterly QETlabs), it provided a wonderful environment for working, both intellectually and socially. Similarly, my journey in the world of quantum information science so far would not have been the same without the learning and laughter shared with Cohort 4 of the Controlled Quantum Dynamics CDT at Imperial College London.

Of course, the real thesis was the friends I made along the way... I will always look back with great happiness at my years in Bristol and this is in large part due to the many close friends and acquaintances with whom many chats, dinners and beers were shared; Alex, Allison, Beccie, Callum, Chris, Hugo, Javier, Lorraine, Nicki, Patrick, Phil, Raf and Will to name but a few.

I would like to thank my whole family for a lifetime of support and providing me with the opportunities and encouragement that have led me here.

Finally, this thesis is dedicated to Kirsten, for her love and for her infinite patience putting up with me while it was undertaken and written up.

Declaration of Originality

I declare that the work in this thesis was carried out in accordance with the requirements of the University's Academic Regulations for Research Degree Programmes. The thesis was written entirely by myself and accurately reflects work carried out during my graduate studies. Work done in collaboration with, or with the assistance of, others, is indicated as such.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

Publications

1. J. Carolan, C. Harrold, **C. Sparrow**, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O’Brien, and A. Laing, ‘Universal Linear Optics’, *Science*, **349**, 711-716, (2015).
2. A. Holleczek, O. Barter, A. Rubenok, J. Dilley, P. B.R. Nisbet-Jones, G. Langfahl-Klabes, G. D. Marshall, **C. Sparrow**, J. L. O’Brien, K. Poulios, A. Kuhn, and J. C.F. Matthews, ‘Quantum Logic with Cavity Photons From Single Atoms’, *Phys. Rev. Lett.*, **117**, 023602, (2016).
3. A. Neville, **C. Sparrow**, R. Clifford, E. Johnston, P. Birchall, A. Montanaro, A. Laing, ‘Classical Boson Sampling Algorithms with Superior Performance to Near-term Experiments’, *Nat. Phys.*, **13**, 1153-1157 (2017).
4. **C. Sparrow**, E. Martín-López, N. Maraviglia, A. Neville, J. Carolan, C. Harrold, N. Matsuda, J. L. O’Brien, Y. Joglekar, D. Tew, A. Laing, ‘Simulating the Vibrational Quantum Dynamics of Molecules with Integrated Photonics’, *Submitted*, (2017).
5. **C. Sparrow***, P. Birchall* *et al.*, ‘Linear Optical Quantum Computing with Partially-Distinguishable Photons’, *in prep*, (2017).

* authors contributed equally to this work

Contents

Abstract	2
Acknowledgements	4
Declaration of Originality	5
Publications	6
1 Introduction	16
1.1 Thesis Outline	18
2 Quantum Information and Computation	20
2.1 Introduction	20
2.2 Quantum Theory	20
2.2.1 Physical states	21
2.2.2 Observables	21
2.2.3 Evolution	22
2.2.4 Measurements	22
2.2.5 Interpretations	23
2.2.6 Operational quantum theory	24
2.3 Quantum Information	26
2.3.1 The Qubit	27
2.3.2 Entanglement	30
2.3.3 Nonlocality	32

2.3.4	Quantum tomography	34
2.3.5	Distance measures	37
2.4	Quantum computing	40
2.4.1	Turing machines	40
2.4.2	Computational Complexity Theory	41
2.4.3	Circuit model quantum computation	43
2.4.4	Quantum error correction	48
2.4.5	Fault tolerance	50
2.4.6	Stabilizer formalism	51
2.4.7	Measurement-based quantum computation	53
2.4.8	Other computational models	55
2.4.9	Physical platforms	57
3	Quantum light	62
3.1	Introduction	62
3.2	Quantisation of the electromagnetic field	62
3.3	Linear optics	67
3.4	Quantum Interference	69
3.4.1	The Hong-Ou-Mandel effect	69
3.4.2	The HOM dip and partial-distinguishability	71
3.5	Coherent states	73
3.6	Linear optical quantum computing	75
3.6.1	KLM	77
3.6.2	MBLOQC	80
3.6.3	Boson sampling	83
3.7	Quantum photonics	89
3.7.1	Photon sources	89
3.7.2	Linear optical networks	93
3.7.3	Detectors	96
3.7.4	Experimental setup	97

4	Universal Linear Optics	100
4.1	Introduction	100
4.2	The Reck <i>et al.</i> Scheme	101
4.3	A Universal Linear Optical Processor	103
4.3.1	Characterisation	104
4.4	Linear Optical Gates	106
4.4.1	Process tomography experiments	106
4.4.2	Heralded integrated gates	111
4.4.3	Characterisation of Linear Optical Networks	117
4.5	Discussion	120
5	Simulating Vibrational Quantum Dynamics of Molecules with Integrated Photonics	122
5.1	Introduction	122
5.2	Molecular Vibrational Dynamics in the Harmonic Approximation	124
5.3	Experimental Procedure	126
5.4	Simulating Four Atom Molecules	127
5.5	Energy transfer and dephasing in NMA	131
5.6	Vibrational relaxation in H ₂ O	134
5.7	Anharmonic Hamiltonian for H ₂ O	137
5.8	Adaptive feedback control in the dissociation of NH ₃	140
5.9	Discussion	144
6	Linear Optical Quantum Computing with Partially-Distinguishable Photons	146
6.1	Introduction	146
6.2	Errors in LOQC	147
6.3	Photon Distinguishability	149
6.3.1	Dual-path states	150
6.3.2	Distinguishability models	151

6.4	Generating entanglement with partially-distinguishable photons	152
6.5	Entangled measurements with partially-distinguishable photons	159
6.6	Measurement-based resource generation	160
6.6.1	Overview and initial state	160
6.6.2	GHZ generation	162
6.6.3	Fusion gates	164
6.7	Error rates on Lattice	167
6.8	Discussion	171
7	Mitigating Distinguishability Errors	174
7.1	Introduction	174
7.2	Hong-Ou-Mandel Filtering	175
7.3	The Distinguishability Tolerance of Bell State Generators	179
7.4	Generating entanglement from distinguishable photons	181
7.5	Discussion	184
8	Conclusion	186
8.1	Key Results	186
8.2	Outlook	188
	Bibliography	190
A	Universal Linear Optics	210
A.1	Transfer Matrix Reconstructions	210
B	Simulating the Vibrational Quantum Dynamics of Molecules with Inte- grated Photonics	212
B.1	Molecule details	212
B.1.1	HFHF	212
B.1.2	HNCO	213

CONTENTS

B.1.3	N_4	213
B.1.4	P_4	214
B.1.5	SO_3	214
B.1.6	NMA	215
B.1.7	H_2O	216
B.1.8	NH_3	216
B.2	Additional data	217
C	Linear Optical Quantum Computing with Partially-Distinguishable Photons	223
C.1	Trace inequality proof	223
C.2	Bound on real term	224
C.3	Additional transfer matrices	226
C.4	Equivalence of pre- and post-fusion corrections	227
C.5	Fusion Operations	228
C.6	Additional measurement operators	231
C.7	Second order fusion error maps	235
D	Mitigating Distinguishability Errors	238

List of Figures

2.1	Bloch sphere representation of pure qubit state	28
2.2	Schematic of CHSH Bell test	33
2.3	Hierarchy of computational complexity classes	43
2.4	Circuit model quantum computation	45
2.5	Measurement based quantum computation	54
3.1	Hong-Ou-Mandel dip	72
3.2	Single and dual-rail photonic qubits	76
3.3	Elementary circuits in the KLM scheme	79
3.4	Type I and II Fusion gates	81
3.5	Schematic of ballistic photonic quantum computation	82
3.6	Boson Sampling schematic	84
3.7	Experimental setup	97
4.1	The Reck <i>et al.</i> scheme	102
4.2	Characterisation of phase shifters	106
4.3	Process tomography for single qubit gates	108
4.4	Process tomography of postselected CNOT gate	111
4.5	KLM CNOT gate	113
4.6	Experimental Bell state generation	115
5.1	Simulating the vibrational dynamics of molecules in the harmonic approx- imation	129
5.2	Evolution of localised vibrational excitations in four atom molecules	130

LIST OF FIGURES

5.3	Quantum energy transfer and dephasing in NMA	132
5.4	Vibrational relaxation in H ₂ O	135
5.5	Anharmonic model for H ₂ O	138
5.6	Optimised success probability for nonlinear phase shift (NPS) gates	139
5.7	Adaptive feedback control algorithm for selective dissociation in NH ₃ . . .	142
6.1	Linear optical networks for Bell state preparation and measurement	154
6.2	Purity and Fidelity of logical Bell states produced with partially-distinguishable photons	157
6.3	Schematic of an all-optical implementation of ballistic LOQC	162
6.4	Summary of LOQC with a distinguishable particle	168
6.5	Output state ρ_G produced in lattice by one distinguishable particle.	169
6.6	Logical error rates on lattice	171
7.1	Hong-Ou-Mandel Filtering	175
7.2	Improvement in dominant eigenvalue for HOM Filter	178
7.3	Linear optical networks for heralded bell generation	179
7.4	Error tolerance of Bell generators	180
7.5	Entanglement generation from spectrally distinguishable photons via temporal resolution	184
B.1	Four atom molecule data	218
B.2	H ₂ CS data	219
B.3	NMA data	220
B.4	H ₂ O data	220
B.5	Cost function and probability of bunching in AFC experiment	221
B.6	Full distributions for initial and final AFC states	222
C.1	Fusion operations including local unitaries	229
C.2	Error maps for second order errors in fusion measurements	237

Chapter 1

Introduction

Classical theories of physics provide sensible and convincing explanations for the world that humans can observe and interact with. Unfortunately, these explanations fall apart at the scale of the smallest physical systems. The discovery and development of quantum theory overturned many of the foundational assumptions underpinning our previous understanding of nature, resulting in decades of debate amongst scientists and philosophers alike. However, despite widely spread opinions on interpretational aspects of the theory, quantum mechanics as a tool for predicting the outcomes of experimental observations remains unparalleled in scientific history.

An investigation of the theory purely in terms of the probabilities of observing events, an information-theoretic perspective, has revealed that the unique phenomena that quantum systems exhibit can allow for fundamentally new ways of processing information. Perhaps the most striking encapsulation of this concept is the quantum computer, a computer built from quantum mechanical components which can perform certain computational tasks exponentially faster than devices operating under the laws of classical physics. The global race to realise the first universal quantum computer represents both a grand engineering challenge with huge potential to solve real world problems, but also a scientific landmark akin to the creation of a new state of matter.

Light has played a key role in the theoretical and experimental development of quantum information science, from the photoelectric effect [1] through to the first loophole-free

Bell tests [2–4]. The ease of manipulation and low noise properties of quantum states of light have meant that optical systems have often been ideal systems to investigate fundamental quantum physics. Recently however, with the development of integrated quantum photonics [5–7], devices which produce, manipulate and detect photons on chips have become a promising platform for future quantum technologies.

The quest to deliver quantum technologies can be tackled via a top-down or bottom-up approach. On the one hand we can produce blueprints for high-level architectures realising universal, fault-tolerant optical quantum computation and investigate the experimental resources and requirements that would be needed to achieve such schemes. From the other side, we can consider the toolbox of the current state-of-art and near-future experimental capabilities and investigate specific information processing tasks which can be targeted for near-term impacts.

In this thesis we will explore both of these approaches towards realising practical quantum computational devices. Experimentally, we demonstrate the first universal linear optical device; an important milestone on the road towards non-universal models of photonic quantum computation. We first illustrate the versatility of this platform by using the single device as an experimental testbed for implementing and characterising a series of quantum logic gates, many for the first time in integrated optics. Although it is known that schematically simple linear optical experiments can solve classically hard computational problems [8], real world applications for such devices are currently limited. Here we show that universal linear optical devices can operate as analog quantum simulators for the dynamics of quantised molecular vibrations. We present experimental simulations with tens of thousands of device configurations, exploring the dynamics of a range of molecules and molecular behaviours including their dephasing, thermalisation, energy transfer and dissociation.

Moving to large-scale universal architectures, a key missing ingredient in our understanding of the practicalities of linear optical quantum computing is the existence of good error models. The link between physical imperfections in realistic devices and errors in the resulting quantum computation is crucial to assessing the fabrication tolerances and error correction overheads necessary to achieve fault-tolerance. A major source of logical

error in photonics is the production of photons which are not perfectly indistinguishable in all their degrees of freedom. Here we develop a model for this error and explore its resulting effects on leading schemes for universal linear optical quantum computing. The insights gained from this study allow us to then propose a series of new techniques designed to suppress or bypass these errors without requiring detailed knowledge of the photonic states.

1.1 Thesis Outline

This thesis is structured as follows:

Chapters 2 and 3 provide the background material. Chapter 2 includes a brief overview and summary of quantum theory, quantum information and computing. Chapter 3 then introduces the quantum description of light, photonic quantum computation and finally the experimental photonic platform we use in the remainder of the thesis.

Chapter 4 describes the demonstration of a universal linear optical processor acting upon six optical modes. This device is used to implement a series of optical quantum logic gates including the first demonstrations of a heralded CNOT and a newly-derived Bell state generator in integrated optics as well as the highest process fidelity recorded for a postselected CNOT gate so far. The reconstruction of the transfer matrices implementing these gates demonstrates that the logical error observed in multi-photon experiments is largely due to imperfect sources and detection.

Chapter 5 proposes and demonstrates universal linear optical devices as a new quantum simulator for molecular vibrational dynamics. The simulation of localised vibrational excitations in a range of molecular systems is presented including investigations of energy transfer, open system dynamics and anharmonic corrections via measurement-induced nonlinearities. We go on to operate the experiment in a closed feedback loop to simulate a quantum control experiment aimed at maximising the photo-dissociation of ammonia.

Chapter 6 investigates the effects of using partially-distinguishable photons in linear optical quantum computers. After presenting a new formalism for calculating the logical information of photonic states containing partially-distinguishable photons, this is then

used to investigate standard linear optical gates and eventually the resulting logical error rates in an architecture for universal computation are computed for the first time.

Chapter 7 seeks to alleviate the difficulty of producing highly indistinguishable photons for quantum technologies by proposing a series of schemes which can result in improved error rates without any direct manipulation of the additional degrees of freedom of the photons. These include a filter based upon Hong-Ou-Mandel interference, new error-tolerant gates and the use of high resolution detectors to erase distinguishing information.

Finally, Chapter 8 summarises the work and provides concluding remarks.

Chapter 2

Quantum Information and Computation

2.1 Introduction

In this chapter we introduce the physical and mathematical framework which will be used in the remainder of the thesis and provide the context in which the work is placed. We begin by introducing the standard postulates of quantum theory. We then introduce quantum information theory and quantum computation. Various parts of this section draw upon notes and presentations from Michael Nielsen & Isaac Chuang, John Preskill, Robert Spekkens, Robin Blume-Kohout, David Jennings and Dan Browne.

2.2 Quantum Theory

By the early 20th century it was becoming increasingly clear that existing theories of physics could not explain many observed phenomena such as black body radiation and the photoelectric effect. Throughout the next century a new theory of physics was developed by Bohr, Einstein, Heisenberg, Schrödinger and many others to explain these experiments built upon the postulate that physical systems and properties can be *quantised*, that is fundamentally discrete. This quantum theory required radical departures

from common-sense assumptions about nature such as the existence of well-defined physical properties prior to and independent of measurement. Furthermore, it introduced new physical phenomena such as entanglement and nonlocality, which had no analogue in the previous classical descriptions. We begin by enumerating the standard postulates of the theory of Quantum Mechanics followed by a brief discussion of the continuing difficulties in its interpretation.

2.2.1 Physical states

A complete description of a physical state in quantum mechanics is given by a ray in a *Hilbert space*, $\mathcal{H}$. A finite-dimensional Hilbert space is simply a vector space over the real or complex numbers with an inner product ¹. They are conventionally represented by a unit vector $|\psi\rangle$ which uniquely defines the state up to a global phase. Systems compose by tensor product, so the composite state of two systems, A and B, is a ray in the space $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$.

2.2.2 Observables

Physically measurable properties, known as observables, are described by Hermitian operators, $A = A^\dagger$. They admit a spectral decomposition

$$A = \sum_k a_k \Pi_k \tag{2.1}$$

where a_k are the eigenvalues of A and $\Pi_k = \sum_k |\phi_k\rangle\langle\phi_k|$ are projectors onto the space spanned by the vectors $|\phi_k\rangle$, the eigenstates of A with corresponding eigenvalue a_k ,

$$A|\phi_k\rangle = a_k|\phi_k\rangle \tag{2.2}$$

¹For infinite-dimensional spaces they must have the additional property of completeness, which loosely means there are no ‘gaps’ in the space.

2.2.3 Evolution

Evolution in quantum mechanics is described by a unitary operator, $U^{-1} = U^\dagger$. In the *Schrödinger picture*, states are time-dependent and the evolution is described by $|\psi(t)\rangle = U|\psi(0)\rangle$. Infinitesimal evolution is via the Schrödinger equation

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = \hat{H}|\psi(t)\rangle \quad (2.3)$$

where $\hat{H}$ is the *Hamiltonian*, the observable corresponding to a system's energy. When the Hamiltonian is time-independent $U = e^{-i\hat{H}t/\hbar}$.

Evolution in quantum mechanics can be equivalently expressed in the *Heisenberg picture*, where the time dependence is instead put into the observables, $A(t) = U^\dagger A(0)U$. In this case infinitesimal evolution is via

$$i\hbar \frac{dA(t)}{dt} = [A(t), \hat{H}] \quad (2.4)$$

assuming no additional time dependence of A .

2.2.4 Measurements

The measurement of a physical observable is a process in which a classical observer learns information about a quantum system. The outcome of a measurement of observable A is one of its eigenvalues $\{a_k\}$. The probability of measuring an outcome a_k when the quantum state prior to measurement is $|\psi\rangle$, is given by the *Born rule*

$$\Pr(k|\psi) = \langle \psi | \Pi_k | \psi \rangle \quad (2.5)$$

The measurement process changes the state, preparing it in an eigenstate (or eigenspace if the eigenvalue is degenerate) corresponding to the outcome via the update map

$$|\psi\rangle \rightarrow \frac{\Pi_k |\psi\rangle}{\sqrt{\langle \psi | \Pi_k | \psi \rangle}}. \quad (2.6)$$

2.2.5 Interpretations

These postulates present some problematic aspects. There appears to be friction between the evolution and measurement postulates. Quantum systems evolve deterministically and continuously via U , until they are measured by a classical observer when they evolve probabilistically and discontinuously. The problem is that we are not told what should be treated as a classical observer and what should be treated as a quantum system, i.e. where should we make the cut. If a measurement device is treated quantum mechanically, the system and measurement device will evolve deterministically to a joint state. In contrast, if we treat the measurement device classically, we ‘collapse’ the system state and receive a probabilistic outcome. This is usually referred to as the measurement problem and remains a subject of intense debate and disagreement to this day [9–11].

There have been a number of approaches to ‘fix’ the measurement problem, though none have garnered universal acceptance, and all seem unsatisfactory to some degree.

One can reject the measurement postulate and posit that all evolution is described by the evolution postulate, e.g. the many worlds interpretation [12]. The cost is of course that one must then accept the existence of a perhaps infinite number of unobservable universes. In addition, there are difficulties in making sense of the role of probabilities and Born’s rule if evolution is entirely deterministic.

Conversely, one can reject the evolution postulate, for instance collapse models where systems undergo spontaneous collapse on some characteristic timescale, though in this case one must accept that energy is no longer completely conserved [13].

Alternatively, one can reject the state postulate, i.e. that $|\psi\rangle$ contains a complete description of the physical state. One such way to do so is to use extra degrees of freedom which describe the ‘real’, or ontological, state. One example is Bohmian mechanics [14], where the quantum state is supplemented with the true positions of all particles, the wavefunction then determines the evolution of these positions in time. The difficulties of this approach include the special status of position over all other observables, and its inherent nonlocality.

Finally, one can reject some other aspect of the formalism, for instance by replacing

the rules of classical logic [15] or rejecting the notion of an objective reality [16].

The remarkable irony of quantum theory is that although there remains little agreement about how it should be understood, it is arguably the most successfully tested physical theory ever devised (e.g. it predicts the fine structure constant to at least within half a part per billion [17]). In the following, we will not concern ourselves with any particular interpretation and instead adopt the “Shut up and calculate!” [18] approach; using the well understood mathematical rules of the theory to translate and predict the results of experiments consisting of state preparations, transformations and measurements on quantum systems. In order to do so we will require a slightly more general formalism, which can additionally describe classical uncertainties and the study of quantum subsystems.

2.2.6 Operational quantum theory

In this section we briefly describe the most general state preparations, measurements and evolution allowed in quantum theory.

States

The most general form of quantum state preparation is an ensemble $\{p_i, |\psi_i\rangle\}$, where p_i is the probability of preparing the *pure* state $|\psi_i\rangle$. This ensemble defines a *density matrix* $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. If $\text{rank}(\rho) > 1$ it is referred to as a *mixed* state. A general ρ has the properties

$$\begin{aligned} \text{Hermiticity: } \rho &= \rho^\dagger \\ \text{Trace one: } \text{Tr}[\rho] &= 1 \\ \text{Positivity: } \langle\psi|\rho|\psi\rangle &\geq 0 \quad \forall |\psi\rangle \in \mathcal{H} \end{aligned} \tag{2.7}$$

where the positivity constraint avoids the possibility of negative probabilities for measurement outcomes and is usually denoted as $\rho \geq 0$. As well as describing a classical probabilistic mixture of pure states, ρ can also describe the reduced state of a subsystem of a composite system. For instance, the state of system A in a joint state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$

can be described by the partial trace over the space $\mathcal{H}_B$,

$$\mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \rho_A. \quad (2.8)$$

Conversely, any ρ can be described as the reduced state of a pure state (known as a *purification* of ρ) in a larger space.

Measurements

The most general form of quantum measurement is a *positive operator valued measure* (POVM). This is a set of operators $\{\mathbf{E}\}$ known as the POVM effects which satisfy the conditions

$$\begin{aligned} \text{Positivity: } \mathbf{E}_k &\geq 0 \quad \forall k \\ \text{Sum to identity: } \sum_k \mathbf{E}_k &= \mathbb{1} \end{aligned} \quad (2.9)$$

The Born rule for these generalised measurements and states becomes

$$\mathrm{Pr}(k|\rho) = \mathrm{Tr}[\mathbf{E}_k \rho] \quad (2.10)$$

and the state update rule for a measurement with outcome k becomes

$$\rho \rightarrow \frac{\mathbf{M}_k \rho \mathbf{M}_k^\dagger}{\mathrm{Tr}[\rho \mathbf{E}_k]} \quad (2.11)$$

where $\mathbf{E}_k = \mathbf{M}_k^\dagger \mathbf{M}_k$. Any POVM can be described by coupling the measured system to an ancilla system prepared in a fixed state and performing a regular projective measurement upon the composite system, a procedure known as a *Naimark extension*.

Evolution

The most general form of quantum evolution is via completely-positive trace-preserving (CPTP) maps, sometimes known as quantum channels. A CPTP map is a superoperator, $\mathcal{E}$, which maps density matrices to density matrices, $\mathcal{E}(\rho) = \rho'$. They must satisfy the

following conditions

$$\begin{aligned}
&\text{Linearity: } \mathcal{E}\left(\sum_k p_k \rho_k\right) = \sum_k p_k \mathcal{E}(\rho_k) \\
&\text{Trace-preserving: } \text{Tr}[\mathcal{E}(\rho)] = 1 \\
&\text{Complete positivity: } [\mathcal{E} \otimes \mathbb{1}](\rho) \geq 0 \quad \forall \rho \in \mathcal{H}_A \otimes \mathcal{H}_B
\end{aligned} \tag{2.12}$$

Where the additional requirement of complete positivity is required to ensure that the output state is valid when the channel is applied to a subsystem of a larger system. All maps satisfying these conditions can be written in *Kraus decomposition* as

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger \tag{2.13}$$

where $\sum_k A_k^\dagger A_k = \mathbb{1}$. $\{A_k\}$ are known as the Kraus operators of $\mathcal{E}$, and they are defined up to a global unitary transformation. Any quantum channel can be described by coupling the system to an ancilla system in a pure state, unitarily evolving the composite system and throwing away the ancilla.

To describe infinitesimal evolution for more general non-unitary evolutions that are nevertheless Markovian ($\rho(t + dt)$ can be determined entirely by $\rho(t)$), the Schrödinger equation can be generalised to the *Lindblad master equation*

$$\dot{\rho} = -i[\hat{H}, \rho] + \sum_i \left(L_i \rho L_i^\dagger - \frac{1}{2} \{L_i^\dagger L_i, \rho\} \right) \tag{2.14}$$

where $\{L_i\}$ are known as the Lindblad operators and $\hat{H}$ is the natural Hamiltonian of the system.

2.3 Quantum Information

In 1948, Claude Shannon at Bell Labs published a seminal paper that established the field of information theory [19]. This work explained how to formalise and quantify the concept of information and its communication via the use of probability theory and entropy. A

‘message’ can be considered as a discrete random variable X which can take one of n symbols $\{x_1, x_2, \dots, x_n\}$. The information content of the message is then determined by the probability mass function $\{\text{Pr}(x_i)\}$ over these symbols and quantified by the *Shannon entropy*

$$H(X) = - \sum_i \text{Pr}(x_i) \log_2 \text{Pr}(x_i) \quad (2.15)$$

Shannon defined the minimal unit of information as a *bit*, the amount of information obtained from a message which has an equal probability of being in one of two symbols, or states, conventionally labelled 0 and 1. Just over half a century later, these contributions have dramatically shaped the modern information age in which we live.

A profound insight, which has become increasingly popular among scientists in the last few decades, is that the concept of information can also be fundamental to our understanding of physics. This position is encapsulated by John Archibald Wheeler’s doctrine ‘it from bit’². Quantum mechanics naturally lends itself an information-theoretic description since, when considered entirely abstractly, it can be thought of as in some sense a generalised probability theory. The development of quantum information theory has lead to a much richer understanding of the theory itself, allowing for the differences between classical and quantum theories to be elucidated and quantified.

2.3.1 The Qubit

As the minimal unit of classical information is the bit, the minimal unit of quantum information is the *qubit*, a quantum system that lives in a two-dimensional Hilbert space. Conventionally, the state of a qubit is often described in the *computational basis*:

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} ; |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (2.16)$$

²“It from bit. Otherwise put, every it - every particle, every field of force, even the space-time continuum itself - derives its function, its meaning, its very existence entirely - even if in some contexts indirectly - from the apparatus-elicited answers to yes-or-no questions, binary choices, bits”

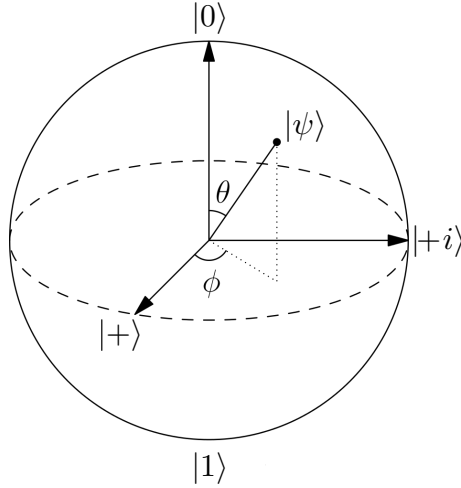


Figure 2.1: Bloch sphere representation of pure qubit state. An arbitrary single qubit pure state $|\psi\rangle$ can be represented by a point on the surface of the unit sphere parameterised by the angles θ and ϕ .

Since an arbitrary pure qubit state can be parameterised as

$$|\psi\rangle = \sin \frac{\theta}{2} |0\rangle + e^{i\phi} \cos \frac{\theta}{2} |1\rangle \quad (2.17)$$

they have an intuitive geometrical representation as points on the surface of the three-dimensional unit sphere. Of particular use in what follows will be the eigenstates of the Pauli matrices $\{\sigma_x, \sigma_y, \sigma_z\}$ (hereafter simply $\{X, Y, Z\}$)

$$Z : \{|0\rangle, |1\rangle\} \quad (2.18)$$

$$X : \{|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\} \quad (2.19)$$

$$Y : \{|+i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), |-i\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\} \quad (2.20)$$

which form the vertices of an octahedron centred on the origin.

Other important aspects of quantum information can be visualised using the Bloch sphere. Points within the Bloch sphere correspond to mixed states. We can therefore

represent an arbitrary qubit state by a vector $\vec{r} = (r_x, r_y, r_z)$ where

$$\rho = \frac{1}{2}(\mathbb{1} + r_x X + r_y Y + r_z Z). \quad (2.21)$$

The pure states correspond to unit vectors $|\vec{r}| = \sqrt{r_x^2 + r_y^2 + r_z^2} = 1$, and the length of the vector describes how mixed the state is. This mixedness is often characterised by the *purity*, $\mathcal{P} = (1 + |\vec{r}|^2)/2$ of the state. More generally,

$$\mathcal{P}(\rho) = \text{Tr}(\rho^2). \quad (2.22)$$

Another important property that is clear from considering the Bloch sphere is that a general decomposition of a mixed state as a convex combination of pure states is not unique. For instance, the maximally mixed state, $\rho = \mathbb{1}/2$, can be decomposed as

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|) \quad (2.23)$$

and, indeed, in an infinite number of ways. However, for all non-maximally mixed states, ρ has a unique spectral decomposition where the constituent pure states are the orthogonal eigenstates of ρ

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle \psi_i| = \frac{1 + |\vec{r}|}{2} |\psi\rangle\langle \psi| + \frac{1 - |\vec{r}|}{2} |\psi_\perp\rangle\langle \psi_\perp|. \quad (2.24)$$

Finally, the concept of information as entropy can be carried over to the quantum realm. The *von Neumann entropy* is defined as

$$S(\rho) = -\text{Tr}(\rho \log_2 \rho) = -\sum_i \lambda_i \log_2(\lambda_i). \quad (2.25)$$

For all pure states $S(|\psi\rangle\langle \psi|) = 0$ which can be understood as the fact that there exists a measurement with a deterministic outcome (the measurement $\{|\psi\rangle\langle \psi|, \mathbb{1} - |\psi\rangle\langle \psi|\}$), whereas $S(\mathbb{1}/2) = 1$ as the result of any measurement is completely random, i.e. it is equivalent to a fair coin flip and therefore a single bit. The Von Neumann entropy

replaces Shannon entropy in quantum information theory and also provides a way to quantify quantum entanglement, as we will see below.

2.3.2 Entanglement

Although much of the quantum information framework can be understood by the single qubit example, a key feature of quantum theory that often defies classical description is *entanglement* [20] between multiple quantum systems. A pure entangled state is one which cannot be written as a *product state*, a tensor product of pure states on each of the subsystems, e.g.

$$|\psi\rangle_{AB} = |\Psi\rangle_A \otimes |\Phi\rangle_B. \quad (2.26)$$

The entangled states, which cannot be written in this way, therefore cannot be understood in terms of independent, *local* states of each subsystem and are instead only fully described by a correlated state of the whole system.

The set of entangled states can be defined as those which are not *separable*. Separable states are those which can be written as

$$\rho_{\text{sep}} = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \otimes |\Phi_i\rangle\langle\Phi_i|. \quad (2.27)$$

Examples of pure entangled states for two qubits that will be used throughout this thesis are the *Bell states*. These are

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad |\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \quad (2.28)$$

and they form a basis for two qubit states.

An alternative way to understand entanglement is to define it as that which cannot be created using only LOCC: local quantum operations on each subsystem and classical communication. This definition helps quantify how much entanglement there is in a state. Although there are many ways to quantify entanglement (see, for example, Ref. [21]), an appealing general property of an entanglement measure is LOCC-monotonicity,

that is functions which do not increase under LOCC. For pure, bipartite systems there is a natural entanglement measure which most such monotones reduce to known as the entropy of entanglement

$$E[|\psi\rangle_{AB}] = S(\text{Tr}_B[|\psi\rangle\langle\psi|]) = -\text{Tr}[\rho_A \log \rho_A] = -\sum_i \lambda_i^A \log \lambda_i^A \quad (2.29)$$

where S is the von Neumann entropy introduced previously. A quick check verifies that product states, with pure reduced states will have no entropy of entanglement (e.g. $E[|00\rangle] = 0$) whereas maximally entangled states whose reduced states are maximally mixed maximise E (e.g. $E[|\phi^+\rangle] = 1$).

Unfortunately when considering mixed state entanglement, things get considerably messier. Consider, for instance, the mixed state formed by taking an equal mixture of Bell states

$$\rho = \frac{1}{2}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|) \quad (2.30)$$

this state has an equivalent decomposition as

$$\rho = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) \quad (2.31)$$

and is therefore a separable, classically correlated state even though it is a mixture of maximally entangled states. The entropy of entanglement can be generalised to mixed states as the *entanglement of formation*

$$E_F(\rho) = \min_{\{p_i, \psi_i\}} \sum_i p_i E[|\psi_i\rangle] \quad (2.32)$$

where the minimum is taken over all pure state decompositions of ρ . The entanglement of formation remains an LOCC-monotone but the form of the minimisation makes its calculation extremely difficult in general. However, for the simple case of two-qubit states there is a closed form solution based on another entanglement measure, the *concurrence*

$$C(\rho) = \max[0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4] \quad (2.33)$$

where λ_i are the ordered square roots of the eigenvalues of the matrix $\rho(Y \otimes Y)\rho^*(Y \otimes Y)$. This can then be related to E_F as

$$E_F(\rho) = h\left(\frac{1 + \sqrt{1 - C^2(\rho)}}{2}\right) \quad (2.34)$$

where

$$h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x). \quad (2.35)$$

For pure states $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, the concurrence takes the particularly simple form

$$C(|\psi\rangle) = 2|ad - bc|. \quad (2.36)$$

2.3.3 Nonlocality

Entanglement was at the core of one of the greatest foundational developments in modern physics, the discovery of Bell nonlocality [22]. In their famous paper of 1935 [23], Einstein, Podolsky and Rosen (EPR) argued ³ that quantum theory could not be a complete description of reality. At the heart of their argument was the implicit assumption of *local realism*; that systems have well-defined states of reality which can only be influenced by their immediate surroundings. If two systems are in an entangled state $|\psi\rangle_{AB}$, then the choice of measurement performed upon A can affect the post-measurement state of system B , $|\phi\rangle_B$. But if $|\phi\rangle_B$ represents the complete reality of system B , then it would appear that what happens at A can instantaneously affect the reality of B , even when the systems are arbitrarily separated. Since such influences violate locality, the conclusion of EPR was that the quantum state was not a complete description of the true underlying reality of a system, and when additional ‘hidden variables’ are included, locality can be restored.

In 1964 John Bell proved the remarkable result that even if the quantum state is an incomplete description of reality, we still cannot describe the outcomes of all quantum

³It has been claimed that Einstein did not in fact read the paper prior to publication and that in correspondence with Schrödinger suggested that it had not completely captured his perspective. See [24] for his own, simpler argument along the lines of what we present here.

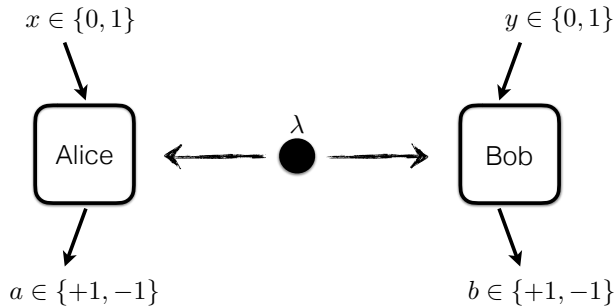


Figure 2.2: Schematic of CHSH Bell test. Two separated parties Alice and Bob, possibly sharing prior correlations λ , are given inputs x and y and output a and b respectively. The conditional probabilities $\Pr(ab|xy)$ allowed in any physical theory which obeys local realism are constrained by inequalities. Quantum theory violates these inequalities, proving that it cannot be described by any local realistic model.

experiments with a local model [25]. The simplest statement of Bell’s theorem can be made by considering the following scenario depicted in Fig. 2.2: two parties, conventionally named Alice and Bob are first separated so that they cannot communicate, they are then each given an input in the form of a single bit, $x, y \in \{0, 1\}$, and told to produce an output $a, b \in \{-1, +1\}$. After many trials, the expectations $\langle a_x b_y \rangle = \sum_{a,b} ab \Pr(ab|xy)$ can be estimated, and the quantity

$$S = \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle \quad (2.37)$$

calculated. Since Alice and Bob are separated and have no prior knowledge of the inputs they will be given, we can express the condition of locality via a factorising of the probabilities of their outcomes, i.e. Alice’s output cannot depend upon Bob’s input. Alice and Bob can share prior correlations however, including the possibility of any underlying ‘hidden variables’ or joint causal influences, and these can be included as an additional variable λ

$$\Pr(ab|xy) = \int d\lambda q(\lambda) \Pr(a|x, \lambda) \Pr(b|y, \lambda) \quad (2.38)$$

where in any given run, the value of λ is determined by some distribution $q(\lambda)$. Now, substituting Eq. (2.38) into Eq. (2.37) results in the bound

$$S \leq 2 \quad (2.39)$$

known as the Clauser-Horne-Shimony-Holt (CHSH) inequality [26]. Note that so far we have not made any reference to physics, the CHSH inequality is simply a general bound upon all physical theories whose statistics satisfy Eq. (2.38).

Now, consider the following scenario in quantum theory: Alice and Bob each take one qubit of the entangled Bell state $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ and according to their input bit, perform measurements of the following observables

$$x_0 \rightarrow Z, x_1 \rightarrow X, y_0 \rightarrow -(Z + X)/\sqrt{2}, y_1 \rightarrow (-Z + X)/\sqrt{2}. \quad (2.40)$$

The expectations in Eq. (2.37) in this case become $\langle a_0 b_0 \rangle = \langle a_0 b_1 \rangle = \langle a_1 b_0 \rangle = 1/\sqrt{2}$ and $\langle a_1 b_1 \rangle = -1/\sqrt{2}$ resulting in

$$S = 2\sqrt{2} > 2. \quad (2.41)$$

This then proves that no local, realistic model can reproduce the correlations in this experiment.

Since the first convincing experimental demonstration by Aspect *et al.* [27], Bell inequalities have been violated in many laboratory settings and recently three experiments [2–4] were performed in a loophole-free manner, meaning that Alice and Bob were separated so that their measurement settings were chosen outside of each others light cones and the detection was efficient enough that no conspiratorial loss could have changed the results.

Finally we note that entanglement and nonlocality, although related, are not equivalent concepts and there exist subtleties in their relationship that are still not well understood. For instance, there are known entangled states which still admit a local hidden variable model [28] and, in some scenarios, the states which maximally violate Bell inequalities are not the most entangled states [29].

2.3.4 Quantum tomography

Although any quantum system is described by its state ρ , learning the state of an unknown system poses fundamentally new challenges.

Firstly, in contrast to classical theories, it is not always possible to simultaneously measure a set of quantum observables. This is because two physical observables $\hat{A}$ and $\hat{B}$ will generally not commute $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0$, i.e. they will have different eigenstates and so cannot be jointly measured whilst remaining consistent with the measurement postulate. Furthermore, since measurement collapses the state, in order to infer the results of incompatible measurements on a system, we require multiple copies of the initial state.

A second reason why multiple copies of a state are required is that our knowledge is restricted by the inherently statistical nature of quantum measurements. Even if we believe the state to be $|\psi\rangle$ and perform the two-outcome measurement $\{|\psi\rangle\langle\psi|, \mathbb{1} - |\psi\rangle\langle\psi|\}$, a positive result is still consistent with any state not orthogonal to $|\psi\rangle$.

The task of quantum state tomography is to estimate an unknown state ρ by performing measurements on a set of identically prepared copies of it⁴. If the system has a d -dimensional Hilbert space, then ρ is a $d \times d$ Hermitian matrix containing d^2 independent parameters. One of these parameters is fixed by the trace condition $\text{Tr}(\rho) = 1$, and so $d^2 - 1$ parameters must be estimated.

The simplest example of state tomography is for a single qubit. A natural choice of measurements are the Pauli observables (plus identity) as they form an orthonormal basis for 2×2 hermitian matrices. From Eq. (2.21) we see that this case is equivalent to estimating the vector $\vec{r}$ where

$$r_x = \langle X \rangle = \text{Tr}(X\rho), \quad r_y = \langle Y \rangle = \text{Tr}(Y\rho), \quad r_z = \langle Z \rangle = \text{Tr}(Z\rho). \quad (2.42)$$

These probabilities are naturally estimated by sample expectation values. If we measure Z on N copies of ρ , collapsing n_0 copies onto $|0\rangle\langle 0|$ and n_1 to $|1\rangle\langle 1|$, then we can estimate r_z by

$$\frac{n_0 - n_1}{N} = r_z^{\text{est}} \approx \langle Z \rangle \quad (2.43)$$

and the state can then be estimated directly via Eq. (2.21). A problem that is apparent

⁴Of course in reality the copies ϱ_i will not be completely identical. However, as long as the choice of measurement we perform on a given copy is not correlated with that position then we can faithfully reconstruct the average state $\rho = \sum_i^N \varrho_i / N$.

in the above is that we can only ever estimate $\langle Z \rangle$ to some degree of precision due to statistical fluctuations. These fluctuations can cause further problems, as if the state that is being tomographed is approximately pure, the estimates we get $\{r_x^{\text{est}}, r_y^{\text{est}}, r_z^{\text{est}}\}$ can result in a reconstructed state which contains negative eigenvalues, i.e. is not a valid density matrix. To overcome this problem there are several methods of statistical inference that can be used [30] which restrict to only valid state estimates, the most commonly used approach is maximum likelihood estimation [31].

The likelihood principle loosely states that all information from a set of data relevant to inferences about the value of some parameter θ is contained in the *likelihood function*, $\mathcal{L}(\theta) \equiv \text{Pr}(\text{data}|\theta)$. Therefore, by using a parameterised ρ as our model, a tomographic estimate can be obtained by maximising the likelihood function $\mathcal{L}(\rho) = \text{Pr}(\text{counts}|\rho)$.

Process tomography is a similar task in which from a series of uses of a quantum channel the transformation it performs is estimated. We begin with an aside regarding representations of quantum processes. In Sec. 2.2.6 we introduced CPTP maps described in Kraus form. For performing process tomography it is usually more convenient to use the *process matrix* or Choi representation. Here we simply express the channel in a basis of operators $\{B\}$

$$\mathcal{E}(\rho) = \sum_{i,j} \chi_{ij} B_i \rho B_j. \quad (2.44)$$

The Kraus representation is recovered by diagonalising the matrix χ . The complete positivity constraint is also straightforwardly expressed as the condition $\chi \geq 0$.

The Choi-Jamiołkowski (CJ) isomorphism states that there is a one-to-one correspondence between the set of d -dimensional CPTP maps and a subset of d^2 -dimensional states. It is given by

$$\rho_{\mathcal{E}} = (\mathcal{E} \otimes \mathbb{1})[|\Psi\rangle\langle\Psi|] \quad (2.45)$$

where $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle|i\rangle$ is a maximally entangled state in d^2 -dimensions. When represented in the computational basis, the process matrix χ is directly related to the CJ state via $\rho_{\mathcal{E}} = \chi/d$. This link means that process tomography can be thought of as state tomography upon the CJ state. The probability for an experiment in which a preparation

ρ_i is input, and a POVM element E_j is measured can be calculated as

$$\Pr(j|i) = d\text{Tr}[(\rho_i^\top \otimes E_j)\rho_{\mathcal{E}}]. \quad (2.46)$$

This expression can then be used to perform maximum likelihood process tomography. The only difference is that the trace constraint $\text{Tr}(\rho) = 1$ becomes the trace-preserving constraint $\text{Tr}_A(\rho_{\mathcal{E}}) = \mathbb{1}/d$. This new constraint means that a CJ state (or CPTP map) has only $d^4 - d^2$ parameters, and not the $d^4 - 1$ of a general state.

2.3.5 Distance measures

In order to characterise the quality or utility of a given state or process, it is important to have some notion of the similarity of quantum states and channels. Here we give the main distance measures that are used in this thesis to compare states and processes. Given that most quantum experiments also involve estimating probabilities from relative experimental frequencies, we also introduce the distance measures for classical probability distributions that will be used.

The *trace distance* (sometimes known as the L^1 or total variation distance) between two probability distributions $\{p_i\}$ and $\{q_i\}$ is defined as

$$D(\vec{p}, \vec{q}) \equiv \frac{1}{2} \sum_i |p_i - q_i|. \quad (2.47)$$

The trace distance is a metric on probability distributions⁵ and has an operational meaning as the probability of successfully distinguishing between the two distributions given that the optimal event for this task is received. The *statistical fidelity* is defined as

$$\mathcal{F}_s(\vec{p}, \vec{q}) \equiv \sum_i \sqrt{p_i q_i}. \quad (2.48)$$

Unlike the trace distance, the statistical fidelity is not a metric, but it gives a measure of the overlap of two distributions and is particularly useful to consider given its quantum

⁵I.e. it satisfies the conditions $D(x, x) = 0$, $D(x, y) = D(y, x)$ and the triangle inequality: $D(x, z) \leq D(x, y) + D(y, z)$

generalisation.

Analogous measures can be defined for quantum states. The trace distance between quantum states ρ and σ becomes

$$D(\rho, \sigma) \equiv \frac{1}{2} \|\rho - \sigma\|_1 \quad (2.49)$$

where the trace norm is $\|X\|_1 = \text{Tr}(\sqrt{X^\dagger X})$. If both states are pure, this reduces to $D(|\psi\rangle, |\phi\rangle) = \sqrt{1 - |\langle\psi|\phi\rangle|^2}$. As in the classical case, it is a metric and has an operational meaning related to a discrimination task. In the scenario in which, with equal probability, an agent is given either ρ or σ , then the maximum probability of the agent correctly guessing the state given as single copy (when they perform the optimal measurement) is given by $\text{Pr}_{\max} = \frac{1}{2} + \frac{1}{2}D(\rho, \sigma)$ [32].

The *quantum state fidelity* of two states ρ and σ is

$$\mathcal{F}(\rho, \sigma) \equiv \|\sqrt{\rho}\sqrt{\sigma}\|_1^2. \quad (2.50)$$

If $\sigma = |\psi\rangle\langle\psi|$ then $\mathcal{F}(\rho, |\psi\rangle) = \langle\psi|\rho|\psi\rangle$. This is often the case when comparing a realistic experiment to its ideal counterpart. In such scenarios the fidelity can be understood as the probability of measuring an ideal target state $|\psi_t\rangle$ (via the measurement $\{|\psi_t\rangle\langle\psi_t|, \mathbb{1} - |\psi_t\rangle\langle\psi_t|\}$) given an experimental state ρ ⁶. Furthermore, it can be shown that the fidelity can more reliably indicate the number of copies required to distinguish two states with high accuracy, a highly relevant task in tomography experiments [35].

The fidelity and trace distance can be related to each other via the Fuchs-van de Graaf inequalities [36]

$$1 - \sqrt{\mathcal{F}(\rho, \sigma)} \leq D(\rho, \sigma) \leq \sqrt{1 - \mathcal{F}(\rho, \sigma)} \quad (2.51)$$

where the second bound becomes an equality when ρ and σ are both pure.

These measures can also be promoted to distance measures on channels [34]. Two natural generalisations of the state fidelity are the *process fidelity* and *average gate fidelity*.

⁶Sometimes fidelity is defined as $\sqrt{\mathcal{F}}$, e.g. in [33], however we choose this definition so that it can be interpreted as a probability in the manner described [34].

The process fidelity between two CPTP maps $\mathcal{E}$ and Λ is simply the state fidelity between their CJ states

$$\mathcal{F}_{\text{pro}}(\mathcal{E}, \Lambda) = \mathcal{F}(\rho_{\mathcal{E}}, \rho_{\Lambda}). \quad (2.52)$$

The average gate fidelity is defined as the state fidelity between the channel outputs, averaged over all pure input states

$$\mathcal{F}_{\text{ave}}(\mathcal{E}, \Lambda) = \int d\mu(\psi) \mathcal{F}[\mathcal{E}(|\psi\rangle\langle\psi|), \Lambda(|\psi\rangle\langle\psi|)] \quad (2.53)$$

where $d\mu$ is the uniform measure (the *Haar measure*) on state space. In the case that one of the channels is unitary, $\Lambda = \mathcal{U}$, these measures can be related via

$$(d+1)\mathcal{F}_{\text{ave}}(\mathcal{E}, \mathcal{U}) = d\mathcal{F}_{\text{pro}}(\mathcal{E}, \mathcal{U}) + 1 \quad (2.54)$$

where d is the Hilbert space dimension. The main reason for using these fidelity measures is largely experimental convenience. They can be estimated with fewer measurements than are required for process tomography [37] and the average gate infidelity $(1 - \mathcal{F}_{\text{ave}})$ is the quantity calculated in *randomised benchmarking* protocols [38], a standard characterisation procedure in matter-based qubits.

Although it is possible to define average and process versions of the trace distance in a similar manner, a more commonly used measure is the *diamond distance* [39]. The diamond distance between maps $\mathcal{E}$ and Λ is

$$\|\mathcal{E} - \Lambda\|_{\diamond} = \max_{\rho} \|(\mathbb{1} \otimes \mathcal{E})[\rho] - (\mathbb{1} \otimes \Lambda)[\rho]\|_1 \quad (2.55)$$

and has an operational meaning which is analogous to the trace distance for states: it is related to the maximum probability of successfully distinguishing between the channels $\mathcal{E}$ and Λ in a single-shot using the optimal input state (which in general can be half of an entangled state on a larger system) and making the optimal measurement. Although less straightforward to calculate than the previous measures, the diamond distance can be numerically computed via semidefinite programming [40].

In the context of quantum computing, the average gate fidelity represents an average-case error rate whereas the diamond distance represents a worst-case error rate. These measures can be related to each other via the inequalities

$$\frac{(d+1)(1 - \mathcal{F}_{\text{ave}}(\mathcal{E}, \mathbb{1}))}{d} \leq \frac{1}{2} \|\mathcal{E} - \mathbb{1}\|_{\diamond} \leq \sqrt{d(d+1)(1 - \mathcal{F}_{\text{ave}}(\mathcal{E}, \mathbb{1}))}. \quad (2.56)$$

2.4 Quantum computing

Having introduced the idea of quantum information and demonstrated ways in which it can deviate from classical information, we now turn to the central objective of this thesis, using quantum systems to perform computational tasks.

2.4.1 Turing machines

In a landmark paper of 1937 [41], Alan Turing developed a theory of computation based on what is now known as a *Turing machine*. This is a hypothetical device which is comprised of two elements. A tape which is divided into cells, each of which contains a symbol, S_i , from some finite alphabet and a controller which is in an internal state Q_i from some finite set. The controller can perform the following actions; move one cell to the left or right, read the current symbol, overwrite the current symbol, change internal state and stop running and output an answer. A computer program can then be understood as the set of rules which determine the action of the controller given $\{S_i, Q_i\}$, and the initial configuration of the tape defines the instance or input of the problem.

Turing showed that there is a universal Turing machine, i.e. there is a Turing machine which can take as input a description of any other Turing machine M and its input x and produce the same output as M would have if given x . This result confirmed the existence of programmable computers, an important notion that means that you can use the same computer to check your emails and play DOOM.

In a famous proof, Turing then went on to show that the halting problem, the problem of deciding if a given program will ever halt, cannot be solved with any Turing machine. This led to the *Church-Turing thesis*, which states that any problem that is computable

is computable by a Turing machine. Formally, a decision problem (one with the two possible outputs ‘accept’ and ‘reject’) can be considered as a *language* $L \subseteq \{0,1\}^*$, a collection of binary strings. For a computable problem there must exist a Turing machine which accepts all strings in L and rejects all strings not in L .

A variation of the Church-Turing thesis that connects this abstract theory to physics is the *Physical Church-Turing Thesis*, which states that all physically computable functions are computable on a Turing machine. However, this statement does not address the efficiency of computing. Certain problems can often be ‘practically uncomputable’ given the resources in terms of time and memory that would be required to solve them. To address this question, and understand where quantum computers enter this story, we need the machinery of computational complexity theory.

2.4.2 Computational Complexity Theory

Complexity theory aims to classify computational problems into classes related to the relationship between the size of the problem instance and the minimal resources required to solve it. This can be quantified by considering the resources of time and space (i.e. memory). $\text{TIME}(f(n))$ is the class of languages for which an instance of size n requires an amount of time (or number of computational steps) that grows as $\mathcal{O}(f(n))$. Similarly, $\text{SPACE}(f(n))$ is the class of languages for which an instance of size n requires an amount of space (i.e. number of bits of memory) that grows as $\mathcal{O}(f(n))$. Using these definitions, the following broader complexity classes can be considered:

- P (Polynomial) is the class of languages that are decidable by a Turing machine in polynomial time, i.e. it is the class $\cup_{k \in \mathbb{N}} \text{TIME}(n^k)$.
- NP (Nondeterministic Polynomial) is the class of languages for which, if a Turing machine accepts then there is a polynomial size witness that allows you to verify that answer in polynomial time. Formally, a language L is in NP if there exists a polynomial time Turing machine M , a string $x \in \{0,1\}^*$ and a witness $w \in \{0,1\}^{\text{poly}(|x|)}$ such that $\forall x \in L, M(x,w)$ accepts and $\forall x \notin L, M(x,w)$ rejects.

- PSPACE (Polynomial Space) is the class of languages that are decidable by a Turing machine in polynomial space, i.e. it is the class $\cup_{k \in \mathbb{N}} \text{SPACE}(n^k)$.
- EXP (Exponential) is the class of languages that are decidable by a Turing machine in exponential time, i.e. it is the class $\cup_{k \in \mathbb{N}} \text{TIME}(2^{n^k})$.

and these classes obey the containments

$$P \subseteq NP \subseteq PSPACE \subset EXP. \quad (2.57)$$

where only the final one is currently provably strict, although all are suspected to be so. The question $P \stackrel{?}{=} NP$ remains one of the great open questions in mathematics but it is widely believed, and often assumed, that $P \neq NP$. Problems in P include basic numerical operations, linear programming and primality testing. Problems in NP include factoring, the travelling salesman problem and Donkey Kong [42]. The former of these is quite special however. To see why we need to introduce some extra concepts. A problem is said to be NP-hard if any NP problem can be efficiently reduced to it. A problem is then said to be NP-complete if it is both a) in NP and b) NP-hard. This means that if an efficient algorithm was known for any NP-complete problem, then all NP problems would be efficiently solvable. Factoring integers into their prime factors is a rare example of a natural problem that is believed to be in NP but not NP-complete.

We can also consider a generalisation of P when the additional resource of a source of randomness is allowed.

- BPP (Bounded-error Probabilistic Polynomial) is the class of languages for which there exists a polynomial time Turing machine M, a string $x \in \{0, 1\}^*$ and a string of perfectly random bits $r \in \{0, 1\}^{\text{poly}(|x|)}$ such that $\forall x \in L, \Pr[M(x, r) \text{ accepts}] \geq \frac{2}{3}$ whereas $\forall x \notin L, \Pr[M(x, r) \text{ accepts}] \leq \frac{1}{3}$.

The probabilities are chosen by convention and in principle could be any values p and $1 - p$ where $p \neq 1/2$. The hierarchy of all these complexity classes is depicted in Fig. 2.3.

This finally leads us to the *Extended Church-Turing thesis* (ECT). It states [8] that

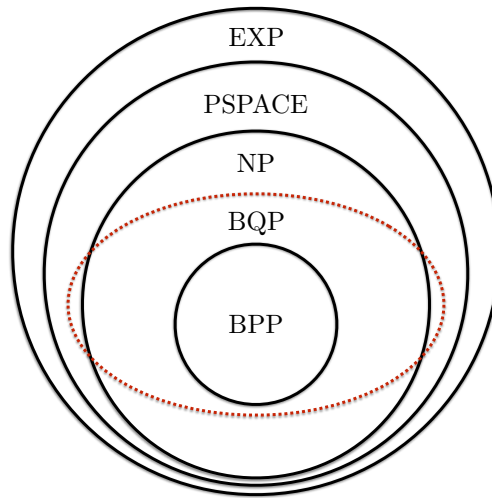


Figure 2.3: Hierarchy of computational complexity classes. Solid black lines indicate known containments. Dashed red line indicates suspected relationship of BQP to other classes.

“All computational problems that are efficiently solvable by realistic physical devices, are efficiently solvable by a probabilistic Turing machine.”

The ECT is a statement about the relationship between physics and computation and its claim is that no problem outside of BPP can be efficiently solved by a physical device. With the introduction of new ideas from quantum information theory however, this statement now appears highly likely to be false.

2.4.3 Circuit model quantum computation

The stage is now set to introduce the concept of a quantum computer. The first description of an analog quantum computing device is usually attributed to Feynman [43] in 1982, where he suggested that the exponentially large state space of quantum systems could only be simulated efficiently by a computer which works according to quantum theory too. This was put onto digital ground in 1985 when David Deutsch defined a universal Quantum Turing machine [44]. This is a generalisation of a Turing machine where now the tape and controller can be quantum systems.

A perhaps simpler way to understand a universal quantum computer is in terms of the

quantum circuit model. In this model we can think of any computation as three stages:

Initialisation

We start with a system of n qubits. These qubits are prepared in a simple to prepare state which can be dependent upon the input string x and a set of ‘ancilla’ qubits prepared in the all zero state, e.g. $|x\rangle \otimes |0\dots 0\rangle$.

Gates

A universal quantum computer should be able to apply an arbitrary unitary transformation to the n qubit input state. For instance, for some computable function f there will be a unitary U which prepares the state $U|x\rangle|0\rangle = |x\rangle|f(x)\rangle$. However, this procedure will generally not be efficient. Instead, in analogy with classical computing where circuits can be built up by composition of a small gate set of logic gates, we are interested in quantum circuits which can be constructed by composing a polynomial number of gates from a small gate set that act on just one or a few qubits.

Examples of common single qubit quantum gates are the Pauli matrices and the Hadamard and T gate, which can be represented in the computational basis as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \quad (2.58)$$

and the most common two qubit gates are the controlled-NOT (CNOT) and controlled-Z gates (CZ)

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \text{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.59)$$

These gates can generate entanglement when acted upon two separable qubits and so are sometimes referred to as ‘entangling gates’. They can be understood as gates that perform conditional logic in the computational basis, for instance a controlled-Z gate performs a

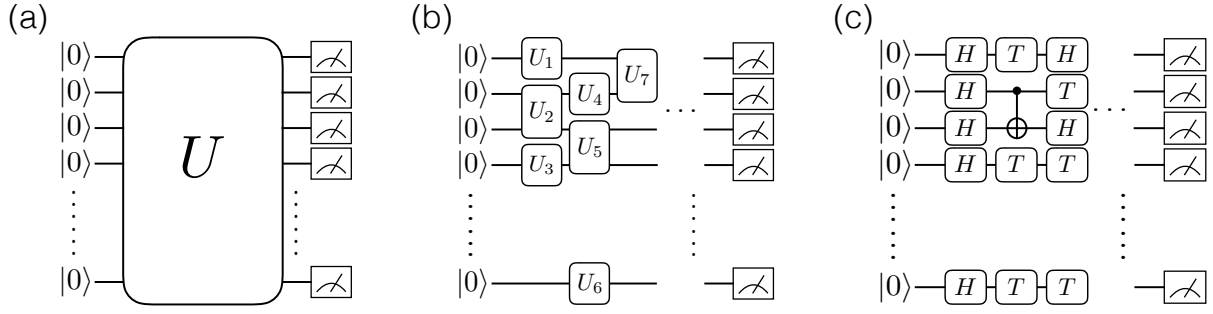


Figure 2.4: Circuit model quantum computation. All qubits are initialised in a known state e.g. $|0\rangle^{\otimes n}$ and measured in the computational basis. a) The computation is encoded in an n -qubit unitary U . b) This is decomposed into single and two-qubit unitaries. c) Each of these are approximated via a universal gate set e.g. H, T, CNOT .

Pauli Z unitary upon the second (target) qubit if the first (control) qubit is in the state $|1\rangle$; $\text{CZ} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Z$.

A universal quantum gate set is a finite set of quantum gates $\mathcal{G}$ such that $\forall \epsilon > 0$ there is a gate sequence of elements of $\mathcal{G}$ that approximates (since the space is continuous) any n -qubit unitary $U \in SU(2^n)$ to an accuracy ϵ . The overhead required to ϵ -approximate a non-gate set element on a fixed number of qubits $U \in SU(d)$ using gate set elements in $SU(d)$ can be shown by the *Solovay-Kiteav theorem* (SK) [45] to be $\mathcal{O}(\log^c(1/\epsilon))$ with c some constant. This is an important result as otherwise any potential quantum advantage could be nullified by the inefficiency of decomposing gates.

Although almost any two qubit gate is universal [46], a conventional gateset that is often considered is the combination of H , T and CNOT . Via SK all universal gatesets can be reduced to each other with at most a polynomial overhead in any case. Once we fix a universal gateset $\mathcal{G}$ then we are interested in the set of circuits $\{C_n\}$ which contain a polynomial number of gates from $\mathcal{G}$.

Measurement

To read out, all qubits are simply measured in the computational basis. Although more complicated measurement schemes could be used, they can always be subsumed somehow into the circuit and ancilla qubits. The outcome of the measurement will therefore be a bitstring for which the quantum computer then either accepts or rejects.

See Fig. 2.4 for a schematic summary of the circuit model of quantum computation. With this model in place, we can now define the complexity class of problems that are efficiently solvable on a quantum computer

- BQP (Bounded-error Quantum Polynomial) is the class of languages for which there exists a family of polynomial-size quantum circuits $\{C_n\}$ and a string $x \in \{0, 1\}^n$ such that $\forall x \in L, \Pr[C_n(x) \text{ accepts}] \geq \frac{2}{3}$ whereas $\forall x \notin L, \Pr[C_n(x) \text{ accepts}] \leq \frac{1}{3}$.

Any proof that $\text{BQP} \neq \text{BPP}$ would then overthrow the ECT and prove the power of quantum computers exceeds that of a probabilistic Turing machine.

The first result to challenge the ECT was courtesy of Deutsch and Josza [47]. In their problem, we are given a black-box known as an oracle that implements some function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ and the task is to decide if it is balanced or constant. They showed an exponential separation between a deterministic quantum algorithm using this oracle and the corresponding deterministic classical algorithm ⁷. However, this problem turns out to be solvable in BPP when we allow a probabilistic classical algorithm. Inspired by this, Bernstein and Vazirini [48] and Simon [49] then found new algorithms which did show $\text{BQP} \neq \text{BPP}$ relative to an oracle with a quasipolynomial and exponential separation respectively. These problems both look quite contrived however, and rely on this black-box oracle model which cast doubts on their practicality.

In 1994, building upon these previous results, Peter Shor then described an algorithm which remains the best-known in quantum computation [50]. He showed that a quantum computer can be used to solve the factoring problem (and discrete logarithm) in polynomial time. As introduced previously, this problem is believed to be in NP, and has immediate real-world applications. Specifically, the algorithm can be used to find the prime factors a and b of a given large composite number $N = ab$ in $\mathcal{O}(\log N)^3$ time. The best known classical algorithm, the general number field sieve, runs in time roughly $\mathcal{O}(2^{(\log N)^{1/3}})$. A scalable implementation of Shor's algorithm would break much of the world's current public key encryption such as the RSA cryptosystem [51].

⁷This determinism means that the algorithm actually defines an oracle relative to which EQP (Exact Quantum Polynomial) is different to P.

The key quantum part of Shor's algorithm does not actually directly solve the factoring problem. Instead, it can be shown using number theory that with classical pre- and post-processing, the problem of factoring can be efficiently reduced to order finding. This is the problem in which we are given integers x and N ($x < N$), which are coprime (they share no common factors) and we wish to find the smallest integer r such that $x^r \bmod N = 1$.

The quantum algorithm for order finding uses two registers of qubits, a work register and a control register. The algorithm proceeds by a series of controlled unitary gates acting on the work register, conditionally controlled by the control register, followed by a quantum Fourier transform (QFT) and measurement of the control register. The QFT consists of a network of single and two-qubit gates which performs the quantum analogue of a fast Fourier transform, i.e. the mapping $|j\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp[i2\pi jk/N] |k\rangle$.

When considered in the language of Group theory, the order finding routine in Shor's algorithm is a special case of the more general hidden subgroup problem (HSP). By considering the HSP over different groups, similar efficient quantum algorithms have been found which break other cryptosystems such as those based on the discrete logarithm or elliptic-curve discrete logarithm problems (see for more details)

Since this breakthrough, many new quantum algorithms have been discovered (see the Quantum Algorithm Zoo [52]), though often these algorithms fall into a few categories. For instance, algorithms built around performing the quantum Fourier transform unitary such as Shor's algorithm and oracular algorithms such as Simon's and the unstructured search algorithm of Grover [53].

Perhaps the most important group of quantum algorithms for future applications are those which perform quantum simulations. Simulating the dynamics of many-body quantum systems is a computational bottleneck across many crucial areas of physics, chemistry and material science. Therefore, the ability to efficiently simulate these systems promises to contribute to a revolution in modern science. Following the spirit of Feynman's original proposal, a digital Hamiltonian simulation algorithm was found by Lloyd [54]. In general, simulating the evolution of an arbitrary Hamiltonian applied for a time t simply requires applying the quantum circuit that implements $U(t) = e^{-iHt/\hbar}$, however this will not be efficient unless it can be performed with a polynomial number of gates. Thankfully,

many-body quantum systems which occur in nature that we would wish to simulate can usually be decomposed into a sum of a polynomial number of Hamiltonians which act locally on just a few systems, $H = \sum_i H_i$. If all the H_i commute, then we can perform the evolution via

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = \prod_i e^{-iH_i t}|\psi(0)\rangle = \prod_i U_i(t)|\psi(0)\rangle \quad (2.60)$$

and every $U_i(t)$ can be efficiently implemented with our universal gateset. However, in general the local Hamiltonians will not commute. To get around this, the solution proposed by Lloyd is to use the Trotter formula

$$e^{-iHt} = \lim_{n \rightarrow \infty} \left(\prod_i e^{-iH_i t/n} \right)^n \quad (2.61)$$

so that evolution can be achieved by a series of repeated small time evolutions of each local Hamiltonian. Understanding the number of Trotter steps n that would be in practice required to achieve an accurate simulation of useful systems is the focus of ongoing research [55, 56] as well as new approaches that use a Taylor series truncation instead of the Trotter decomposition [57]. There is a subtlety involved in these algorithms, which is that although the simulation may offer an exponential speed up over current classical methods for performing the evolution, the result of the algorithm as described is not a bitstring, but the full state $|\psi(t)\rangle$. To make the algorithm useful in practice we then need to identify some physical property of $|\psi(t)\rangle$ that can be efficiently estimated from a polynomial number of measurements of the state.

2.4.4 Quantum error correction

Although we have seen that processing information that is encoded in pure quantum states can provide unique advantages, it is important to remember that this information has to live in the physical world. As such, it is crucial that the information is stable and robust to the natural complexity and uncontrollability of realistic environments.

Classical bits, which are either 0 or 1 are often naturally robust as they are encoded

in macrostates of physical systems (e.g. for a CMOS gate with 5V supply, a logical 0 is assigned to any signal with $< 1.5V$ and logical 1 is any signal $> 3.5V$ [58]) and if not it is straightforward to make this information robust via simple error correction schemes. If there is a probability p of a bit flip error, then as long as $p < 1/2$, by redundantly encoding the bit in many copies and taking a majority vote the errors can be straightforwardly suppressed.

To make quantum information robust, there are many new challenges to consider. Firstly, quantum states form a continuum and so one might be sceptical that the infinite set of possible error mechanisms can all be correctable. Secondly, there is a famous result in quantum information theory which states that quantum information cannot be perfectly cloned [59] (i.e. there is no unitary transformation which can always perform $U|\psi\rangle|0\rangle = |\psi\rangle^{\otimes 2}$ for an unknown state $|\psi\rangle$) as we did in the classical example. Finally, making measurements of quantum systems to identify if an error has occurred disturbs their evolution, potentially itself corrupting the information. For all these reasons it is remarkable that quantum error correction is even possible, though for the same reasons strategies to keep quantum information robust are necessarily more complex and resource heavy than their classical counterparts.

In order to understand quantum error correction, we need the concept of a *codespace* $\mathcal{C}$. This is a subspace of a larger Hilbert space $\mathcal{H}$ in which we can encode a ‘logical’ qubit $\alpha|\bar{0}\rangle + \beta|\bar{1}\rangle$ redundantly in a $\log_2(\dim(\mathcal{H}))$ qubit physical state. The simplest example is Shor’s repetition code [60], the quantum analogue of the classical strategy described above.

Using three physical qubits, the codespace is $\mathcal{C} = \text{span}(|000\rangle, |111\rangle)$ and $|\bar{0}\rangle = |000\rangle$ and $|\bar{1}\rangle = |111\rangle$. If this state is subject to a Pauli X error (i.e. a bitflip) on a physical qubit, by measuring the operators Z_1Z_2 and Z_2Z_3 , where Z_i denotes Pauli Z on qubit i , we can protect against these errors. This is because the measurements do not collapse the state, but any -1 outcome informs us that the state has left $\mathcal{C}$ and the specific outcome string tells us which qubit has been flipped and therefore requires the application of a correcting X gate.

A general condition for a codespace $\mathcal{C}$ to be correctable under some noise process

described by the CPTP map $\mathcal{E}$ is the PEEP condition [33]

$$PE_i^\dagger E_j P = \alpha_{ij} P \quad (2.62)$$

where P is the projector onto $\mathcal{C}$, $\{E_i\}$ are the Kraus operators of $\mathcal{E}$ and α is a Hermitian matrix. If this condition is satisfied then there exists a recovery map $\mathcal{R}$ such that for all states ρ in the codespace, $\mathcal{R}(\mathcal{E}(\rho)) \propto \rho$, i.e. the noise model is correctable on $\mathcal{C}$. A key insight is that if the PEEP condition is satisfied for a channel with $\{E_i\}$, it is also satisfied for any channel whose Kraus operators are linear combinations of $\{E_i\}$. This means that if the depolarising channel $\{\sqrt{1-p}\mathbb{1}, \sqrt{\frac{p}{3}}X, \sqrt{\frac{p}{3}}Y, \sqrt{\frac{p}{3}}Z\}$ is correctable on $\mathcal{C}$, then *any* single qubit channel is correctable on $\mathcal{C}$. For instance, if we promote the three qubit Shor code into a nine qubit version where $|\bar{0}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)^{\otimes 3}$ and $|\bar{1}\rangle = \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)^{\otimes 3}$ which also protects against phase flip (Z) errors, then this can protect against any single qubit error. Error correction codes can be classified by $[[n, k, d]]$ where n qubits are used to encode k logical qubits with a distance d (the minimum number of errors which can perform a logical error in the codespace) and so the Shor code is called a $[[9, 1, 3]]$ code.

2.4.5 Fault tolerance

Of course, simply protecting static quantum information from corruption is not enough. In order to perform *fault-tolerant* quantum computation we need to be able to prepare states, implement gates and perform measurements without errors building up in our computation or losing any potential speed up by resource overhead. Thankfully, in 1996 it was shown that fault-tolerant computation is indeed possible thanks to the existence of threshold theorems [45, 61, 62].

To sketch out the spirit of these results we can consider a hypothetical architecture where it is possible to prepare, transform and measure states with schemes which are able to correct an error on a single qubit (like the Shor code). If errors occur on the physical qubits with a probability p , then the probability that the encoding fails is the sum of all cases where two or more errors occur, which grows like $\mathcal{O}(cp^2)$. The question is then

whether this error rate can be further decreased with a reasonable resource overhead? One way to do so is via *code concatenation*, where new codes are constructed by iteratively performing the same encoding on each logical qubit of the previous code. If we concatenate n times, then the probability of the fault-tolerant component failing will become $c^{-1}(cp)^{2^n}$, so if $p < \frac{1}{c}$ then the error rate will be double-exponentially suppressed with n . Say that we want to be able to survive t timesteps with a probability of error ϵ , then we see that only $n = \log(\log(c\epsilon))$ levels of concatenation are needed and the size of the circuit required to implement a fault tolerant component given the original circuit depth is d is $d^n = \mathcal{O}(\text{poly}(\log(\frac{t}{c\epsilon})))$. We therefore see that if the probability of errors on physical qubits p is less than a ‘threshold’, $\frac{1}{c}$, then increasing the size of the code will continue to suppress errors in the computation with only a poly-logarithmic overhead. This is the fundamental result which makes scalable digital quantum computers a practical possibility in principle.

The challenge that remains is therefore to find the codes on the physical qubits which can protect against errors in state preparation, measurement and gates. In practice, the latter is often the most challenging since performing gates upon encoded information has the potential to spread errors to ‘clean’ qubits. A highly desirable property of a code is therefore that it admits *transversal* gates, that is gates which can be performed by operations acting on only a single qubit in each codeblock. Unfortunately, it was proven that no code can perform a full, universal gateset transversally [63]. Instead, more complicated schemes are required, such as performing gates by distilling large numbers of uncorrected, noisy ancilla states into a few pure copies [64].

Although achieving fault tolerance appears a daunting task, the error thresholds required continue to improve with the current best proven threshold for arbitrary noise 6.7×10^{-4} [65] and up to $\approx 1\%$ for a depolarising noise model with topological codes [66], which will be briefly discussed below.

2.4.6 Stabilizer formalism

To complete our discussion of quantum error correction and for its use later on, we now take a short digression to introduce the *stabilizer formalism*. The heart of this approach is

to describe a state not by its state vector, $|\psi\rangle$, but by a set of operators for which $|\psi\rangle$ is the +1 eigenvalue eigenstate. For instance, we can describe the state $|0\rangle$ by the operator Z and $|+\rangle$ by X . The operators which are considered will be members of the Pauli group. The Pauli group on a single qubit is given by $\mathcal{P}_1 := \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$ and on N qubits $\mathcal{P}_N := \mathcal{P}_1 \otimes \mathcal{P}_1 \dots \otimes \mathcal{P}_1$. A stabilizer group S is then a subgroup of the Pauli group which is said to ‘stabilize’ a subspace V_S if

$$V_S = \{|\psi\rangle : G_i|\psi\rangle = |\psi\rangle \ \forall G_i \in S\} \quad (2.63)$$

where the G_i are called the stabilizers of V_S . An n qubit stabilizer state is one which can be efficiently represented by a generating set of stabilizers $S = \langle G_1, \dots, G_n \rangle$. More generally, if $S = \langle G_1, \dots, G_k \rangle$ with $k < n$ then this defines a stabilized subspace (i.e. a codespace) with a dimension $\dim(V_S) = 2^{n-k}$. Using this subspace as a code, errors can be detected by measuring $G_1, \dots, G_k$, producing a string $\vec{s} \in \{1, -1\}^k$ of outcomes known as *syndrome* measurements where any collection of -1 outcomes indicate an error. The remaining $n-k$ stabilizers can then be used as logical operators on the encoded codewords.

A particularly elegant application of this approach is in topological codes, the current leading approach to implementing a fault-tolerant quantum computer. Consider a lattice of qubits which are stabilized by k -local *plaquette* operators defined at each lattice site. The key idea is that if there is a small stabilized subspace then information can be encoded in the topological features of the whole lattice such that a logical error can only occur if there is a long chain of errors across the lattice, whereas local clusters of errors will be detectable by the plaquette syndromes. A physical error threshold can then be identified, below which, increasing the size of the lattice will reduce the logical error rate. In order to promote topological codes from quantum memories to computational platforms also requires extra features which we will not describe, such as encoding qubits in lattice defects [66] or performing lattice surgery [67]. Due to its high threshold ($\sim 1\%$), a popular architecture is the *surface code*, a 2D array of qubits involving 4-local stabilizer generators. However, the scale of the task in achieving useful fault-tolerant computation when error rates are not far below threshold can be seen from estimates of

performing Shor’s algorithm to factor a 2000 bit number requiring 10^7 physical qubits⁸ [68]. Therefore, along with the experimental efforts in reducing per-component error rates, the development of better fault-tolerant codes and approaches with higher thresholds and lower resource overheads are an area of intensive current research.

Another important insight that is gained from the stabilizer formalism involves the efficient simulation of quantum circuits. The *Clifford group* of unitaries is the set of unitaries which map Pauli group elements to Pauli group elements,

$$\{U : UGU^\dagger \in \mathcal{P}_N \ \forall G \in \mathcal{P}_N\}, \quad (2.64)$$

and so can continue to be described within the stabilizer formalism. The Clifford group on n qubits can be generated by $H, \sqrt{Z}, \text{CNOT}$ gates [33] and so it includes many standard quantum circuits. Since an n qubit stabilizer state can be efficiently represented by its n stabilizer generators (rather than its 2^n amplitudes), and applying Clifford unitaries simply updates these generators, any quantum computation which begins in a stabilizer state, undergoes evolution through a Clifford circuit and Pauli observables are measured is efficiently simulatable. This is the content of the *Gottesman-Knill theorem* [69].

A surprising consequence of the Gottesman-Knill theorem is that even circuits which can generate highly entangled states remain classically tractable (as we will see in the next section). Quantum computational speed up is therefore a much more subtle resource than simply having lots of entanglement.

2.4.7 Measurement-based quantum computation

In the circuit model of quantum computation we begin with a product state and then apply a series of few-qubit gates which typically prepares a large entangled state which is then measured in the fixed computational basis.

In 2001, Raussendorf and Briegel introduced a completely new way of performing universal quantum computation [70]. Measurement-based quantum computation (MBQC)

⁸And in fact ten times as many just to perform the ‘magic’ state distillation that is required to implement a universal gateset

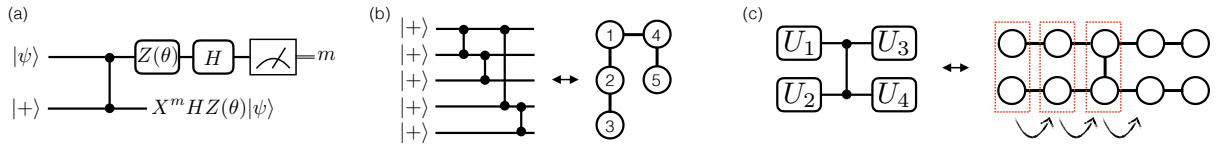


Figure 2.5: Measurement based quantum computation. a) Simple teleportation circuit demonstrating the principle of applying transformations via entanglement and measurement. b) The graph state representation. Each vertex represents a qubit in the $|+\rangle$ state and connected vertices are subject to a CZ gate. c) Converting between the circuit and measurement based models. Single and two-qubit gates are applied in the measurement based model via successive measurements upon a graph state. The outcomes of one layer of measurements inform the measurement basis choices in the next layer.

begins with the preparation of a fixed, highly entangled initial state and quantum circuits are executed by performing adaptive local measurements on this state. In this new paradigm, instead of creating entanglement, we consume it as a resource to move information forward in the computation.

At the heart of MBQC is the notion of quantum teleportation. Teleportation is a seminal result in quantum information which consists of the mapping of a quantum state from one Hilbert space to another via the use of entanglement [71]. This allows the transmission of quantum information without having to transmit physical systems themselves. The standard teleportation protocol consists of an unknown qubit and a Bell state $|\psi\rangle \otimes |\phi^+\rangle$. Performing a measurement in the Bell basis upon the first two qubits⁹ then projects the remaining qubit into the state $X^{s_1} Z^{s_2} |\psi\rangle$, where $s = \{s_1, s_2\}$ is a classical two bit representation of the outcome of the measurement. Therefore, if we begin with entanglement shared between two locations then, with the addition of a classical signal, quantum information can be communicated between them. Teleportation experiments have been performed between locations separated over 100km [72] and even between Earth and an orbiting satellite [73].

A simplified version of the teleportation scheme is shown in Fig. 2.5a where $Z(\theta) = \exp(-i\theta Z/2)$. Entangling a logical qubit with an ancilla and measuring it in a basis determined by θ teleports the logical state to the ancilla with a gate, again determined by θ , applied to it. To generalise this approach of implementing gates via entanglement

⁹One way of doing this is via the circuit consisting of a CNOT followed by H.

and teleportation we introduce a family of multi-qubit states known as *graph states*.

A graph state is defined by a graph $G = (V, E)$ with vertices V and edges E . Every vertex represents a single qubit prepared in the $|+\rangle$ state and every edge connecting two vertices represents a CZ gate applied between the corresponding qubits. Since these are all stabilizer operations, graph states have an efficient description in terms of the stabilizer generators $S_i = X_i \prod_{j \in N(i)} Z_j$, where $N(i)$ is the neighbourhood (i.e. all connected vertices) of i .

The *cluster state* is the graph state corresponding to a square lattice. It can be shown that this state is universal for quantum computation [74]. A horizontally connected linear graph can be considered as a single qubit channel with vertical connections between used to implement two qubit gates. In this way it is possible to map a measurement-based computation to a standard circuit based implementation. Measurement based computation is particularly attractive in implementations where performing large numbers of gates on the same physical systems is harder than producing many physical systems which are all measured after a fixed depth circuit. As we shall see, photonics is one such architecture. Finally, we note that the cluster state is by no means unique, various other states have been found to be universal for MBQC [75], though all share a similar locality and efficient description in terms of local operators. In fact, it has been found that most arbitrary pure states are too entangled to be measurement based resources [76, 77].

2.4.8 Other computational models

Here we briefly mention other models of quantum computation to give a fuller picture of the landscape of potential approaches

Adiabatic quantum computation

Yet another paradigm for implementing quantum computing that does away entirely with quantum gates is via the *adiabatic theorem* [78]. A computational problem is mapped such that its solution is encoded in the ground state of a Hamiltonian $\hat{H}$. A physical system is then prepared in the ground state of an easy-to-implement Hamiltonian $\hat{H}_{\text{in}}$. This

initial Hamiltonian is then tuned to $\hat{H}$ adiabatically, such that the system should remain in the ground state. The speed at which this transition can be performed is related to the spectral gap of the final Hamiltonian, the energy difference between the ground and first excited state. Perhaps surprisingly, this model can be shown to be polynomially equivalent to standard circuit based quantum computing [79].

DQC1

There are other known models of quantum computation that likely cannot exploit the full power of BQP. In 1998, motivated by the physics of current NMR experiments, Knill and Laflamme proposed the Deterministic quantum computation with one pure qubit (DQC1) or one clean qubit model [80]. In this model, just a single pure qubit is prepared in the $|+\rangle$, along with many copies of the maximally mixed state. Performing a controlled- U between the pure qubit and the mixed register, followed by measurement in the X and Y bases allows an estimate of the unnormalized trace of the matrix U . Although non-universal, DQC1 raises some fundamental questions. Since there is no entanglement generated in this process, this further questions the nature of quantum computational speedup. Some have attributed the power of this model to quantum discord, a different form of non-classical correlation [81].

Boson Sampling, IQP, Random circuit sampling

Recently, more non-universal computational models have been the subject of great interest in relation to the concept of ‘*quantum computational supremacy*’ [82]. Since the resource overheads for fault tolerant, universal quantum computing are so large, as sketched out above, there are few prospects for such a device performing supra-classical computation soon. This motivates non-universal models which are adapted to current experimental hardware so that a quantum speed up may be demonstrable in the near term, without full error correction overheads. Whilst likely to not be computationally useful and not a definitive disproof of the ECT, such experiments would be convincing evidence that it is possible to build machines which can exhibit controllable quantum coherence on

large enough scales to not be simulable on classical devices. Several such approaches have recently been proposed, all consisting of short quantum circuits for which classical sampling of the output distribution is likely to be inefficient. Instantaneous Quantum Computing (IQP) [83] circuits consist of a set of qubits prepared in $|+\rangle$, all subject to gates which are diagonal in the Z basis, then measurement in the X basis. Random circuit sampling [84] is an approach in which random single qubit unitaries from a universal gate set are interspersed with CZ gates and the output is sampled. Boson sampling [8] is an algorithm which is suited to linear optical experiments and will be discussed in greater length in the next section.

2.4.9 Physical platforms

After our journey through the mathematical and computer science foundations of quantum computing, we finally turn back toward physics. In order to encode and manipulate quantum information we need physical quantum systems. In 1996 DiVincenzo formulated his famous criteria which must be met by physical implementations of quantum computing. These are [85] (note that these criteria were formulated specifically for circuit model computation and more general, architecture independent, formulations have since been presented [86]):

1. A scalable physical system with well characterized qubits.
2. The ability to initialize the state of the qubits to a simple fiducial state, such as $|000\dots\rangle$.
3. Long relevant decoherence times, much longer than the gate operation time.
4. A universal set of quantum gates.
5. A qubit-specific measurement capability.

These all seem necessary requirements however there is an inherent friction at work here. The ability to maintain long coherence times (3) implies that the system is very well isolated from its environment, yet we need to be able to precisely interact with the system

to prepare, manipulate and measure it (2,4,5). The simultaneous suppression of noise and increase in control is the key challenge in the quest to realise useful quantum information processors. One other thing to note is the inclusion of scalability (1). It is often claimed that like the first classical computers the first quantum computers will be very large, non-practical machines in laboratories. Although this is likely to be true for the first non-universal machines which demonstrate quantum advantages, if it is true that millions or more physical qubits will be needed to perform useful, faster-than-classical computation, then scalability, miniaturisation and manufacturability are likely to be hugely important properties of the eventual preferred platform.

NMR

Many of the early demonstrations of experimental quantum information protocols were performed with Nuclear Magnetic Resonance (NMR) systems. NMR uses the Zeeman levels of spin-1/2 nuclei within a molecule in a magnetic field as its qubits. In contrast to other platforms, NMR uses large ensembles of molecules to prepare ‘pseudo-pure’ states e.g. of the form $\epsilon|00\rangle\langle 00| + (1 - \epsilon)\mathbb{1}/4$ where ϵ is the polarization of the whole sample. Gates and readout can then be performed using standard NMR techniques. There has been much discussion on the legitimacy of this approach, due to the seeming lack of entanglement [87] (this is what motivated the DQC1 model described previously) and the lack of large-scale scalability. Either way, NMR computing has provided many insights and tricks which have been subsequently transferred to other architectures.

Ion Traps

Ion traps remain perhaps the most sophisticated platform for quantum computing experiments. In the standard Paul trap setup, an AC RF field and fixed electrostatic potentials are used to confine single ions and laser cooling techniques are used to cool them to the quantum regime. An ion qubit is then realised by its electronic energy levels (either ground and excited or both ground hyperfine states) and can be manipulated by laser control fields. Two-qubit gates are mediated via the quantised collective vibrations of

multiple ions and readout is achieved via detecting fluorescence from a separate transition. Ion trap experiments have demonstrated the highest reported fidelities [88] and the maximum number of entangled qubits [89].

A major challenge for ion trap experiments is how to scale up beyond a single linear trap. Two approaches include using photonic links [90] and shuttling ions around in a 2D array [91].

Superconducting Qubits

Superconducting qubit systems are a promising platform for large-scale quantum information processing. Using integrated circuits with superconducting Josephson Junctions cooled down to tens of millikelvins, the collective electronic degrees of freedom (such as current and voltage) enter the quantum regime. Two level systems can be realised in a number of different configurations based around control of the quantised charge and phase in such circuits and their states are readout via microwave cavity resonators.

Although initially blighted by low coherence times, superconducting qubits have recently been demonstrated to perform high fidelity one and two-qubit gates [92]. The current engineering challenges involve the realisation of a two-dimensional, scalable architecture and all the control and readout circuitry that this entails.

Silicon Quantum Dots

Another architecture which has been making great progress in recent years is based on the spins of single electrons in quantum dots fabricated in silicon. Following the previously stated vision that useful quantum technology will need to be immediately scalable, such devices can be fabricated via standard CMOS techniques [93]. Recently, high fidelity qubits have been demonstrated and the first two-qubit gates achieved [94].

Linear optics

Finally we turn to photonics, the platform which will be investigated in this thesis. Encoding quantum information in the degrees of freedom of optical photons rather than

matter qubits has the advantage that they do not generally interact with their thermal environment and are therefore less susceptible to decoherence. The drawback is that photons also do not interact strongly with each other, and therefore two-qubit gates using linear optics must be implemented indirectly via measurement and feedforward and are inherently probabilistic. With the development of integrated quantum photonic devices, this technology platform is also moving towards scalability and miniaturisation.

The physics of Linear Optical Quantum Computing (LOQC) will be fully described in the following chapter, beginning with a description of quantum optical states and their interference properties.

Chapter 3

Quantum light

3.1 Introduction

In this chapter we move from the language of quantum information and computational complexity to the physical platform that this thesis will investigate, quantum photonics. We begin with a description of the quantum theory of light and its application in linear optical quantum computational schemes. We finish with a review of current experimental progress in quantum photonics and describe the experimental setup used in the experiments described in Chapters 4 and 5.

3.2 Quantisation of the electromagnetic field

The classical electromagnetic field in a vacuum can be described fully by Maxwell's equations

$$\nabla \cdot \mathbf{E} = 0 \tag{3.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{3.4}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields respectively and μ_0 and ϵ_0 are the permeability and permittivity of free space (and $\mu_0\epsilon_0 = 1/c^2$, where c is the speed of light in vacuum). These fields can be described in terms of a vector potential $\mathbf{A}$, and a scalar potential Φ , such that

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (3.5)$$

$$\mathbf{E} = -\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t}. \quad (3.6)$$

Working in the Lorenz gauge, where $\nabla\Phi = 0$, and the Coulomb gauge, where $\nabla\cdot\mathbf{A} = 0$, substitution into Eq. (3.4) results in the wave equation for $\mathbf{A}$,

$$\nabla^2\mathbf{A} = \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}. \quad (3.7)$$

This equation can be solved via separation of variables to produce the general solution

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}, s} \left(a(\mathbf{k}, s) e^{-i\omega_{\mathbf{k}}t} u_{\mathbf{k}}(\mathbf{r}) \boldsymbol{\epsilon}(\mathbf{k}, s) + a^*(\mathbf{k}, s) e^{i\omega_{\mathbf{k}}t} u_{\mathbf{k}}^*(\mathbf{r}) \boldsymbol{\epsilon}^*(\mathbf{k}, s) \right) \quad (3.8)$$

such that

$$\mathbf{E}(\mathbf{r}, t) = i \sum_{\mathbf{k}, s} \omega_{\mathbf{k}} \left[a(\mathbf{k}, s) e^{-i\omega_{\mathbf{k}}t} u_{\mathbf{k}}(\mathbf{r}) \boldsymbol{\epsilon}(\mathbf{k}, s) - a^*(\mathbf{k}, s) e^{i\omega_{\mathbf{k}}t} u_{\mathbf{k}}^*(\mathbf{r}) \boldsymbol{\epsilon}^*(\mathbf{k}, s) \right] \quad (3.9)$$

where $\{u_{\mathbf{k}}(\mathbf{r})\}$ is an orthonormal set of mode functions labelled by $\mathbf{k}$ which will depend upon the boundary conditions used, $a(\mathbf{k}, s)$ are the complex field amplitudes and $\boldsymbol{\epsilon}(\mathbf{k}, s)$ is a complex polarisation vector where s labels the two independent polarisations.

Using the change of variables

$$\begin{aligned} a(\mathbf{k}, s) &= \sqrt{\frac{1}{4\epsilon_0\omega_{\mathbf{k}}}} (\omega_{\mathbf{k}}x(\mathbf{k}, s) + ip(\mathbf{k}, s)) \\ a^*(\mathbf{k}, s) &= \sqrt{\frac{1}{4\epsilon_0\omega_{\mathbf{k}}}} (\omega_{\mathbf{k}}x(\mathbf{k}, s) - ip(\mathbf{k}, s)) \end{aligned} \quad (3.10)$$

it can then be shown that the classical Hamiltonian

$$H = \frac{1}{2} \int d^3r \left(\epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{E}(\mathbf{r}, t) + \frac{1}{\mu_0} \mathbf{B}(\mathbf{r}, t) \cdot \mathbf{B}(\mathbf{r}, t) \right), \quad (3.11)$$

where the integral is over all space defined by the boundary conditions, can be reduced to the simple form

$$H = \frac{1}{2} \sum_{\mathbf{k}, s} (p^2(\mathbf{k}, s) + \omega_{\mathbf{k}}^2 x^2(\mathbf{k}, s)). \quad (3.12)$$

We therefore see that the Hamiltonian is a sum of independent simple harmonic oscillators for each mode.

A formal quantisation of the field can be found in e.g. [95, 96], here we will simply apply the standard quantisation rules by replacing the position and momentum¹ by the quantum operators $\hat{x}(\mathbf{k}, s)$ and $\hat{p}(\mathbf{k}, s)$ obeying the commutation relations

$$[\hat{x}(\mathbf{k}, s), \hat{x}(\mathbf{k}', s')] = [\hat{p}(\mathbf{k}, s), \hat{p}(\mathbf{k}', s')] = 0 \quad (3.13)$$

$$[\hat{x}(\mathbf{k}, s), \hat{p}(\mathbf{k}', s')] = i\hbar \delta_{\mathbf{k}\mathbf{k}'} \delta_{ss'}. \quad (3.14)$$

The quantum Hamiltonian $\hat{H}$ can then be diagonalised by returning to the now quantised form of Eq. (3.10) (for a single mode)

$$\hat{a} = \sqrt{\frac{1}{2\hbar\omega}} (\omega \hat{x} + i\hat{p}) \quad (3.15)$$

$$\hat{a}^\dagger = \sqrt{\frac{1}{2\hbar\omega}} (\omega \hat{x} - i\hat{p}) \quad (3.16)$$

which satisfy

$$[\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger] = 0, \quad [\hat{a}, \hat{a}^\dagger] = 1. \quad (3.17)$$

Applying this transformation results in the Hamiltonian operator describing a sum of

¹Note these do not correspond to physical position or momentum but define an abstract phase space for describing the state of the optical mode.

independent quantum harmonic oscillators,

$$\hat{H} = \sum_{\mathbf{k},s} \hbar\omega_{\mathbf{k}} \left[\hat{a}^\dagger(\mathbf{k},s) \hat{a}(\mathbf{k},s) + \frac{1}{2} \right]. \quad (3.18)$$

Since the sum $\sum_{\mathbf{k},s} \hbar\omega_{\mathbf{k}}/2$ just sets the zero-point energy, and so just appears as a constant offset in all expectation values, conventionally it is omitted

$$\hat{H} = \sum_{\mathbf{k},s} \hbar\omega_{\mathbf{k}} a^\dagger(\mathbf{k},s) a(\mathbf{k},s) \quad (3.19)$$

and from now on we will drop the hats on a operators for convenience. The Hilbert space of a quantum harmonic oscillator is an infinite dimensional space whose eigenstates are the *Fock states*, $\{|n\rangle\}$, which correspond to states of a fixed number, n , of quantised excitations. In the case of the electromagnetic field, these excitations are known as photons. The operator $\hat{n} = a^\dagger a$ is known as the number operator since its expectation counts the number of photons in the oscillator's state

$$a^\dagger a |n\rangle = n |n\rangle \quad (3.20)$$

and $a^\dagger$ and a are known as *creation* and *annihilation* operators since they add or subtract excitations from the oscillator's state

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad (3.21)$$

$$a |n\rangle = \sqrt{n} |n-1\rangle. \quad (3.22)$$

A Fock state of n photons can therefore be written as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |\mathbf{o}\rangle \quad (3.23)$$

where we will use $|\mathbf{o}\rangle$ to denote the vacuum state, i.e. $a|\mathbf{o}\rangle = 0$.

A basis state for a general multimode quantum optical state is therefore a tensor

product of Fock states for each mode

$$\bigotimes_{\mathbf{k},s} |n_{\mathbf{k},s}\rangle \quad (3.24)$$

A fundamental difference between classical and quantum particles is the notion of indistinguishability. In principle, a set of identical classical particles can always be individually labelled such that their trajectories can be described independently. This is not the case for quantum particles. If the particles are identical bosons e.g. photons, swapping two particles does not change the global quantum state. Therefore, these particles can not be individually labelled, and we say they are *indistinguishable*.

The field quantisation formalism we have considered here has this indistinguishability implicitly defined via the use of creation and annihilation operators on Fock space. However, we can also describe the state of multiple bosons in a basis of the tensor product of the single-particle states $\{|\psi_{\mathbf{k}s}\rangle\}$ of the modes. In order to satisfy the indistinguishability condition, the full wavefunction in this picture must be symmetric under exchange of particles, i.e.

$$\prod_{\mathbf{k},s} a^\dagger(\mathbf{k},s)|\mathbf{o}\rangle = \bigotimes_{\mathbf{k},s} |1\rangle_{\mathbf{k},s} \equiv \frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} \sigma \left[\bigotimes_{\mathbf{k},s} |\psi_{\mathbf{k}s}\rangle \right] \quad (3.25)$$

where σ is the operator which permutes the systems and S_n is the symmetric group (whose elements are all the permutations of $(1,2,\dots,n)$). In contrast, for a system of identical fermions, swapping particles results in a π phase shift to the global wavefunction and therefore fermionic wavefunctions must be antisymmetric under exchange.

In what follows, we will often consider a series of separated but identical locations for which we can assign the same Hamiltonian of the form Eq. (3.19) to each location (e.g. a waveguide within an integrated photonic device). We will label these the paths of the system. The operator which creates a photon in the i th path, in a state defined by the superposition of complex amplitudes $\phi(\mathbf{k},s)$ in each mode will then be denoted as

$$a_i^\dagger[\phi] \equiv \sum_{\mathbf{k},s} \phi(\mathbf{k},s) a_i^\dagger(\mathbf{k},s) \quad (3.26)$$

with $[a(\mathbf{k}, s), a^\dagger(\mathbf{k}', s')] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{ss'}$. When the set of allowed modes is continuous, $\sum_{\mathbf{k}} \rightarrow \int d\mathbf{k}$ and $\delta_{\mathbf{k}\mathbf{k}'} \rightarrow \delta^n(\mathbf{k} - \mathbf{k}')$ becomes a Dirac delta function. If it is assumed that the function ϕ is the same for all paths, then simply $a_i^\dagger$ will be used.

3.3 Linear optics

We now introduce the transformations that will largely be the focus of this thesis. Linear optical transformations are those whose Hamiltonians describe a simple linear coupling of optical modes, i.e. for a system of n paths as described in the previous section

$$\hat{H} = \sum_{i,j=1}^n H_{ij} a_i^\dagger a_j \quad (3.27)$$

where H is an $n \times n$ hermitian matrix.

The most standard linear optical component is a beamsplitter. Considering two spatial paths with associated creation operators $a_1^\dagger$ and $a_2^\dagger$, a beamsplitter between them can be described by the Hamiltonian

$$\hat{H} = a_1^\dagger a_2 + a_2^\dagger a_1. \quad (3.28)$$

This Hamiltonian generates a unitary operator $\mathcal{U} = e^{-i\theta\hat{H}}$, where θ determines the reflectivity of the beamsplitter. To show how this operator acts on photonic states we need the Baker-Hausdorff lemma [97]

$$e^{-i\theta\hat{A}}\hat{B}e^{i\theta\hat{A}} = \hat{B} - i\theta[\hat{A}, \hat{B}] - \frac{\theta^2}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + i\frac{\theta^3}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots \quad (3.29)$$

Since $[\hat{H}, a_1^\dagger] = a_2^\dagger$ and $[\hat{H}, a_2^\dagger] = a_1^\dagger$ we can use Eq. (3.29) to show that

$$\mathcal{U}a_1^\dagger\mathcal{U}^\dagger = e^{-i\theta\hat{H}}a_1^\dagger e^{i\theta\hat{H}} = a_1^\dagger \cos \theta - a_2^\dagger i \sin \theta \quad (3.30)$$

$$\mathcal{U}a_2^\dagger\mathcal{U}^\dagger = e^{-i\theta\hat{H}}a_2^\dagger e^{i\theta\hat{H}} = -a_1^\dagger i \sin \theta + a_2^\dagger \cos \theta. \quad (3.31)$$

We can summarise these relations as

$$\mathcal{U} \begin{pmatrix} a_1^\dagger \\ a_2^\dagger \end{pmatrix} \mathcal{U}^\dagger = \begin{pmatrix} \cos \theta & -i \sin \theta \\ -i \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_1^\dagger \\ a_2^\dagger \end{pmatrix}. \quad (3.32)$$

When $\theta = \pi/4$ this transformation describes a 50:50 beamsplitter operation.

This approach can be generalised to a general m -mode linear optical transformation of the form of Eq. (3.27), resulting in the identity

$$\mathcal{U} \begin{pmatrix} a_1^\dagger \\ a_2^\dagger \\ \vdots \\ a_m^\dagger \end{pmatrix} \mathcal{U}^\dagger = T^\top \begin{pmatrix} a_1^\dagger \\ a_2^\dagger \\ \vdots \\ a_m^\dagger \end{pmatrix} \quad (3.33)$$

Where $T = e^{-i\theta H}$ is known as the *transfer matrix* and is the same matrix which describes the transformation of classical electric fields through the component. Using this identity, we can describe the Schrödinger evolution of an arbitrary m -mode p -photon state $|\psi\rangle$ under a linear optical transfer matrix T by

$$|\psi(T)\rangle = e^{-i\theta \hat{H}} |\psi\rangle \quad (3.34)$$

$$\begin{aligned} &= \mathcal{U}(T) \sum_{\vec{n} \in N_p} \alpha_{\vec{n}} \prod_{i=1}^m \frac{[a_i^\dagger]^{n_i}}{\sqrt{n_i!}} |\mathbf{o}\rangle \\ &= \sum_{\vec{n}} \alpha_{\vec{n}} \prod_i \frac{[\mathcal{U}(T) a_i^\dagger \mathcal{U}^\dagger(T)]^{n_i}}{\sqrt{n_i!}} \mathcal{U}(T) |\mathbf{o}\rangle \\ &= \sum_{\vec{n}} \alpha_{\vec{n}} \prod_i \frac{[\sum_j T_{ji} a_j^\dagger]^{n_i}}{\sqrt{n_i!}} |\mathbf{o}\rangle \end{aligned} \quad (3.35)$$

where N_p is the set of all m element tuples $(n_1, \dots, n_m)$ which sum to p and the third line follows from p uses of $\mathbb{1} = \mathcal{U}^\dagger \mathcal{U}$. Therefore, the most concise description of a linear optical transformation is via the update rules for the creation and annihilation operators

$$a_i^\dagger \rightarrow \sum_j T_{ji} a_j^\dagger, \quad a_i \rightarrow \sum_j T_{ji}^* a_j. \quad (3.36)$$

3.4 Quantum Interference

3.4.1 The Hong-Ou-Mandel effect

Although the quantum operators evolve in accordance with classical fields as shown above, due to the notion of indistinguishable particles explained earlier, single photons can demonstrate uniquely quantum behaviour, even when there are no interactions between the photons themselves.

The most famous example of such an effect is the quantum interference demonstrated in the Hong-Ou-Mandel (HOM) effect [98]. Here, a single photon is incident upon each input port of a 50:50 beamsplitter. Classically, one would expect that half the time the photons will be detected at different output ports due to the two possibilities of this event occurring (both reflect or both transmit). However since these two possibilities are indistinguishable, their probability amplitudes destructively interfere, resulting in no coincidental detections at all. We can see this using the transfer matrix for a beamsplitter²

$$\text{BS} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \quad (3.37)$$

and the update rule for creation operators,

$$\begin{aligned} & \mathcal{U}(\text{BS})|11\rangle \\ &= \mathcal{U}(\text{BS})a_1^\dagger a_2^\dagger |\mathbf{o}\rangle \\ &= \mathcal{U}(\text{BS})a_1^\dagger \mathcal{U}^\dagger(\text{BS})\mathcal{U}(\text{BS})a_2^\dagger \mathcal{U}^\dagger(\text{BS})\mathcal{U}(\text{BS})|\mathbf{o}\rangle \\ &= \frac{1}{2}(a_1^\dagger + ia_2^\dagger)(ia_1^\dagger + a_2^\dagger)|\mathbf{o}\rangle \\ &= \frac{1}{2}i[(a_1^\dagger)^2 + (a_2^\dagger)^2]|\mathbf{o}\rangle \\ &= \frac{1}{\sqrt{2}}(|20\rangle + |02\rangle) \end{aligned} \quad (3.38)$$

²Note that physical beamsplitters are often instead described by the transformations $a_1^\dagger \rightarrow (a_1^\dagger + a_2^\dagger)/\sqrt{2}$ and $a_2^\dagger \rightarrow (a_1^\dagger - a_2^\dagger)/\sqrt{2}$. This asymmetry is due to the dielectric coating used on one side of the beamsplitter. Here we will be considering a symmetric beamsplitter transfer matrix which represents for example, a directional coupler in integrated optics.

where we have ignored the global phase in the last line. This interference between indistinguishable single photons is what we will describe as *quantum interference*. The phenomenon of the photons always exciting the same port of a balanced beamsplitter is often referred to as *photonic bunching*.

We can also understand the HOM effect in the single particle representation described by Eq. (3.25). Here, the effect appears much less mysterious, since we see explicitly that the particles do not interact, but their correlation is related to the initial symmetrisation of the wavefunction [99]

$$\mathcal{U}(\text{BS}) \frac{1}{\sqrt{2}} [|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle] \quad (3.39)$$

$$= (\text{BS} \otimes \text{BS}) \frac{1}{\sqrt{2}} [|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle] \quad (3.40)$$

$$= \frac{1}{\sqrt{2}} [|\psi_1\rangle|\psi_1\rangle + |\psi_2\rangle|\psi_2\rangle] \quad (3.41)$$

where $|\psi_i\rangle$ is the single photon wavefunction in path i .

Finally, we can also think of quantum interference in terms of a Feynman path integral-like approach. Here we consider the trajectories of distinguishable particles undergoing evolution under a transfer matrix U . The amplitudes of trajectories which are the same up to the labelling of the particles are then summed, resulting in the quantum amplitude for that transition. For the HOM setup, the amplitude for the anti-bunched term can therefore be expressed as the sum of the amplitudes of the trajectory in which the ordering of the photons is preserved and the one for which they are swapped

$$\langle 11 | \mathcal{U}(U) | 11 \rangle = U_{11}U_{22} + U_{12}U_{21} \quad (3.42)$$

$$= \text{Per}(U) \quad (3.43)$$

$$\xrightarrow{U=\text{BS}} = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}} \times \frac{i}{\sqrt{2}} \quad (3.44)$$

$$= 0 \quad (3.45)$$

where the *permanent* of an $n \times n$ matrix A is defined as

$$\text{Per}(A) = \sum_{\sigma \in S_n} \prod_i A_{i, \sigma(i)} \quad (3.46)$$

and is equal to the more common matrix determinant without the alternating sign. This approach generalises to the result [8, 100]

$$\langle S | \mathcal{U}(U) | T \rangle = \frac{\text{Per}(U_{S,T})}{\prod_i \sqrt{s_i!} \sqrt{t_i!}} \quad (3.47)$$

where $|T\rangle = \bigotimes_i |t_i\rangle$, $|S\rangle = \bigotimes_i |s_i\rangle$ are the input and output Fock states respectively and $U_{S,T}$ is the submatrix of U formed by taking t_i copies of the i th column of U and s_i copies of the i th row of U . This picture again makes it clear that it is not really the photons themselves that interfere but rather global trajectories of the photons which cannot, in principle, be distinguished by any measurement.

3.4.2 The HOM dip and partial-distinguishability

In our discussion of quantum interference, we have been implicitly assuming that the photons are in the same optical modes relative to each path, i.e. $a_1^\dagger[\psi]$, $a_2^\dagger[\psi]$. In a practical scenario however, their internal states (spectral, polarisation etc.) will not be completely identical. If we assume that the beamsplitter performs the same transfer matrix on all internal states, we can analyse the HOM setup again, this time where the photons begin in arbitrary internal states

$$|\psi_{\text{HOM}}\rangle = \mathcal{U}(\text{BS}) a_1^\dagger[\psi_1] a_2^\dagger[\psi_2] |\mathbf{o}\rangle \quad (3.48)$$

$$= \frac{1}{2} \left[i(a_1^\dagger[\psi_1])^2 + i(a_2^\dagger[\psi_2])^2 + a_1^\dagger[\psi_1] a_2^\dagger[\psi_2] - a_2^\dagger[\psi_1] a_1^\dagger[\psi_2] \right]. \quad (3.49)$$

Clearly when $\psi_1 = \psi_2$ we see that the final two terms cancel and we obtain the standard bunching behaviour. However, if the functions ψ_1 and ψ_2 are not perfectly overlapped, $\langle \psi_1 | \psi_2 \rangle \neq 1$, then a non-zero probability of a coincidental detection will occur. Using a model in which single-photon detectors can measure the number of photons in each path

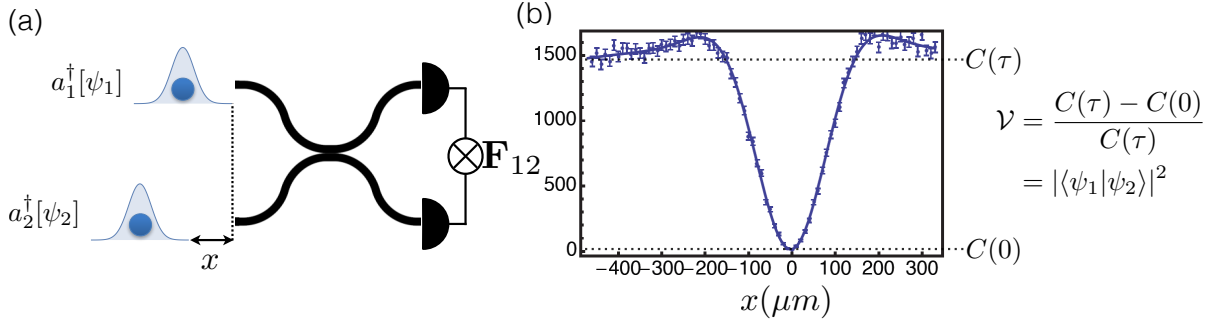


Figure 3.1: Hong-Ou-Mandel dip. a) Standard HOM setup where two photons with internal states $|\psi_1\rangle, |\psi_2\rangle$ are impinged upon a balanced beamsplitter and the coincidental countrate is measured, represented by the POVM element $\mathbf{F}_{12}$, as a function of the displacement of one photon relative to the other. b) Example experimental HOM dip. From this experiment, the visibility $\mathcal{V}$ of the two photons can be measured. Data courtesy of J. Carolan.

but cannot otherwise resolve the mode, the POVM corresponding to a joint detection is $\mathbf{F}_{12} = \mathbf{F}_1 \otimes \mathbf{F}_2$, where

$$\mathbf{F}_i = \sum_k a_i^\dagger[\xi_k] |\mathbf{o}\rangle \langle \mathbf{o}| a_i[\xi_k]. \quad (3.50)$$

and $\{\xi_k\}$ is a basis of functions on the internal state space. The probability of this outcome is then

$$\text{Pr}(11) = \langle \psi_{\text{HOM}} | \mathbf{F}_{12} | \psi_{\text{HOM}} \rangle = \frac{1 - |\langle \psi_1 | \psi_2 \rangle|^2}{2} \quad (3.51)$$

so we see that the effective indistinguishability of two photons is determined by the overlap $|\langle \psi_1 | \psi_2 \rangle|^2$ and can be measured by the probability of anti-bunching in the HOM setup.

In order to perform a typical HOM experiment, one photon is displaced relative to the other such that at long delays τ , the countrate of the joint detection is proportional to the probability of distinguishable particles, $C(\tau) \propto 1/2$. When instead the displacement is set to 0, $C(0) \propto \text{Pr}(11)$. As the relative delay is changed, a dip in countrate is traced out, known as the *HOM dip*, see Fig. 3.1. From these counrates, the HOM dip *visibility* can then be calculated

$$\mathcal{V} = \frac{C(\tau) - C(0)}{C(\tau)} = |\langle \psi_1 | \psi_2 \rangle|^2. \quad (3.52)$$

We can generalise this result to mixed internal states by considering ensembles of pure states. By introducing superoperators $A_j[\varrho]$ which create a photon in the j th path with

an internal state described by $\varrho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, the state in Eq. (3.49) becomes

$$\rho_{\text{HOM}} = \mathcal{U}(BS) \left(A_1^\dagger[\varrho_1] \circ A_2^\dagger[\varrho_2] \circ |\mathbf{o}\rangle\langle\mathbf{o}| \right) \mathcal{U}^\dagger(BS) \quad (3.53)$$

resulting in

$$\mathcal{V} = 1 - 2\text{Pr}(11) \quad (3.54)$$

$$= 1 - 2\text{Tr}[\mathbf{F}_{12}\rho_{\text{HOM}}] \quad (3.55)$$

$$= \text{Tr}[S(\varrho_1 \otimes \varrho_2)] \quad (3.56)$$

$$= \text{Tr}[\varrho_1 \varrho_2] \quad (3.57)$$

where S is the swap operator $S|\psi_i\rangle \otimes |\psi_j\rangle = |\psi_j\rangle \otimes |\psi_i\rangle$. It is interesting to note that for pure internal states, the notion of photon indistinguishability as measured by the HOM visibility is equivalent to the previously considered scenario of state discrimination in Sec. 2.3.5, where the distinguishability of states was determined by $|\langle\psi_1|\psi_2\rangle|^2$. However for mixed states there is no longer an equivalence, since two photons that are in identical - but mixed - states ϱ have a HOM visibility equal to their purity $\mathcal{V} = \text{Tr}[\varrho^2]$ whereas their trace distance is clearly zero. The notion of photon indistinguishability here is instead related to the symmetry of the internal states under exchange as can be explicitly seen in Eq. (3.56). In fact Eq. (3.56) generalises to arbitrary joint internal states, $\mathcal{V} = \text{Tr}[S\varrho_{12}]$, which means that for instance for a symmetric state of the form $|\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle)$, even if $\langle\psi_1|\psi_2\rangle = 0$, a perfect HOM dip will be observed.

3.5 Coherent states

Although in this thesis we will largely consider photonic states with a fixed number of photons such that the Hilbert space can be truncated, more general optical states live in the full infinite dimensional Fock space. An important class of optical states are the

coherent states. These states are the eigenstates of the annihilation operator

$$a|\alpha\rangle = \alpha|\alpha\rangle \quad (3.58)$$

and they can be described in Fock space as

$$\hat{D}(\alpha)|\mathfrak{o}\rangle \quad (3.59)$$

$$= \exp(\alpha a^\dagger - \alpha^* a)|\mathfrak{o}\rangle \quad (3.60)$$

$$= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (3.61)$$

where $\hat{D}(\alpha)$ is known as a displacement operator. Coherent states are considered the ‘most classical’ pure states as their expectation for the electric field is equivalent to the classical result with only the fluctuations of the vacuum. Coherent states are used to represent laser light³ and the evolution of a set of coherent states $|\boldsymbol{\alpha}\rangle = |(\alpha_1, \alpha_2, \dots, \alpha_m)^T\rangle \equiv \bigotimes_{i=1}^m |\alpha_i\rangle$ input to the m modes of a transfer matrix T evolve as

$$\mathcal{U}(T)|\boldsymbol{\alpha}\rangle = \bigotimes_{i=1}^m \left| \sum_j T_{ij} \alpha_j \right\rangle = |T\boldsymbol{\alpha}\rangle. \quad (3.62)$$

As such, coherent states do not demonstrate the quantum interference of the highly non-classical Fock states. It is possible however, to see an analogue of the HOM dip, with a maximum visibility of 0.5, by injecting two coherent states with a fast-varying relative phase between them upon a beamsplitter [102].

Coherent states are a special case of a more general class of states known as Gaussian states, which are states that can be described by Gaussian quasi-probability functions in optical phase space [103]. Although not explored in this thesis, the two-mode squeezed vacuum state we will describe later in this chapter is another example of a Gaussian state.

³Physically this is not likely to be true, since phase diffusion processes with optical frequency coherence times in the laser mean that coherence between the Fock states is quickly suppressed i.e. the state is better described by $\rho = e^{-|\alpha|^2} \sum_n \frac{|\alpha|^{2n}}{n!} |n\rangle\langle n|$ [101]. However, if all measurements on a laser beam are performed in relation to a reference beam from the same source, then they will produce the same result as $|\alpha\rangle$.

3.6 Linear optical quantum computing

Optical systems make natural quantum information carriers due to their accessibility and ease of manipulation in experiments. Photonic states tend to stay highly pure as they do not generally interact with their thermal environment, and can be manipulated via classical optical components. As such, photonic experiments do not require cooling to low temperatures and can be constructed and calibrated in the classical regime. In principle, a photonic qubit in an optical device can be defined using any two-level subspace of the full multi-mode Fock space. The most common encodings however are known as single-rail and dual-rail and are shown in Fig. 3.2. Single-rail qubits [104, 105] are defined by the vacuum and single photon states of a single optical mode,

$$|0\rangle_L \equiv |\mathbf{o}\rangle, \quad |1\rangle_L \equiv a^\dagger |\mathbf{o}\rangle. \quad (3.63)$$

This approach however suffers from the difficulty of preparing and measuring superpositions of the vacuum and single photon states. This is because the phase relationship between these levels $|\mathbf{o}\rangle + e^{i\omega t} |1\rangle$ oscillates rapidly and so coherence between these states cannot be practically measured without an additional phase reference, resulting in an effective photon super-selection rule [106]. A more common approach therefore is to use a dual-rail encoding [33], where the qubits are defined by a single photon across two modes (here we shall assume these are the paths considered earlier, though in general polarisation or more exotic modes could be used)

$$|0\rangle_L \equiv a_1^\dagger |\mathbf{o}\rangle = |10\rangle, \quad |1\rangle_L \equiv a_2^\dagger |\mathbf{o}\rangle = |01\rangle. \quad (3.64)$$

The advantage of this encoding is that universal transformations on a single dual-rail qubit are straightforward. A linear optical component described by the transfer matrix U acts on a dual-rail qubit $|\psi\rangle$ as $\mathcal{U}(U)|\psi\rangle = U|\psi\rangle$. It is possible to implement any $U \in SU(2)$ with just a variable-reflectivity beamsplitter and two phase shifters, which perform the evolution $e^{i\theta\hat{n}} a^\dagger e^{-i\theta\hat{n}} = e^{i\theta} a^\dagger$ on a single mode, via an Euler angle decomposition.

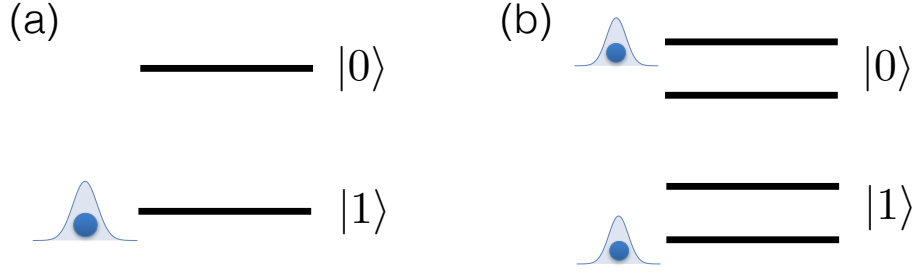


Figure 3.2: Single and dual-rail photonic qubits. a) Single-rail qubits are defined by the vacuum and single photon Fock state of a single optical mode. b) Dual-rail qubits are defined by a single photon spread across two optical modes.

Since a universal set of single qubit gates can be implemented straightforwardly with linear optical components, universality can be achieved via just the addition of a CZ gate. In the dual-rail picture, the only non-trivial evolution that is required is a π phase shift, conditional upon there being a photon in both $|1\rangle$ modes. Although not possible with just linear optical transformations, as explained in the proceeding section, this evolution can be achieved via a nonlinear optical process in a Kerr medium. A cross-Kerr Hamiltonian of the form

$$\hat{H}_{\text{kerr}} = \phi a_2^\dagger a_2 a_4^\dagger a_4 \quad (3.65)$$

results in the operator evolution

$$a_1^\dagger \rightarrow a_1^\dagger, \quad a_3^\dagger \rightarrow a_3^\dagger, \quad a_2^\dagger \rightarrow e^{i\phi a_4^\dagger a_4} a_2^\dagger, \quad a_4^\dagger \rightarrow e^{i\phi a_2^\dagger a_2} a_4^\dagger \quad (3.66)$$

and so the logical qubit states evolve as

$$a_1^\dagger a_3^\dagger |\mathbf{o}\rangle \rightarrow a_1^\dagger a_3^\dagger |\mathbf{o}\rangle, \quad a_1^\dagger a_4^\dagger |\mathbf{o}\rangle \rightarrow a_1^\dagger a_4^\dagger |\mathbf{o}\rangle, \quad a_2^\dagger a_3^\dagger |\mathbf{o}\rangle \rightarrow a_2^\dagger a_3^\dagger |\mathbf{o}\rangle, \quad a_2^\dagger a_4^\dagger |\mathbf{o}\rangle \rightarrow e^{i\phi} a_2^\dagger a_4^\dagger |\mathbf{o}\rangle. \quad (3.67)$$

When $\phi = \pi$, this performs the correct CZ gate [107, 108]. Although an attractive proposal, this approach to optical quantum computation has not seen significant experimental progress. This is because Cross-Kerr nonlinear coefficients are generally so weak that to produce a π phase shift from a single photon requires an impractical amount of material (for instance, it would require $\sim 10^5 \text{km}$ of optical fibre [109]). A more recent

approach is to use light-matter interaction, such as cavity-QED to achieve large nonlinearities though this results in a far higher experimental complexity. In addition, it has been argued that the above single-mode analysis fails to account for multimode effects that preclude high-fidelity gates when more fully investigated [110]. However, although still not demonstrated, recent experimental [111] and theoretical results [112] do suggest a more promising outlook for a nonlinear optical quantum computer.

3.6.1 KLM

For a long time it was believed that scalable universal quantum computation was not possible with just single photons, linear optics and photon detection. This is because without nonlinear optical transformations, there are no photon-photon interactions and thus useful entanglement cannot be generated. Consider, for instance, the evolution $|00\rangle \rightarrow (|00\rangle + |11\rangle)/\sqrt{2}$. In the dual-rail picture it is easy to see that no linear optical transformation can perform this transformation, since the Bell state cannot be factored into a product of linear combinations of creation operators [113]

$$\mathcal{U}(U)a_1^\dagger a_3^\dagger \mathcal{U}^\dagger(U) = \left(\sum_i U_{i1}a_i^\dagger\right)\left(\sum_j U_{j3}a_j^\dagger\right) \neq (a_1^\dagger a_3^\dagger + a_2^\dagger a_4^\dagger)/\sqrt{2}. \quad (3.68)$$

However, in 2000 Knill, Laflamme and Milburn (KLM) [114] showed the surprising result that photonic quantum computation was in fact possible without nonlinear optics. The key insight was that the combination of quantum interference and photon detection can be used to implement ‘measurement-induced’ nonlinearities which can be used to perform a probabilistic CZ gate. They then provided constructions which can convert these probabilistic gates into an asymptotically deterministic scheme with only a polynomial resource overhead, therefore proving that universal linear optical quantum computation (LOQC) was in fact possible using only single photon sources, linear optics (including quantum memories) and adaptive measurements (the ability to feedforward measurement outcomes to decide future measurements).

The building block of the KLM CZ is the nonlinear sign-shift (NS) gate. This gate performs a π phase shift upon the state in a single optical mode conditional upon it

containing two photons

$$U_{\text{NSS}}(c_0|0\rangle + c_1|1\rangle + c_2|2\rangle) = c_0|0\rangle + c_1|1\rangle - c_2|2\rangle. \quad (3.69)$$

The innovation was to use ancilla single photons and measurement such that if a desired measurement outcome is registered, then the correct operation of the gate is performed. This process of measuring additional photons to project the remaining photons into a known logical state is known as ‘heralding’. Using a single ancilla photon, two additional optical modes, the transfer matrix

$$T_{\text{NS}} = \begin{pmatrix} 1 - 2^{1/2} & 2^{-1/4} & (3/2^{1/2} - 2)^{1/2} \\ 2^{-1/4} & 1/2 & 1/2 - 1/2^{1/2} \\ (3/2^{1/2} - 2)^{1/2} & 1/2 - 1/2^{1/2} & 2^{1/2} - 1/2 \end{pmatrix} \quad (3.70)$$

and heralding upon detection of a photon in the second mode and vacuum in the third mode results in the correct application of the NSS gate,

$$\langle 10|\mathcal{U}(T_{\text{NS}})|\psi_{\text{in}}\rangle|10\rangle \propto U_{\text{NS}}|\psi_{\text{in}}\rangle \quad (3.71)$$

where states are represented in the Fock basis. The correct measurement pattern is obtained with a probability of $1/4$. Interestingly, this scheme has provably the highest probability of success for an NS gate even if arbitrary ancilla resources are available [115].

Using two NS gates in the configuration shown in Fig. 3.3 can then implement the required CZ gate. When there are 0 or 1 photons in modes 2 and 4, then the NS gates act as identity such that the beamsplitters (which are taken to be self-inverse) result in a trivial evolution. However, when the initial state is $a_2^\dagger a_4^\dagger$, a HOM bunching occurs after the beamsplitter such that both NS gates implement π phase shifts, and the overall state picks up the desired -1 factor.

The immediate problem with this gate is that it only succeeds with a probability of $(1/4)^2 = 1/16$, such that the probability of successfully applying n two-qubit gates is exponentially small in n . When the NS gates do not succeed, the logical information is

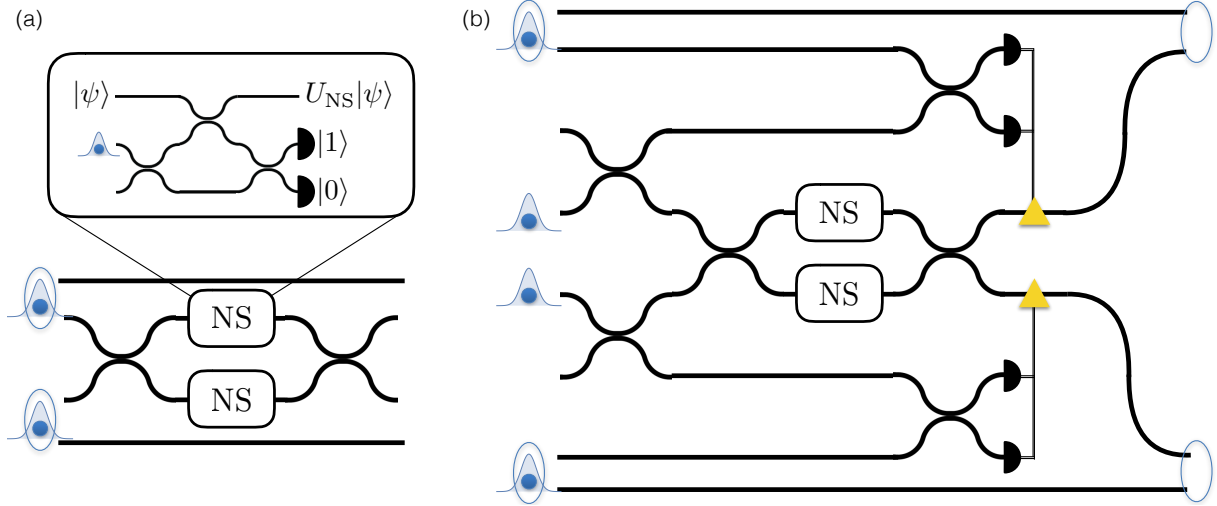


Figure 3.3: Elementary circuits in the KLM scheme. a) The nonlinear sign (NS) gate performs a π shift on the $|2\rangle$ state conditional upon the correct heralding measurement. Two of these gates can then be used to perform a probabilistic but heralded CZ gate between two dual-rail qubits. b) This gate can be performed via teleportation such that the logical qubits do not enter the gate unless the NS gates are successful. Conditional upon the detection outcomes, phase shifts on the output may need to be applied.

lost. The second important result shown in the KLM paper was that the probability of this gate can be boosted towards unity using quantum teleportation. The trick, as shown in Fig. 3.3(b), is that a CZ can be commuted through a teleportation circuit such that the probabilistic gates can be performed ‘offline’. This means the NS gates can be repeated until successful, and only then are the logical qubits teleported to the output.

Unfortunately, the Bell measurements required for teleportation are themselves probabilistic using only linear optics. KLM therefore gave a generalisation of the teleportation circuit which uses an increasing number of ancilla resources to boost the probability of success of the circuit towards one. In addition, they also provided error correction protocols which allowed information to be preserved even when the teleportation fails.

Unfortunately, although a striking proof that LOQC is possible, the resource overheads in the original KLM proposal are highly impractical. To implement a single two-qubit gate with a 95% success rate would require tens of thousands of optical components [116]. Since the original proposal, a series of improvements and adaptations to the scheme have been proposed [113], though all still suffer from a steep resource overhead for practical

implementations.

3.6.2 MBLOQC

Two developments which have pushed LOQC closer to a practical reality are the shift to measurement-based quantum computation and the development of ballistic schemes. Photonics is particularly well-suited to the MBQC model as the entangled resource state generation process can be performed using probabilistic gates offline, leaving the computation to proceed deterministically via adaptive single qubit measurements.

Using cluster state computation for linear optics was first suggested by Nielsen[116] whilst still using the existing KLM schemes, and later improved by Browne and Rudolph [117] by replacing the KLM gates with simpler gates referred to as *fusion* gates. The Type-I fusion gate, shown in Fig. 3.4(a), takes two photonic qubits as input and measures one, performing a logical projection which is unitarily equivalent to the form

$$F_I = \frac{1}{\sqrt{2}}(|0\rangle\langle 00| + |1\rangle\langle 11|) \quad (3.72)$$

with a probability of $1/2$. This transformation can be used to ‘fuse’ together graph states. Consider two graph states where one qubit from each 1 and $1'$, connected to qubits i and j respectively, are input to the gate. The initial states can be written as

$$|G\rangle = \frac{1}{\sqrt{2}}(|0\rangle|G_1\rangle + \prod_i Z_i|1\rangle|G_1\rangle), \quad |G'\rangle = \frac{1}{\sqrt{2}}(|0\rangle|G_{1'}\rangle + \prod_j Z_j|1\rangle|G_{1'}\rangle) \quad (3.73)$$

where $|G_k\rangle$ is the graph state remaining when qubit k is removed. Now, consider applying the fusion gate F_I

$$F_I|G\rangle|G'\rangle = \frac{1}{\sqrt{2}}(|0\rangle|G_1\rangle|G_{1'}\rangle + \prod_{i,j} Z_i Z_j|1\rangle|G_1\rangle|G_{1'}\rangle) \quad (3.74)$$

which is a new graph state joined at the location of one of fused qubits. If a different detection pattern is measured, then the resulting state will be the same graph state with

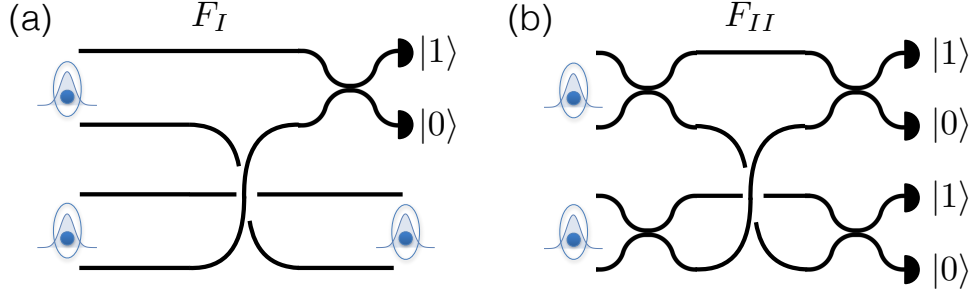


Figure 3.4: Type I and II Fusion gates. a) The type-I gate F_I takes two dual-rail photonic qubits and measures one, preserving the remaining qubit with a probability $1/2$. b) The type-II gate F_{II} takes two dual-rail qubits and measures both, projecting them onto a Bell state with a probability of $1/2$.

local Pauli rotations which can be corrected for.

The Type-II fusion gate, shown in Fig. 3.4(b) is unitarily equivalent to the linear optical Bell state analyser [118]. This gate takes two photonic qubits as input and measures both, projecting onto a logical bell state, e.g. $F_{II} = |\phi^+\rangle\langle\phi^+|$, with a probability of $1/2$. Again consider the fusion of two qubits 1 and 1' which are connected to qubits i and j respectively. The stabilizer generators associated with 1, 1', i and j are given by

$$S_1 = \prod_i X_1 Z_i, \quad S_{1'} = \prod_j X_{1'} Z_j, \quad S_i = X_i Z_1 A_i, \quad S_j = X_j Z_{1'} B_j \quad (3.75)$$

Where A_i and B_j represent Z on all other qubits connected to i and j . To fuse the clusters, we first perform a Hadamard rotation on qubit 1' which results in

$$S_{1'} = \prod_j Z_{1'} Z_j, \quad S_j = X_j X_{1'} B \quad (3.76)$$

Projecting this state onto the Bell basis of qubits 1 and 1' then leads to, for the case of a $|\phi^+\rangle$ outcome,

$$\bar{S}_1 = Z_1 Z_{1'}, \quad \bar{S}_{1'} = X_1 X_{1'}, \quad \bar{S}_i = \prod_j X_i Z_j A_i, \quad \bar{S}_j = \prod_i X_j Z_i B_j \quad (3.77)$$

forming a new graph state with qubits 1 and 1' measured out and bonds between all qubits i and j .

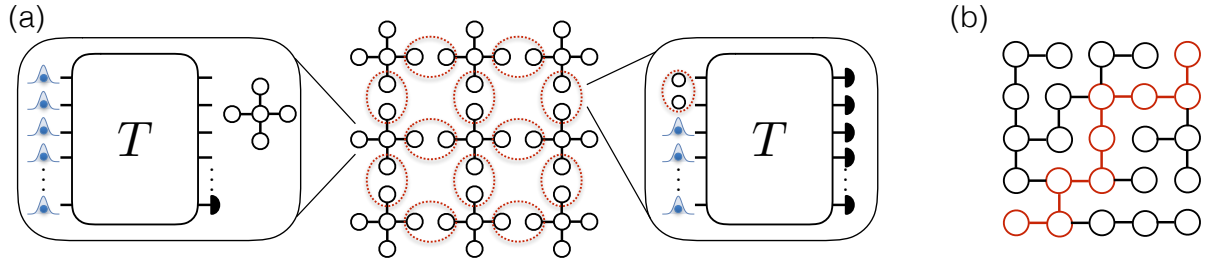


Figure 3.5: Schematic of ballistic photonic quantum computation. a) A supply of single photons is used to generate a supply of few-qubit entangled states. These entangled states are then fused together probabilistically. b) The resulting state is a cluster state with missing bonds. If the probability of success of the fusions is high enough, then there is guaranteed to be a spanning path with high probability.

Using a combination of these two fusion gates and a supply of dual-rail Bell states, a cluster state capable of universal computation can be constructed. The Bell states can be generated from four single photons with a probability of $3/16$. This circuit will be explored in detail in Chapter 6.

The second paradigmatic shift from the KLM approach was first suggested in [119], which is to use a ballistic instead of repeat-until-success strategy. Even though the entanglement can be generated offline in the MBQC model, when a probabilistic gate fails it will generally be necessary to re-try the gate, sometimes many times. This results in an architecture which requires sophisticated fast-switching networks and delay lines. In contrast, ballistic schemes use each gate just once and then use adaptive post-processing to achieve universal computation on the resulting non-ideal cluster states. This is achieved by using ideas from percolation theory. Once the probabilistic fusion gates have been implemented, the resulting global state is a graph state of a target lattice with a fraction of missing bonds (entanglement) and sites (qubits). If the probability of missing bonds is below a critical threshold, then with high probability there will be a spanning path through the lattice, providing a universal resource state. For a schematic of such schemes, see Fig. 3.5. One way to then perform MBQC is to ‘renormalise’ blocks of imperfect lattice into single qubits of a perfect lattice [119].

The best known current scheme for MBQC in a ballistic approach are the protocols from [120, 121], where ‘boosted’ fusion gates, which trade-off extra ancilla resources for

higher success probabilities, are used to ballistically generate graph states from a supply of three-photon entangled states. These schemes will be described in more detail in Chapter 6.

3.6.3 Boson sampling

Despite the huge effort and resources being poured into the global race to build quantum computers in the hope of solving real world problems, the more fundamental scientific challenge of demonstrating the inherent computational advantages of quantum systems over classical systems is yet to be convincingly shown. An obvious route to this goal would be to use the best known quantum algorithm, Shor’s algorithm, to factor a number that is not possible with current classical technology. However, this approach has two drawbacks. Firstly, experimental progress has been limited to a handful of qubits, whereas Shor’s algorithm requires hundreds of logical qubits (and therefore likely millions of physical qubits) to outperform current algorithms. Furthermore, it is yet to be conclusively proven that factoring is a classically hard problem and the existence of an efficient classical algorithm for factoring would not imply any dramatic theoretical consequences such as the collapse of complexity classes. A recent trend in quantum information science has therefore been to develop models of quantum computation which, although non-universal, perform classically hard tasks under strong complexity arguments and are tailored to current physical platforms in order to demonstrate a quantum advantage with near-term experimental resources [83, 84, 122]. The prototypical example of such an algorithm is Boson Sampling.

Boson Sampling is the following computational problem:

- Take as input an $m \times m$ unitary matrix $U \in SU(m)$. Consider the $m \times n$ matrix A formed from its first n columns (where we will take $m = n^2$).
- Sample from the distribution over tuples $S = (s_1, s_2, \dots, s_m)$, where $s_i \in \{0, 1\}$ and $\sum_i s_i = n$, with the probability mass function $\Pr(S) = |\text{Per}(A_S)|^2$, where A_S is the matrix formed by taking the rows of A for which $s_i = 1$.

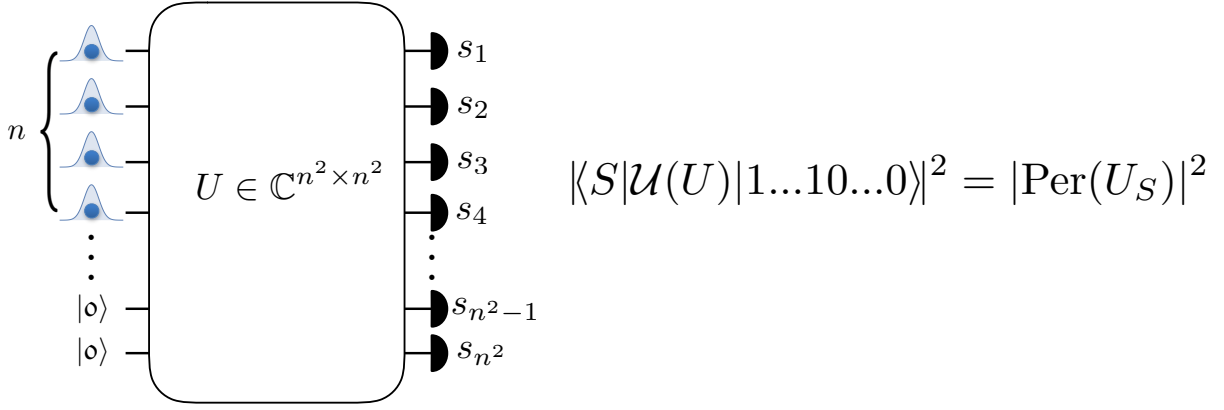


Figure 3.6: Boson Sampling schematic. n single photons are input into a linear optical network with transfer matrix U acting on n^2 modes. All collision-free detection patterns $S = (s_1, s_2, \dots, s_{n^2})$ are measured, sampling from a distribution whose probability mass function is related to the permanent of submatrices of U .

It can be straightforwardly seen, via Eq. (3.47), that this computational problem maps onto the following physical problem (see Fig. 3.6):

- Take as input an $m \times m$ unitary matrix $U \in SU(m)$. Construct a linear optical network which implements this transfer matrix.
- Sample from the distribution of n -fold photon detection patterns obtained at the output of the linear optical network if the input is prepared in the Fock state $|1_n\rangle = |1\rangle^{\otimes n} |0\rangle^{\otimes m-n}$.

In 2010, Aaronson and Arkhipov [8] showed that if there is an efficient classical algorithm for Boson sampling, then the polynomial hierarchy would collapse to its third level, a complexity implication thought highly unlikely. In contrast, the quantum experiment described above scales polynomially in experimental resources and avoids some of the impracticalities of LOQC, such as probabilistic gates, fast switching and feed-forward. However, any realistic, non-error corrected experiment will not solve the Boson sampling exactly. The crucial result proven in [8] was therefore that, modulo some highly plausible conjectures, this sampling problem is still classically hard for distributions $\mathcal{D}'$ that are merely close to the Boson sampling distribution $\mathcal{D}_{\text{BS}}$ in variation distance $\|\mathcal{D}' - \mathcal{D}_{\text{BS}}\| \leq \epsilon$. We now give a brief summary of this proof.

We first need to introduce some new computational complexity classes. $\#P$ is the class of function problems of the form “compute $f(x)$ ” where f gives the number of accepting paths of an NP machine. Note that this is not a decision problem as we were considering in Sec. 2.4.2 but is instead a counting problem. If an NP problem is of the form ‘are there any Hamiltonian paths with cost less than c ’ then the corresponding $\#P$ problem asks ‘how many Hamiltonian paths have cost less than c ’.

The *polynomial hierarchy*, PH, is the union of a set of complexity classes, the 0th level of which is P and the higher levels take the form $NP, NP^{NP}, NP^{NP^{NP}}, \dots$ ⁴ where NP^{NP} is the class of problems that can be solved by an NP machine given an NP oracle. Similar to $P \neq NP$, It is a fundamental assumption in complexity science that this full hierarchy is not contained within any specific level, i.e. it does not collapse to this level. These two complexity classes can be linked via Toda’s theorem [123] which shows that $PH \subseteq P^{\#P}$, i.e. that all problems in the entire polynomial hierarchy can be solved in polynomial time with access to a $\#P$ oracle.

The reason that we have introduced the $\#P$ class is that the matrix permanent is the prototypical problem in this class. In 1979, Valiant [124] proved that the problem of computing the permanent of a matrix of 0’s and 1’s is $\#P$ -complete. The fact therefore that the amplitudes of transitions of Fock states through linear optical networks are proportional to permanents (see Eq. (3.47)) implies a striking connection between complexity theory and the physics of fundamental particles. Interestingly, the same transmission expression for a system of non-interacting fermions is given by the determinant, a function that can be computed in polynomial time.

Unfortunately, although the evolution of photons in linear optical networks is governed by permanents, it is not possible to use this fact directly to compute a $\#P$ -hard problem using a photonic experiment. This can be understood by considering the problem of estimating the permanent of an arbitrary $n \times n$ matrix B . In order for B to be embedded into a unitary transfer matrix, we can place a rescaled version $B/||B||$, where $||B||$ is the largest singular value of B , in the top left corner of a $2n \times 2n$ matrix U . In Chapter 4 we

⁴PH also contains the similar set of classes with at the first level CoNP. CoNP is the complementary class to NP i.e. if an NP problem asks whether there exists a string satisfying certain properties then the problem in CoNP asks whether all strings satisfy certain properties.

will explore how this is done, and further applications of this concept. The probability of measuring the photons exciting in the first n modes is

$$|\langle 1_n | \mathcal{U}(U) | 1_n \rangle|^2 = \|B\|^{-2n} |\text{Per}(B)|^2. \quad (3.78)$$

This probability however, is exponentially small for $\|B\| > 1$, meaning that an exponential number of trials would be needed to accurately estimate it. On the other hand, when $\|B\| < 1$, there exists a classical algorithm, Gurvits' algorithm [125], for approximating the permanent to additive error $\pm\epsilon$ in time $\mathcal{O}(\frac{n^2}{\epsilon^2})$ which does just as well as the quantum experiment.

The insight of Aaronson and Arkhipov was to consider a more general scenario and define the Boson sampling problem as above to perform exactly the task that quantum experiments actually solve. Their argument then relies on two facts; firstly, that approximating $|\text{Per}(X)|^2$ to a *multiplicative* factor is still $\#P$ -hard. Secondly, there exists an algorithm by Stockmeyer [126] that says that a BPP^{NP} machine can always estimate the probability that a polynomial-time randomised algorithm accepts to within a multiplicative factor, even when it is exponentially small. Therefore, if a polynomial-time algorithm exists for exact Boson sampling, then this could be fed to Stockmeyer's algorithm to estimate a $\#P$ quantity in BPP^{NP} , resulting in $P^{\#P} \subseteq BPP^{\text{NP}}$. Then, via Toda's theorem and known containments, it can be shown that

$$P^{\#P} \subseteq BPP^{\text{NP}} \quad (3.79)$$

$$PH \subseteq BPP^{\text{NP}} \quad (3.80)$$

$$PH \subseteq \text{NP}^{\text{NP}^{\text{NP}}} \quad (3.81)$$

and so the polynomial hierarchy collapses to its third level.

Although a surprising complexity result this result is hard to connect to real experiments. This is because realistic experiments can only approximately sample from the exact Boson sampling distribution. Thankfully, Aaronson and Arkhipov were also able to show, modulo two conjectures, that the polynomial hierarchy also collapses if there is an

efficient algorithm for *approximate* Boson sampling.

Their proof works via the following steps. The first important observation is that the $n \times n$ submatrices of a Haar-random $m \times m$ unitary matrix, with $m \geq n^6$ (though it is expected to also be true for $m \geq n^2$) are approximately equal to matrices of i.i.d Gaussians. Therefore, they first consider the following problem

$$\begin{aligned} |\text{GPE}|_{\pm}^2 : & \text{Given a matrix } X \sim \mathcal{N}(0, 1)_{\mathbb{C}}^{n \times n} \text{ of i.i.d Gaussians, estimate } |\text{Per}(X)|^2 \text{ to} \\ & \text{within additive error } \pm \epsilon n! \text{ in } \text{poly}(n, 1/\epsilon) \text{ time.} \end{aligned} \quad (3.82)$$

and prove that if there exists a classical algorithm which can sample from a probability distribution $\mathcal{D}'$ such that $\|\mathcal{D}' - \mathcal{D}_{\text{BS}}\| \leq \epsilon$ in $\text{poly}(n, 1/\epsilon)$ time then $|\text{GPE}|_{\pm}^2$ is solvable in BPP^{NP} . The next step is then to show that if the Permanent Anti-Concentration Conjecture (PACC) holds, which roughly says that the distribution of Gaussian permanents is not highly concentrated around zero, then $|\text{GPE}|_{\pm}^2$ is poly-time equivalent to another problem $\text{GPE}_{\times}$

$$\begin{aligned} \text{GPE}_{\times} : & \text{Given a matrix } X \sim \mathcal{N}(0, 1)_{\mathbb{C}}^{n \times n} \text{ of i.i.d Gaussians, estimate } \text{Per}(X) \text{ to} \\ & \text{within multiplicative error } \pm \epsilon |\text{Per}(X)| \text{ in } \text{poly}(n, 1/\epsilon) \text{ time.} \end{aligned} \quad (3.83)$$

A final conjecture, the Permanent-of-Gaussians conjecture (PGC), then states that $\text{GPE}_{\times}$ is $\#P$ -hard. Then, via Toda's theorem and the same arguments as before we have shown that if an efficient classical algorithm for approximate Boson sampling exists, then the polynomial hierarchy again collapses to its third level.

Boson sampling is a particularly attractive approach for current quantum photonic technologies. It does not require either the measurement-induced nonlinearity or adaptive measurement schemes of KLM or MBQC, simply quantum interference between photons. Furthermore, it seems possible that Boson sampling experiments can be scaled without the full machinery of fault-tolerant error correction towards the quantum supremacy threshold.

A number of adaptations and follow-up results have been suggested since the original

proposal. These include using inherently probabilistic sources in a so-called ‘scattershot’ approach [127], rather than the $|1_n\rangle$ state, more generally using Gaussian states or measurements [128, 129], and allowing for a constant number of lost photons [130].

A number of proof of principle experimental demonstrations of Boson sampling have been performed [131–135], with the current record 5 photons in 9 modes [136]. The regime in which quantum experiments would outperform classical machines was initially predicted to be in the region 20-30 [8] and recently even as low as 7 [137].

However it was recently shown by co-authors and myself [138] that problem sizes of up to 50-60 can likely be solved by supercomputers with practical runtimes.

Metropolised independence sampling (MIS) is a Markov chain Monte Carlo method used to sample from distributions for which direct sampling is difficult. The algorithm works via constructing a Markov chain on the state space of possible outcomes where the probability of transitioning between a pattern S and S' is split into a proposal and acceptance stage. The proposal distribution $g(S)$ is chosen such that efficient sampling from this distribution is possible. A new state S' is then drawn from this proposal distribution, and the algorithm accepts this state with a probability given by

$$T(S'|S) = \min \left[1, \frac{f(S')}{f(S)} \frac{g(S)}{g(S')} \right] \quad (3.84)$$

where $f(S)$ is the distribution one wishes to sample from. Using the Boson Sampling distribution as $f(S)$, a good proposal $g(S)$ was found to be the equivalent distribution for classical distinguishable particles whose probability mass function is given by

$$\Pr(S) = \text{Per}(|A_S|^2). \quad (3.85)$$

If one can show that the resulting Markov chain has converged to its stationary distribution, and using a thinning procedure to eliminate autocorrelation between samples (since, for instance, if a proposal is not accepted the chain will remain in the same state for multiple timesteps), this algorithm can be compared to experimental approaches. Strong numerical evidence suggests that convergence and thinning requires only ~ 100 steps,

independent of n , and so can output a sample from the Boson Sampling distribution with a constant number of permanent evaluations. This is a vast improvement over the direct sampling approach, which would require the calculation of $\binom{n^2}{n}$ permanents.

30 photon Boson Sampling was shown to be possible in tens of minutes on a standard laptop and, assuming that the algorithm continues to perform equally well on larger instances, that ~ 50 photon experiments could be simulated on a supercomputer. This result was recently reinforced by Ref. [139], where the classical complexity of exact Boson sampling was shown to be $\mathcal{O}(n2^n + \text{poly}(m, n))$.

3.7 Quantum photonics

In this section we shall turn to the current experimental state-of-the-art in quantum photonics, along with a description of the experimental setup used in the experiments described in this thesis.

3.7.1 Photon sources

Although sources of laser light that can be described quantum optically have been widely available for many decades, quantum technologies require the generation of optical states which are ‘truly’ nonclassical such as single photons. Single photon sources (SPS) are perhaps the most pressing engineering challenge for current quantum photonic technologies. This is because they require nonlinear processes with trade-offs which are hard to simultaneously satisfy. We can list the most important properties of single photon sources as

- Source efficiency: the probability that a single photon is generated by our source on every attempt.
- Single photon purity: the purity of the full photonic state describing the single photon (e.g. its spectral purity).
- Multi-photon contamination: the probability of the state to be found in higher

photon number spaces.⁵

- Indistinguishability: how identical the single photons produced are, both from subsequent uses of the same source and between different sources.
- Repetition rate: the rate at which photon generation is attempted.

There are two general approaches to realising single photon sources, all-optical and matter-based. They each have their own strengths and weaknesses. All-optical sources tend to produce purer and more indistinguishable photons at the cost of a low efficiency due to their probabilistic nature and to keep multi-photon contamination minimised. On the other hand, matter-based sources can generally operate with higher efficiencies and do not suffer from multi-photon contamination to the same degree. However, producing pure and indistinguishable photons from multiple sources is generally a greater challenge. Both sources can operate at high repetition rates however, into the GHz regime, one of the reasons LOQC and photonic technologies are appealing for information processing.

The most common SPS which has been the workhorse of quantum optical experiments for several decades is based upon the principle of spontaneous parametric down-conversion (SPDC). SPDC is an optical process which occurs within nonlinear media. When light is propagating in a medium, the electromagnetic Hamiltonian from Eq. (3.11) has an additional term in terms of the macroscopic polarization $\mathbf{P}$

$$H = \frac{1}{2} \int d\mathbf{r}^3 \left(\epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{E}(\mathbf{r}, t) + \frac{1}{\mu_0} \mathbf{B}(\mathbf{r}, t) \cdot \mathbf{B}(\mathbf{r}, t) + \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{P}(\mathbf{r}, t) \right). \quad (3.86)$$

$\mathbf{P}$ can be expanded in a power series of the electric field amplitudes

$$\mathbf{P}(\mathbf{r}, t) = \mathbf{P}^{(1)}(\mathbf{r}, t) + \mathbf{P}^{(2)}(\mathbf{r}, t) + \mathbf{P}^{(3)}(\mathbf{r}, t) + \dots \quad (3.87)$$

Usually, the terms in this expansion are considered in the frequency domain (rather than

⁵Note that often this is considered to be included in the purity. Here we describe these as separate imperfections and consider them independently.

time), and their components are given by [140]

$$P_j^{(n)}(\omega) = \chi_{j\alpha_1 \dots \alpha_n}^{(n)}(\omega_\sigma; \omega_1, \dots, \omega_n) E_{\alpha_1}(\mathbf{r}, \omega_1) \dots E_{\alpha_n}(\mathbf{r}, \omega_n) \quad (3.88)$$

where $\omega_\sigma = \sum_i^n \omega_i$, $\chi^{(n)}$ is the nonlinear susceptibility tensor of rank $n + 1$ and $\alpha \in \{x, y, z\}$.

SPDC is a $\chi^{(2)}$, three-wave mixing process in which a single photon from a pump field is converted to two photons of lower frequency. The standard setup consists of pumping a nonlinear crystal with a strong laser. In this configuration the pump can be described by a classical field with a spectral amplitude $\alpha(\omega_p)$ and assuming a frequency-independent $\chi^{(2)}$ in the spectral range of interest, it can be shown that the interaction Hamiltonian of the SPDC process is given by [141]

$$\hat{H}_I \propto \int d\omega_s \int d\omega_i \alpha(\omega_s + \omega_i) \Phi(\omega_i, \omega_s) a_1^\dagger(\omega_s) a_2^\dagger(\omega_i) + \text{h.c.} \quad (3.89)$$

where s, i label the signal and idler modes respectively, the function Φ describes the phase-matching conditions and the constant of proportionality is a function of $\chi^{(2)}$ and the intensity of the pump field. The function $f(\omega_i, \omega_s) = \alpha(\omega_s + \omega_i) \Phi(\omega_i, \omega_s)$ is known as the *joint spectrum* of the signal/idler fields. Ideally, the joint spectrum is separable, $f(\omega_i, \omega_s) = \psi(\omega_i) \phi(\omega_s)$. In this case, the Hamiltonian simplifies to

$$H_I \propto a_1^\dagger[\psi] a_2^\dagger[\phi] + \text{h.c.} \quad (3.90)$$

This process results in a unitary operator

$$\hat{S}(\xi) \equiv \exp(-\xi a_1^\dagger[\psi] a_2^\dagger[\phi] + \xi^* a_1[\psi] a_2[\phi]). \quad (3.91)$$

Acting on the vacuum, this generates the two-mode squeezed vacuum state

$$\hat{S}(\xi)|\mathbf{o}\rangle = \text{sech}(r) \sum_{n=0}^{\infty} (-1)^n e^{in\theta} \tanh^n(r) \frac{1}{n!} (a_1^\dagger[\psi] a_2^\dagger[\phi])^n |\mathbf{o}\rangle \quad (3.92)$$

where $\xi \equiv r e^{i\theta}$. When the squeezing parameter r is small this state is given by $\sim \sum_n \xi^n |n\rangle_1 |n\rangle_2$. Choosing r such that $r^2 \gg r^4$ means that if a detector for one mode clicks, with high probability a single photon has been prepared in the remaining mode. The cost is that most of the time no photons are generated at all. One promising approach to overcoming this inherent inefficiency is to *multiplex* many SPDC sources [142–144], such that with high probability at least one source clicks. The generated photon can then be switched to the desired mode via feedforward. The main barrier for current realisations of such schemes is the lack of availability of high speed, low-loss switches.

Another challenge in SPDC is producing highly pure single photons. This is because in practice the joint spectrum is rarely separable. Under fairly general conditions it can be represented in Schmidt decomposition as [145]

$$f(\omega_1, \omega_2) = \sum_{i=1}^N \sqrt{p_i} \psi_i(\omega_1) \phi_i(\omega_2). \quad (3.93)$$

This results in a Hamiltonian which is a sum of terms like Eq. (3.90) and a state which is a tensor product of independent two-mode squeezers

$$\bigotimes_{i=1}^N \hat{S}_i(\xi_i) |\mathbf{o}\rangle. \quad (3.94)$$

In the single pair subspace, at low pump powers, this is then

$$\sum_i \sqrt{p_i} |\psi_i\rangle_1 |\phi_i\rangle_2 \quad (3.95)$$

where $|\chi\rangle_i \equiv a_i^\dagger[\chi]|\mathbf{o}\rangle$. This means that the generated photons are spectrally entangled. If the photon in mode 2 is measured, the resulting photon in mode 1 is then described by the state

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (3.96)$$

and so the purity of the source is simply $\sum_i p_i^2$.

For some platforms, such as integrated Silicon photonics, there is no $\chi^{(2)}$ coefficient

due to the centro-symmetry of the crystal lattice. Therefore sources following very similar principles can be implemented using higher-order nonlinearities such as spontaneous four-wave mixing (SFWM) [146].

The other main approach to producing single photon sources is to use a single quantum emitter. This is an (approximately) two-level quantum system with a radiative optical transition that can be used to generate single photons. Usually such sources work via the coherent excitation of the system, followed by its spontaneous decay and the emission of a single photon. If all non-radiative decay is suppressed then such sources can be highly efficient with minimal multi-photon contamination. Emitters that have been used as SPSs include single atoms [147], ions [148], molecules [149], defects in crystal structures such as nitrogen-vacancy centres in diamond [150] and quantum dots [151].

The main shortcomings of these approaches have usually been twofold. On the one hand, producing high efficiency sources in practice requires highly efficient coupling of the light from the source into a single optical mode. Secondly, producing high purity and indistinguishable photons is often difficult due to for instance phonon dissipation and local environmental effects [152].

Recent improvements in quantum dot sources have demonstrated increased efficiencies and high indistinguishability between photons produced in successive time-bins [153, 154], and seem to be the most promising route to producing tens of photons in the near-term. Integrating such sources into a miniaturised architecture, as all-optical sources can be, remains an engineering challenge although recent progress coupling integrated gallium arsenide dots to waveguides is promising [155].

3.7.2 Linear optical networks

As we will see in the next chapter, arbitrary linear optical networks (LON) can be composed by a set of beamsplitter and phaseshifter operations. For example, if a dual-rail qubit is encoded in the polarisation degree of freedom $\{|H\rangle, |V\rangle\}$ of a single photon propagating in free-space then these operations can be implemented via waveplates. Waveplates are birefringent crystals oriented such that they can produce a phase shift in one polar-

isation component relative to the other without changing the propagation of the light. The combination of so-called half-wave plates and quarter-wave plates (differing in their thickness relative to the wavelength of light) allows for arbitrary unitary transformations of polarisation encoded qubits.

However, in order to achieve large-scale photonic networks, one needs to construct complex interferometers between many optical modes. This is difficult to achieve using bulk optical components such as waveplates and half-silvered mirrors due to the practical size of the optical setup required and the problem of maintaining precise phase-stability across the network in the face of fluctuating ambient conditions.

Integrated optical circuits [5, 6] have provided a platform allowing dramatic improvements in experimental capabilities due to their inherent phase stability and small device footprint. Waveguides which spatially confine light in integrated devices are formed via the refractive index contrast between a cross-section of a core material and its surrounding cladding. By solving Maxwell's equations within this region, including the boundary conditions,⁶ TE (where there is no electric field in the direction of propagation) and TM (where there is no magnetic field in the direction of propagation) modes can be found. Waveguide structures are then designed such that in the frequency range of operation only a single TE and TM mode can propagate, so the mode in a single waveguide corresponds to a single path of a linear optical network, or a single rail of a photonic qubit.

Beamsplitters in integrated devices are often achieved by two principal mechanisms. Directional couplers are formed by bringing two waveguides close together such that there is an overlap between the single waveguide modes. This coupling can be described by the Hamiltonian from Eq. (3.27). By coupling for a length L , an arbitrary reflectivity beamsplitter can in principle be achieved. An alternative approach to perform the same linear optical transformation is to use a multi-mode interference (MMI) device. This is a device in which two waveguides are fed into a larger region in which many more modes can propagate. The dimensions of this region are then selected such that light can be coupled back into single mode waveguides but interferences have changed their relative

⁶Interestingly, for a rectangular waveguide, no analytical solutions to Maxwell's equations exist and modes must be found by discretising the space and using a numerical mode solver.

intensities. In practice, MMIs often provide a more precise reflectivity at the cost of being lossier.

Phaseshifts can be implemented by modulation of the phase in one waveguide relative to others. There are a number of possible ways of implementing this, such as using electro-optic modulators which exploit the Pockels or Kerr effect. To date however, most experiments have used thermo-optic modulators which are much slower, but lower loss. Increasing the temperature of a waveguide changes its refractive index. This can be understood by considering the linear term of the polarization, $\mathbf{P}^{(1)} = \epsilon_0(\epsilon_r - 1)\mathbf{E}$, where $n = \sqrt{\epsilon_r}$, and noting that the dipole moment of the medium, and so ϵ_r , is increased as the temperature is increased. The temperature-dependent phase shift due to the thermo-optic effect relative to another waveguide is

$$\Delta\phi = 2\pi \frac{L}{\lambda} \frac{dn}{dT} \Delta T, \quad (3.97)$$

where L is the length, λ the frequency, ΔT the temperature shift and dn/dT the thermo-optic coefficient. Thermo-optic phaseshifters are usually achieved by patterning resistive heaters above the waveguides and applying currents to change the temperature.

Popular material platforms for integrated quantum photonics include Silica, Silicon Nitride, Indium Phosphide, Gallium Arsenide, Lithium Niobate and Silicon. All these platforms have strengths and weaknesses in terms of their size, loss, nonlinearity and performance of passive and active optical components. Most current multi-photon experiments, including those reported in this thesis, use a Silica (SiO_2) cladding and silica-doped core, similar to optical fibres [156–160]. These circuits have the advantage of being low loss in both propagation but also in coupling into and out of the chip. However, the size of the waveguides and integration density means that such circuits are not appropriate for large-scale applications. More potentially long-term scalable platforms are Silicon [7], where high component densities can be achieved and classical CMOS-compatible silicon photonics fabrication techniques can be used and Gallium arsenide [161], which can support deterministic quantum dot photon sources.

3.7.3 Detectors

The final key component of a quantum photonics experiment are the single photon detectors [162]. Again, we can list the most important properties which define the quality of a single photon detector

- Efficiency: the probability of detecting a photon if one arrives at the detector
- Dark count rate: the probability of measuring a false positive, i.e. a photon when none is present.
- Timing jitter: the variation in time interval between the absorption of a photon and the registering of a classical electrical pulse signifying its detection.
- Dead time: the time it takes for the detector to be ready for subsequent measurement following a detection event.
- Photon number resolving: the ability of the detector to measure the exact number of photons arriving at the detector.

Again, there are often trade-offs, for instance between efficiency and dark count rate.

The most commonly used detectors in current experiments are single photon avalanche photodiode (SPAD) detectors. These are semiconductor devices biased such that a single photon absorption process triggers a self-sustaining avalanche creating a sharp rise in current. SPADs tend to have efficiencies of around $\sim 50 - 60\%$.

More recently, detectors with improved performance have been demonstrated and are commercially available using superconducting technology. Transition edge sensor (TES) detectors [163] work via the use of a superconducting film at its critical temperature. When a photon is absorbed, the temperature rises, increasing its resistance and this change can be measured, allowing for number resolution. Superconducting nanowire single photon detectors (SNSPDs) [164] are another type of superconducting detector which, although not offering intrinsic number resolving⁷, are faster and can be placed onto integrated waveguide structures for scalability. SNSPDs work below the critical temperature

⁷Though pseudo-number resolving can be achieved e.g. by the use of multiple nanowires on the same optical waveguide

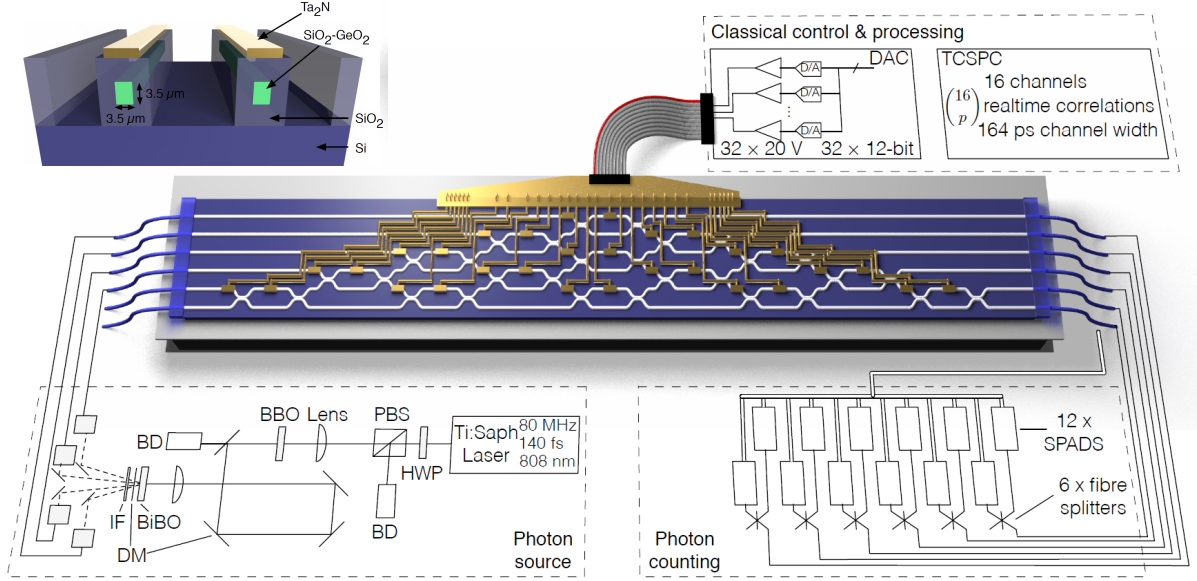


Figure 3.7: Experimental setup. The multiphoton source uses a pulsed Ti:Saph laser to perform SPDC, preparing two two-mode squeezed states which are coupled into fibre. The integrated photonic chip consists of six germanium doped silica waveguides with a silicon cladding with 30 directional couplers and thermo-optic phase shifters. Photon counting is performed via an array of single photon avalanche photodiodes. The experiment is interfaced with a computer for real-time photon coincidence counting and setting of all phase shifters. Figure adapted from [165]. Reprinted with permission from AAAS.

but biased such that the current is close to its critical current. A photon incident upon the nanowire results in the formation of a ‘hotspot’, a small non-superconducting region with nonzero resistance, this then produces a voltage pulse which can be measured. Superconducting detectors have shown efficiencies of over 90% as well as improving timing jitter and dark count rates. The main obstacle to scalability using these detection systems is that they both require low temperatures ($<10\text{K}$), whereas the ability to operate at room temperature is one of the most appealing aspects of quantum photonic technologies.

3.7.4 Experimental setup

Here we introduce the experimental setup shown in Fig. 3.7 which was used in the experiments reported in this thesis.

Source

A pulsed, Titanium:Sapphire laser (Coherent Chameleon) is used to generate 140fs pulses at $\sim 808\text{nm}$ at a repetition rate of 80MHz. A half-wave plate and polarising beamsplitter (PBS) are used to attenuate the power. Next, a β -barium borate (BBO) crystal is used to perform second harmonic generation. Dichroic mirrors are then used to remove the remaining 808nm light and a 0.5mm Bismuth Triborate BiB_3O_6 (BiBO) crystal is used to perform SPDC from the up-converted 404nm pulse. Down-converted photons are emitted in a cone at opening angle $\theta = 6^\circ$ and pass through a 3nm interference filter at 808nm. Prisms are aligned to couple light from four equidistant points on the SPDC cone (at 12,3,6 and 9 o'clock) into polarisation maintaining fibres (PMF). Ideally, this source produces two independent two-mode squeezed vacuum states (at 12,9 and 3,6 o'clock respectively). Two of the collection stages have motorised actuators which can be computer controlled for performing HOM-dip type experiments.

When either pair is connected directly to the detectors, the ratio between the coincidence detection rate and the single detection rate is $\sim 12\%$. Taking into account the detector efficiency, this translates to a heralding efficiency of around 24%. The indistinguishability, measured by Hong-Ou-Mandel dip visibility, for photons generated in the same pair generation event is ~ 0.98 while for photons generated in different pair generation events, the HOM visibility is ~ 0.91 .

These non-unit visibilities result from a number of mechanisms. Since each generated photon is incident upon the 3nm filter at different positions and angles, their spectral properties will all be slightly different. When performing interference between photons from the two independent sources there is the additional effect of non-unit purity caused by spectral entanglement as discussed in Sec. 3.7.1. Also in this case a further temporal jitter is introduced from group velocity mismatch resulting from generation of the photon pairs at different locations in the crystal [166]. Since the downconverted photons travel at a different speed through the crystal than the pump, the two photon pairs will experience a different amount of temporal walk-off on a shot-by-shot basis.

Circuit

The silica-based integrated photonic chip was fabricated at the Nippon Telegraph and Telephone company (NTT) in Japan. Flame hydrolysis deposition followed by photolithographic and reactive ion etching was used to fabricate germanium doped silica ($\text{SiO}_2\text{-GeO}_2$) waveguides with dimensions $3.5\mu\text{m} \times 3.5\mu\text{m}$ with a silica cladding onto a silicon substrate. Thin-film Tantalum Nitride (Ta_2N) thermo-optic heaters were then fabricated on top of the circuit with dimensions $1.5\text{mm} \times 50\mu\text{m}$. The circuit is formed of a cascaded array of 30 directional couplers (each with a length of $500\mu\text{m}$) and 30 phase shifters designed to perform a universally reconfigurable transfer matrix on six waveguide modes. The design of this circuit will be explored further in the following chapter. The coupling losses have been estimated as $\sim 9\%$ per facet and the directional couplers at $< 2.3\%$. The average loss fibre-to-fibre was measured to be $\sim 42\%$. The device is actively cooled via a Peltier cooling unit.

The thermo-optic modulators are driven by electronic heater driver boards designed in-house which can deliver up to 20V with 4.9mV resolution and current up to 100mA. These are then interfaced with a computer to set all the heaters to implement a given transfer matrix.

Detectors

The detection system uses 12 SPADs (Perkin Elmer SPCM-AQRH-14), each with efficiencies 50 – 60%, a dark count rate of $\sim 100\text{Hz}$, timing jitter of $\sim 350\text{ps}$ and a dead time of 32ns. A coincidence counting card time-tagging all simultaneous channels in a time window usually set to be around 2ns is used to register detection events (see the thesis of Peter Shadbolt for more details [167]). For each channel it is possible to set a specific time delay that is used by the counting card to compensate for the discrepancy in the signals arrival time introduced in the experiment by optical fibres, detectors, electronics and coaxial cables.

Chapter 4

Universal Linear Optics

4.1 Introduction

Universality is a crucial concept in computation. The ability to reprogram a single device to solve any possible instance of a problem is a key capability which has driven the modern computing revolution. For linear optical devices, this notion of universality is captured by the ability to perform any unitary transformation upon a set of optical modes. The standard approach to constructing quantum optical experiments for over twenty years has involved assembling and realigning new setups (or more recently, fabricating new integrated devices) to implement a given optical circuit. However, it was shown in 1994 by Reck *et al.* [168] that there is a constructive scheme for implementing an arbitrary unitary transfer matrix via a set of beamsplitters and phase shifters. If these components are all fully reconfigurable, then such a design provides a platform for universal linear optical transformations. In this chapter we report the implementation of the world's first universal linear optical chip on six optical modes and describe its use as a testbed for experimentally exploring a wide range of linear optical quantum information processing tasks.

4.2 The Reck *et al.* Scheme

It has been known since the 19th century [169] that any $m \times m$ unitary matrix in $U(m)$ can be decomposed into a series of $\sim m^2/2$ two-dimensional unitaries. The scheme of Reck *et al.* uses a circuit of tunable beamsplitters and phase shifters to physically realise this decomposition in linear optics. The building blocks are two-mode rotations acting on modes i and $i + 1$ of the form

$$T_i(\theta, \phi) = \begin{pmatrix} 1 & 0 & & & & \\ 0 & 1 & & & & \\ & & \ddots & & & \\ & & & e^{i\phi} \sin \frac{\theta}{2} & e^{i\phi} \cos \frac{\theta}{2} & \\ & & & \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} & \\ & & & & \ddots & \\ & & & & & 1 & 0 \\ & & & & & 0 & 1 \end{pmatrix}. \quad (4.1)$$

These transformations can be achieved (up to a global phase which can be corrected for) using two controllable phase shifters via the circuit on the left-hand side of Fig. 4.1(b) where the Mach-Zehnder interferometer (MZI) with a phase θ acts as a variable reflectivity beamsplitter and an additional phaseshifter (PS) implements ϕ .

$$\text{MZI}(\theta) = \text{BS} \text{PS}(\theta) \text{BS} \quad (4.2)$$

$$= \frac{1}{2} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \quad (4.3)$$

$$= e^{-i\frac{(\theta+\pi)}{2}} \begin{pmatrix} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \\ \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \end{pmatrix} \quad (4.4)$$

$$\text{PS}(\phi) \text{MZI}(\theta) = e^{-i\frac{(\theta+\pi)}{2}} T(\theta, \phi) \quad (4.5)$$

We now consider the norm 1 complex vector $(U_{11}, U_{21}, \dots, U_{m1})^\top$ which forms the first column of the unitary matrix U that we wish to implement. This vector of amplitudes

4. Universal Linear Optics

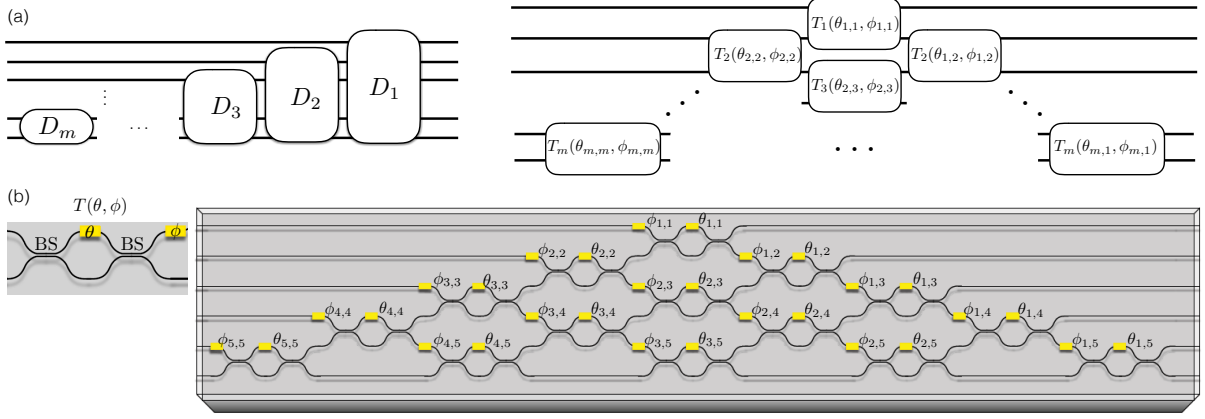


Figure 4.1: The Reck *et al.* scheme. (a) Left: an arbitrary linear optical unitary U can be decomposed into a product of unitaries D_i acting on modes $i, \dots, m$. Right: These can then be decomposed into a cascade of two-mode rotations T . (b) Left: Physical implementation of a T rotation as a MZI with phase shift θ followed by a phase shift ϕ . Right: Full scheme for a 6×6 unitary.

can be prepared from an input in the first mode via the circuit

$$\begin{aligned} D_1(1, 0, \dots, 0)^\top &= T_{m-1}(\theta_{1,m-1}, \phi_{1,m-1}) \dots T_2(\theta_{1,2}, \phi_{1,2}) T_1(\theta_{1,1}, \phi_{1,1})(1, 0, \dots, 0)^\top \\ &= (U_{11}, U_{21}, \dots, U_{m1})^\top. \end{aligned} \quad (4.6)$$

With $U_{ij} = u_{ij}e^{i\alpha_{ij}}$, the θ 's are used to set the amplitudes u according to the relation

$$u_{i1} = \cos \frac{\theta_{1,1}}{2} \cos \frac{\theta_{1,2}}{2} \dots \sin \frac{\theta_{1,i}}{2} \quad i = 1, \dots, m-1 \quad (4.7)$$

and ϕ 's set the phases (relative to the phase on the m th element, which is not tunable) via $\phi_{1,i} = \alpha_{i1} - \alpha_{m1}$. We can now consider the new matrix

$$D_1^\dagger U = \begin{pmatrix} e^{i\alpha_{m1}} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U^{m-1} & \\ 0 & & & \end{pmatrix} \quad (4.8)$$

If the same procedure is performed to prepare the first column of U^{m-1} given an input

in the second mode, we obtain another diagonal cascade of transformations

$$D_2 = T_{m-1}(\theta_{2,m-1}, \phi_{2,m-1}) \dots T_3(\theta_{2,3}, \phi_{2,3}) T_2(\theta_{2,2}, \phi_{2,2}). \quad (4.9)$$

Continuing in the same way for the remaining columns we produce $D_{m-1}^\dagger \dots D_2^\dagger D_1^\dagger U = \Lambda$ where $\Lambda = \text{Diag}[(e^{i\alpha_{m1}}, e^{i\alpha_{m2}}, \dots, e^{i\alpha_{mm}})]$ and can be implemented by m additional phase shifts. Rearranging we therefore find $U = D_1 D_2 \dots D_{m-1} \Lambda$. Fig. 4.1(a) shows this circuit schematically and Fig. 4.1(b) shows the physical layout of the circuit for $m = 6$ ¹. There are $m(m-1)/2$ pairs $\{\theta_{ij}, \phi_{ij}\}$ in the two-mode rotations plus the m input phase shifts results in a full parameterisation of $U(m)$.

We note that the Reck scheme results in a triangular network of components. Recently, a new scheme has been proposed by Clements *et al.* [170] which uses a new decomposition of $U(m)$ into a square network with the same number of components. The Clements scheme benefits from a smaller maximum *circuit depth*, the number of layers of beam-splitters, and is less susceptible to path-dependent losses due to its symmetric structure. On the other hand, the Reck scheme is likely to be preferable when states are only to be input into a subset of modes due to its reduced effective circuit depth in this case. There is also some evidence that it has improved resilience to fabrication errors [171].

4.3 A Universal Linear Optical Processor

The device introduced in Sec. 3.7.4 is an experimental realisation of a universal linear optical network. The circuit is comprised of 30 directional couplers and 30 thermo-optic phase shifters forming 15 MZIs, plus an additional 15 phase shifters to implement the ϕ parameters in the Reck scheme. Silica is used as the platform of choice due to the low insertion and propagation losses which allow for multi-photon experiments with practical count rates. To implement a 2π phase shift for a given heater requires a temperature shift from room temperature to around 74°C, applying a power of $\sim 0.8W$ for $\sim 10ms$.

¹The observant reader will note that this circuit is reversed relative to (a). Due to experimental practicalities the device is operated this way. The Reck decomposition is therefore performed for $U^\top$. Phases on the output modes are also neglected since photon counting measurements are insensitive to them

This switching time is likely not fast enough for the adaptive feedforward required for true LOQC, but suffices for applications such as testing linear optical circuitry and gates and the bosonic simulations presented in the next chapter.

4.3.1 Characterisation

Since each heater is slightly different, and the path lengths of the Mach-Zehnder arms are not perfectly matched, each phase shifter must be calibrated such that its phase-voltage relationship is known. This phase-voltage relationship is modelled using the parameterisation

$$\phi(V) = \text{mod}[a + bV^2 + cV^3, 2\pi] \quad (4.10)$$

where a is the phase difference when the heater is switched off i.e. at 0V, b is the ohmic response of the resistor and c is a correction for non-ohmic response ($b \gg c$). Phase shifters in MZIs can then be calibrated by sweeping the voltage when a heralded single photon² is input into its top port and the count rate is measured at the top output port. This count rate is then modelled by

$$C = A - B \cos [\phi(V)] \quad (4.11)$$

and A, B and a, b, c are estimated via a least squares fit. The a values were found to be approximately uniformly distributed between 0 and 2π , and typical values for b and c were 0.05 and 0.001. The procedure for a full calibration of the device proceeds as follows.

A calibration curve is measured using $T_1(\theta_{1,1}, \phi_{1,1})$ to estimate $\theta_{1,1}(V)$. Then, setting $\theta_{1,1}$ to divert light to the input of $T_2(\theta_{1,2}, \phi_{1,2})$, $\theta_{1,2}(V)$ can be estimated in the same way. This procedure is then continued to calibrate all $\{\theta_{i,j}\}$. In order to calibrate the $\{\phi_{i,j}\}$ heaters, the calibrated $\theta_{i,j}$'s are set to form larger MZIs. For instance, the arrangement

$$T_2(\pi/2, 0)T_3(\pi, 0)T_1(\pi, 0)T_2(\pi/2, \phi_{2,2}) \quad (4.12)$$

²This can be done equivalently with classical laser light and intensity measurements, though given the operation will be to some degree mode-dependent, the best calibration in practice uses the same optical states that are to be used for the quantum experiments.

can be used to measure $\phi_{2,2}(V)$. See Fig. 4.2(b) for a schematic of this procedure and example calibration curve. Once a calibration is obtained for all heaters, any unitary can be ‘dialed up’ on the device via its Reck decomposition. This calibration was found to continue to perform well over a long period of many months without any noticeable drop in experimental fidelities.

The uncertainty in the phase set at a given voltage can be estimated via the uncertainties in a, b and c . Averaging across the device we find $\Delta\theta \approx 0.03$ rad, with the uncertainty predominantly coming from Δa . The experimental fringes also have a visibility $V = (C_{\max} - C_{\min})/C_{\max}$ which is given by $2B/(A + B)$. The main reason that $V < 1$ is due to the non-ideal reflectivity of the directional couplers. If the reflectivities of the two directional couplers are given by $1/2 + \delta_1$ and $1/2 + \delta_2$ then

$$V = \frac{2\sqrt{(1 - 4\delta_1^2)(1 - 4\delta_2^2)}}{1 + 4\delta_1\delta_2 + \sqrt{(1 - 4\delta_1^2)(1 - 4\delta_2^2)}} \quad (4.13)$$

and if $\delta_1 = \delta_2$, which may be the case if fabrication errors result in a constant offset across a device, then $V = 1 - 4\delta^2$. The average visibility seen across all heaters is 0.995, so we can loosely estimate that $\delta \sim 0.03$. This non-unit visibility in each MZI means that it is not possible to achieve full transmission or reflection, even if the phase shift θ was set perfectly.

A final experimental consideration is determining any sources of crosstalk. The two most likely mechanisms are thermal and electrical. Thermal crosstalk occurs when heat diffusion from one phase shifter can change the phase in another interferometer. It is mitigated in the design of the device by the fabrication of trenches between waveguides (see Fig. 3.7) resulting in no measurable thermal crosstalk. Electrical crosstalk can occur when several heaters share a common ground. When multiple heaters are driven, the non-zero resistance of the ground ($\sim 10m\Omega$) results in a raising of the ground voltage and the wrong phase shifts being set. To solve this problem, all resistances were measured, and extra classical pre-processing was used to calculate a set of adjusted voltages incorporating the expected ground shift.

Finally, it was measured that the polarisation of the input light was being rotated at

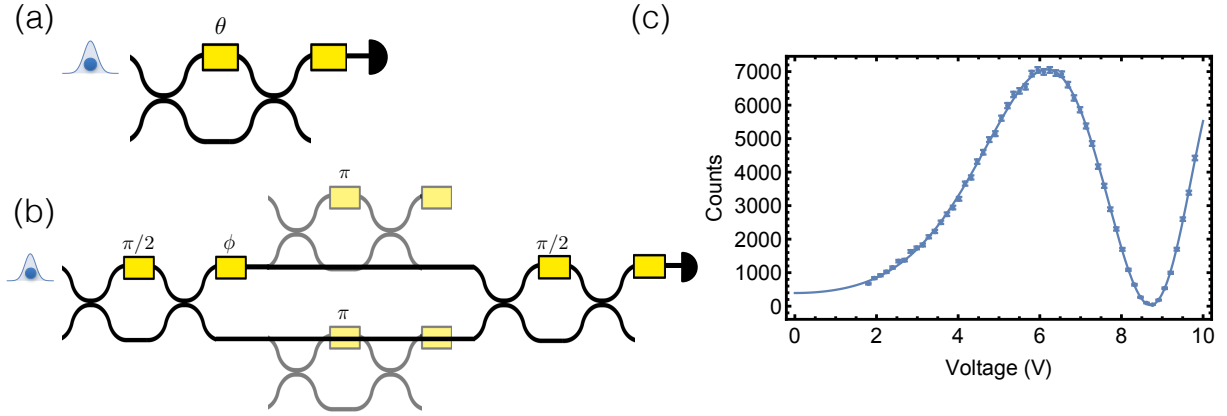


Figure 4.2: Characterisation of phase shifters. (a) Phases in MZIs can be characterised by monitoring the countrate modulation when the voltage is swept. (b) For additional phases, larger MZIs can be formed using four calibrated MZIs, with two acting as beam-splitters and two as identity transformations. (c) Example of calibration curve for one heater measuring the countrate vs voltage applied.

certain places through the circuit. This would result in non-ideal multi-photon interference and so is filtered out via the use of fibre PBSs on the output channels.

4.4 Linear Optical Gates

Once calibrated, the device can be used as a testbed for linear optical quantum logic. As a benchmark of device performance, we demonstrated a set of process tomography experiments on standard linear optical gates. Then, going beyond previous demonstrations, we implemented the first heralded logic gates in integrated photonics, performing a KLM CNOT gate and a heralded Bell state generator.

4.4.1 Process tomography experiments

The simplest mapping of quantum logic to the universal device is to use each MZI and associated phase shifts to implement arbitrary single qubit logic on a single photon in two modes. In order to perform a process tomography experiment, a complete set of initial states and measurements must be made. This can be achieved using a chain of three MZIs, the first to prepare a state, the second to perform the gate and the third to

State	θ	ϕ	Observable	ϕ	θ
$ 0\rangle$	π	0	Z	0	π
$ 1\rangle$	0	0	X	0	$\frac{\pi}{2}$
$ +\rangle$	$\frac{\pi}{2}$	0	Y	$\frac{\pi}{2}$	$\frac{\pi}{2}$
$ -\rangle$	$\frac{\pi}{2}$	π			
$ +i\rangle$	$\frac{\pi}{2}$	$\frac{\pi}{2}$			
$ - i\rangle$	$\frac{\pi}{2}$	$\frac{3\pi}{2}$			

Table 4.1: On-chip phase shifts to prepare Pauli eigenstates and measure Pauli observables.

change the measurement basis (see Fig. 4.3). Using the three $T_3(\theta, \phi)$ transformations, process tomography was performed on the set of single qubit gates $\{X, Y, Z, H, T\}$ using a set of preparations and measurements consisting of the six Pauli eigenstates. The phases required to prepare Pauli eigenstates and measure Pauli observables are displayed in Table 4.1.

An estimate of the Choi-Jamiołkowski state $\rho_{\mathcal{E}}^{\text{est}}$ corresponding to the gates were found using maximum-likelihood estimation (MLE). As described in Sec. 2.3.4, the theoretical probability of measuring an outcome corresponding to the POVM effect E_j when the state ρ_i is prepared is given by $\text{Pr}_{ij}(\rho_{\mathcal{E}}) = d\text{Tr}[(\rho_i^{\top} \otimes E_j)\rho_{\mathcal{E}}]$. Therefore, if n_{ij} events are registered in a total of N_{ij} trials, the Likelihood function has the binomial form

$$\mathcal{L}(n_{ij}|\rho_{\mathcal{E}}) = \binom{N_{ij}}{n_{ij}} \text{Pr}_{ij}(\rho_{\mathcal{E}})^{n_{ij}} [1 - \text{Pr}_{ij}(\rho_{\mathcal{E}})]^{N_{ij}-n_{ij}}. \quad (4.14)$$

The MLE estimate is then given by the maximum over the product of likelihoods for each experiment labelled by i and j subject to physicality constraints

$$\rho_{\mathcal{E}}^{\text{est}} = \max_{\rho_{\mathcal{E}}} \prod_{i,j} \mathcal{L}(n_{ij}|\rho_{\mathcal{E}}) \quad (4.15)$$

$$\text{subject to: } \rho_{\mathcal{E}} \geq 0, \text{Tr}_A(\rho_{\mathcal{E}}) = \mathbb{1}/d \quad (4.16)$$

Unfortunately performing this maximisation is generally not computationally practical, especially when moving to two or more qubits. To alleviate this problem we can make a

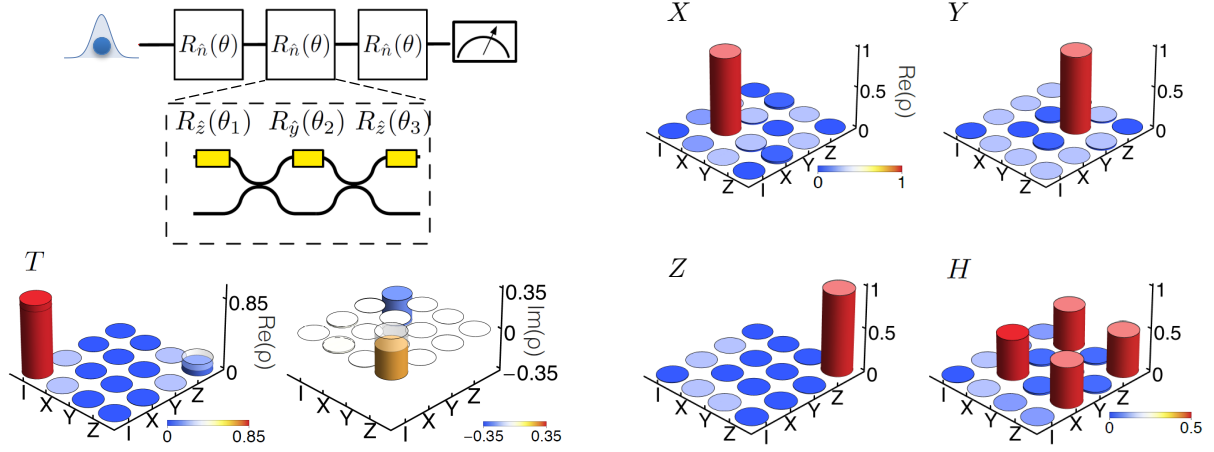


Figure 4.3: Process tomography for single qubit gates. Using three consecutive two-mode unitaries, single qubits can be prepared, manipulated and measured. The real part of the estimated process matrices χ^{est} for the single qubit Pauli and H gates are shown along with both the real and imaginary part for T . Hollow bars represent the ideal values. Figure adapted from [165]. Reprinted with permission from AAAS.

Gaussian approximation to Eq. (4.14)

$$\mathcal{L}(n_{ij}|\rho_{\mathcal{E}}) \approx \exp \left\{ \frac{-N_{ij} [f_{ij} - \text{Pr}_{ij}(\rho_{\mathcal{E}})]^2}{f_{ij}(1 - f_{ij})} \right\} \quad (4.17)$$

where $f_{ij} = n_{ij}/N_{ij}$ is the experimental frequency. Since the log of the likelihood has the same maximum, we can simplify the problem to the minimisation of the negative log-likelihood

$$\begin{aligned} \rho_{\mathcal{E}}^{\text{est}} &= \min_{\rho_{\mathcal{E}}} \left\{ \sum_{i,j} \frac{N_{ij} [f_{ij} - \text{Pr}_{ij}(\rho_{\mathcal{E}})]^2}{f_{ij}(1 - f_{ij})} \right\} \\ &\text{subject to: } \rho_{\mathcal{E}} \geq 0, \text{Tr}_A(\rho_{\mathcal{E}}) = \mathbb{1}/d \end{aligned} \quad (4.18)$$

We now have obtained an optimisation which can be phrased as a semidefinite program (SDP), a standard form of convex optimisation problem for which many practical solvers are available. Here we use the CSDP solver [172].

One problematic aspect of MLE is that if the direct linear inversion estimate of $\rho_{\mathcal{E}}$ is not positive (as described in Sec. 2.3.4), then the estimate $\rho_{\mathcal{E}}^{\text{est}}$ will contain zero eigenvalues

Gate	$\mathcal{F}_{\text{pro}}$	$ \cdot _{\diamond}$
X	0.997 ± 0.001	0.023 ± 0.001
Y	0.993 ± 0.001	0.031 ± 0.001
Z	0.994 ± 0.001	0.031 ± 0.001
H	0.995 ± 0.001	0.030 ± 0.001
T	0.978 ± 0.001	0.068 ± 0.001

Table 4.2: Process fidelities and diamond distances between ideal single qubit gates and experimentally reconstructed gates.

[173]. Zero eigenvalues suggest that certain outcomes can never occur, which is not a justified conclusion given a small sample of data. Furthermore, as can be seen from Eq. (4.18), when $f_{ij} = 0$ or 1 the objective function is not well-defined. A statistically well-motivated approach to deal with these problems is to use *hedged* MLE [174], where the experimental frequencies are adjusted to

$$f_{ij} = \frac{n_{ij} + \beta}{N_{ij} + K\beta} \quad (4.19)$$

where K is the number of outcomes in the measurement of f_{ij} and β is a hedging parameter (we set $\beta = 0.1$ as suggested in [175]). The experimentally reconstructed process matrices in the Pauli basis are shown in Fig. 4.3 and the process fidelities and diamond distances to the ideal gates are displayed in Table 4.2. Errors in these quantities were estimated by a bootstrapping method, where the raw experimental counts were resampled 100 times drawn from a distribution with means given by the experimental data and the statistics of the resulting distributions analysed.

A standard test of linear optical quantum logic is the *postselected* CNOT gate from Ref. [176]. This is a gate which, instead of using ancilla photons and heralding, performs the correct evolution conditional upon the output state remaining within the computational subspace. The limitation of postselection is that with destructive single photon detection the gate can only be confirmed to have worked by measuring the computational photons so that no further processing of the qubits is possible. This approach is therefore only scalable if highly challenging single photon non-demolition measurements can be made. On the other hand, these gates demonstrate the basic building blocks of an LOQC

architecture, relying on phase stable linear optical networks and the quantum interference of indistinguishable photons.

Fig. 4.4 shows the gate schematic and results. The CNOT gate is applied to two photonic qubits (defined by modes 2,3 and 4,5) conditional upon finding a single photon in each of the qubits at the output, which occurs ideally with a probability $1/9$. The transfer matrix for this gate is given by

$$T_{\text{post}} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \sqrt{2} & 0 & 0 & 0 & 0 \\ \sqrt{2} & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \quad (4.20)$$

which are mapped to the following on-chip phases (see Fig. 4.1(b) for reference)

$$R = \begin{pmatrix} (0.00, 1.23) & (\theta_{12}, \phi_{12}) & (0.00, 3.14) & (\theta_{14}, \phi_{14}) & (0.00, 3.14) \\ & (0.00, \phi_{22}) & (\theta_{23}, 1.23) & (0.00, 1.57) & (3.14, 3.14) \\ & & (0.00, 3.14) & (3.14, 1.57) & (0.000, 1.23) \\ & & & (0.00, \phi_{44}) & (\theta_{45}, 3.14) \\ & & & & (0.00, 3.14) \end{pmatrix}. \quad (4.21)$$

where $R_{ij} = (\theta_{i,j}, \phi_{i,j})$. Phases $\{\theta_{22}, \phi_{23}\}$ and $\{\theta_{44}, \phi_{45}\}$ can be set to prepare arbitrary pure states of qubits 1 and 2 respectively, and $\{\theta_{12}, \phi_{12}\}$ and $\{\theta_{14}, \phi_{14}\}$ can be set to perform arbitrary projective measurements on qubits 1 and 2 respectively.

Process tomography was performed via the same method as before, this time preparing and measuring states in a basis formed by the tensor product of Pauli eigenstates. This results in a total of 1296 experimental frequencies (though only 324 unitary configurations since four projectors can be measured simultaneously). The estimated process was found to have a process fidelity of $\mathcal{F}_{\text{pro}} = 0.909 \pm 0.001$, average gate fidelity $\mathcal{F}_{\text{ave}} = 0.927 \pm 0.001$ and diamond distance $\|\cdot\|_{\diamond} = 0.115 \pm 0.002$. These are better than those previously

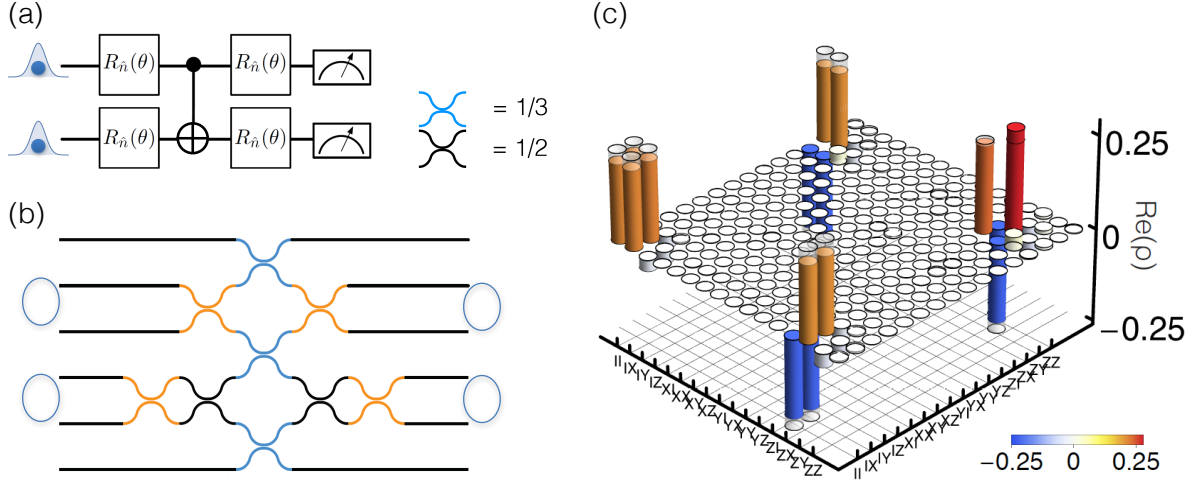


Figure 4.4: Process tomography of postselected CNOT gate. (a) The postselected CNOT gate works via the interference of two photons followed by measurement of a pattern within the logical subspace which occurs $1/9$ of the time. (b) Optical circuit for this gate. Couplers represent two mode rotations and their colour represents their reflectivity according to the key. Orange transformations are arbitrary two mode rotations which are picked via Table 4.1 to perform process tomography. (c) Real part of the experimentally estimated process matrix χ^{est} of the gate. Hollow bars represent ideal CNOT gate values. Figure adapted from [165]. Reprinted with permission from AAAS.

reported[157, 158, 177] in bulk and integrated experiments specifically designed to perform this gate. This demonstrates a high level of performance of the device along with its versatility.

4.4.2 Heralded integrated gates

The key mechanism which enables LOQC is the effective interaction between photons implemented indirectly via quantum interference and photon measurement. This is the approach which is used in the original KLM CZ gate described in Sec. 3.6.1. Here we implement a version of this gate with simplified NS circuits courtesy of Ralph *et al.* [178]. This gate can then be used to implement a CNOT via the simple circuit identity $(\mathbb{1} \otimes H)\text{CZ}(\mathbb{1} \otimes H) = \text{CNOT}$.

A full implementation of this gate would generally require eight optical modes, however we can note that the $|0\rangle$ rail of the control qubit, c_0 , does not need to couple to any others and so can be implemented via an additional straight alignment waveguide on the chip.

Also, the final ancilla mode for the NS gate on the control qubit can be implemented via simply changing the effective detector efficiency of that mode in post-processing. Therefore, although full tomography as before is not possible, the gate's action in the computational basis can be demonstrated, the first demonstration of this form of photonic logic in an integrated device.

The unitary transfer matrix required on our six mode circuit to perform the KLM CNOT gate is given by

$$T_{\text{KLM}} = \begin{pmatrix} 0.476 & -0.622 & -0.440 & 0.440 & 0 & 0 \\ -0.622 & -0.476 & 0 & 0 & 0.622 & 0 \\ -0.383 & 0 & 0.293 & 0.707 & -0.383 & -0.348 \\ 0.383 & 0 & 0.707 & 0.293 & 0.383 & 0.348 \\ 0 & 0.622 & -0.440 & 0.440 & 0.476 & 0 \\ 0.306 & 0 & 0.166 & -0.166 & 0.306 & -0.870 \end{pmatrix} \quad (4.22)$$

corresponding to a set of phases

$$R = \begin{pmatrix} (0.00, 0.99) & (4.71, 1.57) & (4.54, 4.95) & (5.38, 1.79) & (3.82, 0.00) \\ & (0.00, 1.57) & (1.57, 0.99) & (2.19, 0.00) & (4.71, 0.00) \\ & & (0.00, 1.57) & (5.50, 0.00) & (1.57, 2.23) \\ & & & (0.00, 3.14) & (3.14, 0.00) \\ & & & & (0.00, 0.00) \end{pmatrix}.$$

Control and target qubits are implemented via the modes $\{0, 2\}, \{3, 4\}$ respectively (mode 0 representing the additional straight waveguide), and the heralding modes are $\{1, 5\}$. One final experimental difficulty is the preparation of the $|1111\rangle$ state required at the input (the two computational basis states along with the ancilla photons in modes 1,5). Restricting to the four-photon subspace, the state produced by the source is actually $\sim \frac{1}{\sqrt{3}}(|1111\rangle + |2200\rangle + |0022\rangle)$. A configuration of the input modes is therefore chosen for which, in ideal operation, these extra terms in the input cannot give rise to a correct output pattern.

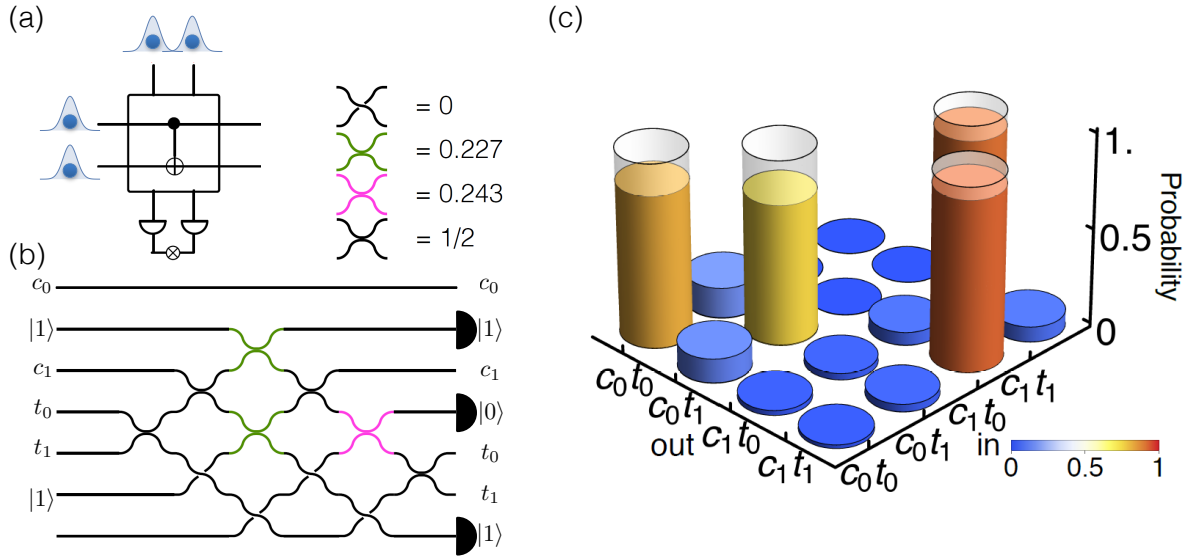


Figure 4.5: KLM CNOT gate. (a) Two ancilla photons are detected to herald a CNOT acting upon two photonic qubits. (b) Circuit mapped onto the universal linear optical circuit, colours represent the reflectivities of two mode rotations according to the key. The control qubit is defined by modes c_0 and c_1 and target qubit t_0, t_1 . (c) Logical truth table measured in the computational basis. Hollow bars represent the ideal CNOT operation. Figure adapted from [165]. Reprinted with permission from AAAS.

Fig. 3.3 shows the truth table obtained from the experiment. Each measurement took around 5 hours to collect an average of 120 four-photon coincidences. The mean statistical fidelity between the experimental and ideal probability distributions, averaged over the four input states was found to be $\bar{\mathcal{F}}_s = 0.930 \pm 0.003$.

As described in Sec. 3.6, although the KLM scheme provides a technically scalable architecture for LOQC it has been superseded by more efficient approaches such as those based on MBQC. The main resources that these schemes require is a supply of entangled photonic qubits. The most efficient circuit for producing heralded Bell states from single photon states, which we will label B_8 , is that of Zhang *et al* [179] which uses four single photons and measures two to herald a Bell state with a probability³ of $p_{\text{succ}} = 3/16$. Again however, this circuit requires eight optical modes. Instead, we developed a new scheme for Bell state generation using just six modes, which we will label B_6 . Although less efficient, succeeding with a probability of $p_{\text{succ}} = 2/27$, this circuit has some interesting properties

³This can be boosted to $p_{\text{succ}} = 1/4$ with additional circuitry and distillation [180]

when compared to the scheme of [179]. These will be explored in Chapter 6.

The circuit schematic is shown in Fig. 4.6 and the transfer matrix for this gate is given by

$$B_6 = \begin{pmatrix} 0.707 & 0.707 & 0 & 0 & 0 & 0 \\ 0.408 & -0.408 & -0.577 & 0.577 & 0 & 0 \\ 0.408 & -0.408 & 0.289 + 0.289i & -0.289 + 0.289i & -0.408i & -0.408i \\ 0.408 & -0.408 & 0.289 - 0.289i & -0.289 - 0.289i & 0.408i & 0.408i \\ 0 & 0 & 0.333 - 0.471i & 0.333 - 0.471i & 0.236 - 0.333i & 0.236 - 0.333i \\ 0 & 0 & 0 & 0 & 0.707 & 0.707 \end{pmatrix}$$

corresponding to a set of phases

$$R = \begin{pmatrix} (0.00, 1.57) & (0.00, 1.23) & (0.00, 1.57) & (0.00, 3.14) & (0.00, 3.14) \\ & (0.00, 3.14) & (0.00, 1.57) & (1.57, 1.23) & (0.00, 3.14) \\ & & (0.00, 3.14) & (0.00, 3.14) & (0.00, 1.57) \\ & & & (0.00, 3.14) & (0.00, 3.14) \\ & & & & (0.00, 3.14) \end{pmatrix}.$$

The qubits are defined by modes $\{1, 2\}$ and $\{5, 6\}$. In order to verify the output of the circuit, we would ideally perform tomography upon the heralded state. However, separating the tomography measurements from the state generation is not possible for the first qubit. The natural measurements that can be made are therefore of operators $Z \otimes \{X, Y, Z\}$, the output distributions of which are shown in Fig. 4.6c. In order to verify entanglement, a measurement of the first qubit is needed in another basis.

To obtain statistics for an X measurement on qubit 1, we created a fibre Sagnac loop between the output of modes 1 and 2 which injects the photon back into the chip with its phase coherence maintained but with an X gate applied to the logical qubit. Due to the configuration of the circuit, the photon then sees a $1/3$ reflectivity MZI between modes 2 and 3, before a $1/2$ reflectivity MZI between modes 1 and 2. This circuit therefore

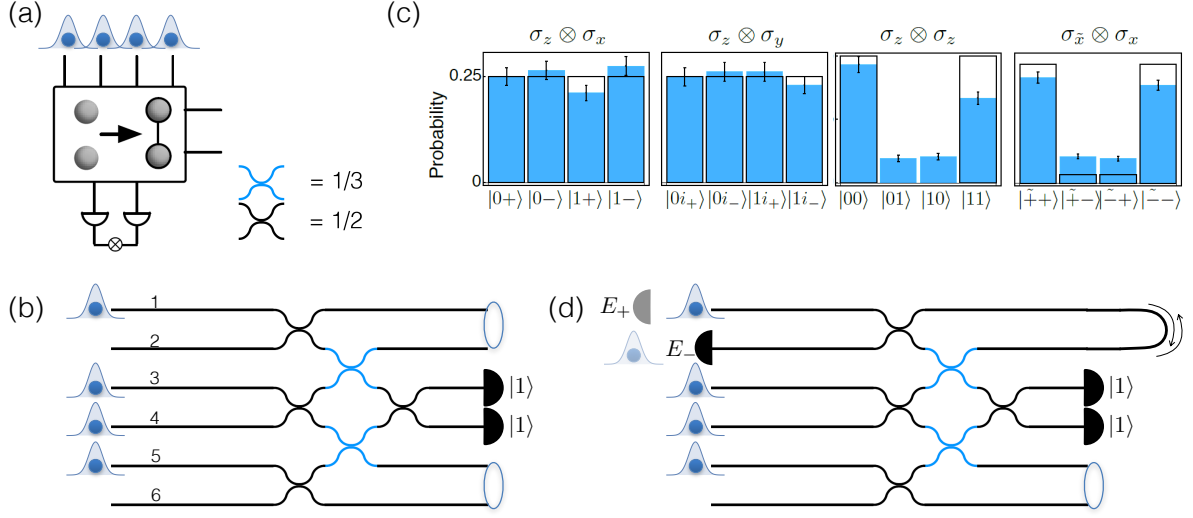


Figure 4.6: Experimental Bell state generation. (a) four input photons are used to generate a bell pair via measurement of two photons. (b) Circuit for new Bell state generation scheme. (c) Experimentally measured distributions at the output for Pauli measurements, furthest right uses the Sagnac loop setup. (d) Sagnac loop setup to measure top qubit in conjugate basis. The top photon is re-injected into the chip and, in two experiments, measured in either in mode 2 E_- or mode 1 E_+ . Figure adapted from [165]. Reprinted with permission from AAAS.

implements a single qubit POVM with the three elements, $\{E_+, E_-, E_1\}$ where

$$E_+ = \frac{1}{6}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2\sqrt{3}}(|0\rangle\langle 1| + |1\rangle\langle 0|) \quad (4.23)$$

$$E_- = \frac{1}{6}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| - \frac{1}{2\sqrt{3}}(|0\rangle\langle 1| + |1\rangle\langle 0|) \quad (4.24)$$

$$E_1 = \frac{2}{3}|0\rangle\langle 0| \quad (4.25)$$

corresponding to detection in modes $\{1, 2, 3 \text{ or } 4\}$ respectively. Since it is not possible to detect in modes 3 and 4 (these are inputs), we restrict to events where the photon is detected in modes 1 and 2, which occur with relative probabilities given by

$$\text{Pr}_+ = \frac{\text{Tr}(E_+\rho)}{\text{Tr}(E_+\rho) + \text{Tr}(E_-\rho)}, \quad \text{Pr}_- = \frac{\text{Tr}(E_-\rho)}{\text{Tr}(E_+\rho) + \text{Tr}(E_-\rho)} \quad (4.26)$$

allowing for the estimation of the value

$$\text{Pr}_+ - \text{Pr}_- = \frac{\sqrt{3}\langle X \rangle}{\langle 0|\rho|0\rangle + 3\langle 1|\rho|1\rangle} \quad (4.27)$$

such that, with additional knowledge from the Z measurement, we are able to estimate the expectation of X .

Since the experiment would require both input and detection in mode 1, we implement it in two stages, first inputting in mode 1 and detecting four-fold coincidences in modes $\{2,3,4,5\}$ and $\{2,3,4,6\}$, then inputting in mode 2 and detecting four-fold coincidences in modes $\{1,3,4,5\}$ and $\{1,3,4,6\}$ (note that by inputting in mode 2, the state would pick up a relative π phase shift so this is offset by setting $\theta_{3,5} = \frac{3\pi}{2}$). The two data sets are then normalised with respect to each other by the ratio of total counts in each experiment. Measurements with the Sagnac loop were taken for 60 hours, whilst measurements without were taken for 16 hours, collecting in total around 200 four-fold coincidences for each experiment.

Estimating the expectations $\langle Z \otimes Z \rangle$ and $\langle X \otimes X \rangle$ allows the entanglement of the generated state to be verified. Since

$$\mathcal{F}_{\min} = \frac{1}{2}(\text{Tr}[\rho Z \otimes Z] + \text{Tr}[\rho X \otimes X]) \quad (4.28)$$

$$= \langle \phi^+ | \rho | \phi^+ \rangle - \langle \psi^- | \rho | \psi^- \rangle \quad (4.29)$$

$$= \mathcal{F}(\rho, |\phi^+\rangle) - \mathcal{F}(\rho, |\psi^-\rangle) \quad (4.30)$$

$$\leq \mathcal{F}(\rho, |\phi^+\rangle) \quad (4.31)$$

and any state with $\mathcal{F}(\rho, |\phi^+\rangle) > 1/2$ is entangled [181], $\mathcal{F}_{\min} > 1/2$ is an entanglement witness. We find that $\mathcal{F}_{\min} = 0.673 \pm 0.031$ and therefore certify to a high confidence that the experiment does indeed generate heralded entanglement.

4.4.3 Characterisation of Linear Optical Networks

Comparing the four-photon heralded gates and the single and two photon experiments, it is clear that the quality of the quantum logic is significantly diminished in the higher photon number experiments. One unique property of photonic gates however is that they can be characterised independently of the photonic qubits they act upon and, in principle, with only classical states of light. Full process tomography of a linear optical gate, though necessarily inefficient, can be performed using multimode coherent states as probes and homodyne measurement [182, 183]. However, if the assumption is made that the linear optical network perfectly preserves coherence and applies the same transfer matrix M to all relevant optical modes, then this problem can be reduced to the far simpler task of estimating this transfer matrix.

The transfer matrix can be estimated in a number of ways. It is possible to do so using coherent state inputs and intensity measurements [184]. These approaches require additional interferometry before or after the gate however. Here we use an approach which is motivated by current experimental regimes, where light is generated and detected off-chip and all interferometry is performed within the integrated circuit itself. Using single photons and pairs of photons at the input ports the transfer matrix can be reconstructed.

The protocol we use, based on Ref. [185], first estimates the absolute value of each element of M . In principle $|M_{ij}|^2$ can be directly estimated by inputting a single photon in mode j and measuring the probability of it exiting from mode i . However, this estimate includes several sources of loss which we may wish to separate. Firstly, there may be input-mode dependent loss arising from fluctuations in intensity in the probe state or coupling into the device. Similarly, there may be output-mode dependent loss arising from coupling out of the device and differing detector efficiencies. Finally, there is the possibility of path-dependent loss through the circuit itself.

If the path-dependent loss is negligible in comparison to the input and output losses, i.e. the LON is well-described by a transformation proportional to a unitary U , then it is possible to infer the absolute values $|U_{ij}|$ from the matrix of single photon counts. Sinkhorn's theorem [186] states that any nonnegative matrix can be decomposed as the

product of two diagonal matrices and a doubly stochastic matrix ⁴

$$|M|^2 = L_{\text{out}}|U|^2L_{\text{in}} \quad (4.32)$$

where L_{in} and L_{out} are diagonal matrices quantifying the input and output losses respectively. The Sinkhorn-Knopp algorithm [187], which consists of iteratively normalising the columns and rows of $|M|^2$ until convergence, can then be used to obtain $|U|$ from the measured $|M|$.

However, when using circuits where different paths experience differing circuit depths, such as in the Reck scheme, this assumption is likely not to be valid. Therefore, we must attempt to characterise the input and output losses independently. This is done by estimating the relative detector efficiencies by the intensities measured when configuring the device to send all light to each detector in turn. The input losses can be estimated by monitoring the single photon generation rate, e.g. via beamsplitting part of the input beam directly to detectors. Adjusting the single photon count matrix to take these measured relative losses into account and then renormalising the result by a single scalar equal to the largest total countrate across all experiments then provides an estimate of $|M|$.

In order to learn the phases $\{\alpha_{ij}\}$ of the matrix elements $M_{ij} = |M_{ij}|e^{i\alpha_{ij}}$, pairs of photons can be used in a HOM-type configuration. Since the input and output phases of a LON are only defined relative to any pre- or post-interferometry, these phases can be set such that the first column and row of M are real without loss of generality. Inputting and detecting one photon in the first mode results in experimental probabilities given by

$$\text{Pr}(1i|1j) = |M_{11}|^2|M_{ij}|^2 + |M_{i1}|^2|M_{1j}|^2 + 2|M_{11}||M_{ij}||M_{i1}||M_{1j}|\cos(\alpha_{ij}) \quad (4.33)$$

where the photons are input in modes 1 and j and detected in modes 1 and i . Performing HOM dips in this configuration then allows $|\alpha_{ij}|$ to be estimated from the visibility $\mathcal{V}$ of

⁴A matrix for which all columns and rows sum to 1

the dip (or anti-dip) and the known $|M|$ matrix via

$$|\alpha_{ij}| = \arccos \left[\frac{-(|M_{11}|^2 |M_{ij}|^2 + |M_{i1}|^2 |M_{1j}|^2) \mathcal{V}(1i|1j)}{2|M_{11}||M_{ij}||M_{i1}||M_{1j}|} \right]. \quad (4.34)$$

Finally, the sign of the phases can be deduced from a further set of two-photon experiments. For an experiment in which photons are input in modes j, j' and detected in i, i' the quantity $|\alpha_{i'j'} - \alpha_{ij'} - \alpha_{i'j} + \alpha_{ij}|$ can be estimated. The signs can then be calculated as

$$\text{sgn}(\alpha_{ij}) = \text{sgn} \left(||\alpha_{i'j'} - \alpha_{ij'} - \alpha_{i'j} + \alpha_{ij}| - |\alpha_{i'j'} - \alpha_{ij'} - \alpha_{i'j} - |\alpha_{ij}|| \right) \quad (4.35)$$

$$- ||\alpha_{i'j'} - \alpha_{ij'} - \alpha_{i'j} + \alpha_{ij}| - |\alpha_{i'j'} - \alpha_{ij'} - \alpha_{i'j} + |\alpha_{ij}|| \right). \quad (4.36)$$

We are free to choose $\text{sgn}(\alpha_{22}) = +1$ and then the set of experiments which measure $\mathcal{V}(2i|12), \mathcal{V}(12|2j)$ and $\mathcal{V}(2i|2j)$ where i, j run $[3, m]$ are sufficient. The full procedure requires m^2 single photon measurements (for which only m experimental configurations are required) and $2(m-2)(m-1)$ two-photon measurements (for which $2m-3$ experimental configurations are required).

We performed this characterisation protocol on the three two-qubit gates described in the above (these reconstructions can be found in appendix A.1). To compare these we use the *circuit fidelity* [170]

$$F_c = \left| \frac{\text{Tr}(U^\dagger M_{\text{est}})}{\sqrt{m \text{Tr}(M_{\text{est}}^\dagger M_{\text{est}})}} \right|^2 \quad (4.37)$$

to assess the distance between the desired unitary transformation U and implemented transfer matrix. This can be thought of as the fidelity between the state produced when light is input in a single port of the device and the target state, averaged over all ports. We find $F_c = 0.996 \pm 0.004, 0.969 \pm 0.018$ and 0.988 ± 0.040 for the Bell state generator, the KLM CNOT gate and the postselected CNOT gate respectively. Using these reconstructions, it is possible to estimate the logical error introduced via the optical circuitry. Considering for instance, the Bell state generator circuit, assuming only the coherent error

in the transfer matrix

$$\langle 11|_{23}\mathcal{U}(M_{\text{est}})|011110\rangle \propto |\psi_M\rangle \quad (4.38)$$

we find a state fidelity of $\mathcal{F}(|\psi_M\rangle, |\phi^+\rangle) = 0.988$ and a concurrence $C(|\psi_M\rangle) = 0.998$. This shows that the circuit itself is capable of producing highly entangled states if acting upon ideal input states. We further note that all of these circuits were obtained via the individual phase shifter calibrations and performing further optimisation on the applied voltages is likely to increase these fidelities [188]. This highlights that the logical error observed in the four-photon experiments can be largely attributed to a reduction in the quality of the quantum interference, as well as errors introduced due to accidental counts and imperfect detection. For instance the HOM visibility between off-pair photons is ~ 0.9 , which we will see in Chapter 6 already limits the concurrence of the Bell state generator to ~ 0.8 .

4.5 Discussion

In this chapter we have demonstrated the experimental operation of a universal linear optical circuit and its use as a testbed for photonic quantum logic schemes, including new entanglement generation schemes and the first demonstration of measurement-induced nonlinearity in an integrated platform. The versatility of the device allows us to switch between all these experiments in seconds, providing a platform for rapidly prototyping new and existing schemes. We demonstrated high performance of the linear optical transformation over different configurations via transfer matrix characterisation, showing the potential for the device to produce highly entangled heralded states and identifying imperfect quantum interference and detection as the main sources of experimental error.

Statement of Work

Much of the work reported in this chapter was published in Carolan et. al. *Science*, 349, 711 (2015) [165]. The universal linear optical device was conceived by A. Laing and fabricated by N. Matsuda and colleagues at NTT. Characterisation, calibration and

data taking were largely performed by J. Carolan and C. Harrold with input from N. Matsuda and myself. The Process tomography experiments were designed and analysed by myself. The KLM gate experiment was designed by myself and analysed by C. Harrold and J. Carolan. The Bell state generator circuit was conceived by myself, the experiment designed by myself, J. Carolan and J. Matthews and analysed by myself. The transfer matrix characterisation methods were initially conceived by A. Laing, developed by J. Carolan and myself and analysed by myself.

Chapter 5

Simulating Vibrational Quantum Dynamics of Molecules with Integrated Photonics

5.1 Introduction

In the previous chapter we demonstrated a powerful new capability for quantum photonics, universal linear optical circuits. We went on to demonstrate their use as a testbed for experimentally investigating linear optical logic gates. Here we propose and experimentally demonstrate another application for such devices as a new platform for quantum simulations.

Computer simulations that model physical systems have revolutionised modern science, technology and engineering. Before the advancement of digital computers, early analog computers that mapped the controllable equations of motion of one system to mimic another helped solve many real world problems from simulating ballistics and rocket science to modelling tidal systems and power networks [189].

Analog quantum computers are devices which can directly simulate the quantum dynamics of systems in nature via tuning an equivalent Hamiltonian of controllable quantum systems in the lab. This analogical approach aims to maximise the complexity and utility

of naturally evolving quantum systems without the vast challenges and overheads required to engineer them into individually addressable qubits and gate sequences. Although necessarily limited in scale and versatility, such devices may find uses in the near term as well as representing an important milestone on the road to universal digital quantum computation. Here we demonstrate programmable integrated quantum photonics as a simulation platform for models of vibrational molecular quantum dynamics using the analogies of optical modes in integrated waveguides for vibrational modes in molecules and photons for quantised vibrational excitations.

Progress in modelling, controlling, and observing the ultrafast dynamics of molecules, has revealed the underpinning mechanism of interference among vibrational modes in important behaviour such as bond selective chemistry [190]. Increasing levels of control over vibrational wavepackets has allowed selective dissociation governed by a single quantum of vibrational energy [191], manipulation of individual molecules at ambient conditions [192], preparations of coherent superpositions on sub-femtosecond timescales [193], and single vibrational states of ultracold molecules [194, 195]. Molecular dynamics are now observable on their ultrafast intrinsic timescale [196, 197].

The prospect of more sophisticated control with quantum states of light and larger molecules brings additional challenges for modelling however. Evolving a multi-excitation state across many vibrational modes described by independent quantum harmonic oscillators is equivalent to the Boson sampling model and therefore computationally inefficient. More sophisticated molecular models, for example with anharmonic corrections to the potentials, are also likely to be computationally complex.

In this chapter we first demonstrate that integrated photonics makes a platform for the efficient simulation of vibrational molecular dynamics in the harmonic approximation. We show how to go beyond this model by simulating thermal relaxation using additional ancilla optical modes and anharmonicities using measurement-induced nonlinearities. Finally, we use adaptive feedback control with our simulator in a closed loop for finding optimised initial states for the dissociation of NH_3 , demonstrating the setup as a testbed for control theory.

5.2 Molecular Vibrational Dynamics in the Harmonic Approximation

A quantum mechanical description of molecular vibrations begins with the Born-Oppenheimer approximation (BOA) [198], the assumption that the nuclear and electronic dynamics can be separated due to their difference in masses and therefore timescales. The full molecular Hamiltonian can be written as

$$\hat{H}_{\text{mol}} = T_n + T_e + V_{n,n} + V_{n,e} + V_{e,e} \quad (5.1)$$

where T refers to kinetic energy, V potential energy, n the nuclei and e the electrons. The BOA separates the wavefunction into a nuclear part $\chi(\mathbf{R})$, where $\mathbf{R}$ is a vector of the nuclear positions and an electronic part $\psi(\mathbf{r}, \mathbf{R})$ where $\mathbf{r}$ is a vector of the electron positions. An adiabatic approximation is then made whereby the electronic time-independent Schrödinger equation is solved for fixed configurations of the nuclei $\mathbf{R}$ resulting in the electronic potential energy surface (PES) $E_e(\mathbf{R})$. The nuclear Schrödinger equation then reduces to

$$[T_n + E_e(\mathbf{R})]|\chi(\mathbf{R})\rangle = E_{\text{total}}|\chi(\mathbf{R})\rangle. \quad (5.2)$$

To solve this equation, the dependence of the PES is first expressed in mass-weighted cartesian coordinates $\tilde{x}_i = \sqrt{m_i}x_i$ where i labels both the nucleus $\in 1, \dots, N$ and coordinate $\in \{x, y, z\}$. The PES is then expressed as a Taylor series around the nuclear equilibrium structure $\tilde{\mathbf{R}} = \mathbf{0}$,

$$E_e(\tilde{\mathbf{R}}) = E_e(\mathbf{0}) + \sum_{i=1}^{3N} \left(\frac{\partial E_e(\tilde{\mathbf{R}})}{\partial \tilde{x}_i} \right)_{\tilde{\mathbf{R}}=\mathbf{0}} \tilde{x}_i + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 E_e(\tilde{\mathbf{R}})}{\partial \tilde{x}_i \partial \tilde{x}_j} \right)_{\tilde{\mathbf{R}}=\mathbf{0}} \tilde{x}_i \tilde{x}_j + \dots \quad (5.3)$$

Since the energy gradients are zero at the equilibrium (minimum energy) structure, setting $E_e(\mathbf{0}) = 0$ and using the harmonic approximation in which we truncate to second order,

the Hamiltonian in Eq. 5.2 becomes

$$-\frac{1}{2} \sum_{i=1}^{3N} \frac{\partial^2}{\partial \tilde{x}_i^2} + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 E_e(\tilde{\mathbf{R}})}{\partial \tilde{x}_i \partial \tilde{x}_j} \right)_{\tilde{\mathbf{R}}=\mathbf{0}} \tilde{x}_i \tilde{x}_j. \quad (5.4)$$

The Hessian matrix with elements $\mathbf{H}_{ij} = \left(\frac{\partial^2 E_e(\tilde{\mathbf{R}})}{\partial \tilde{x}_i \partial \tilde{x}_j} \right)_{\tilde{\mathbf{R}}=\mathbf{0}}$ can be found using standard numerical electronic structure methods from quantum chemistry for large molecules with hundreds of vibrational modes [199]. Diagonalising the Hessian produces a set of $3N$ mass-weighted normal coordinates $\{Q_k\}$ for which the Hamiltonian reduces to that of $3N$ independent quantum harmonic oscillators

$$\sum_k \frac{\partial^2}{\partial Q_k^2} + \omega_k Q_k^2. \quad (5.5)$$

The unitary matrix $\mathbf{Q}$ which diagonalises $\mathbf{H}$ has six eigenvectors with zero eigenvalues corresponding to molecular translation and rotation. The remaining $3N - 6$ eigenvectors correspond to the normal modes of the molecule and their eigenvalue the vibrational frequency ω_i . Finally, we can express Eq. (5.5) with creation operators that create quantised vibrational excitations in the vibrational modes; $\hat{H} = \sum_i \hbar \omega_i a_i^\dagger a_i$ (omitting zero point energy). We can finally then map this system of non-interacting bosonic modes to our linear optical device.

The spatial localisation of vibrational energy is important for understanding many molecular phenomena such as energy transport and dissociation. We therefore consider a basis transformation, $a_i^\dagger \rightarrow \sum_j U_{ji}^L a_j^\dagger$, from the collective normal modes to a set of modes which are localised around a single atomic site or chemical bond. These localised modes are found numerically using a Pipek-Mezey type method [200]. The kinetic energy contribution from a nucleus p in a normal mode i is given by $K_{p,i} = \sum_{\alpha=x,y,z} Q_{p\alpha,i}^2$. The localised basis is then found by maximising the sum of squares of these kinetic energy contributions over all unitary transformations $Q'_{k\alpha,i} = \sum_j U_{ji}^L Q_{p\alpha,j}$

$$\max_{U^L} \left[\sum_i^{3N-6} \sum_k^{3N} K_{p,i}'^2 \right]. \quad (5.6)$$

Once the localised basis is found, dynamics in this basis can then be tracked by simulating the Hamiltonian $\hat{H}_L = \sum_{k,j} H_{kj}^L a_k^\dagger a_j$ where $H_{kj}^L = \sum_i \hbar \omega_i U_{ki}^L \overline{U_{ji}^L}$. Since the matrix H^L can be exponentiated efficiently in m (roughly in time $\mathcal{O}(m^3)$), reconfiguring a universal linear optical device to implement the transfer matrix $U(t_i) = \exp(-iH^L t_i/\hbar)$ for a series of timesteps $\{t_i\}$ allows the efficient simulation of the Hamiltonian $\hat{H}_L$ upon arbitrary multi-mode vibrational input states.

5.3 Experimental Procedure

The procedure for collecting experimental data begins with a set of experiments to measure detector efficiencies. A heralded single photon is input to the chip which is set to route the photon to each output mode. Collecting detection events for a fixed period of time for each output mode then allows us to estimate the relative detector efficiencies by dividing all the count rates by the largest count rate. The input state is then prepared and the experiment is performed by logging all coincidental detections at the output with the integrated chip set to perform the set of unitaries required for the simulation. These experimental counts are then corrected using the measured efficiencies by dividing the total counts in a detection pattern by the product of the relative detector efficiencies of the detectors in that pattern. These corrected counts are then used to estimate the expectation value of projecting the output state onto a detection pattern as the frequency of corrected counts in that pattern to the total number of corrected counts.

Finally, the statistical noise in the estimates of these probabilities resulting from the finite counting statistics is estimated via a bootstrapping approach, where 1000 sets of data are resampled from Poissonian distributions with means given by the experimentally measured counts, and the standard deviation of the probabilities estimated from these sets are used to provide error bars on plots.

Many of the experiments require the ability to resolve multiple occupancy in the output modes. Although number-resolving detectors were not available in this experiment we are able to circumvent this requirement by performing probabilistic number resolving between one and two photons using auxiliary fibre beamsplitters (FBS) and detectors.

This is achieved by inserting a FBS at each of the relevant output fibres and connecting both output modes to detectors; when there are two photons incident at the FBS there is then a probability that both the detectors generate a signal. For a beamsplitter with reflectivity r connected to detectors with efficiencies η_1 and η_2 , the probability of detecting a coincidental detection given a two-photon input is $2r(1 - r)\eta_1\eta_2$. Performing detector efficiency measurements as described above, the count rates in each detector will be proportional to their effective efficiencies, $r\eta_1$ and $(1 - r)\eta_2$. As before, relative efficiencies are then calculated by dividing by the maximum count rate and all coincidence counts are re-normalised by the product of these relative effective efficiencies. A final correction is then applied to account for multiple occupancy, for instance in a two photon experiment counts corresponding to the same output mode are halved, whereas counts corresponding to different output modes are averaged over the four possible ways of detecting such an output.

5.4 Simulating Four Atom Molecules

As a first example we consider Thioformaldehyde (H_2CS), a molecule with four atoms and six vibrational modes which is a prototype for many spectroscopic experiments [201]. For this, and all four atom molecules below, the equilibrium geometries and Hessian matrices were computed using CC2, an approximate coupled-cluster correlated wavefunction method, and the TVZPP atomic orbital basis sets [199, 202, 203].

The computed wavenumbers of the normal modes of H_2CS are

$$\{3195.45, 3102.74, 1501.77, 1061.45, 1017.08, 1008.04\}\text{cm}^{-1} \quad (5.7)$$

and the transformation between the normal modes and localised modes is found to be

$$U_L = \begin{pmatrix} 0.563 & 0.572 & 0.332 & 0.251 & 0. & 0.427 \\ -0.563 & 0.572 & 0.332 & 0.251 & 0. & -0.427 \\ -0.427 & -0.392 & 0.572 & 0.138 & 0. & 0.564 \\ 0. & -0.196 & -0.355 & 0.914 & 0. & 0. \\ 0. & 0. & 0. & 0. & 1. & 0. \\ -0.427 & 0.392 & -0.572 & -0.138 & 0. & 0.564 \end{pmatrix} \quad (5.8)$$

Fig. 5.1(a) visualises these normal modes as well as the localised modes which consist of two CH stretch modes, two CH bend modes, a wagging mode (in which both H atoms move out of the plane) and a CS stretch mode which are mapped in this order to the six waveguide modes of the photonic chip. We note that since the wagging mode is not coupled to the other modes in U_L , the resulting system is effectively five mode.

Injecting the two-mode squeezed vacuum state produced by one SPDC source into the two waveguides corresponding to the CH stretch modes, we track the combined probability of detecting all excitations in these CH stretch modes, in the CH bend modes and split between these four for the two- and four-photon subspaces. The left panel of Fig. 5.1(b) shows this evolution restricted to the two-photon subspace, we observe dynamics in which both excitations oscillate between stretches and bends via the intermediate states where one is in each. The L^1 distance between the theoretical and experimentally obtained distributions, averaged over all timesteps is $\bar{\mathcal{D}} = 0.064 \pm 0.028$. The right panel of Fig. 5.1(b) displays the evolution of the four quanta term, where now each of the stretch modes are initially doubly occupied, again combining the probabilities of detecting all excitations in stretches, bends and combinations for measurement patterns with up to two excitations per mode. We see a similar oscillatory behaviour between stretches and bends, but with an apparent suppression of the probabilities attributable to the state extending into the combinatorially growing space of multiple excitations. The full four photon distribution has a distance $\bar{\mathcal{D}} = 0.159 \pm 0.074$ to the ideal distribution.

One of the features of this analog simulator is that time is a programmable parameter.

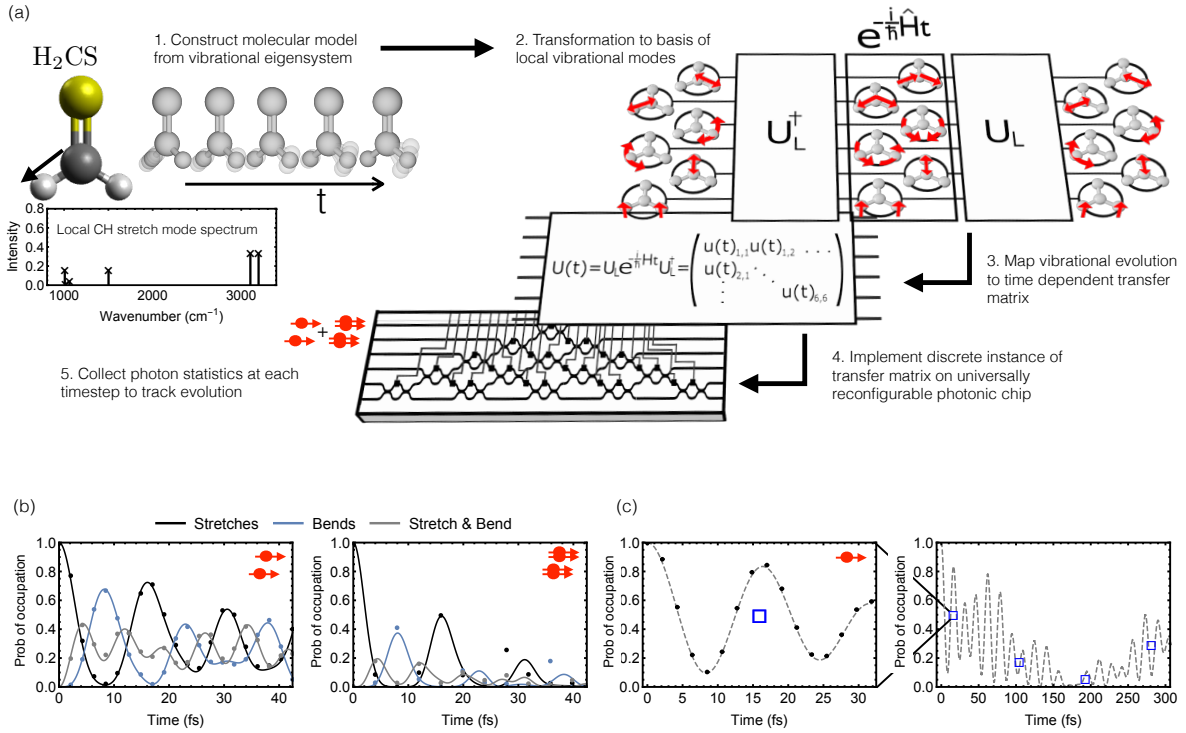


Figure 5.1: Simulating the vibrational dynamics of molecules in the harmonic approximation. (a) Steps for implementing a simulation using H_2CS as an example. (b) Evolution of a vibrational two-mode squeezed vacuum state in the localised CH stretch modes of H_2CS . Left panel displays the theoretical and experimentally estimated total probability within the two-excitation subspace of measuring both excitations in the CH stretch modes (black), CH bend modes (blue) and a combination of both (grey). Right panel shows the same for the four-excitation subspace with up to two excitations in each mode. (c) Short and long timescale evolution of a single vibrational excitation in a CH stretch mode of H_2CS . Measuring the probability of finding the excitation in the stretch mode by taking the mean (blue boxes) of measurements taken in narrow windows (left panel) spread over a longer timescale (right panel).

This allows us to probe dynamics at arbitrary evolution times. In Fig. 5.1(c) we provide an example of this by observing the probability for detecting a localised excitation in a CH stretch vibration given it is prepared in this mode at $t = 0$. As shown in Fig. 5.1(a), this stretch mode is a superposition of a set of lower and higher frequency components. Using the source to produce heralded single photons and measuring time steps at an increased rate in narrow windows across a longer evolution period we are able to observe both the high frequency behaviour within each window as well as the low frequency periodicity

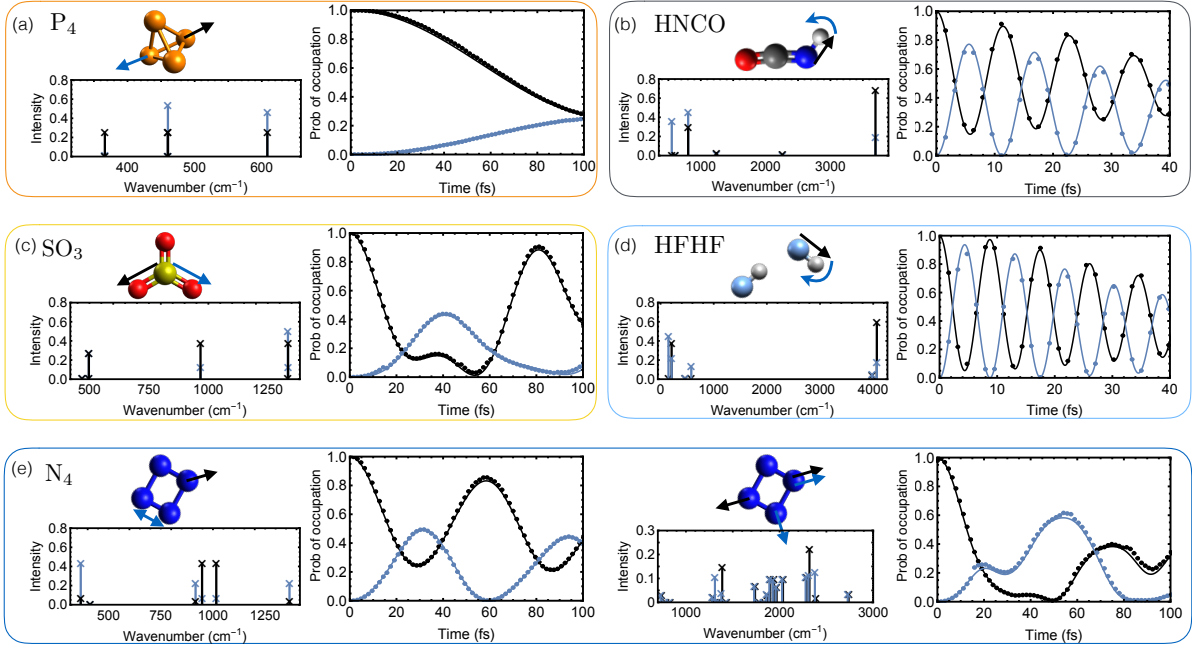


Figure 5.2: Evolution of localised vibrational excitations in four atom molecules. (a-d) Evolution of a single excitation in P_4 , SO_3 , HNCO and HFHF between a stretch mode (black) and another coupled local mode (blue). The modes involved are represented diagrammatically with arrows as well as by their spectral intensities alongside the theoretical evolution and experimental data. (e) Evolution of one and two-vibrational excitations in N_4 . Left, the relevant single excitation spectrum and data. Right, the relevant two excitation spectrum and data. In all plots error bars resulting from shot-noise in the experimental relative frequencies are too small to see.

by tracking the average probability of each window. The mean distance across these experiments is $\bar{D} = 0.014 \pm 0.006$. Additional one, two and three excitation data can be found in Appendix B.2.

Since general four atom molecules have $3 \times 4 - 6 = 6$ vibrational modes¹, our simulator can explore the dynamics of any such system. To demonstrate this we simulated the coherent dynamics of a series of four atom molecules which represent a wide range of molecular structures and energy scales, P_4 , SO_3 , HNCO , HFHF and N_4 . Figs. 5.2(a-d) show the time evolution of a single excitation initially prepared in a local stretch mode. The change in the occupation probability to a second, spectrally overlapped (‘coupled’) local mode is plotted. We observe dynamics with varying characteristic times governed

¹Except for linear molecules which have $3N - 5$ vibrational modes, as they have one fewer rotational mode.

by the vibrational spectra of the molecules. Due to the geometry and bonding structure of P_4 , it has the longest period oscillations between opposing stretches with SO_3 showing similar stretch coupling on shorter timescales. HNCO and HFHF in contrast display faster dynamics with increased mode coupling between hydrogen bond stretches and bends. The average L^1 distance over all these experiments is $\bar{D} = 0.022 \pm 0.007$. Fig. 5.2(h) shows the dynamics of both a single excitation and two excitations initially prepared in stretch modes of N_4 . The additional structure in the vibrational spectrum and the introduction of multi-photon quantum interference result in a more complex time dependence of the detection probabilities.

The vibrational spectra, U_L and additional data for these four atom molecules can be found in Appendix B.

5.5 Energy transfer and dephasing in NMA

The flow of vibrational energy in molecules is a fundamental process for chemical reaction rates [204] and functionality in biomolecules [205], while control of intramolecular energy transfer could lead to new nanotechnologies such as molecular electronics [206]. The vibrational quantum dynamics of a molecule within an environment can be described by the interplay of coherent unitary evolution and incoherent dephasing resulting from random fluctuations of the vibrational frequencies - a process referred to as spectral diffusion. N-methylacetamide (NMA) is the simplest molecular model (Fig. 5.3(a)) of the peptide bond in proteins, where quantum coherence may play a role in energy transfer [207]. In this section, we simulate a model for intramolecular energy transport in NMA in the presence of dephasing.

We consider a subspace spanning 6 (of its 30) backbone vibrational modes ² (see Appendix B for full details), which support a basis of approximately localised vibrations, including two rocking modes (curved arrows in Fig. 5.3) and two stretch modes (straight arrows in Fig. 5.3). These six local modes are mapped to our photonic chip, such that

²For this case density functional theory with the B3LYP functional and a split-valence atomic orbital basis set was used to calculate the Hessian

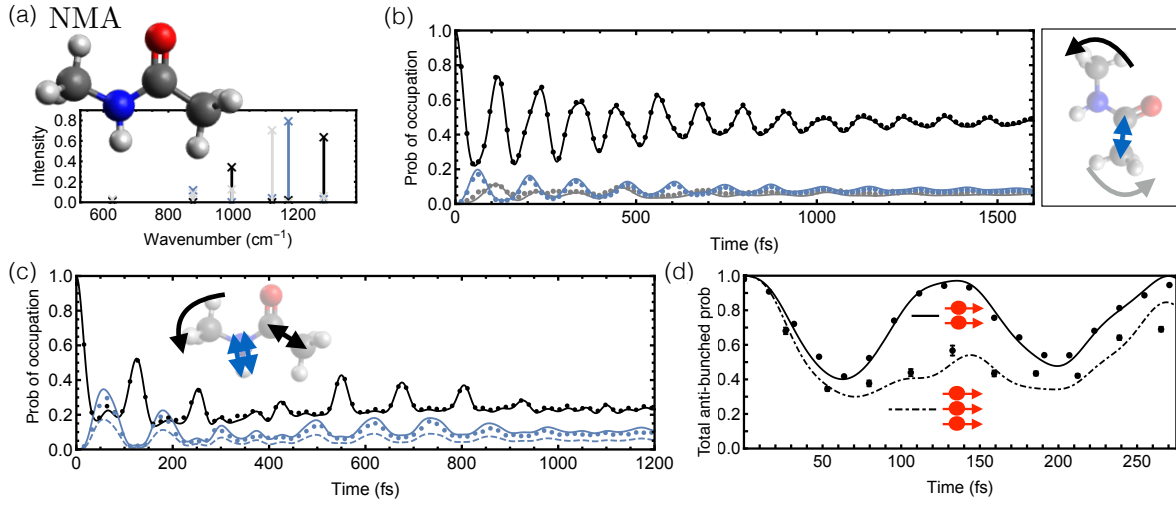


Figure 5.3: Quantum energy transfer and dephasing in NMA. (a) A six-mode vibrational subspace of the NMA molecule is considered with the spectral components of three localised modes (used in (b)) shown in the right panel. (b) Energy transfer of a single local vibrational excitation. Simulation of a single excitation initially in a local rocking mode at one end (black) and its transfer to local modes at the opposite ends (blue and grey) when subject to a dephasing channel with $T_2 = .53\text{ps}$. Solid lines represent theory and points represent experimentally estimated probabilities. (c) Quantum interference and transport of multiple vibrational excitations. Evolution of a two-excitation state initially in separate local modes (black) and its probability to be found bunched in the NH stretch mode for indistinguishable bosons (solid blue) and under the same dephasing channel. Also included is the theory curve for distinguishable particles (dashed blue). (d) The total probability for measuring a fully anti-bunched state (solid black) and the same for an initial state of three excitations (dot-dashed black) initialised in the modes from (b).

the coherent time evolution of an excitation can be tracked, following the same procedure as described in the previous section.

To simulate uniform dephasing across the six vibrational modes we consider an ad hoc model described by the Lindblad master equation

$$\dot{\rho} = \mathcal{L}(\rho) = -i[\hat{H}, \rho] + \frac{\Gamma}{6} \sum_{k=1}^5 \left[\hat{Z}(k)\rho\hat{Z}^\dagger(k) - \frac{1}{2}\{\hat{Z}^\dagger(k)\hat{Z}(k), \rho\} \right] \quad (5.9)$$

where $\hat{Z}(k) = \sum_j \exp(i2\pi jk/6)a_j^\dagger a_j$, i.e. $Z(k)$ are six dimensional Heisenberg-Weyl matrices and Γ^{-1} , usually referred to as the T_2 time, describes the characteristic coherence time between the modes. Solving this equation results in the channel

$\rho(t) = \sum_k M_k(t) e^{-i\hat{H}t/\hbar} \rho(0) e^{i\hat{H}t/\hbar} M_k^\dagger(t)$ with time-dependent Kraus operators

$$\begin{aligned} M_0(t) &= \sqrt{\frac{1 + 5e^{-\Gamma t}}{6}} \sum_{j=1}^6 a_j^\dagger a_j \\ M_{k>0}(t) &= \sqrt{\frac{1 - e^{-\Gamma t}}{6}} \sum_{j=1}^6 \exp\{i2\pi jk/6\} a_j^\dagger a_j \end{aligned} \quad (5.10)$$

Including the change of basis to the local vibrational modes, we can therefore simulate these dephasing dynamics for a set of vibrational modes with arbitrary T_2 times by statistical averaging the expectation values obtained from experiments implementing the transfer matrices $U(t, j) = U_L Z(j) e^{-iHt/\hbar} U_L^\dagger$ over j .

Using a single photon, we simulated the probability for a single excitation initialised in a local rocking mode at one end of the molecule to be transferred to two localised modes (a rocking mode and a C-C stretch mode) at the opposite end of the molecule. The experimental results shown in Fig. 5.3b show dynamics that are initially oscillatory, with vibrational energy transfer between the rocking modes at either end of the molecule via an intermediate C-C stretch. The increasing effect of the suppression of coherences from dephasing results in evolution toward a steady state with reduced peak probability for the excitation to be found in these localised modes. The T_2 time constant of coherence decay used is 0.53ps, though any time constant can be simulated by changing only the post-processing of data.

Generalising the simulation for initial states of multiple excitations allows us to investigate the combination of dephasing and quantum interference for multi-excitation energy transport. Injecting one photon into the waveguide corresponding to the rocking mode and another in the waveguide corresponding to the C-C stretch mode, which are each localised at opposite ends of the NMA molecule (black arrows in Fig. 5.3(c)), we simulated the probabilities for finding the excitations in this state, and the state in which both excitations bunch in an NH stretch mode (double blue arrows in Fig. 5.3(c)). The results in Fig. 5.3(c) show more complex oscillatory transfer between these separated and localised states, that again tends toward a steady state. However, after full dephasing has occurred, the probability for two excitations to be bunched in the NH stretch mode

is twice as high for excitations that behave as indistinguishable bosonic particles than for excitations that behave as distinguishable classical particles whose theoretical probability is also plotted.

For a given molecule, the probability that no bunching occurs (multiple excitations are not localised) generally decreases as the number of excitations increases [208]³. In Fig. 5.3(d) the probability for the subspace of no-bunching events is simulated for two and three excitations under fully coherent dynamics. The initial state for the two-excitation evolution is the same as in the previous example; the initial state for the three-excitation evolution comprises an excitation in each of the local modes shown in Fig. 5.3(b).

The average distances across all single, two and three excitation distributions in these examples are given by 0.017 ± 0.005 , 0.046 ± 0.014 and 0.139 ± 0.068 respectively.

5.6 Vibrational relaxation in H₂O

Although our model of a closed system of uncoupled harmonic oscillators accounts for much of the physics of molecular vibrations, we now consider extensions to better describe realistic systems by including energy dissipation and potential anharmonicities.

When a vibrational state is excited in a molecule, exchange of this energy via intra- and inter-molecular coupling to other degrees of freedom will eventually lead to thermalisation - a process known as vibrational relaxation. The vibrational relaxation pathways of H₂O remain an area of current investigation [209–211].

Here we simulate the relaxation of H₂O via an amplitude damping model, depicted in Fig. 5.4a. Consider the Lindblad master equation

$$\begin{aligned} \dot{\rho} = \mathcal{L}(\rho) = & -i[\hat{H}, \rho] + \sum_{j=1,2} \Gamma_1 \left(|3\rangle\langle j| \rho |j\rangle\langle 3| - \frac{1}{2} \{ |j\rangle\langle j|, \rho \} \right) + \Gamma_2 \left(|0\rangle\langle 3| \rho |3\rangle\langle 0| - \frac{1}{2} \{ |3\rangle\langle 3|, \rho \} \right) \\ & + \sum_{j=1,2} \Gamma_3 \left(|0\rangle\langle j| \rho |j\rangle\langle 0| - \frac{1}{2} \{ |j\rangle\langle j|, \rho \} \right) \end{aligned} \quad (5.11)$$

Where $|j\rangle \equiv a_j^\dagger |\mathbf{o}\rangle$ represent a single excitation in one of the three vibrational modes

³Averaged over all possible unitaries this probability goes as $\binom{m}{n} / \binom{m+n-1}{n}$ for n excitations in m modes.

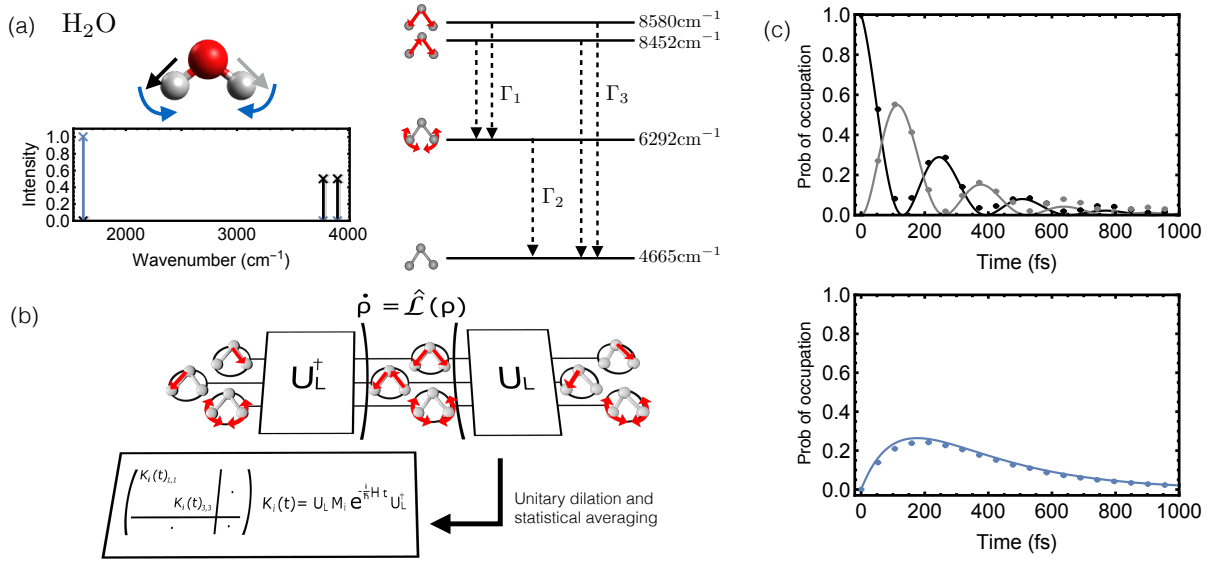


Figure 5.4: Vibrational relaxation in H₂O. (a) Energy level diagram for single excitation harmonic levels and ground state of H₂O along with the spectral components of the two local OH stretch modes (black and grey) and the symmetric bend normal mode (blue). (b) The open system dynamics resulting from a vibrational relaxation process can be simulated by statistical averaging of the evolution under a set of linear operators implemented via unitary dilation. (c) Evolution of probability for a single excitation initially in an OH stretch mode (black) to be found in the stretch modes (top panel) and symmetric bend (bottom panel) under relaxation dynamics.

and $|0\rangle \equiv |\mathbf{o}\rangle$ represents no vibrational excitations. Solving Eq. (5.11) results in time dependent Kraus operators

$$M_1(t) = \sqrt{1 - e^{-\Gamma_1 t}} |3\rangle\langle 1|, \quad M_2(t) = \sqrt{1 - e^{-\Gamma_1 t}} |3\rangle\langle 2| \quad (5.12)$$

$$M_3(t) = \sqrt{1 - e^{-\Gamma_2 t}} |0\rangle\langle 3|, \quad M_4(t) = \sqrt{1 - e^{-\Gamma_3 t}} |0\rangle\langle 1| \quad (5.13)$$

$$M_5(t) = \sqrt{1 - e^{-\Gamma_3 t}} |0\rangle\langle 2|, \quad M_0(t) = \sqrt{\mathbb{1} - \sum_{i=1}^5 M_i^\dagger(t) M_i(t)} \quad (5.14)$$

The evolution between the non-vacuum states can then be simulated by the set of transfer matrices K

$$K_{i<3}(t) = U_L M_i(t) e^{-iHt/\hbar} U_L^\dagger. \quad (5.15)$$

Since this channel is non-unital (it doesn't map $\mathbb{1} \rightarrow \mathbb{1}$), it cannot be simulated as a

convex sum of unitary evolutions as before. However, we can implement general linear transformations using the concept of unitary dilation at the cost of a non-unit probability of success. Any $m \times m$ matrix A with spectral norm $\|A\|_2 \leq 1$ can be embedded as the top-left block of a $2m \times 2m$ unitary matrix U_A . All possible unitary dilations of A can be parameterised using the Cosine-Sine decomposition [212]. By considering a singular value decomposition of A

$$A = U \cos(\vec{\theta}) V^\dagger$$

where $\cos(\vec{\theta}) = \text{Diag}(\cos \theta_1, \dots, \cos \theta_m)$ and $U, V^\dagger$ unitaries, we can express the possible dilations of A into $2m$ modes as

$$U_A = \begin{pmatrix} U & 0 \\ 0 & X \end{pmatrix} \begin{pmatrix} \cos(\vec{\theta}) & -\sin(\vec{\theta}) \\ \sin(\vec{\theta}) & \cos(\vec{\theta}) \end{pmatrix} \begin{pmatrix} V^\dagger & 0 \\ 0 & Y^\dagger \end{pmatrix}$$

where $X, Y \in U(m)$ are arbitrary unitaries. In this experiment we used the choice $X = Y = \mathbb{1}$. Using this technique, we are then able to simulate each of the set of transformations K using their dilations U_K , and postselecting upon detection in the top three modes of the device.

The terms involving M_{3-5} can be simulated via introducing relative losses to the modes. This could be achieved by coupling to further ancilla modes and this time a detection in these modes would correspond to the projection $|0\rangle\langle 0|$. Here instead, we simply apply a time-dependent effective efficiency to the detected counts on each mode. Finally, $M_0(t)$ is diagonal and commutes with U_L , meaning that it can also be achieved via a rescaling of the detected counts from the ideal unitary evolution $U_L e^{-iHt/\hbar} U_L^\dagger$.

Statistically mixing these experiments results in the correct expectation values and states $\rho(t)$ for the thermalisation process. We used $\Gamma_1 = 0.24\text{ps}$, $\Gamma_2 = 0.26\text{ps}$ and $\Gamma_3 = 1.36\text{ps}$, which are the experimentally measured relaxation times from Ref. [210].

In Fig. 5.4c we simulate the thermalisation of a local OH stretch excitation via the symmetric bend normal mode to its ground state. The probability of measuring the excitation in the two local stretch (left panel) modes and the symmetric bend (right panel) demonstrates the transfer of population from the high energy stretch modes via the lower

energy bend mode to the ground state. The average L^1 distance in these experiments was $\bar{D} = 0.024 \pm 0.007$.

5.7 Anharmonic Hamiltonian for H₂O

Going beyond the harmonic approximation we now consider the simulation of vibrational modes described by coupled anharmonic oscillators. Again considering H₂O, including all third derivatives and the semi-diagonal quartic derivatives in the PES and applying vibrational perturbation theory [213, 214] yields an adjusted Hamiltonian

$$\hat{H}_a = \hat{H} + \sum_{i \leq j} \hbar x_{ij} \sqrt{\omega_i \omega_j} / 2 (a_i^\dagger a_i + a_j^\dagger a_j + 2a_i^\dagger a_j^\dagger a_i a_j) \quad (5.16)$$

where $\hat{H}$ is the harmonic approximation Hamiltonian and x_{ij} are the perturbation theory coefficients.

In principle the higher order nonlinear terms in $\hat{H}_a$ can be simulated directly via weak optical nonlinearities such as self-phase modulation and cross-Kerr effects. Here instead we demonstrate an approach based on measurement-induced nonlinearity.

It is possible to implement a conditional phase shift on a two-photon Fock state using an additional ancilla photon and postselection upon detection of a specific heralding signal in additional optical modes. If the phase is π , this is the KLM NS gate [114]. Using generalised nonlinear phase-shift (NPS) gates we are able to implement arbitrary phase shifts between the zero, one and two Fock states of an optical mode.

In order to implement the NPS gates, we wish to perform diagonal unitaries in the Fock basis characterised by the evolution

$$U_{\text{NPS}}(\phi_1, \phi_2) [a_0|0\rangle + a_1|1\rangle + a_2|2\rangle] = a_0|0\rangle + a_1 e^{i\phi_1}|1\rangle + a_2 e^{i\phi_2}|2\rangle \quad (5.17)$$

for a single optical mode. We will search for gates which use a single ancilla photon, injected in an additional mode. In this case it is possible to then restrict the search to arbitrary 2×2 transfer matrices M which can be extended to unitaries using the principle

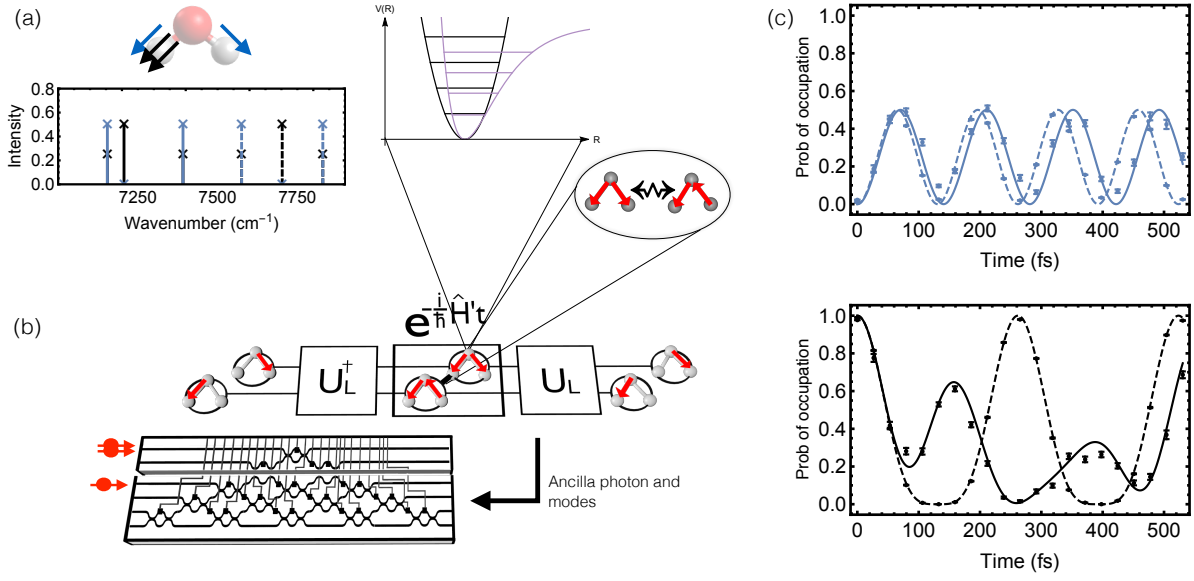


Figure 5.5: Anharmonic model for H₂O. (a) Spectrum of two excitations in bunched (black) and anti-bunched (blue) local stretch modes for a harmonic (dashed) and anharmonic model (solid). (b) The anharmonic evolution is implemented via measurement-induced nonlinearity using an ancilla photon and modes. (c) Evolution of the two bunched local stretch excitations to be found in the anti-bunched state (top panel) and the bunched state (bottom panel) under both models.

of dilation explained above. By considering the equations that must be satisfied

$$\begin{aligned} \text{Per}[M_{(0,1),(0,1)}] &= e^{-i\phi_1} \text{Per}[M_{(1,1),(1,1)}] \\ 2\text{Per}[M_{(0,1),(0,1)}] &= e^{-i\phi_2} \text{Per}[M_{(2,1),(2,1)}] \end{aligned} \quad (5.18)$$

the set of gates can be parameterised as

$$M = \begin{pmatrix} 1 - \sqrt{1 - e^{i\Phi}} & x \\ \frac{y}{x} \sqrt{1 - e^{i\Phi}} & y \end{pmatrix} \begin{pmatrix} e^{i\phi_1} & 0 \\ 0 & 1 \end{pmatrix} \quad (5.19)$$

where $\Phi = \phi_2 - 2\phi_1$. Optimal gates can then be found by performing the optimisation

$$\begin{aligned} \max P_{\text{succ}} \\ \text{subject to: } ||M||_2 \leq 1 \end{aligned} \quad (5.20)$$

where $P_{\text{succ}} = |y|^2$ is the probability that the gate succeeds in performing the desired

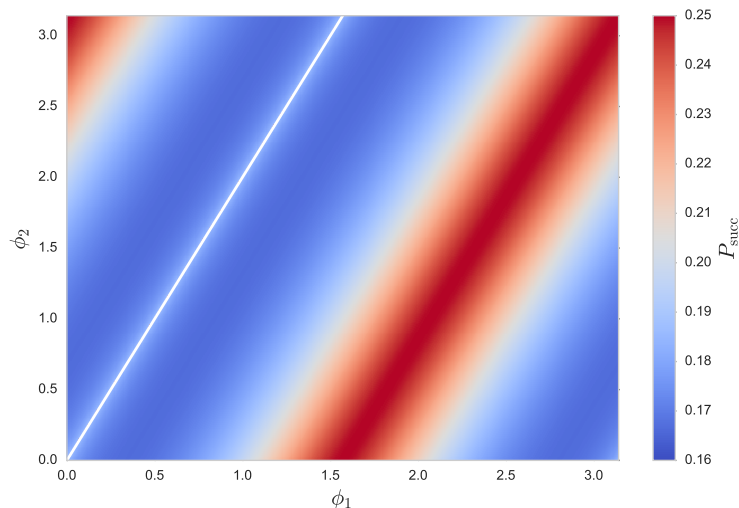


Figure 5.6: Optimised success probability (P_{succ}) for a nonlinear phase shift (NPS) gate mapping performing the transformation using two ancilla modes and one ancilla photon. The white line is the curve $\phi_2 = 2\phi_1$, along which $P_{\text{succ}} = 1$.

transformation. The optimisation is performed by using the sequential least squares programming (SLSQP) [215] algorithm included in the SciPy [216] python package.

Fig. 5.6 shows the success probabilities for the optimised gates for pairs of phases $\{\phi_1, \phi_2\}$. Beyond the trivial case of $\{\phi_1, 2\phi_1\}$, where $P_{\text{succ}} = 1$, globally optimal solutions occur for $\{\phi_1, \phi_1 + \pi\}$ where $P_{\text{succ}} = 1/4$. These solutions can be seen to be generalisations of the KLM NS gate. For the set of worst case $\{\phi_1, \phi_2\}$ pairs, we find that $P_{\text{succ}} = 1/6$.

Ref. [115] establishes an upper bound on the probability of success of NPS gates of the form $\{0, \phi\}$ with arbitrary ancilla states of $P_{\text{succ}} \leq (3 - \cos(\pi - \phi))^2/16$. For the case when $\phi = \pi$ this bound is saturated by the KLM NS gate. However, from our numerical search we found that in restricting to such single photon ancillas, this upper bound was not generally achievable. Furthermore, we observed that these probabilities were not improved by considering separable ancilla states containing two photons. We leave it as an important open problem to establish whether the upper bound from [115] is achievable and what form of ancilla state is required. If the bound is achievable with reasonable additional physical resources, this measurement-induced approach may be preferable to a genuine analog nonlinear photonic simulator.

We simulated both harmonic and anharmonic evolution, again using H_2O , this time

restricting to the stretch mode subspace. Two photons injected together into the top mode of the chip serves to simulate two excitations initialised in a superposition of the harmonic model eigenstates corresponding to a local OH stretch. As shown in Fig. 5.5(b), when simulating the anharmonic model, this input state can now be understood as the same superposition of the new energy eigenstates of $\hat{H}_a$, while a third photon injected into the third optical mode serves as the ancillary system.

Fig. 5.5(c) shows the results for simulating the probabilities for these two vibrational excitations to remain bunched or anti-bunch, under both the harmonic model, as before, and the anharmonic model using the NPS gate. The difference in the detection patterns between the two models, $\hat{H}$ and $\hat{H}_a$, is a function of their different spectra, shown in Fig. 5.5(a). The probability to detect a single excitation in each of the modes (i.e. anti-bunched), shown in the top panel of Fig. 5.5(c), acquires a simple frequency shift for the anharmonic evolution corresponding to the adjusted energy levels, as seen in Fig. 5.5(a). In contrast, the probabilities for the state to remain doubly occupied display markedly different dynamics between the harmonic and anharmonic cases, as shown in the bottom panel of Fig. 5.5(c). This is a result of the three vibrational eigenstates no longer being equally spaced in energy, as seen in Fig. 5.5(a), introducing new frequencies in the evolution. For this set of experiments, the average distances between the ideal and experimental distributions for the harmonic and anharmonic cases are 0.02 ± 0.01 and 0.06 ± 0.02 respectively.

5.8 Adaptive feedback control in the dissociation of NH_3

We finally consider the application of our photonic simulator as a platform for testing approaches to quantum control. Adaptive feedback control (AFC) has been proven to be a practical laboratory technique in finding optimal control fields for selective chemistry. In AFC, a pulse-shaped control field is iteratively refined based on the evaluation of results from molecular measurements. An algorithm evaluates each control setting based on its

measured outcome with respect to a pre-defined goal, and searches through the available space of control parameters to move towards a locally optimal solution. This procedure naturally incorporates the realistic laboratory environment and constraints.

Our example molecule is ammonia, NH_3 , a prototype for studying dissociation, including vibrationally mediated pathways, in which the states of its products, $\text{NH}_2 + \text{H}$ depend on the prior vibrational state in the ground electronic state [191]. We therefore consider the following model; the molecule is initially in its ground electronic state with some vibrational energy among its normal modes. Through interaction with a light field an electronic state is excited. We can obtain the resulting vibrational state by projecting the vibrational ground state modes onto the vibrational modes of the excited state. This transformation can in general be achieved via single mode squeezing, displacement and linear optic transformations [217], here we approximate the projection by a unitary transformation between the modes U_{GE} . The transformation used to approximately model the projection of the vibrational modes of the ground state Q_G to those of the excited state Q_E is the one that maximises the overlap

$$U_{\text{GE}} = \min_U ||Q_E - Q_G U||_2. \quad (5.21)$$

The vibrational state then evolves in time under the Hamiltonian corresponding to the excited state normal modes⁴. Probing the resultant state in a localised basis as before allows us to identify three local NH stretch modes.

The specific aim of this simulation, schematically depicted in Fig. 5.7(c), is to let an AFC algorithm find initial states of two vibrational excitations (in the electronic ground state molecule) which result in a maximal total probability to bunch in any of the three NH stretch modes (of the electronically excited molecule) over the first 10fs of evolution, which we associate with a preferred dissociation pathway, whilst suppressing other bunched events which we associate with other pathways.

The algorithm begins with an initial state of two photons in orthogonal optical modes.

⁴The planar equilibrium geometry and Hessian matrix of the first electronic excited state of NH_3 were computed at the CC2/TZVPP level of theory, where the Hessian was computed by finite difference using the nuclear gradients of the excited electronic state [218]

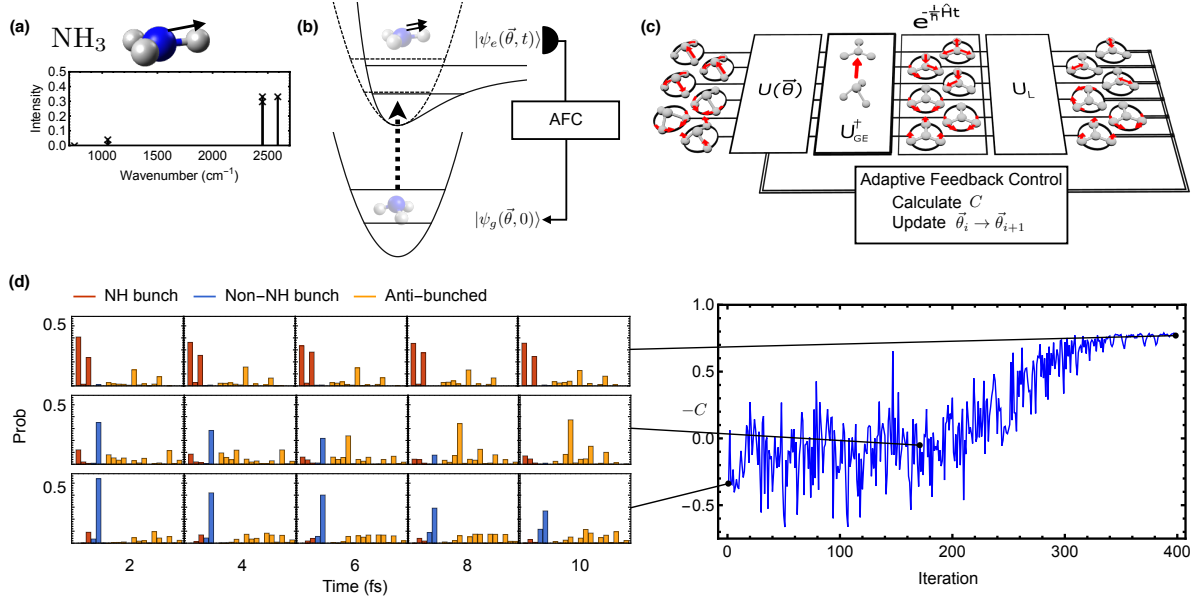


Figure 5.7: Adaptive feedback control algorithm for selective dissociation in NH_3 . (a) A two-excitation vibrational state, parameterised by $\vec{\theta}$, is initialised in the ground state vibrational modes of NH_3 . The electronic state is excited and the localisation of vibrational energy is measured over time. These measurements are then used to feedback to the state preparation in order to increase energy localisation in NH stretch modes, promoting a break-up of the molecule. (b) This scenario is simulated via the state preparation $U(\vec{\theta})$, a transformation between the ground and excited modes $U_{\text{GE}}^\dagger$, evolution under the excited state modes and measurement via U_L . (c) An example dataset showing (left panel) the initial probabilities with a random input state for bunching in the NH stretch modes (red), bunching in the remaining three localised modes (blue) and detection in anti-bunched patterns (yellow) for five timesteps for three iteration numbers. The right panel shows the measured cost function at every iteration.

We then imagine we are able to prepare a set of states defined by this state subject to a unitary mixing of the vibrational modes, $U(\vec{\theta})$.

We ignore the sixth mode as this remains uncoupled from the others (see Appendix B.1) and perform an arbitrary unitary transformation upon the remaining five modes. There are 14 free parameters, θ_i , in this state preparation, which are mapped to the phase shifts and beamsplitter splitting ratios within a linear optical network. These are the parameters which are optimised in the experiment. This parameterised initial state is then used as input to a simulation where a vibrational state in the ground state of NH_3 is electronically excited, evolves for a time t and we measure the outcome distribution in the local basis. An AFC algorithm then updates the parameters $\vec{\theta}$ in order to maximise

the probability of bunching events in any of the three NH stretches within the first 10fs of evolution whilst suppressing other bunched events.

The feedback algorithm is a type of random restart hill climbing optimisation with an objective function which depends on observed data, and is implemented as follows. We begin with a set of uniformly random parameters. We then collect all two-photon coincidences at the output for a series of time steps separated by Δt . After correcting for detector efficiencies and the inefficient detection of multiple photon mode-occupancies, the probabilities of bunched events

$$p_{\text{bunch}}^i(\vec{\theta}, t) = |\langle 2_i | \psi(\vec{\theta}, t) \rangle|^2 \quad (5.22)$$

can be estimated, where

$$|\psi(\vec{\theta}, t)\rangle = \sum_{j,k} U_{j1}(\vec{\theta}, t) U_{k2}(\vec{\theta}, t) a_j^\dagger a_k^\dagger |\mathbf{o}\rangle \quad (5.23)$$

and $U(\vec{\theta}, t) = U_L e^{-iHt/\hbar} U_{GE}^\dagger U(\vec{\theta})$. Since we wish to selectively promote dissociation via the three NH stretch modes (corresponding to the first three modes in the localised basis), the quantity we wish to maximise is $p_1(\vec{\theta}, t) = \sum_{i=1}^3 p_{\text{bunch}}^i(\vec{\theta}, t)$, whereas we wish to minimise bunching in (and therefore dissociation via) the remaining localised modes $p_2(\vec{\theta}, t) = \sum_{i=4}^6 p_{\text{bunch}}^i(\vec{\theta}, t)$. The objective function that we use is

$$\min_{\vec{\theta}} C \quad (5.24)$$

$$C = -\alpha \sum_i w_i \Delta p_i \quad (5.25)$$

with

$$w_1 = 1, w_{i>1} = \prod_{j=1}^{i-1} (1 - [p_1(\vec{\theta}, j\Delta t) + p_2(\vec{\theta}, j\Delta t)]/N) \quad (5.26)$$

$$\Delta p_i = [p_1(\vec{\theta}, i\Delta t) - p_2(\vec{\theta}, i\Delta t)]/N \quad (5.27)$$

$$\alpha = \frac{1}{1 - (\frac{N-1}{N})^N} \quad (5.28)$$

where later times are weighted in order to correct for the possibility of dissociation occurring previously, N is the total number of timesteps and the factor α is used to renormalise the objective function such that it is constrained to the interval $[-1, 1]$.

The updating of the values of θ_i is performed using the Nelder-Mead simplex method [219] contained in the SciPy [216] python package. At each iteration, the unitaries are set, data is collected and the objective function C evaluated. The parameters $\vec{\theta}$ are then updated and the procedure is repeated until no further progress is detected, or 2000 evaluations of C have occurred for a given starting set of parameters. The Nelder-Mead method is preferred to gradient based methods in this situation, as the presence of experimental noise is likely to cause C to be nonsmooth. The optimisation process is repeated for many random initial parameter sets (random restarts) in order to increase the probability of finding the globally optimal solution.

Fig. 5.7c shows an example experimental trial. We obtain a final value of $C = -0.771$ starting from a random initial state with $C = +0.337$. In all, this procedure involved the collection of measurement statistics for almost 2000 unique settings of the photonic chip. The experiment was repeated with six random initial states, all achieving similar final values of the cost function with a mean $\bar{C} = -0.845$ (see Appendix B.2 for full data set). These results show that it is possible to highly localise vibrational energy in NH stretches of electronically excited NH_3 by selective initial state preparation in its ground state in addition to demonstrating the viability of our simulation scheme in this setting.

5.9 Discussion

In this chapter we have demonstrated that quantum states of light propagating in reconfigurable integrated photonic devices provide a new platform for the simulation of molecular vibrational dynamics. In order to scale the experimental demonstration presented here to tackle real world problems will require some additional questions to be investigated and settled.

The losses and inefficiencies in quantum light sources, linear optical networks and detectors will limit the rates at which the photonic simulator can be operated and in

turn its advantage over classical technology [138]. One approach to mitigating losses may be to incorporate quantum memories, for instance by heralding and storing the successful transformation of photons at each input optical mode sequentially. Considering the simulation of different classes of input and output configurations such as Gaussian states [128] and measurements [129] may also help reduce the experimental demands required for demonstrating a quantum speedup.

Although even approximately sampling from the output distribution of linear optical experiments like the ones we present can be classically inefficient [8], in order to produce a useful simulation the linear optical network must be able to perform the correct transfer matrix to a good degree of precision. Although the precision in the circuit parameters must necessarily increase with dimension [220, 221], linear optical elements with extinction ratios of over 60dB have been demonstrated [222, 223].

Finally, the development of programmable nonlinear optics at the quantum level is a key functionality for quantum technologies and remains a major challenge for the field, however, we have shown that weak nonlinearities provide a sufficient augmentation to simulate richer molecular models. With modest progress in these areas, our approach could yield an early class of practical quantum simulations that operate beyond current classical limits. It might then be possible to discover new possibilities for molecular control with quantum states of light, including new dissociation pathways and molecular conformations.

Statement of Work

The simulation of localised molecular vibrations with universal linear optics was conceived and developed by A. Laing and E. Martín-López. Computation of the Hessian matrix, normal modes and local basis transformations for all molecules were provided by D. Tew. Data taking was performed by N. Maraviglia, A. Neville and myself. All experiments were designed and analysed by myself. Methods for simulating open system dynamics were conceived and developed by myself. Implementation of NPS gates and the AFC algorithm were developed by myself, A. Neville and N. Maraviglia.

Chapter 6

Linear Optical Quantum Computing with Partially-Distinguishable Photons

6.1 Introduction

The preceding chapters investigated current and near-term experimental applications of photonic quantum information processing. Here we now turn our attention to the longer term goal of achieving universal quantum computation with linear optics.

As discussed in Sec. 2.4.5, the fundamental challenge in constructing a useful quantum computer is in implementing a set of universal quantum gates which can preserve and manipulate the information in a quantum codespace with an error rate that is below the fault-tolerance threshold of that code. Understanding how imperfections in device fabrication and control result in corruption of the quantum information is therefore essential for assessing the requirements of physical components to achieve fault-tolerance as well as guiding the development of bespoke error correction strategies. Furthermore, the natural error rate of a platform can make a significant difference to its practical scalability; a quantum computer running just below the error threshold will require orders of magnitude more physical qubits than one running significantly below it. Since LOQC

architectures already require large resource overheads due to their inherently probabilistic gates, it is particularly important to suppress their logical error rates.

It is not always clear how physical errors emerge as logical errors however. In LOQC, HOM interference, when coupled with photon detection, provides a mechanism to generate entanglement which is unreliant on direct interaction between photons. However, ideal quantum interference between photons only occurs if the photons are completely indistinguishable in all degrees of freedom but the one being interfered. Various forms of filtering are commonly used to achieve improved photon indistinguishability at a cost of added photon loss. In practice however, no two photons will be perfectly indistinguishable and therefore it is important to understand how using partially-distinguishable photons will affect the quantum information within the linear optical quantum computer and to find out how close to perfectly indistinguishable photons are required to be for a fault-tolerant computation.

In this chapter we study the effect of using sources of partially-distinguishable photons to generate and process photonic entanglement. We describe how states of partially-distinguishable photons are transformed by linear-optical networks and measured by photon-counting detectors and how to calculate the logical content of these states. We then apply our methods to simple examples of entanglement generation and measurement with linear optics which help elucidate the key features of the general problem. Finally, we use these tools to calculate the resulting error rates in a recent proposal for LOQC given in Ref. [120]. This analysis provides several insights into general features of error models in these architectures as well as specific properties of distinguishability error models which are explored further in the next chapter.

6.2 Errors in LOQC

The main ways in which errors can occur in the operation of a linear optical quantum computer are the following:

- **Photon loss.** This can occur through a number of mechanisms such as the scattering of light into undetected modes, absorption in its propagation medium or

inefficiency in detection.

- **Circuit error.** Errors introduced by imperfect linear optical components. This can be split into coherent and stochastic errors. Coherent errors are unitary errors which preserve purity, e.g. an offset in the reflectivity of a directional coupler. Stochastic errors reduce the purity of the state, e.g. dephasing caused by electrical noise in an electro-optic phase shifter.
- **Photon distinguishability.** Errors introduced by the non-ideal interference of partially-distinguishable photons.
- **Multiphoton terms.** Errors introduced by a photon source generating multiple photons instead of a single one.
- **Dark counts.** Detectors clicking when there is no photon present.

Of course in any realistic device all these errors and more will occur simultaneously and likely interact in non-trivial ways. It is important however to identify the most dominant of these errors in terms of their likely influence on the overall error rates in current and future devices.

Photon loss is perhaps the most considerable practical challenge in LOQC. However, since it can generally be heralded without affecting the logical state, loss can be considered largely a problem for scalability rather than fault-tolerance. Furthermore, it is known that fault tolerant computation can still be achieved in the presence of very high loss rates [224, 225], in fact it has been argued that in the limit of vanishing logical error, heralded loss rates approaching 100% could be tolerated in the resource state generation [226].

Circuit errors in linear optical components, such as phase shifters or beamsplitters, are predominantly of the coherent form due to fabrication imprecision [227]. As we demonstrated in Chapter 4, experiments can already perform highly accurate multimode transformations. Furthermore, since coherence is not lost, these errors can be corrected by simple calibration procedures resulting in interferometers with high extinction ratios e.g. 99.999% in Ref. [222].

Multiphoton errors and dark counts will contribute to the logical error but in principle can be greatly mitigated if reliable photon-number resolving detectors with high timing resolution are available. Any detection events involving too many or few photons can simply be treated as a heralded loss. In this case, the only errors are caused by second-order processes where a combination of multiphoton emission and photon loss or photon loss and dark counts occurs.

The effects of partial-distinguishability are important to consider as they are the greatest source of logical error in current experiments as we saw in Chapter 4. As discussed in Sec. 3.7.1, there are practical difficulties in producing identical photon sources using both all-optical and matter-based systems. Although the deleterious effects of partial-distinguishability can be reduced via filtering, for instance using narrowband frequency filters to reduce spectral entanglement in pair generation sources, this process is generally inefficient and to achieve perfect indistinguishability requires the full control of the photon's additional degrees of freedom. The highest HOM-dip visibility between independent sources with at least modest efficiencies is $\sim 0.95\%$ [160], emphasising the technical challenge.

6.3 Photon Distinguishability

Here we introduce the approach that will be used to model partial-distinguishability errors. The main assumption that will be made is that the distinguishability originates from the generation of the photons and linear optical networks manipulate only the paths leaving all additional, internal, degrees-of-freedom unchanged.

Evolution of a photonic state under the LON T is therefore governed by the mapping

$$a_r^\dagger[\psi_j] \rightarrow \sum_k T_{kr} a_k^\dagger[\psi_j]. \quad (6.1)$$

where, as before, $a_i^\dagger[\psi_j]$ is the operator which creates a photon in path r with an internal

state specified by the state $|\psi_j\rangle \in \mathcal{H}_{\text{int}}$. Similarly, we can consider mixed states

$$A_r^\dagger[\varrho_{\text{int}}] \circ \rho = \sum_i p_i a_r^\dagger[\xi_i] \rho a_r[\xi_i] \quad (6.2)$$

where ϱ_{int} is an operator on $\mathcal{H}_{\text{int}}$ with eigenvalues and eigenstates p_i and $|\xi_i\rangle$ respectively.

A single coarse-grained photon-number-resolving detector (CG-PNRD) which does not resolve any information except for the number of photons in a path, r , may be described by a POVM $\{\mathbf{F}_r^p\}$ with elements associated with detecting p photons

$$\mathbf{F}_r^p = \sum_{\mathbf{n} \in N_p} \left[\prod_{i=1}^k \frac{1}{\sqrt{n_i!}} (a_r^\dagger[\xi_i])^{n_i} \right] |0\rangle\langle 0| \left[\prod_{i=1}^k \frac{1}{\sqrt{n_i!}} (a_r[\xi_i])^{n_i} \right] = \sum_{\mathbf{n} \in N_p} \bigotimes_{i=1}^k |n_i\rangle\langle n_i|_{r,i}. \quad (6.3)$$

Here N_p is the set of all tuples $(n_1, \dots, n_k)$ which sum to p , $\{\xi_i\}$ is an orthonormal set of functions spanning $\mathcal{H}_{\text{int}}$ and $k = \dim(\mathcal{H}_{\text{int}})$. For brevity we define operators associated to detection patterns: $\mathbf{F}_{\mathbf{r}}^{\mathbf{p}} \equiv \mathbf{F}_{(r_1, \dots, r_j)}^{(p_1, \dots, p_j)} \equiv \mathbf{F}_{r_1}^{p_1} \otimes \dots \otimes \mathbf{F}_{r_j}^{p_j}$. We discuss in the next chapter the potential advantages of using PNRDs which provide additional information, e.g. timing information.

6.3.1 Dual-path states

To describe the dual-rail logical state of a partially-distinguishable photonic state we can use $a_b^\dagger[\xi_e]|\mathbf{o}\rangle$ as a basis for a single photon within one of two paths, labeled by a bit $b \in \{0, 1\}$, with an internal state $|\xi_e\rangle$ from an orthonormal basis $\{|\xi_i\rangle\}$ of $\mathcal{H}_{\text{int}}$. This basis state can then be expressed as a *dual-path state* whose isomorphic Hilbert space is a tensor product between the logical qubit and internal qudit spaces: $a_b^\dagger[\xi_e]|\mathbf{o}\rangle = |b\rangle \otimes |\xi_e\rangle$. Here the internal state can be traced out to yield a dual-rail qubit. Similarly, the dual-path encoding of a state of n photons can be obtained by mapping the corresponding Fock space to the tensor product of logical and internal spaces, $\mathcal{H}_F \rightarrow \mathcal{H}_L \otimes \mathcal{H}_{\text{int}}$, in an analogous manner to the single-photon case.

If there is more than one or no photons within the two rails, then we describe all such states by a third level $|2\rangle$ which is outside of the computational space. Such cases

will be considered as heralded losses as we will assume unit number-resolved detection efficiency. We note that in order to correctly describe the final logical state in any scheme, the tracing out procedure can only be performed upon photons which are detected or can no longer interfere with any other photons.

6.3.2 Distinguishability models

It will often be helpful in what follows to consider photonic states with specific distinguishability structures. The two models that will be predominantly used are the *orthogonal bad-bits model* (OBBM) and the *random source model* (RSM).

In the OBBM we consider a set of photons with pure internal states $|\psi_i\rangle = \mathcal{V}^{\frac{1}{4}}|\xi_0\rangle + (1 - \sqrt{\mathcal{V}})^{\frac{1}{2}}|\xi_i\rangle$ with a common overlapping part $|\xi_0\rangle$ and a contribution from $|\xi_i\rangle$ which is orthogonal to all other internal states, $\langle\xi_j|\xi_i\rangle = 0 \ \forall i \neq j$. This ensures that the HOM-visibility of any pair of photons is equal, $\mathcal{V}_{ij} = |\langle\psi_i|\psi_j\rangle|^2 = \mathcal{V} \ \forall i \neq j$.

In the RSM we consider a photon source factory which produces independent and identically distributed photon sources characterised by a set of variables $\underline{x}$ which produce photons characterised by these variables $\varrho(\underline{x})$. Since the sources are independent, on a single-shot basis it is possible to define a common average internal state for all the photons, $\varrho = \sum_{\underline{x}} \text{Pr}(\underline{x})\varrho(\underline{x}) = \sum_j p_j |\xi_j\rangle\langle\xi_j|$. The average pairwise HOM-visibility of any pair of sources is then given by $\mathcal{V} = \text{Tr}(\varrho^2) = \sum_j p_j^2$. The RSM is particularly relevant when using single-photon sources based on multiplexing a large number of non-deterministic sources since each photon will originate from a different physical origin in each shot.

Under certain conditions, these models can become equivalent. Since each term in the n -photon OBBM input state contains a different combination of internal states, they do not interfere with each other and their coherences do not contribute to the logical state. As a result, all logical quantities in the OBBM are equivalent to a model with the initial state $\bigotimes_i \varrho_i = \bigotimes_i \sqrt{\mathcal{V}}|\xi_0\rangle\langle\xi_0| + (1 - \sqrt{\mathcal{V}})|\xi_i\rangle\langle\xi_i|$.

This equivalent dephased model can then be understood as the high-dimensional limit of a model consisting of i.i.d mixed states, $\varrho = \sqrt{\mathcal{V}}|\xi_0\rangle\langle\xi_0| + (1 - \sqrt{\mathcal{V}}) \lim_{d \rightarrow \infty} \sum_i^d |\xi_i\rangle\langle\xi_i|/d$ such that the probability of any two bad photons being in the same internal state is

negligible. The OBBM is therefore equivalent to a specific case of the RSM described above with one dominant eigenvalue $p_0 = \sqrt{\mathcal{V}}$ and a large number of vanishing ones $p_{i \neq 0} \sim 0$.

Finally, we note that all RSM inputs become equivalent to first-order in the limit of high purity, $p_0 = 1 - \epsilon$. For n photon sources, the probability of one photon not being in the dominant mode $|\xi_0\rangle$ is $\mathcal{O}(\epsilon)$ whereas the case where two photons are not in the dominant mode will occur with probability $\mathcal{O}(\epsilon^2)$. To describe the n -photon state to first-order in ϵ it is therefore sufficient to describe the single photon internal states as $\varrho = (1 - \epsilon)|\xi_0\rangle\langle\xi_0| + \epsilon|\xi_1\rangle\langle\xi_1|$, since the distinguishable photon will not undergo interference with any other photons whichever specific $|\xi_{k \neq 0}\rangle$ it may be. This high purity limit will be relevant when considering architectures that are approaching fault-tolerance.

6.4 Generating entanglement with partially-distinguishable photons

In this section we provide explicit calculations of the logical state produced by some standard entanglement-generating linear optical gates as a function of the distinguishability structure of the input photons. The simplest case to consider is when one is allowed to perform postselection. We consider the simple four path interferometer, shown in Fig 6.1a, which ideally generates a postselected Bell state of two indistinguishable photons. The transfer matrix of this interferometer is

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \end{pmatrix} \quad (6.4)$$

and we consider an input of two photons in pure internal states in paths one and three $a_1^\dagger[\psi_1]a_3^\dagger[\psi_2]|\mathbf{o}\rangle$. Evolving under the LON produces the output state

$$\begin{aligned} \left(\sum_i T_{i1}a_i^\dagger[\psi_1]\right)\left(\sum_j T_{j3}a_j^\dagger[\psi_2]\right)|\mathbf{o}\rangle &= \frac{1}{2}(a_1^\dagger[\psi_1] + a_4^\dagger[\psi_1])(a_2^\dagger[\psi_2] + a_3^\dagger[\psi_2])|\mathbf{o}\rangle \\ &\xrightarrow{\text{post.}} \frac{1}{\sqrt{2}}(a_1^\dagger[\psi_1]a_3^\dagger[\psi_2] + a_2^\dagger[\psi_2]a_4^\dagger[\psi_1])|\mathbf{o}\rangle \end{aligned} \quad (6.5)$$

where we have applied the postselection to remove non-logical terms and renormalised on the second line. By re-writing this postselected state in the logical-internal basis

$$\frac{1}{\sqrt{2}}(|00\rangle \otimes |\psi_1\psi_2\rangle + |11\rangle \otimes |\psi_2\psi_1\rangle) \quad (6.6)$$

the logical postselected state can then be recovered by tracing out the additional degrees of freedom, resulting in the postselected state

$$\begin{aligned} \rho_L^{\text{post}} &= \frac{1}{2} \left[|00\rangle\langle 00| + |11\rangle\langle 11| + |\langle\psi_1|\psi_2\rangle|^2 (|00\rangle\langle 11| + |11\rangle\langle 00|) \right] \\ &= \frac{1 + \mathcal{V}_{12}}{2} |\phi^+\rangle\langle\phi^+| + \frac{1 - \mathcal{V}_{12}}{2} |\phi^-\rangle\langle\phi^-|. \end{aligned} \quad (6.7)$$

Where $\mathcal{V}_{12} = |\langle\psi_1|\psi_2\rangle|^2$ is the HOM visibility of the two photons. This result can be readily generalised to independent mixed input states in the additional DOF, $A_1^\dagger[\varrho_1] \circ A_3^\dagger[\varrho_2] \circ |\mathbf{o}\rangle\langle\mathbf{o}|$, by considering ensemble averages. In this case the state keeps the form of (6.7) with the HOM visibility $\mathcal{V}_{12} = \text{Tr}(S(\varrho_1 \otimes \varrho_2)) = \text{Tr}(\varrho_1\varrho_2)$ where S is the swap operator (see Sec. 3.4.2). We note that in this setting, dual-rail ‘entanglement’ is generated by photons with any non-zero overlap in their internal states, and partial distinguishability simply results in an effective dephasing channel applied to one qubit of the desired maximally-entangled postselected state.

One observation we can make from this simple example is that the effect of partial-distinguishability is to reduce the coherence of the logical state. We therefore see that although photonic states are not as susceptible to traditional decoherence via the loss of information into a thermal environment, the inability of detectors to resolve individual optical modes necessarily means that information is lost and the logical states become

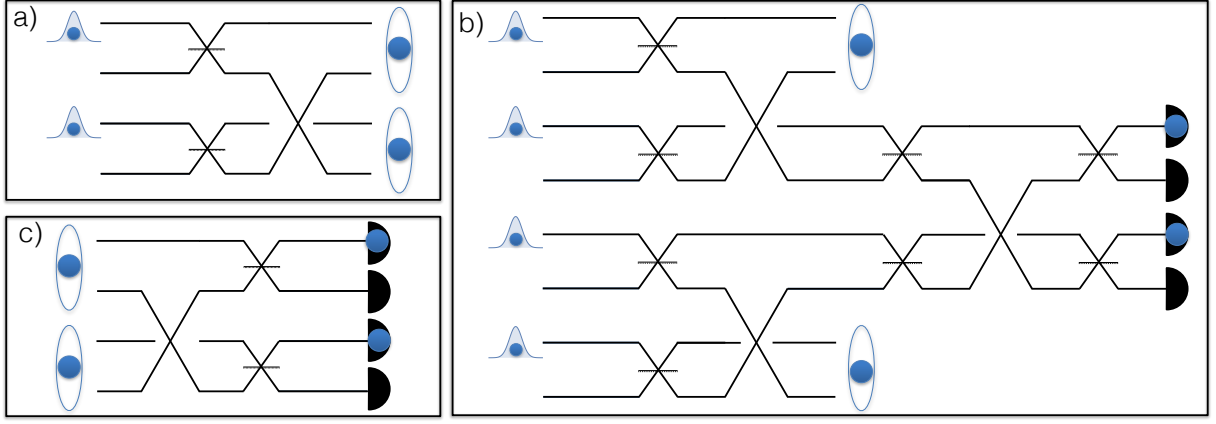


Figure 6.1: Linear optical networks for Bell state preparation and measurement. a) Post-selected bell state generator. Generates a bell state with two single photons as input and postselecting on logical states at the output. b) Heralded bell state generator. Generates a bell state with four single photons as input, heralding on measurement of two. c) Probabilistic bell state analyser. Given a logical two qubit input, performs a projective measurement onto the $|\phi^+\rangle$ and $|\phi^-\rangle$ bell states, otherwise gives a heralded failure outcome.

mixed if their internal states are not identical.

In order to perform scalable computation with linear optics and single-photon detectors, the entanglement generation must be heralded via the detection of additional ancilla photons. The most efficient interferometer known for heralded generation of two-qubit entangled states from single photons is the B_8 interferometer discussed in Sec. 4.4.2 and shown in Fig. 6.1b [179]. This circuit has the transfer matrix

$$B_8 = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \end{pmatrix}. \quad (6.8)$$

The network heralds the Bell state $|\phi^+\rangle$, up to permutations of the output paths, from

four photons conditional upon the measurement of two of them in a set of six possible heralding patterns. We consider an input state of the form:

$$A_1^\dagger[\varrho_1] \circ A_3^\dagger[\varrho_2] \circ A_5^\dagger[\varrho_3] \circ A_7^\dagger[\varrho_4] \circ |\mathbf{o}\rangle\langle\mathbf{o}|,$$

after evolution via B_8 , projection onto one of the six heralding patterns and performing any output path permutations required to produce the bell state $|\phi^+\rangle$, tracing out the internal states and ignoring non-computational terms produces the logical state:

$$\rho_L^o = \frac{1}{32 \times p^o} \times \begin{pmatrix} 1 + \mathcal{V}_{jk} & 0 & 0 & \mathcal{V}_{ij}\mathcal{V}_{kl} + \text{Tr}(\varrho_i\varrho_j\varrho_l\varrho_k) \\ 0 & 1 - \mathcal{V}_{jl} & \mathcal{V}_{ij}\mathcal{V}_{kl} - \text{Tr}(\varrho_i\varrho_j\varrho_k\varrho_l) & 0 \\ 0 & \mathcal{V}_{ij}\mathcal{V}_{kl} - \text{Tr}(\varrho_i\varrho_l\varrho_k\varrho_j) & 1 - \mathcal{V}_{ik} & 0 \\ \mathcal{V}_{ij}\mathcal{V}_{kl} + \text{Tr}(\varrho_i\varrho_k\varrho_l\varrho_j) & 0 & 0 & 1 + \mathcal{V}_{il} \end{pmatrix} \quad (6.9)$$

with $p^o = [4 + (\mathcal{V}_{jk} + \mathcal{V}_{il}) - (\mathcal{V}_{jl} + \mathcal{V}_{ik})]/32$ and where $o = 1, 2, 3$ corresponds to the state produced when an element from one of three different sets of measurement outcomes $\{\mathbf{F}_{(3,4,5,6)}^{(1,0,1,0)}, \mathbf{F}_{(3,4,5,6)}^{(0,1,0,1)}\}$, $\{\mathbf{F}_{(3,4,5,6)}^{(1,0,0,1)}, \mathbf{F}_{(3,4,5,6)}^{(0,1,1,0)}\}$, $\{\mathbf{F}_{(3,4,5,6)}^{(1,1,0,0)}, \mathbf{F}_{(3,4,5,6)}^{(0,0,1,1)}\}$ is obtained respectively. For $o = 1, 2, 3$ the state is specified by $(i, j, k, l) = (1, 2, 3, 4); (2, 1, 3, 4); (1, 4, 3, 2)$ respectively such that all ρ_L^o are equivalent up to permutations of the input photons' internal states.

We see that the logical state ρ_L^o is no longer entirely determined by the pairwise HOM-dip visibilities. Instead, it is a function of higher-order trace overlaps such as $\text{Tr}(\varrho_1\varrho_2\varrho_3\varrho_4)$ corresponding to the expectation of permutation operators on the internal space. However, it is important to understand in what ways the state is constrained by just the dip visibilities, since this is the experimental quantity that is most readily accessible and also determines the complexity of models used to simulate the effects of distinguishability.

The state fidelity between ρ_L^o and $|\phi^+\rangle$ is:

$$\begin{aligned}\mathcal{F}(\rho_L^o, |\phi^+\rangle) &= \frac{1}{32 \times p^o} \left\{ 1 + \frac{1}{2} [\mathcal{V}_{il} + \mathcal{V}_{jk}] + \mathcal{V}_{ij}\mathcal{V}_{kl} + \Re[\text{Tr}(\varrho_i \varrho_j \varrho_l \varrho_k)] \right\} \\ &\leq \frac{1}{32 \times p^o} \left[1 + \frac{1}{2} (\mathcal{V}_{il} + \mathcal{V}_{jk}) + \mathcal{V}_{ij}\mathcal{V}_{kl} + \sqrt{\mathcal{V}_{ij}\mathcal{V}_{jl}\mathcal{V}_{kl}\mathcal{V}_{ik}} \right] \equiv \mathcal{F}_{\text{ub}}.\end{aligned}\quad (6.10)$$

where, to provide an upper bound in terms of visibilities, we have used $|\text{Tr}(\prod_{i=1}^n \varrho_i)| \leq |\sqrt{\mathcal{V}_{12}\mathcal{V}_{23}\dots\mathcal{V}_{n1}}|$ as proven in appendix C.1. Pure states with entirely real positive inner products $\langle \psi_i | \psi_j \rangle = \sqrt{\mathcal{V}_{ij}}$ saturate this bound. We note that different visibilities have different roles in $\mathcal{F}_{\text{ub}}$. For example, the value of $\mathcal{V}_{ij}$ affects $\mathcal{F}_{\text{ub}}$ more than the value of $\mathcal{V}_{jl}$ does. This suggests that in some cases the fidelity of the output may be optimised via a prudent ordering of input photons. In the case that all visibilities are equal $\mathcal{V}_{ij} = \mathcal{V}$, $\mathcal{F}_{\text{ub}}$ simplifies to $\mathcal{F}_{\text{ub}}^\mathcal{V} = (1 + \mathcal{V} + 2\mathcal{V}^2)/4$. This bound is saturated by the OBBM introduced in Sec. 6.3.2.

Similarly, the purity can be upper bounded:

$$\begin{aligned}\mathcal{P}(\varrho_L^o) &= \frac{1}{(32 \times p^o)^2} \left\{ 4 + \mathcal{V}_{jk}(\mathcal{V}_{jk} + 2) + \mathcal{V}_{il}(\mathcal{V}_{il} + 2) + \mathcal{V}_{ik}(\mathcal{V}_{ik} - 2) + \mathcal{V}_{jl}(\mathcal{V}_{jl} - 2) + 4\mathcal{V}_{ij}^2\mathcal{V}_{kl}^2 \right. \\ &\quad \left. + 2|\text{Tr}(\varrho_i \varrho_j \varrho_l \varrho_k)|^2 + 2|\text{Tr}(\varrho_i \varrho_j \varrho_k \varrho_l)|^2 + 4\mathcal{V}_{ij}\mathcal{V}_{kl}\Re[\text{Tr}(\varrho_i \varrho_j \varrho_l \varrho_k) - \text{Tr}(\varrho_i \varrho_j \varrho_k \varrho_l)] \right\} \\ &\leq \frac{1}{(32 \times p^o)^2} \left\{ 4 + \mathcal{V}_{jk}(\mathcal{V}_{jk} + 2) + \mathcal{V}_{il}(\mathcal{V}_{il} + 2) + \mathcal{V}_{ik}(\mathcal{V}_{ik} - 2) + \mathcal{V}_{jl}(\mathcal{V}_{jl} - 2) + 4\mathcal{V}_{ij}^2\mathcal{V}_{kl}^2 \right. \\ &\quad \left. + 2\mathcal{V}_{ij}\mathcal{V}_{jl}\mathcal{V}_{kl}\mathcal{V}_{ik} + 2\mathcal{V}_{ij}\mathcal{V}_{jk}\mathcal{V}_{kl}\mathcal{V}_{il} + 4\mathcal{V}_{ij}\mathcal{V}_{kl}\sqrt{\mathcal{V}_{ij}(1 - \mathcal{V}_{ij})\mathcal{V}_{kl}(1 - \mathcal{V}_{kl})} \right\} \equiv \mathcal{P}_{\text{ub}}.\end{aligned}\quad (6.11)$$

For the proof of this upper bound see appendix C.2. This bound is not generally saturable and $\mathcal{P}_{\text{ub}}$ is not generally a monotonic function of the visibilities. However, in the region $\mathcal{V}_{ij} \geq 0.8 \forall i \neq j$, $\mathcal{P}_{\text{ub}}$ is monotonic. In the case that all the pairwise overlaps are equal $\mathcal{P}_{\text{ub}}$ becomes $\mathcal{P}_{\text{ub}}^\mathcal{V} = (1 + \mathcal{V}^2 + \mathcal{V}^3 + \mathcal{V}^4)/4$.

Given a set of photons we can also calculate a lower bound on the average purity and fidelity by considering the case corresponding to the RSM, where each photon shares a common average internal state ϱ . One could achieve this by for instance simply randomly permuting the input photons on every trial.

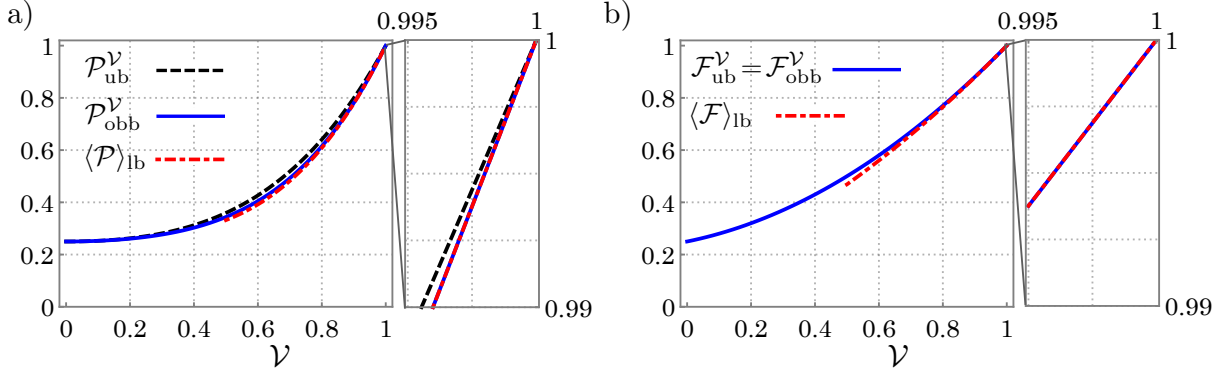


Figure 6.2: Purity and Fidelity of logical Bell states produced with partially-distinguishable photons. a) Purity of the output of the B_8 circuit for differing distinguishability models as a function of the pairwise visibilities $\mathcal{V}$. $\mathcal{P}_{ub}^\mathcal{V}$ is a generally unsaturable upper bound, $\langle \mathcal{P} \rangle_{lb}$ is the lower bound for average purity and $\mathcal{P}_{obb}^\mathcal{V}$ is for the OBBM. b) Fidelity of the output as a function of $\mathcal{V}$. In this case the upper bound is saturated by the OBBM $\mathcal{F}_{ub}^\mathcal{V} = \mathcal{F}_{obb}^\mathcal{V}$ and the average lower bound $\langle \mathcal{F} \rangle_{lb}$ is also shown. All these lines are extremely similar, suggesting that the pairwise visibilities highly constrain the resulting logical states in this circuit.

$$\begin{aligned} \mathcal{P}(\rho_L^o(\varrho)) &= \frac{1}{4} [1 + \mathcal{V}^2 + \mathcal{V}^4 + \text{Tr}^2(\varrho^4)] \\ &\stackrel{(\mathcal{V} \geq 1/2)}{\geq} \frac{1}{16} (5 - 4\mathcal{V} + 6\mathcal{V}^2 + 4\mathcal{V}^3 + 5\mathcal{V}^4) \equiv \langle \mathcal{P} \rangle_{lb} \end{aligned} \quad (6.12)$$

Where the inequality uses the lower bound $\text{Tr}(\varrho^n) \geq \frac{1}{2^n} [(1 - \sqrt{2\mathcal{V} - 1})^n + (1 + \sqrt{2\mathcal{V} - 1})^n]$ which is valid and saturable when $\mathcal{V} > 1/2$.

$\mathcal{P}_{ub}^\mathcal{V}$, $\mathcal{P}_{obb}^\mathcal{V}$ and $\langle \mathcal{P} \rangle_{lb}$ are plotted in Fig. 6.2a). The OBBM is also included since it provides a saturable value for $\mathcal{P}$ which is not an average case. In the same way, $\langle \mathcal{F} \rangle_{lb} = (1 + 4\mathcal{V} + 3\mathcal{V}^2)/8$ can be found using the RSM. $\langle \mathcal{F} \rangle_{lb}$ and $\mathcal{F}_{ub}^\mathcal{V}$ are plotted in Fig. 6.2b).

Fig. 6.2 shows that the achievable purity and fidelities using B_8 are heavily constrained by the pairwise visibilities. This is likely to be a property of the network itself, since it is known that some interferometers can result in detection events which are highly dependent upon the internal state configuration [228]. This idea will also be explored in the following chapter.

Using either the OBBM or RSM, the logical state ρ_L in Eq. (6.9) can be conveniently

represented by an ideal preparation of the Bell state $|\phi^+\rangle$ subject to a Pauli error channel applied to either qubit. In the case of the OBBM this single qubit error map is

$$\mathcal{E}^{\text{obb}}(\rho) = \frac{1}{4} [(1 + \mathcal{V} + 2\mathcal{V}^2)\rho + (1 - \mathcal{V})(X\rho X + Y\rho Y) + (1 + \mathcal{V} - 2\mathcal{V}^2)Z\rho Z] \quad (6.13)$$

so that $\rho_L^{\text{obb}} = (\mathcal{E}^{\text{obb}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|] = (\mathbb{1} \otimes \mathcal{E}^{\text{obb}})[|\phi^+\rangle\langle\phi^+|]$. Similarly for the RSM,

$$\begin{aligned} \mathcal{E}^{\text{rs}}(\rho) = \frac{1}{4} [& \{1 + \mathcal{V} + \mathcal{V}^2 + \text{tr}(\varrho^4)\}\rho + \{1 - \mathcal{V} + \mathcal{V}^2 - \text{tr}(\varrho^4)\}X\rho X \\ & + \{1 - \mathcal{V} - \mathcal{V}^2 + \text{tr}(\varrho^4)\}Y\rho Y + \{1 + \mathcal{V} - \mathcal{V}^2 - \text{tr}(\varrho^4)\}Z\rho Z]. \end{aligned} \quad (6.14)$$

Since both states are Bell-diagonal their entanglement properties can be straightforwardly evaluated [229]. The concurrence of the states are given by the simple expressions $C(\rho_L^{\text{obb}}) = \text{Max} [0, \frac{\mathcal{V}-1}{2} + \mathcal{V}^2]$ and $C(\rho_L^{\text{rs}}) = \text{Max} [0, \frac{1}{2}(\mathcal{V} - 1 + \mathcal{V}^2 + \text{tr}(\varrho^4))]$. In contrast to the postselected case, there is a minimum overlap condition to generate any entanglement at all given by $\mathcal{V} > \frac{1}{2}$.

The simple form of these error models can be explained by the symmetries of both the transfer matrix T and the internal states. All four-photon trajectory amplitudes $\prod_i T_{\sigma(r_i), c_i}$ where $c = (1, 3, 5, 7)$ and e.g. $r = (1, 2, 3, 5, 7, 8)$, are invariant under exchanging columns (1, 3) and (5, 7) and rows (1, 2) and (7, 8) of T . This corresponds to swapping internal states (ϱ_1, ϱ_2) and (ϱ_3, ϱ_4) as well as the output modes corresponding to both qubits. This means that if the internal state has the permutational symmetry $(S \otimes S)\varrho_{\text{int}}(S \otimes S) = \varrho_{\text{int}}$, then the resultant logical state will be invariant under the logical $X \otimes X$ operation (the mode-swaps), $(X \otimes X)\rho_L(X \otimes X) = \rho_L$. It can also be shown that the trajectories are invariant under the logical transformation $Z \otimes Z$ at the output without any permuting of the input columns, therefore $(Z \otimes Z)\rho_L(Z \otimes Z) = \rho_L$. Since $(\pm XX, \pm ZZ)$ are the stabilizer generators of the Bell states, any input with the $S \otimes S$ internal state symmetry must result in a logical state that is a mixture of the four Bell states. This explains why the resulting logical states are Bell-diagonal for both the RSM and OBBM and the effective error maps are Pauli diagonal in Eqs. (6.13, 6.14).

6.5 Entangled measurements with partially-distinguishable photons

The other operations at the heart of LOQC are the measurements that probabilistically fuse multiple small entangled states into a single larger one. Here we briefly describe how to calculate the resulting reduced states and effective measurement operators when partially-distinguishable photons are used in these gates.

A general linear optical measurement device is characterised by its LON transfer matrix T , and a set of successful measurement patterns, $\{\mathbf{F}\}$, registered by a collection of PN-RDs. The POVM element associated with each pattern is $\mathbf{E}_{\mathbf{m}}^{\mathbf{p}} = \mathbf{E}_{(m_1, \dots, m_j)}^{(p_1, \dots, p_j)} = \bigotimes_i^j \mathbf{E}_{m_i}^{p_i}$ where $\mathbf{E}_m^p = \mathcal{U}^\dagger(T) \mathbf{F}_m^p \mathcal{U}(T)$.

Consider an initial dual-path state of n photons $\rho \in \mathcal{H}_L^{\otimes n} \otimes \mathcal{H}_{\text{int}}^{\otimes n}$. The measurement device is then performed upon the first d photons which encode logical qubits and include any ancillary states. Since this input state will generally be non-separable between the logical and internal spaces, it is not possible to describe an effective POVM acting on the logical Hilbert space which is independent of the state in the internal Hilbert space. However, an effective logical POVM can be described if the state on the internal Hilbert space is fixed. In the following we will therefore consider an initial state of the form:

$$\rho = \rho_L^{1\dots n} \otimes \varrho_{\text{int}}^{1\dots d} \otimes \varrho_{\text{int}}^{d+1\dots n}. \quad (6.15)$$

In this case, the reduced state on the remaining systems is given by

$$\rho^{d+1\dots n}(\vec{r}) = \frac{1}{\mathcal{N}} \text{Tr}_{1\dots d} [\mathbf{F}_{\mathbf{m}}^{\mathbf{p}} \mathcal{U}(T) \rho \mathcal{U}^\dagger(T) \mathbf{F}_{\mathbf{m}}^{\mathbf{p}}] \quad (6.16)$$

$$= \frac{1}{\mathcal{N}} \text{Tr}_{1\dots d} [\mathbf{E}_{\mathbf{m}}^{\mathbf{p}} (\rho_L^{1\dots n} \otimes \varrho_{\text{int}}^{1\dots d})] \otimes \varrho_{\text{int}}^{d+1\dots n} \quad (6.17)$$

$$= \frac{1}{\mathcal{N}} \text{Tr}_{L_{1\dots d}} [\text{Tr}_{\text{int}_{1\dots d}} [\mathbf{E}_{\mathbf{m}}^{\mathbf{p}} (\mathbb{1} \otimes \varrho_{\text{int}}^{12}) (\rho_L^{1\dots n} \otimes \mathbb{1})]] \otimes \varrho_{\text{int}}^{d+1\dots n} \quad (6.18)$$

$$= \frac{1}{\mathcal{N}} \text{Tr}_{L_{1\dots d}} [E_{\varrho_{\text{int}}} \rho_L^{1\dots n}] \otimes \varrho_{\text{int}}^{d+1\dots n} \quad (6.19)$$

$$= \rho_L^{d+1\dots n} \otimes \varrho_{\text{int}}^{d+1\dots n} \quad (6.20)$$

where $E_{\varrho_{\text{int}}}(\vec{r}) = \text{Tr}_{\text{int}1\dots d}[\mathbf{E}_{\mathbf{m}}^{\mathbf{p}}(\mathbb{1} \otimes \varrho_{\text{int}}^{1\dots d})]$ is the effective logical POVM element.

A simple example of a commonly used entangling measurement device is the two-photon bell state analyser [118] (up to local unitaries this is also equivalent to the Type-II Fusion operation described in Sec. 3.6.2). As can be seen from Fig. 6.1c, such a measurement device is simply the inverse process of the postselected Bell state generator in Fig. 6.1a, and the LON is given by $T^\dagger$ (see Eq. (6.4)). The measurements $\mathbf{E}_{(1,2,3,4)}^{(1,0,1,0)}$ and $\mathbf{E}_{(1,2,3,4)}^{(0,1,0,1)}$ ideally project onto the $|\phi^+\rangle$ Bell state. Both of these detection patterns result in the same effective logical operator

$$E_{\varrho_{\text{int}}} = \frac{1 + \mathcal{V}_{12}}{4} |\phi^+\rangle\langle\phi^+| + \frac{1 - \mathcal{V}_{12}}{4} |\phi^-\rangle\langle\phi^-|, \quad (6.21)$$

with $\mathcal{V}_{12} = \text{Tr}[S\varrho_{\text{int}}]$. Up to normalisation, this is the same as Eq. (6.7) which is unsurprising, as it is simply the time-reversed process of the postselected Bell state generator. When $\varrho_{\text{int}} = |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2|$, this reproduces a known result from Ref. [230].

6.6 Measurement-based resource generation

Having introduced the basic methods for obtaining the logical content of partially distinguishable photonic states we now turn our attention to understanding the implications of partial-distinguishability in a large-scale LOQC architecture.

6.6.1 Overview and initial state

As touched on in Sec. 3.6.2, it was recently shown that LOQC can be performed ballistically [120, 121], i.e. without the need for a repeat-until-success strategy that would require fast switching networks or quantum memories. These schemes use the Bell state measurement circuits of [231, 232] in order to probabilistically fuse a supply of entangled 3-GHZ states with a ‘boosted’ success probability of 75%. In the following we will be assuming an all-optical implementation which begins with a large array of independent, deterministic single photon sources. The best known approach to generating the 3-GHZ states with single photons and linear optics is to multiplex many probabilistic GHZ generator circuits

such that with near unit probability at least one is successful. Once the fusion gates have been implemented, the resulting global state is a three-dimensional graph state of a target lattice with a fraction of missing bonds. Using ideas from percolation theory, since the fusion succeeds with a probability far above the critical threshold of the diamond [120] and pyrochlore [121] lattices it is guaranteed that there will be a spanning path through the lattice providing a universal resource state. The computation can then take place via a series of adaptive single qubit measurements upon this resource state.

Our calculations begin with the ansatz that in order to achieve fault tolerance the photons in LOQC must be highly indistinguishable. If the photons are highly indistinguishable, the probability of finding any photon in the dominant internal mode $|\xi_0\rangle$ must be close to one, $p_0 = 1 - \epsilon$. As described above, in this case the OBBM and RSM coincide to first order and it is sufficient to consider the joint internal state $\bigotimes_i^n \varrho_i = \varrho^{\otimes n}$ where $\varrho = (1 - \epsilon)|\xi_0\rangle\langle\xi_0| + \epsilon|\xi_1\rangle\langle\xi_1|$. As such, the initial state of n single photons is

$$\begin{aligned} \rho &= [\bigcirc_i^n A_i^\dagger(\varrho)] \circ |\mathbf{o}\rangle\langle\mathbf{o}| \\ &= (1 - \epsilon)^n |\Psi_0\rangle\langle\Psi_0| + \epsilon(1 - \epsilon)^{n-1} \sum_{j=1}^n |\Psi_j\rangle\langle\Psi_j| + \mathcal{O}(\epsilon^2) \end{aligned} \quad (6.22)$$

where $|\Psi_0\rangle = |\xi_0\rangle^{\otimes n}$ is the fully indistinguishable part and $|\Psi_j\rangle = |\xi_0\rangle^{\otimes j-1} \otimes |\xi_1\rangle \otimes |\xi_0\rangle^{\otimes n-j}$ labels the position of the distinguishable photon. The parameter ϵ can be physically understood in terms of the average pairwise visibility between any two sources $\mathcal{V} = \text{Tr}(\varrho^2) = (1 - \epsilon)^2 + \mathcal{O}(\epsilon^2) \Rightarrow \epsilon + \mathcal{O}(\epsilon^2) = (1 - \mathcal{V})/2$.

Within this model, the probability of one or fewer photons being distinguishable decreases exponentially with an increasing number of photons. However we will consider only the number of photons required to obtain an error rate on a single logical qubit, a quantity which does not scale with the size of the computation. The schematic in Figure 6.3a shows the scheme for generating a single qubit in the full resource state lattice in Ref [120] from 34 single photon sources. The generation of the resource state is analysed in two steps; the production of 3-GHZ states from single photons followed by the fusion operations.

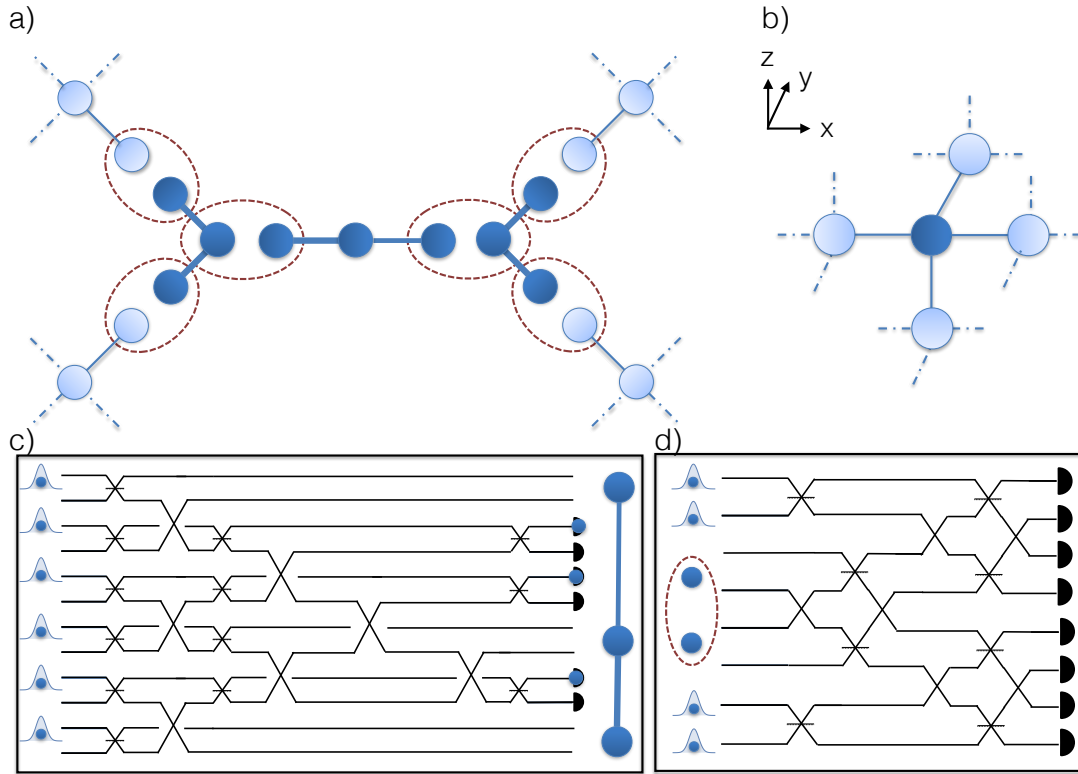


Figure 6.3: Schematic of an all-optical implementation of ballistic LOQC from [120]. The outputs of three GHZ generators (dark blue) are initially fused into a five qubit microcluster by two boosted bell measurements and the local rotations depicted. This is then fused with similar microclusters (light blue) via four further boosted fusion operations. This process generates a single photon in the final lattice from 34 initial single photons. b) A configuration of these qubits in the final diamond lattice in the case that all fusion measurements are successful. c) Optical circuits. Left is the circuit from [233] which generates a 3-GHZ state from 6 single photons with a probability of $1/32$. Right is the circuit from [231] which performs a bell-basis measurement using 4 single photon ancillas with a success probability of $3/4$.

6.6.2 GHZ generation

The circuit described in Ref. [233] is a generalisation of the heralded Bell state scheme B_8 which ideally generates the three qubit linear graph state $|G_3\rangle$ from six indistinguishable single photons with a probability $1/32$ and is shown in path encoding in Figure 6.3c). The evolved state of six photons from the input state (6.22) through this circuit can be rewritten in the logical-internal basis. Three of these photons are then measured in one of the successful heralding patterns which allows us to immediately trace out the additional

DOF information of these photons. The resulting state, prior to normalisation, of the remaining three photons is then

$$\rho_{\text{ghz}} = \frac{1}{32}(1 - \epsilon)^6 |G_3\rangle\langle G_3| \otimes |\Psi_0\rangle\langle\Psi_0| + \frac{6}{64}\epsilon(1 - \epsilon)^5 [2\rho_{\text{dist}} + |\mathbf{2}\rangle\langle\mathbf{2}|] + \mathcal{O}(\epsilon^2) \quad (6.23)$$

where $|\mathbf{2}\rangle\langle\mathbf{2}|$ collects all illogical terms (i.e. terms where at least one qubit is in the $|2\rangle\langle 2|$ logical state) and ρ_{dist} is the remaining (normalised) logical-internal state containing one distinguishable photon. We therefore find that when there is a distinguishable photon, although the overall heralding probability is increased ($3/64$), this is caused by leakage error into non-computational states and the probability of heralding a valid logical state is unchanged. Due to the symmetry of the interferometer and input state, ρ_{dist} is the same for any detection pattern up to a conditional swapping of modes.

To analyse the logical part of the state, we represent it in the graph-state basis [234]. As described in Sec. 2.4.7 a graph-state, $|G\rangle$, can be defined by a set of n stabilizer generators $S_i \equiv X_i \prod_{j \in N(i)} Z_j$, where $N(i)$ is the neighbourhood of vertex i in the corresponding graph, such that $S_i |G\rangle = |G\rangle \forall i$. The graph state basis is an orthonormal basis whose elements are the common eigenstates of these stabilizer generators. Each common eigenstate $|G^{\mathbf{s}}\rangle$ can be indexed by a bit-string $\mathbf{s} \in \{0, 1\}^n$ where $s_i = 1$ corresponds to performing a Z operation upon the i^{th} qubit of the graph state $|G\rangle$. In the case of the three qubit linear graph, this basis contains the ideal state $|G^{(0,0,0)}\rangle = |G_3\rangle$ and states which can be obtained by acting Pauli unitaries on a single qubit of $|G_3\rangle$.

After tracing out the internal Hilbert space, we find that the resultant logical state $\text{Tr}_{\text{int}}(\rho_{\text{dist}})$ is diagonal in the graph state basis and hence differs from the ideal state by random applications of single qubit Pauli operators. Moreover, even without tracing out the internal Hilbert space, ρ_{dist} effectively admits a decomposition with terms where the location of the distinguishable particle is well defined and the logical state is diagonal in the graph state basis,

$$\rho_{\text{dist}} = \sum_{j=0}^3 \left(\sum_{\mathbf{s}} p_{\mathbf{s}}^j |G^{\mathbf{s}}\rangle\langle G^{\mathbf{s}}| \right) \otimes |\Psi_j\rangle\langle\Psi_j| + \sum_{i \neq j} R_{i,j} \otimes |\Psi_i\rangle\langle\Psi_j|. \quad (6.24)$$

Terms relating to coherence between different positions of the distinguishable particle i.e. $|\Psi_i\rangle\langle\Psi_j|$, $\forall i \neq j$ will not lead to observable effects in the type of device we are studying. This is evident as these contributions do not lead to further interference after the generation of the three-qubit resource state. The coefficients $p_{\mathbf{s}}^j$ are:

$$\begin{aligned} (p_{000}^0, p_{001}^0, \dots) &= \frac{1}{24} (3, 1, 3, 1, 1, 1, 1, 1), & (p_{000}^1, \dots) &= \frac{1}{12} (1, 0, 0, 0, 1, 0, 0, 0), \\ (p_{000}^2, \dots) &= \frac{1}{12} (1, 0, 0, 0, 0, 1, 0, 0), & (p_{000}^3, \dots) &= \frac{1}{12} (1, 1, 0, 0, 0, 0, 0, 0). \end{aligned} \quad (6.25)$$

Therefore given that one of the input photons is distinguishable, the total probability of a logical error is $\Pr(\text{error}|\text{dist}) = 1 - \sum_{j=0}^3 p_{000}^j = 5/8$.

6.6.3 Fusion gates

We now consider the boosted Bell measurement approach from Ewert *et al.* [231] which uses four single ancilla photons to perform a probabilistic Bell basis measurement on a two photon logical state. The path-encoded circuit for this operation is shown in Fig. 6.3c. Using ancilla photons in the initial state Eq. (6.22), the effective logical measurement operators, as per Sec. 6.5, can be calculated for three input states; the ideal case where all photons involved are indistinguishable, the case where one of the logical photons is distinguishable and finally the case where one of the ancilla photons is distinguishable.

The scheme of Ref. [231] has a total of 572 detection patterns $\{\mathbf{F}_i\}$, 500 associated with successful Bell state projection and 72 associated with failure. Any detection events in other patterns are detectable errors and will be treated as a heralded leakage. Ideally, the logical measurement operators are given by $E_i = \text{Tr}_{\text{int}}[\mathbf{E}_i(\mathbb{1} \otimes |\Psi_0\rangle\langle\Psi_0|)] = \langle\Psi_0|\mathbf{E}_i|\Psi_0\rangle$. Many of the E_i are proportional to each other so that when combining these equivalent

patterns the full measurement implements the logical POVM $\{E\}$ with elements

$$\begin{aligned} E^1 &= |\psi^+\rangle\langle\psi^+|, & E^2 &= |\psi^-\rangle\langle\psi^-|, & E^3 &= \frac{1}{2}|\phi^+\rangle\langle\phi^+|, \\ E^4 &= \frac{1}{2}|\phi^-\rangle\langle\phi^-|, & E^5 &= \frac{1}{2}|00\rangle\langle 00|, & E^6 &= \frac{1}{2}|11\rangle\langle 11|. \end{aligned} \quad (6.26)$$

The probability of a successful Bell state projection on a maximally mixed input is $\sum_{d=1}^4 \text{Tr}(E^d/4) = 3/4$ and failure $\sum_{d=5}^6 \text{Tr}(E^d/4) = 1/4$.

In the case where either of the logical photons at the input is distinguishable, the circuit implements the effective operators $\{\langle\Psi_3|\mathbf{E}_i|\Psi_3\rangle\}, \{\langle\Psi_4|\mathbf{E}_i|\Psi_4\rangle\}$ which both result in the measurement $\{E_L\}$ given by

$$\begin{aligned} E_L^1(\phi^+) &= \frac{1}{8}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), \\ E_L^2(\phi^-) &= \frac{1}{8}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), \\ E_L^3(\psi^+) &= \frac{1}{8}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{4}|01\rangle\langle 01|, \\ E_L^4(\psi^+) &= \frac{1}{8}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{4}|10\rangle\langle 10|, \\ E_L^5(\psi^-) &= \frac{1}{16}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{8}|01\rangle\langle 01|, \\ E_L^6(\psi^-) &= \frac{1}{16}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{8}|10\rangle\langle 10|, \\ E_L^7(\psi^-) &= \frac{1}{128}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{16}|01\rangle\langle 01|, \\ E_L^8(\psi^-) &= \frac{1}{128}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{16}|10\rangle\langle 10|, \\ E_L^9(\psi^-) &= \frac{11}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|), \\ E_L^{10}(00) &= \frac{1}{8}|00\rangle\langle 00|, \\ E_L^{11}(11) &= \frac{1}{8}|11\rangle\langle 11|. \end{aligned} \quad (6.27)$$

where the bracket indicates the outcome corresponding to the measured patterns in the ideal case. The probability of success is $5/8$, failure is $1/16$ and leakage is $5/16$. When a failure outcome is received, no logical error occurs. Using the same approach we can find the measurement which is performed in the case where a randomly-chosen ancilla

photon is distinguishable. That is, we consider the internal state of the input state to be $\varrho_{\text{int}} = \sum_{i \in \{1,2,5,6\}} |\Psi_i\rangle\langle\Psi_i|/4$ and calculate $\text{Tr}_{\text{int}}[\mathbf{E}_i(\mathbb{1} \otimes \varrho_{\text{int}})]$. Collecting all equivalent terms this measurement $\{E_A\}$ is given by

$$\begin{aligned}
 E_A^1(\phi^+) &= \frac{3}{32}|\phi^+\rangle\langle\phi^+| + \frac{1}{32}|\phi^-\rangle\langle\phi^-| + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, \\
 E_A^2(\phi^+) &= \frac{3}{32}|\phi^+\rangle\langle\phi^+| + \frac{1}{32}|\phi^-\rangle\langle\phi^-|, \\
 E_A^3(\phi^-) &= \frac{3}{16}|\phi^-\rangle\langle\phi^-| + \frac{1}{16}|\phi^+\rangle\langle\phi^+| + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, \\
 E_A^4(\psi^+) &= \frac{5}{32}|\psi^+\rangle\langle\psi^+| + \frac{1}{8}|00\rangle\langle 00|, \\
 E_A^5(\psi^+) &= \frac{5}{32}|\psi^+\rangle\langle\psi^+| + \frac{1}{8}|11\rangle\langle 11|, \\
 E_A^6(\psi^+) &= \frac{3}{32}|\psi^+\rangle\langle\psi^+| + \frac{1}{8}|00\rangle\langle 00|, \\
 E_A^7(\psi^+) &= \frac{3}{32}|\psi^+\rangle\langle\psi^+| + \frac{1}{8}|11\rangle\langle 11|, \\
 E_A^8(\psi^-) &= \frac{1}{2}|\psi^-\rangle\langle\psi^-|, \\
 E_A^9(00) &= \frac{3}{16}|00\rangle\langle 00|, \\
 E_A^{10}(11) &= \frac{3}{16}|11\rangle\langle 11|.
 \end{aligned}$$

and has a probability of success of 9/16, failure 3/32 and leakage 11/32. Measurement patterns corresponding to E_A^8 perform an ideal projection. This is because $|\psi^-\rangle$ detection relies only on interference between the two logical photons, which in this case interfere perfectly.

We note that a significant portion of the time, if we have a distinguishable photon in the fusion measurement we are able to identify this by receiving an incorrect leakage measurement pattern. Therefore this is a mechanism by which logical error is converted into loss. It would be possible to go further by using the information about the specific measurement pattern we detect. The post-fusion state is dependent upon the outcome of the measurement. In the ideal case, local unitaries are conditionally applied such that the final state remains a graph state with a joined bond (success) or no bond (failure) between the neighbours of the fused qubits. When, instead, the measurement $\{E_L\}$ or $\{E_A\}$ is

implemented, the errors on the final state are conditional upon the specific measurement pattern obtained, e.g. for patterns corresponding to the ideal outcome (ϕ^+), the outcome E_A^1 would result in a logical error 1/4 of the time, whereas outcome E_A^2 would result in an error 5/8 of the time. This suggests that by considering only the E_A^1 outcome as successful we can further suppress the deleterious effects of the distinguishability. Such an approach is a trade-off in the ballistic scheme however, since a decrease in the fusion success probability will push the scheme back towards the percolation threshold. The precise nature of this trade-off is a subject for future work.

In the following we will imagine we treat all successful patterns equally and evaluate average error rates by considering the averaged post-fusion logical state, $\rho_{\text{out}} = \frac{1}{N} \sum_i K_i \text{Tr}_{1,2}[E^i \rho_{\text{in}}] K_i$, where K_i are the local correcting unitaries performed on the remaining qubits. Since the pre-fusion states we are considering are diagonal in the graph-state basis, this can be equivalently expressed as $\rho_{\text{out}} = \frac{1}{N} \text{Tr}_{1,2}[\sum_i \tilde{K}_i E^i \tilde{K}_i \rho_{\text{in}}] = \frac{1}{N} \text{Tr}_{1,2}[\bar{\mathbf{L}} \rho_{\text{in}}]$, where now $\tilde{K}_i$ are unitaries performed on the measured qubits prior to measurement (see appendix C.4 for details). This averaging process significantly simplifies the above measurement operators in the case of successful detection patterns to $\bar{E}_L = 5(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|)/4$ and $\bar{E}_A = (11|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + 3|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|)/8$.

6.7 Error rates on Lattice

Putting together the results from the previous section now allows us to calculate the full output state on the lattice section displayed in Fig. 6.3b as a result of a single distinguishable photon and logical error rates on the full lattice. Combining the GHZ states produced and Bell measurements performed with partially-distinguishable photons from the previous sections, we can calculate the logical output states corresponding to all possible outcomes of the fusion gates involved (success or failure for each).

In order to maximise the connectedness of the post-fusion graph states, additional local rotations are required before and after performing the circuit in 6.3 (see appendix C.5 for details of these rotations). This results in a breaking of the symmetry of the final noisy states between two inequivalent lattice sites. However, this effect is small and so

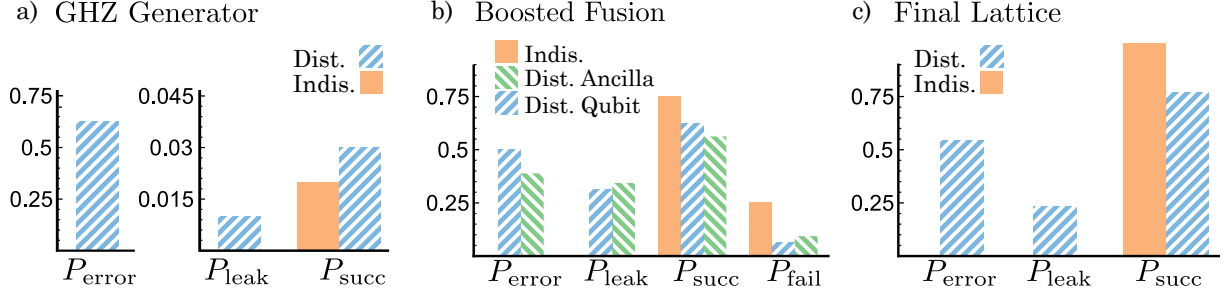


Figure 6.4: Summary of LOQC with a distinguishable particle. a) the probability of logical error, the probability of leakage and the probability of successful heralding of the GHZ generator when using all indistinguishable photons (orange) and a single distinguishable photon (blue). b) the same for the Boosted fusion gate with additionally the probability of the failure outcome for indistinguishable photons (orange), a distinguishable logical photon (blue) and a distinguishable ancilla photon (green). c) Same as a) for the local patch of the final lattice generated via three GHZ states and six fusion gates, averaged over all the measurement outcomes.

for simplicity we will just consider the average output state across all lattice sites.

All resulting states are graph-basis diagonal so that in each case the whole state generation process is equivalent to the application of a Pauli error map parameterised by ϵ , Λ_G , applied to the ideally produced graph state $|G\rangle$.

These maps can be written in terms of the imperfect state generated by a single distinguishable particle ρ_G as

$$\Lambda_G(|G\rangle\langle G|) = \frac{1}{\mathcal{N}} \left[(1 - 34\epsilon) |G\rangle\langle G| + 34\eta_G \epsilon \rho_G \right] + \mathcal{O}(\epsilon^2) \quad (6.28)$$

where G is the graph state produced in the ideal case given the fusion outcomes, η_G is the ratio of the average success probability of the fusions in the distinguishable case versus the ideal case and $\mathcal{N}$ is a normalisation constant. The state ρ_G for the case where all fusion gates are successful is shown in Fig. 6.5.

The simplest figure of merit that can be calculated from these maps is the average probability of any logical error conditional upon having one distinguishable photon across all outcome channels. This value is ~ 0.54 and there is a leakage probability of 0.23 from the fusion gates which leads to the probability of error contributed from a single lattice site error map scaling as $\approx 14\epsilon$ for small ϵ . Fig. 6.4 shows the probabilities of

$$\begin{aligned}
 & \frac{1}{1362} \left(\begin{array}{c}
 \text{584} \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \end{array} + 264 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} + 29 \begin{array}{c} \text{Z} \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \end{array} + 29 \begin{array}{c} \bullet \quad \text{Z} \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \end{array} + 29 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \text{Z} \quad \bullet \end{array} + 29 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} \\
 + 36 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \text{X} \quad \bullet \end{array} + 12 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \text{Y} \end{array} + 31 \begin{array}{c} \text{Z} \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} + 31 \begin{array}{c} \bullet \quad \text{Z} \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} + 31 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \text{Z} \quad \text{Z} \end{array} + 31 \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \text{Z} \quad \text{Z} \end{array} \\
 + 68 \begin{array}{c} \text{Z} \quad \bullet \\ \diagdown \quad \diagup \\ \text{Z} \quad \bullet \end{array} + 68 \begin{array}{c} \bullet \quad \text{Z} \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} + 52 \begin{array}{c} \text{Z} \quad \text{Z} \\ \diagdown \quad \diagup \\ \text{Z} \quad \text{Z} \end{array} + 52 \begin{array}{c} \bullet \quad \text{Z} \\ \diagdown \quad \diagup \\ \bullet \quad \text{Z} \end{array} \end{array} \right)
 \end{aligned}$$

Figure 6.5: Output state ρ_G given a single distinguishable particle within the tile shown in Fig. 6.3. All addition signs refer to an incoherent mixture of the represented graph states.

logical error, leakage and success/failure for the GHZ generator, boosted fusion and final lattice generation. We see that the error per distinguishable photon is better for the Boosted fusion, and therefore final lattice also, than the GHZ generator. This is due to the conversion of errors into leakage within the measurement circuit.

So far we have only considered errors around one lattice site resulting from a single bad photon, assuming all its neighbours to be ideal. If we imagine a perfect lattice formed from all successful fusions, then we can apply the error map at each lattice site $\bigcirc_i \Lambda_G^i$, resulting in a global noisy state. From this full lattice error model we can then calculate the logical error rates at each lattice site. Fig. 6.6 shows this error rate as a function of the parameter ϵ (and $1 - \mathcal{V} + \mathcal{O}(\epsilon^2)$). Due to the first order error assumption in our input state, we only investigate the regime in which this assumption is valid ($\epsilon \leq 10^{-3}$). The probability of a logical error occurring on each qubit scales as $\sim 28\epsilon$ for small ϵ .

In the above calculation we have just considered errors at a single lattice site. Thresholds for quantum error correcting codes (QECCs) are often numerically estimated by assuming such i.i.d local weight-one error models and, as such, are only fully valid for such models (e.g. for local depolarizing noise models, the surface code can tolerate $\approx 1\%$ logical error [66]). Although the physical errors considered here are i.i.d on each photon, due to the use of the fusion gates to entangle non-neighbouring photons we find that there is an appreciable level of correlated, higher weight errors. This is shown in Fig. 6.6

where the probability of logical error on neighbouring lattice sites, next-nearest neighbour sites in the shown configuration (up to horizontal flip) and both nearest and next-nearest neighbours (same) is shown. Also depicted are the corresponding error rates for lattice sites that are separated by at least three bonds. We see that higher weight errors are significantly more likely (scaling as ϵ rather than ϵ^2 or ϵ^3) at these nearby lattice sites than an independent error model would predict.

Many QECCs are able to tolerate localised regions of noise. To assess the impact of distinguishability on a QECC it is important to consider the error maps which arise beyond the first-order approximation which lead to higher weight errors spanning neighbouring sites across the lattice. This can occur when the set of photons used to construct neighbouring lattice sites both contain a bad photon. The resulting error map will not simply be given by Λ_G^i applied to these neighbouring sites since the two distinguishable particles can interfere with each other during the secondary fusion stage resulting in their errors no longer being independent. However, we note that when two photons do meet at a fusion the maximum weight of the error caused by this event is no higher than when only one distinguishable photon enters the fusion operation, corresponding to Pauli-Z errors on all qubits in the neighbourhood of the fused qubits. Therefore, corrections due to photons meeting at a fusion operation, for a given weight of error, will be of the order of ϵ^2 and hence negligible in the $\epsilon^2 \ll \epsilon$ regime (for completeness appendix C.7 provides an explicit calculation of the error maps when two distinguishable photons meet at a fusion). Hence we can say that the rate of an error with a given weight will be given by composing Λ_G^i s independently up to higher orders of ϵ i.e. the probability of weight six errors given by composing Λ_G^i is proportional to ϵ^2 and will be correct up to $\mathcal{O}(\epsilon^3)$ such that $\Xi \approx \bigcirc_i \Lambda_G^i$ where Ξ is the error map for the full lattice.

Finally, we note that our calculations only consider logical errors on the physical resource state generated. As this is not a regular lattice however, it is expected that logical qubits would be defined from large subsets of photons within the resource state, so that they are in a suitable form for implementing standard techniques from measurement-based quantum computing. This can be done using the renormalisation schemes of Ref. [119] for example. However, the resulting error rate on the logical qubits may generally

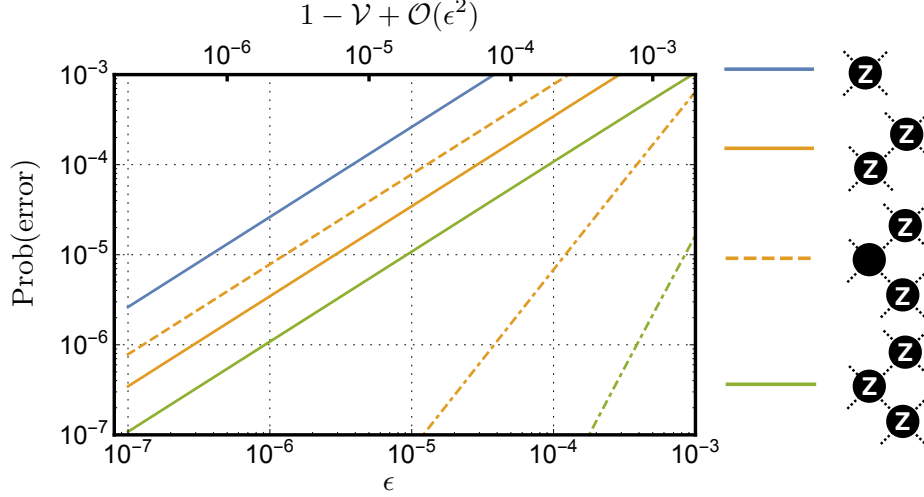


Figure 6.6: Logical error rates on lattice. Probabilities of logical error as a function of ϵ and $1 - \mathcal{V} + \mathcal{O}(\epsilon^2)$. Blue, probability of logical error on any single lattice site. Orange solid, probability of logical errors on both neighbouring lattice sites. Orange dashed, probability of logical errors on both next-nearest neighbours in the configuration shown (up to a 180° rotation). Orange dot-dashed, probability of logical error on both for any other two lattice sites. Green, probability of logical error on three neighbouring sites in the configuration shown (up to a 180° rotation). Green dot-dashed, probability of logical errors on three lattice sites for any other set of lattice sites.

be higher, due to the combined errors on the constituent photons.

6.8 Discussion

In this chapter we have shown how to describe, evolve and measure computational states of partially-distinguishable photons. This has allowed us to develop an error model for a leading proposal for LOQC. Our approach could also be used to investigate errors in other applications involving photonic cluster states such as all-optical quantum repeaters [235, 236].

These calculations suggest that the requirements for single-photon sources to perform fault-tolerant LOQC are indeed very stringent and challenging when considering the state-of-the-art. On the other hand, it is to be expected that the per-photon error rate must be low in LOQC given that tens of photons are consumed to produce a single lattice qubit. In fact the error rate per distinguishable particle in the input state is lower for producing a

final lattice qubit than generating the GHZ state. As noted in section 6.6.3, this is because using the information from detection patterns helps ‘filter’ distinguishability errors in the fusion measurements. Although we showed in section 6.4 that such an effect is not prevalent in the entanglement generation circuits considered in this scheme, it is possible that using different interferometric configurations may allow for significant improvements in the error rates. This effect will largely be the focus of the next chapter.

This work also suggests several other future directions. Can the specific features of this error model be used to develop tailored error correction strategies for LOQC? Finally, can more general results be obtained by developing a resource theory to understand the conversion between the indistinguishability of a set of single photon sources and the maximum achievable dual-rail entanglement?

Statement of Work

This chapter contains joint work with P. Birchall. In particular, the logical-internal representation of photonic states, the RSM model for photon sources and the proof of the bounds in Sec. 6.4 are due to P. Birchall. All calculations of imperfect state preparation, measurements and logical error rates were performed by myself.

Chapter 7

Mitigating Distinguishability Errors

7.1 Introduction

Multiphoton quantum interference is the fundamental physical phenomenon underpinning many photonic technologies from Boson Sampling [8] to device independent quantum security [237]. The production of highly indistinguishable photons is therefore crucial for many applications. We have described both the difficulties in producing highly indistinguishable photons and the errors they can result in. Much of the current work in improving the indistinguishability of photon sources involves in-depth engineering and material science considerations. However, an important observation made in the previous chapter was that the logical errors introduced by partial-distinguishability could be, to some extent, naturally filtered by conditioning upon specific detection patterns.

In this chapter we explore this concept further, proposing a series of schemes which can help improve the interference of a set of non-ideal single photons without any direct manipulation of their internal states. We first introduce a ‘filter’ based upon Hong-Ou-Mandel interference and bosonic bunching which distills a single purer photon from a set of noisy ones. We then consider how these ideas can help influence linear optical gate design and finally show how high timing or spectral resolution in detectors may be exploited to generate heralded entanglement, even from photons with distinguishable internal states.

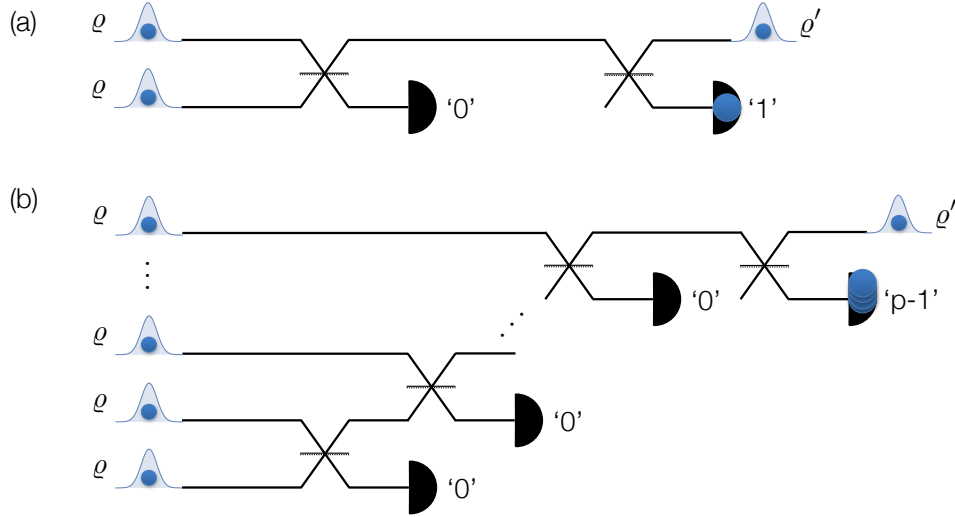


Figure 7.1: Hong-Ou-Mandel Filtering. a) Two photon case, where bunching of two photons in the state ρ is confirmed by a vacuum measurement followed by a heralded subtraction of one bunched photon resulting in a single photon state ρ' . b) p -copy generalisation. All photons are bunched in a single mode and $p - 1$ then subsequently subtracted.

7.2 Hong-Ou-Mandel Filtering

The underlying principle by which we observed a filtering of logical errors in the previous chapter can be simply stated; the detection pattern of a set of photons at the exit of a linear optical network can provide us with information regarding their degree of interference, and therefore the distinguishability of their internal states. The simplest case in which this phenomenon can be observed is for just two photons at a beamsplitter, a HOM setup.

Consider a photon in the mixed internal state $\rho = \sum_i p_i |\xi_i\rangle\langle\xi_i|$ incident at each port of a beamsplitter. If the two photons are in the same mode $|\xi_i\rangle$ then they will interfere, and bunch at the exit with a probability 1, whereas if they are in orthogonal modes then they will not interfere and will bunch with a probability $1/2$. Therefore, if we observe the photons bunch, then via Bayes rule, they were more likely to be in the same mode. Updating our description of the photon's states then increases their purity. This process can be thought of as filtering the internal states of the photons by conditioning on patterns that are most consistent with ideal quantum interference.

The scheme that we consider is shown in Fig. 7.1a. By measuring vacuum at the bottom port of the beamsplitter, we herald the bunching of the photons in the top port

in the state

$$\rho_{\text{bunch}} = \frac{1}{\mathcal{N}} \left(\sum_i p_i^2 a_1^{\dagger 2}[\xi_i] |\mathbf{o}\rangle\langle\mathbf{o}| a_1^2[\xi_i] + \sum_{i \neq j} p_i p_j a_1^{\dagger}[\xi_i] a_1^{\dagger}[\xi_j] |\mathbf{o}\rangle\langle\mathbf{o}| a_1[\xi_i] a_1[\xi_j] \right) \quad (7.1)$$

where $\mathcal{N}$ is a normalisation constant and $\{|\xi_i\rangle\}$ with $i = 0, \dots, m$ is a basis of internal states. A single photon is then obtained by a heralded subtraction of a photon from this mode. This can be achieved using another beamsplitter to split the photons and the heralded detection of one photon in mode 2, described by the measurement operator $\mathbf{F}_2$

$$\varrho' = \text{Tr}_2[\mathbf{F}_2 \mathcal{U}(\text{BS}) \rho_{\text{bunch}} \mathcal{U}^{\dagger}(\text{BS}) \mathbf{F}_2] \quad (7.2)$$

$$= \frac{1}{\mathcal{N}} \sum_i p_i (p_i + 1) a_1^{\dagger}[\xi_i] |\mathbf{o}\rangle\langle\mathbf{o}| a_1[\xi_i] \quad (7.3)$$

$$= \sum_i \frac{p_i (p_i + 1)}{\mathcal{P} + 1} |\xi_i\rangle\langle\xi_i| \quad (7.4)$$

where $\mathcal{P} = \sum_i p_i^2$ is the purity of ϱ . An eigenvalue p_i is then increased if $p_i > \mathcal{P}$. The procedure works with a probability $(\mathcal{P} + 1)/8$ ¹.

When ϱ has a single dominant eigenvalue p_0 , two extremal cases can be identified; when $\mathcal{P}$ is maximised and minimised given p_0 . The purity is maximised by the state $\varrho_{\text{max}} = p_0 |\xi_0\rangle\langle\xi_0| + (1 - p_0) |\xi_1\rangle\langle\xi_1|$, where $p'_0 = p_0(p_0 + 1)/2[p_0(p_0 - 1) + 1]$. The purity is minimised in the case that there are many additional eigenvectors, all with vanishing eigenvalues. In this case $p'_0 \rightarrow p_0(p_0 + 1)/(p_0^2 + 1)$.

When ϱ is almost pure, $p_0 = 1 - \epsilon$, then both cases tend to the relation $p'_0 = 1 - \epsilon/2 + \mathcal{O}(\epsilon^2)$, i.e. the probability of the photon being found in a ‘bad’ mode is halved by the filter.

In principle, this process could be achieved essentially deterministically for high purity inputs. When the input photons are already close to pure, they bunch almost all the time. Photon subtraction can then be achieved with a probability approaching 1 in principle via a repeat-until-success application of beamsplitters with reflectivity $\rightarrow 1$ [238]. This subtraction scheme can be applied adaptively to the states in the two modes after the

¹The scheme would still work with a reduced probability of success by replacing the beamsplitters with non-1/2 reflectivities

initial beamsplitter is applied. Wherever the first photon is detected, the other mode can then be fully measured. A vacuum detection then leads to a heralding of the remaining bunched photon. In an integrated device, such an approach could be achieved via ring resonators with tunable coupling to a single photon detector, e.g. the proposal in Ref. [239].

The scheme readily generalises to the case in which we have p -copies of ϱ (Fig. 7.1b). Using $p - 1$ beamsplitters, all the photons can be heralded to have bunched into a single mode resulting in the state

$$\rho_{\text{bunch}}^p = \frac{1}{\mathcal{N}} \sum_{\mathbf{n} \in N_p} \bigotimes_{i=0}^m p_i^{n_i} a_1^\dagger[\xi_i]^{n_i} |\mathbf{o}\rangle \langle \mathbf{o}| a_1[\xi_i]^{n_i} \quad (7.5)$$

where N_p is the set of all $m + 1$ element tuples $(n_0, n_1, \dots, n_m)$ summing to p , $\sum_i n_i = p$. As before, these photons can then be split up. For convenience we will assume that all photons are then antibunched into p modes via the transformation $a_1^\dagger \rightarrow \sum_j a_j^\dagger / \sqrt{p}$

$$\begin{aligned} \mathcal{U} \rho_{\text{bunch}}^p \mathcal{U}^\dagger &= \frac{1}{\mathcal{N}} \sum_{\mathbf{n} \in N_p} \bigotimes_{i=0}^m p_i^{n_i} \left(\sum_j a_j^\dagger[\xi_i] \right)^{n_i} |\mathbf{o}\rangle \langle \mathbf{o}| \left(\sum_{j'} a_{j'}[\xi_i] \right)^{n_i} \\ \mathbf{F}_{(12\dots p)}^{(11\dots 1)} \mathcal{U} \rho_{\text{bunch}}^p \mathcal{U}^\dagger \mathbf{F}_{(12\dots p)}^{(11\dots 1)} &= \frac{1}{\mathcal{N}} \sum_{\mathbf{n} \in N_p} \sum_{\sigma, \sigma' \in S_p} \bigotimes_{i=0}^m p_i^{n_i} a_{\sigma(j)}^\dagger[\xi_i] |\mathbf{o}\rangle \langle \mathbf{o}| a_{\sigma'(j)}[\xi_i]. \end{aligned} \quad (7.6)$$

Tracing out all but the first mode then results in the reduced state

$$\begin{aligned} \text{Tr}_{23\dots p} [\mathbf{F}_{(12\dots p)}^{(11\dots 1)} \mathcal{U} \rho_{\text{bunch}}^p \mathcal{U}^\dagger \mathbf{F}_{(12\dots p)}^{(11\dots 1)}] &= \frac{1}{\mathcal{N}} \sum_{\mathbf{n} \in N_p} \sum_k \prod_{i=0}^m p_i^{n_i} n_i! \binom{p-1}{n_0, n_1, \dots, n_k-1, \dots, n_m} a_1^\dagger[\xi_k] |\mathbf{o}\rangle \langle \mathbf{o}| a_1[\xi_k] \\ \varrho' &= \sum_k \frac{1}{\mathcal{N}} \sum_{\mathbf{n} \in N_p} \prod_{i=0}^m p_i^{n_i} n_i |\xi_k\rangle \langle \xi_k| \\ &= \sum_k p'_k |\xi_k\rangle \langle \xi_k|. \end{aligned} \quad (7.7)$$

Again, we can consider the two extremal purity cases for a fixed p_0 . The expression for

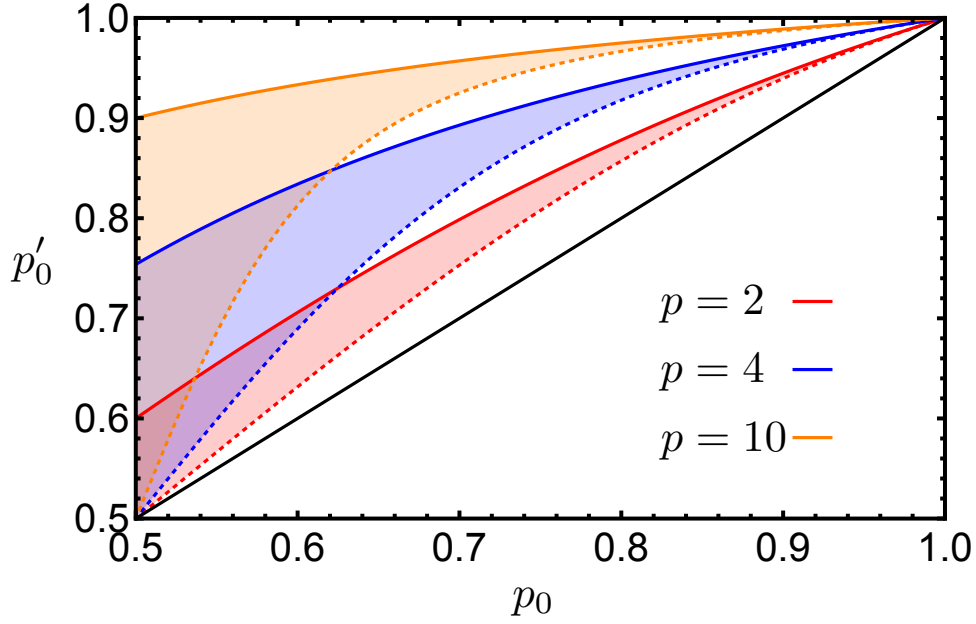


Figure 7.2: Improvement in dominant eigenvalue for HOM Filter. Dominant eigenvalue of initial states p_0 is plotted against that of the output photon p'_0 for $p = 2, 4$ and 10 initial copies of ϱ . Dotted lines correspond to the maximum purity ϱ given p_0 , solid lines to the minimum purity case.

the largest eigenvalue for the maximum purity case $\varrho_{\max}$ is

$$p'_0 = \frac{\sum_{n=0}^p n p_0^n (1-p_0)^{p-n}}{\sum_{n=0}^p n (p_0^n (1-p_0)^{p-n} + p_0^{p-n} (1-p_0)^n)} \quad (7.8)$$

$$= \frac{p_0 [(1-p_0)^{p+1} + p_0^p (p(2p_0-1) + p_0 - 1)]}{p(2p_0-1)(p_0^{p+1} - (1-p_0)^{1+p})} \quad (7.9)$$

and for the low purity case, with $\varrho_{\min} = p_0 |\xi_0\rangle\langle\xi_0| + \sum_{j=1}^d (1-p_0)/d |\xi_j\rangle\langle\xi_j|$,

$$p'_0 = \frac{\sum_{n=0}^p n p_0^n \left(\frac{1-p_0}{d}\right)^{p-n} \binom{d+p-n-1}{p-n}}{\sum_{n=0}^p n p_0^n \left(\frac{1-p_0}{d}\right)^{p-n} \binom{d+p-n-1}{p-n} + d \sum_{n_k=0}^p n_k \sum_{n_0=0}^{p-n_k} p_0^{n_0} \left(\frac{1-p_0}{d}\right)^{p-n_0} \binom{d-1+p-n_0-n_k-1}{p-n_0-n_k}} \quad (7.10)$$

$$\lim_{d \rightarrow \infty} p'_0 = \frac{\sum_{n=0}^p n p_0^n (1-p_0)^{p-n} / (p-n)!}{\sum_{n=0}^p n p_0^n (1-p_0)^{p-n} / (p-n)! + \sum_{n=0}^{p-1} p_0^n (1-p_0)^{p-n} / (p-1-n)!}. \quad (7.11)$$

These expressions for $p = 2, 4$ and 10 are plotted in Fig. 7.2. In the high purity limit, $p_0 = 1 - \epsilon$, both cases tend to $p'_0 = 1 - \epsilon/p + \mathcal{O}(\epsilon^2)$.

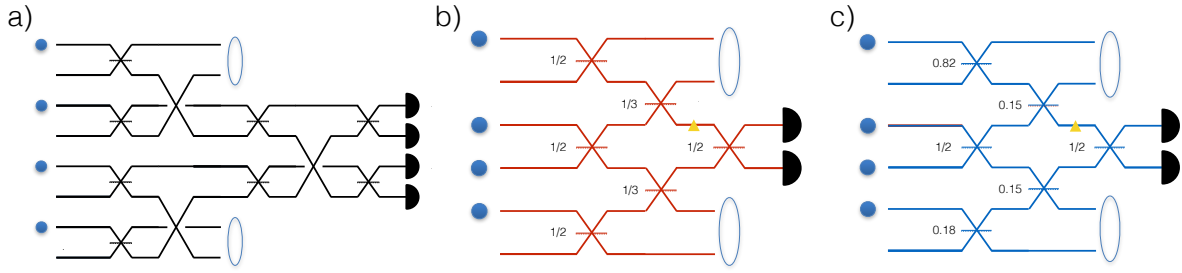


Figure 7.3: Linear optical networks for heralded bell generation. a) B_8 b) B_6 c) $B(0.18, 0.152)$ optimised for distinguishability error tolerance. The phase shifter is a $\pi/2$, beamsplitters are $\{\{\sqrt{\eta}, \sqrt{1-\eta}\}, \{\sqrt{1-\eta}, -\sqrt{\eta}\}\}$.

7.3 The Distinguishability Tolerance of Bell State Generators

Although the HOM filtering scheme introduced in the previous section can in principle produce arbitrarily indistinguishable photons, many photons must be consumed to achieve this. Using the same principle that detection patterns which have a higher probability of occurring with indistinguishable particles produce higher quality states, we can think of the perhaps more practical approach of designing gates themselves which are naturally more tolerant to partial-distinguishability errors. As a first example of this approach we consider the two heralded Bell state generator circuits described in this thesis; B_6 , introduced in Sec. 4.4.2, and B_8 , considered in Sec. 6.4.

The two schemes are depicted together in Fig. 7.3a and b. One difference between these circuits that can be noted is that in B_8 there are no closed interferometers and for the cases when heralding is successful there is no HOM-style interference (i.e. 2 photons meeting at a beamsplitter and bunching) and only interference between different overall trajectories of the photons occurs. In contrast, the B_6 scheme is built around a central closed interferometer and its operation is highly dependent upon HOM-style bunching.

B_6 is in fact a special case of a two-parameter family of unitaries $B(r_1, r_2)$ which are all Bell generator circuits with differing probabilities of success and error-tolerance. When the qubits are defined by the first two and last two modes and photons are inputted in modes 1,3,4 and 5, these transfer matrices are parameterised by

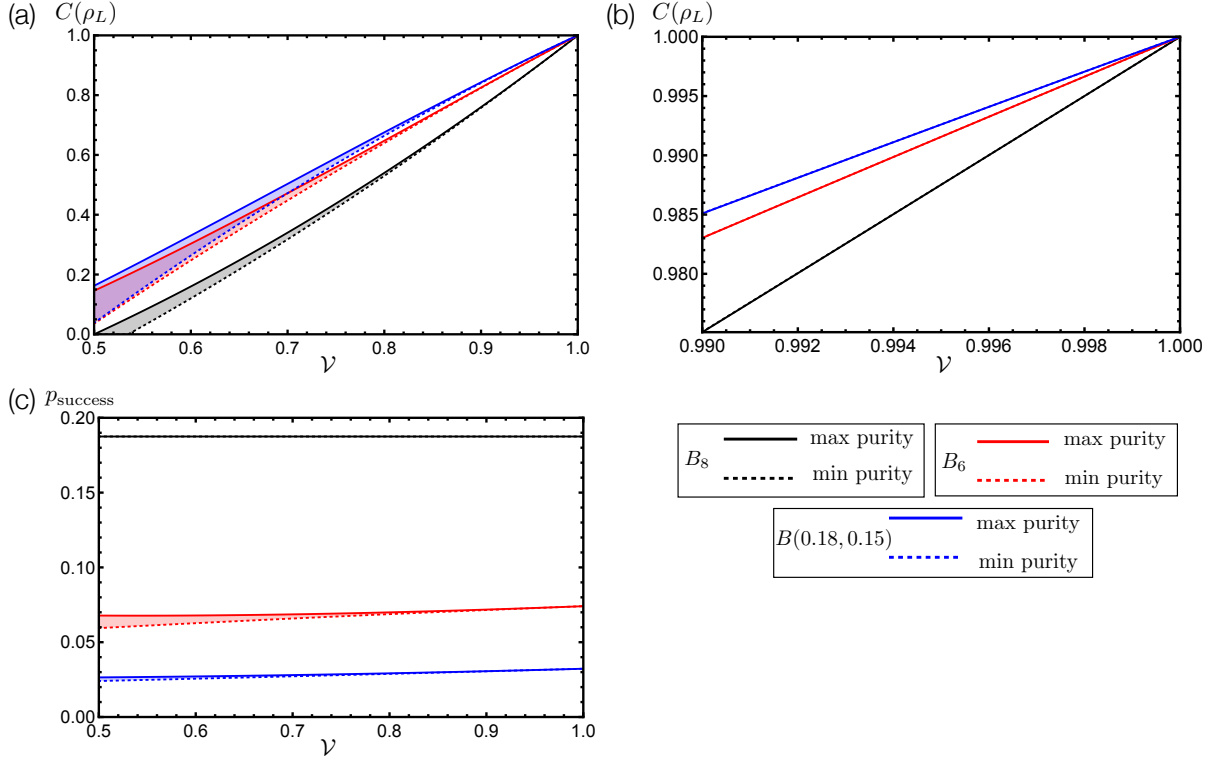


Figure 7.4: Error tolerance of Bell generators. a) Concurrence of resulting logical state produced from circuits B_8 (black) B_6 (red), and $B(0.18, 0.152)$ (blue) and input internal states ϱ . The two limiting cases which maximise (dotted) and minimise (solid) the purity of ϱ given p_0 are shown, and intermediate cases shaded. b) Shows the same plot in the linear region of high purity. c) Shows the probability of success in these schemes.

$$B(r_1, r_2) = \begin{pmatrix} \sqrt{1-r_1} & \sqrt{r_1} & 0 & 0 & 0 & 0 \\ \sqrt{r_1}\sqrt{r_2} & -\sqrt{1-r_1}\sqrt{r_2} & \frac{\sqrt{1-r_2}}{\sqrt{2}} & \frac{\sqrt{1-r_2}}{\sqrt{2}} & 0 & 0 \\ \frac{i\sqrt{r_1}\sqrt{1-r_2}}{\sqrt{2}} & -\frac{i\sqrt{1-r_1}\sqrt{1-r_2}}{\sqrt{2}} & (\frac{1}{2} - \frac{i}{2})\sqrt{r_2} & (-\frac{1}{2} - \frac{i}{2})\sqrt{r_2} & \frac{\sqrt{r_1}\sqrt{1-r_2}}{\sqrt{2}} & \frac{\sqrt{1-r_1}\sqrt{1-r_2}}{\sqrt{2}} \\ \frac{i\sqrt{r_1}\sqrt{1-r_2}}{\sqrt{2}} & -\frac{i\sqrt{1-r_1}\sqrt{1-r_2}}{\sqrt{2}} & (-\frac{1}{2} - \frac{i}{2})\sqrt{r_2} & (\frac{1}{2} - \frac{i}{2})\sqrt{r_2} & -\frac{\sqrt{r_1}\sqrt{1-r_2}}{\sqrt{2}} & -\frac{\sqrt{1-r_1}\sqrt{1-r_2}}{\sqrt{2}} \\ 0 & 0 & \frac{\sqrt{1-r_2}}{\sqrt{2}} & -\frac{\sqrt{1-r_2}}{\sqrt{2}} & -\sqrt{r_1}\sqrt{r_2} & -\sqrt{1-r_1}\sqrt{r_2} \\ 0 & 0 & 0 & 0 & \sqrt{1-r_1} & -\sqrt{r_1} \end{pmatrix} \quad (7.12)$$

The probability of success in the ideal case is given by $p_s = 2r_1r_2(1-r_1)(1-r_2)^2$ and is maximised by B_6 , where $r_1 = 1/2, r_2 = 1/3$ and $p_s = 2/27$.

Using the formalism developed in the previous chapter, we can evaluate the output

state of these circuits in the RS model where all the input photons are described with i.i.d mixed states ϱ in their additional degrees of freedom (see appendix D for the full state). Fig. 7.4a,b plot the concurrence of the resulting Bell states generated via B_8 , $B_6(1/2, 1/3)$ and a numerically optimised $B(0.18, 0.152)$ as a function of the pairwise HOM visibilities.

We find that the B_6 circuit is more error-tolerant than B_8 , and can be further improved by using $B(0.18, 0.152)$. Fig. 7.4c shows the probability of success of these schemes demonstrating a clear trade-off between efficiency and entanglement. Given the strict logical error budget required to achieve fault-tolerance, such improvements, although not dramatic, may still be worth trading for lower efficiencies in some contexts.

7.4 Generating entanglement from distinguishable photons

Finally, our entire discussion of partial-distinguishability thus far has assumed that the single photon detectors cannot resolve the additional degrees of freedom. This assumption is valid in many circumstances, when for instance the spectral response of the detector is uniform over the photon's bandwidth and their coherence length is shorter than the detector's temporal resolution. However, we now consider what can be achieved if we instead use detectors which can resolve additional information.

We first note that obtaining this additional information alone does not solve the problem of distinguishability errors. As an example we can once again consider the Bell state generator B_8 . The state produced by this circuit with input photons with spectral functions $\{\psi_i(\omega)\}$ in the logical-internal basis is

$$\begin{aligned} |\phi(\{\psi_i\})\rangle = \frac{1}{16} & \left[|00\rangle \otimes (|\psi_1\psi_2\psi_3\psi_4\rangle + |\psi_1\psi_3\psi_2\psi_4\rangle) + |01\rangle \otimes (|\psi_1\psi_2\psi_4\psi_3\rangle - |\psi_1\psi_4\psi_2\psi_3\rangle) \right. \\ & \left. + |10\rangle \otimes (|\psi_2\psi_1\psi_3\psi_4\rangle - |\psi_2\psi_3\psi_1\psi_4\rangle) + |11\rangle \otimes (|\psi_2\psi_1\psi_4\psi_3\rangle + |\psi_2\psi_4\psi_1\psi_3\rangle) \right] \end{aligned} \quad (7.13)$$

In what follows, we will analyse the scenario in which the photons are produced with

Gaussian spectral functions

$$\psi_i(\omega) = \left(\frac{1}{2\pi\Delta\omega^2} \right)^{1/4} \exp \left[\frac{-(\omega - \omega_i)^2}{4\Delta\omega^2} \right] \quad (7.14)$$

whose central frequencies are separated by much more than their bandwidth $|\omega_i - \omega_j| \gg \Delta\omega$ so that the condition $\langle \psi_i | \psi_j \rangle \sim 0$ is met, i.e. the photons are all distinguishable.

Detectors which can measure the frequency of measured photons ω' to within a spectral window of width Ω then produce a final state dependent upon overlaps of the form $A(\psi_i, \psi_j)|_{\omega'} = \int_{\omega' - \Omega/2}^{\omega' + \Omega/2} \psi_i(\omega) \psi_j^*(\omega) d\omega$. When $\Omega \gg |\omega_i - \omega_j|$, then $A(\psi_i, \psi_j)|_{\omega'} \sim |\langle \psi_i | \psi_j \rangle|^2 \sim \delta_{ij}$ and the state tends to that calculated in Eq. (6.9) with all visibilities set to zero, i.e. a maximally mixed state $\rho_L = \mathbb{1}/4$.

When we have high spectral resolution however, $\Omega \ll |\omega_i - \omega_j|$, then the four photons can be distinguished by the measurement, $A(\psi_i, \psi_j)|_{\omega'} \sim \delta_{ij} \delta_{\omega_i \omega'}$, and the output is projected onto just one of the terms in Eq. (7.13). For instance, if the photons are detected with frequencies $(\omega_1, \omega_2, \omega_3, \omega_4)$ respectively, then the logical state is collapsed to $|00\rangle$. The state in this case then remains pure, however, no dual rail entanglement has been generated.

One may be tempted to conclude that the overlaps of the internal states entirely determine the amount of entanglement that can be generated from them via transformations on the paths and detection. We now show that this is not always the case.

If we take the same set of photons, but now express their internal state in the temporal domain via a Fourier transform, we find they have the form

$$\Psi_i(t) = \mathcal{F}[\psi_i(\omega)](t) = \left(\frac{2\Delta\omega^2}{\pi} \right)^{1/4} \exp[-\Delta\omega^2 t^2] \exp[i\omega_i t] \quad (7.15)$$

where t is defined relative to a common emission time of all the photons. We see therefore that all the photons have the same temporal envelope $|\Psi_i(t)| = |\Psi_j(t)| = |\Psi(t)| \forall i, j$. If our detectors can detect photons within a timing window T , then the relevant overlaps which determine the output state become $A(\Psi_i, \Psi_j)|_{t'} = \int_{t' - T/2}^{t' + T/2} \Psi_i(t) \Psi_j^*(t) dt$.

When $T \gg \Delta t$, where Δt is the temporal width of the photons, we again find that all overlaps $A(\Psi_i, \Psi_j) \approx \delta_{ij}$ and the logical output state is again $\rho_L = \mathbb{1}/4$. However, when

$T \ll \Delta t$, the overlaps become $A(\Psi_i, \Psi_j)|_{t'} \sim |\Psi(t')|^2 \exp[i(\omega_i - \omega_j)t']$ and so interference between the spectrally distinguishable photons can be observed. This is an example of quantum erasure; the recovery of quantum interference by the removal of which-path (or in this case which-frequency) information.

In the limit of infinite timing precision, and detecting the photons at times (t_1, t_2, t_3, t_4) respectively, produces the pure state output

$$\begin{aligned} |\phi(\{t_i\})\rangle = \frac{1}{\mathcal{N}} & \left[e^{i(t_1\omega_1+t_4\omega_4)} \{e^{i(t_2\omega_2+t_3\omega_3)} + e^{i(t_2\omega_3+t_3\omega_2)}\} |00\rangle \right. \\ & + e^{i(t_1\omega_1+t_4\omega_3)} \{e^{i(t_2\omega_2+t_3\omega_4)} - e^{i(t_2\omega_4+t_3\omega_2)}\} |01\rangle \\ & + e^{i(t_1\omega_2+t_4\omega_4)} \{e^{i(t_2\omega_1+t_3\omega_3)} - e^{i(t_2\omega_3+t_3\omega_1)}\} |10\rangle \\ & \left. + e^{i(t_1\omega_2+t_4\omega_3)} \{e^{i(t_2\omega_1+t_3\omega_4)} + e^{i(t_2\omega_4+t_3\omega_1)}\} |11\rangle \right]. \end{aligned} \quad (7.16)$$

Using the simple formula for the concurrence of a two-qubit pure state (Eq. (2.36)),

$$C(|\phi(\{t_i\})\rangle) = \frac{1}{4 - 4 \cos[(t_3 - t_2)\omega] + \cos[2(t_3 - t_2)\omega]} \quad (7.17)$$

for central frequencies given by $\omega_k = \omega_0 + k\omega$. Although the specific state produced is dependent upon all detection times, entanglement is always generated and the amount is determined just by the time difference $t_3 - t_2$. When $(t_3 - t_2)\omega = 2n\pi$, the state is maximally entangled $C = 1$, whereas when $(t_3 - t_2)\omega = (2n - 1)\pi$ the entanglement is minimised, $C = 1/9$.

Taking into account a finite detection resolution, this entanglement is degraded as the window T is increased, as shown in Fig. 7.5. We note that the same situation can be set up conversely, such that temporally distinguishable photons can be interfered via high precision spectral measurements.

The timing jitter of current single-photon detectors can be in the 10's of picoseconds [240], which may already make this approach practical for photons with lengths of 100's of picoseconds produced in current integrated experiments [241]. High resolution frequency detection could be achieved via wavelength-division multiplexing into many spatial modes

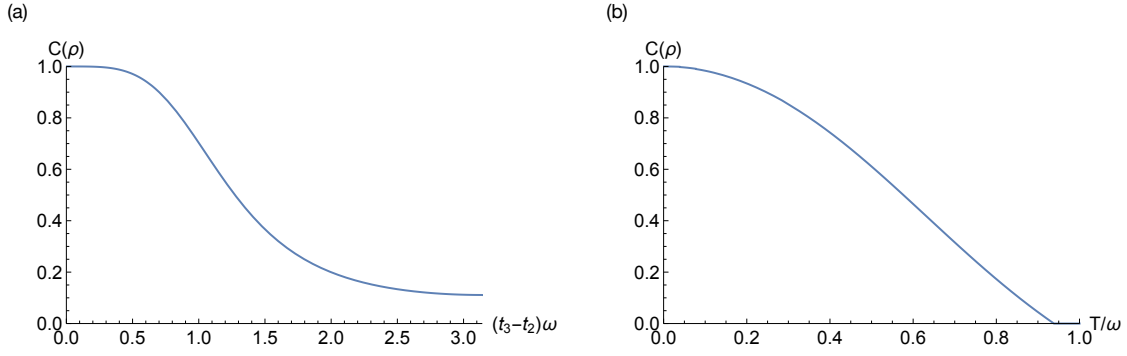


Figure 7.5: Entanglement generation from spectrally distinguishable photons via temporal resolution. (a) Concurrence of the resulting two qubit pure state as a function of the timing difference of the ancilla photons $(t_3 - t_2)\omega$ when perfectly time resolving detectors are used to measure photons with central frequencies separated by ω . (b) Concurrence of resulting two qubit state when all photons are detected simultaneously by detectors with a temporal resolution of T .

which are each fed to a detector.

7.5 Discussion

As we have seen in previous chapters, photon distinguishability is a major source of logical error in current quantum photonics experiments and is likely to be so for large-scale photonic architectures as well. Producing highly indistinguishable photons from photon sources is challenging without full control and manipulation of all degrees of freedom of the photons. Here we have proposed a number of schemes which use just manipulations of the paths and photon detection in order to mitigate the initial distinguishability of a set of photons.

Using a simple circuit which bunches and then anti-bunches a set of photons, we can distill a single purer photon from many mixed photons. We went on to show that certain circuits which generate heralded entanglement from single photons can be more and less error-tolerant in the presence of partial-distinguishability errors. Finally we showed that high resolution detectors can erase the distinguishability of photons in conjugate variable degrees of freedom.

There are many open questions and avenues for future work that are suggested by

this work. It would be interesting to investigate the possibility of similar schemes to the HOM filter approach which achieve higher success probabilities or do not require the heralded subtraction stage. One can also consider the case when the input states are known i.e. $\bigotimes_{i=1}^n \varrho_i$ instead of $\varrho^{\otimes n}$. In this case it seems likely that the best circuit will be one which uses this additional information (e.g. by the optimal pairing of photons at beamsplitters).

Our discussion of error-tolerant gates naturally leads to a further investigation in which new circuit designs are searched for by optimising the maximal entanglement given a set of partially-distinguishable photons, rather than efficiency.

Finally, an experimental exploration of using high resolution detectors and long or broadband photons in photonic information processing tasks such as LOQC or Boson Sampling (where such a scheme has been proposed [242]) may reveal the practicality and scalability such protocols.

Statement of Work

The idea of using quantum interference to filter distinguishability errors was conceived by myself and P. Birchall. The HOM filter in 7.1a was conceived by P. Birchall and the results for the many copy generalisation are due to myself. Using multi-pair emission in sources to purify photons was conceived by P. Birchall and developed by myself. The comparison between error tolerances of Bell generator circuits is due to myself. The analysis of detector timing resolution in Bell state generation is due to myself.

Chapter 8

Conclusion

8.1 Key Results

In this thesis we have explored the opportunities and challenges associated with realising practical photonic quantum technologies. Here I summarise the key results:

1. Prototyping quantum logic gates with universal linear optics.

Using the first universally reconfigurable linear optical processor we performed a series of photonic logic gates. We performed process tomography on a collection of single qubit gates and obtained the highest reported process fidelity for a postselected CNOT demonstrating high performance across hundreds of circuit configurations. We went on to present the first experimental heralded multiphoton quantum gates in integrated optics. We implemented a KLM CNOT gate demonstrating measurement-induced nonlinearity between photons. Introducing a new circuit for Bell state generation, we performed a series of measurements to verify the production of a heralded entangled state. Finally, we characterised the transfer matrices of these gates to show that errors in these experiments are dominated by imperfect sources and detectors.

2. Simulating the quantum dynamics of molecular vibrations with integrated photonics.

We proposed reconfigurable integrated photonics as a platform for the quantum

simulation of molecular vibrational dynamics. We performed a wide array of experiments to highlight the versatility and scope of this approach. A series of four atom molecules were simulated and the coherent and incoherent energy transfer properties of NMA were explored. Thermalisation and anharmonic evolution for H_2O were simulated using additional optical modes and photons, and finally the device was used in an adaptive feedback loop to simulate the optimisation of selective dissociation in ammonia.

3. An error model for linear optical quantum computing with partially-distinguishable photons.

We showed how to evolve, measure and obtain the effective logical state from a collection of partially-distinguishable photons in an optical quantum computer. We used this formalism to obtain the logical states generated and measurements performed by a series of standard linear optical gates for different distinguishability models. Applying these techniques to a large-scale architecture we were then able to calculate the error maps and logical error rates resulting from distinguishable photons for the first time.

4. Schemes for improving the quantum interference of imperfect photon sources.

We proposed several novel protocols for mitigating the errors introduced by photon distinguishability without manipulating the photons' internal states. We first introduced the Hong-Ou-Mandel filter, where the heralded bunching of photons can distill a single purer photon from many impure ones. We demonstrated that different circuits which perform the same logical function can have different tolerances to distinguishability errors and finally that entangled states can be heralded from a set of completely distinguishable photons if high resolution detection is possible in conjugate degrees of freedom.

8.2 Outlook

The development of universally reconfigurable linear optical devices has introduced a new paradigm in photonic quantum information processing. New experiments no longer require a lengthy cycle of design, fabrication and characterisation. With the further integration of photon sources [146, 160] and single photon detectors [240, 243], along with the increased component density and capabilities arising from the classical silicon photonics industry [244, 245], large-scale, reconfigurable linear optical processors have become a near-term prospect. The natural coherence and complexity of these devices promise new possibilities for non-universal models of quantum computation, such as the molecular simulation platform we have introduced here.

Many important questions remain unsettled for this route to supra-classical photonic computation. Considering the apparent classical simulability limit for Boson sampling [138, 139], such a quantum device would require the near-perfect interference of order 50 photons in a universal linear optical device with millions of highly precise, coherence-preserving tunable optical components all with per-photon loss rates significantly better than current experiments [138]. One may then legitimately question whether the construction of such a device would not be more challenging than universal LOQC which can be achieved with a constant-depth architecture, and interference only between photons which are near-neighbours. Another more fundamental open question is whether it will be possible to achieve a genuine quantum advantage for a ‘useful’ computational task without fault-tolerance. All the leading approaches to demonstrating quantum computational advantage consist of sampling from distributions for which individual probabilities cannot be accurately estimated in polytime. If we wish to gain useful information from the physical simulations presented in Chapter 5, the output distribution will therefore require some structure or sparseness which may in turn be exploitable by classical methods [246].

Given some of these difficulties, one potential avenue for future analog photonic computation may move towards more general continuous variable states [128], measurements [129] and the development of programmable nonlinear optical components. The pursuit

of classical nonlinear optical devices to solve computationally hard problems [247, 248] also suggests a potential new shortcut to large-scale quantum photonics may be to push existing complex classical optical systems incrementally into the quantum domain.

The rapid progress in integrated photonic platforms together with improvements in the theoretical resource overheads [120, 121, 249, 250] has reignited the pursuit of universal linear optical quantum computing. A full understanding of the relationship between physical imperfections in photonic devices and logical error rates is therefore a crucial priority. Comparing our calculations for partial-distinguishability to current technology suggests that obtaining the far-below-threshold logical error rates desirable for practical quantum computing may be difficult to achieve. This is particularly problematic for the photonic architecture, where qubit loss and probabilistic gates already result in high resource costs [251] versus matter-based platforms.

However, we have also shown that leveraging capabilities in circuitry and detection can help mitigate these errors, reducing the hardware requirements on photon sources. More generally, in contrast to matter systems coupled to thermal environments, the coherence properties of optical states often allow for the conversion of logical errors into detectable loss. The key question then becomes whether an architecture designed to optimally trade-off errors and loss can result in a practical device blueprint.

The future success of both approaches to photonic quantum computation will therefore rely on the combination of engineering expertise to drive component-level improvements together with insights from quantum physics to develop the architectures and applications which maximise their impact.

Bibliography

- [1] A. Einstein, “Über einen die erzeugung und verwandlung des liches betreffenden heuristischen gesichtspunkt,” *Annalen der Physik* **322**, 132–148 (1905).
- [2] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenbergh, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, “Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres,” *Nature* **526**, 682–686 (2015).
- [3] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla LAMBROCCO, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam, “Strong loophole-free test of local realism,” *Phys. Rev. Lett.* **115**, 250402 (2015).
- [4] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger, “Significant-loophole-free test of bell’s theorem with entangled photons,” *Phys. Rev. Lett.* **115**, 250401 (2015).
- [5] Alberto Politi, Martin J. Cryan, John G. Rarity, Siyuan Yu, and Jeremy L. O’Brien, “Silica-on-silicon waveguide quantum circuits,” *Science* **320**, 646–649 (2008).
- [6] Jeremy L. O’Brien, Akira Furusawa, and Jelena Vuckovic, “Photonic quantum technologies,” *Nat Photon* **3**, 687–695 (2009).
- [7] Damien Bonneau, Joshua W. Silverstone, and Mark G. Thompson, “Silicon quantum photonics,” in *Silicon Photonics III: Systems and Applications* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2016) pp. 41–82.

- [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*, STOC ’11 (ACM, New York, NY, USA, 2011) pp. 333–342.
- [9] Maximilian Schlosshauer, Johannes Kofler, and Anton Zeilinger, “A snapshot of foundational attitudes toward quantum mechanics,” *Stud. Hist. Phil. Mod. Phys.* **44**, 222–230 (2013).
- [10] Christoph Sommer, “Another survey of foundational attitudes towards quantum mechanics,” (2013), arXiv:1303.2719 .
- [11] Travis Norsen and Sarah Nelson, “Yet another snapshot of foundational attitudes toward quantum mechanics,” (2013), arXiv:1306.4646 .
- [12] Hugh Everett, “‘relative state’ formulation of quantum mechanics,” *Rev. Mod. Phys.* **29**, 454–462 (1957).
- [13] G. C. Ghirardi, A. Rimini, and T. Weber, “Unified dynamics for microscopic and macroscopic systems,” *Phys. Rev. D* **34**, 470–491 (1986).
- [14] David Bohm, “A suggested interpretation of the quantum theory in terms of ”hidden” variables. i,” *Phys. Rev.* **85**, 166–179 (1952).
- [15] Samson Abramsky and Bob Coecke, “A categorical semantics of quantum protocols,” in *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science: LICS 2004* (IEEE Computer Society Press, 2004) pp. 415–425.
- [16] Christopher A. Fuchs, N. David Mermin, and Rudiger Schack, “An introduction to qbism with an application to the locality of quantum mechanics,” *American Journal of Physics* **82**, 749–754 (2014).
- [17] D. Hanneke, S. Fogwell, and G. Gabrielse, “New measurement of the electron magnetic moment and the fine structure constant,” *Phys. Rev. Lett.* **100**, 120801 (2008).
- [18] David Mermin, “What’s wrong with this pillow?” *Physics Today* **42**, 9 (1989).
- [19] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal* **27**, 379–423 (1948).
- [20] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.* **81**, 865–942 (2009).
- [21] Martin B. Plenio and Shashank Virmani, “An introduction to entanglement measures,” *Quantum Info. Comput.* **7**, 1–51 (2007).
- [22] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, “Bell nonlocality,” *Rev. Mod. Phys.* **86**, 419–478 (2014).

- [23] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?” *Phys. Rev.* **47**, 777–780 (1935).
- [24] Albert Einstein, “Physik und realitt,” *Journal of the Franklin Institute* **221**, 313 – 347 (1936).
- [25] J. S. Bell, “On the Einstein-Podolsky-Rosen paradox,” *Physics* **1**, 195–200 (1964).
- [26] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, “Proposed experiment to test local hidden-variable theories,” *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [27] Alain Aspect, Philippe Grangier, and Gérard Roger, “Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: A new violation of bell’s inequalities,” *Phys. Rev. Lett.* **49**, 91–94 (1982).
- [28] Reinhard F. Werner, “Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model,” *Phys. Rev. A* **40**, 4277–4281 (1989).
- [29] Antonio Acín, Richard Gill, and Nicolas Gisin, “Optimal bell tests do not require maximally entangled states,” *Phys. Rev. Lett.* **95**, 210402 (2005).
- [30] Christopher Granade, Joshua Combes, and D G Cory, “Practical bayesian tomography,” *New Journal of Physics* **18**, 033024 (2016).
- [31] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White, “Measurement of qubits,” *Phys. Rev. A* **64**, 052312 (2001).
- [32] Stephen M. Barnett and Sarah Croke, “Quantum state discrimination,” *Adv. Opt. Photon.* **1**, 238–278 (2009).
- [33] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, New York, NY, USA, 2011).
- [34] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen, “Distance measures to compare real and ideal quantum processes,” *Phys. Rev. A* **71**, 062310 (2005).
- [35] D. H. Mahler, Lee A. Rozema, Ardavan Darabi, Christopher Ferrie, Robin Blume-Kohout, and A. M. Steinberg, “Adaptive quantum state tomography improves accuracy quadratically,” *Phys. Rev. Lett.* **111**, 183601 (2013).
- [36] C. A. Fuchs and J. van de Graaf, “Cryptographic distinguishability measures for quantum-mechanical states,” *IEEE Trans. Inf. Theor.* **45**, 1216–1227 (2006).
- [37] Michael A Nielsen, “A simple formula for the average gate fidelity of a quantum dynamical operation,” *Physics Letters A* **303**, 249 – 252 (2002).
- [38] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer,

- R. Ozeri, S. Seidelin, and D. J. Wineland, “Randomized benchmarking of quantum gates,” *Phys. Rev. A* **77**, 012307 (2008).
- [39] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Boston, MA, USA, 2002).
- [40] John Watrous, “Semidefinite programs for completely bounded norms,” *Theory of Computing* **5**, 217–238 (2009).
- [41] A. M. Turing, “On computable numbers, with an application to the entscheidungsproblem,” *Proceedings of the London Mathematical Society* **s2-42**, 230–265 (1937).
- [42] Greg Aloupis, Erik D. Demaine, Alan Guo, and Giovanni Viglietta, “Classic nintendo games are (computationally) hard,” in *Fun with Algorithms: 7th International Conference, FUN 2014, Lipari Island, Sicily, Italy, July 1-3, 2014. Proceedings*, edited by Alfredo Ferro, Fabrizio Luccio, and Peter Widmayer (Springer International Publishing, Cham, 2014) pp. 40–51.
- [43] R.P. Feynman, “Simulating Physics with Computers,” *International Journal of Theoretical Physics* **21**, 467–488 (1982).
- [44] D. Deutsch, “Quantum theory, the church-turing principle and the universal quantum computer,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **400**, 97–117 (1985).
- [45] A. Y. Kitaev, “Quantum computations: algorithms and error correction,” *Russ. Math. Surv.* **52**, 1191 (1997).
- [46] Seth Lloyd, “Almost any quantum logic gate is universal,” *Phys. Rev. Lett.* **75**, 346–349 (1995).
- [47] David Deutsch and Richard Jozsa, “Rapid solution of problems by quantum computation,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **439**, 553–558 (1992).
- [48] Ethan Bernstein and Umesh Vazirani, “Quantum complexity theory,” *SIAM Journal on Computing* **26**, 1411–1473 (1997).
- [49] Daniel R. Simon, “On the power of quantum computation,” *SIAM Journal on Computing* **26**, 1474–1483 (1997).
- [50] Peter W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM Review* **41**, 303–332 (1999).
- [51] R. L. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signatures and public-key cryptosystems,” *Commun. ACM* **21**, 120–126 (1978).
- [52] “Quantum algorithm zoo,” .

- [53] Lov K. Grover, “A fast quantum mechanical algorithm for database search,” in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96 (ACM, New York, NY, USA, 1996) pp. 212–219.
- [54] Seth Lloyd, “Universal quantum simulators,” *Science* **273**, 1073–1078 (1996).
- [55] Dave Wecker, Bela Bauer, Bryan K. Clark, Matthew B. Hastings, and Matthias Troyer, “Gate-count estimates for performing quantum chemistry on small quantum computers,” *Phys. Rev. A* **90**, 022305 (2014).
- [56] David Poulin, Matthew B. Hastings, Dave Wecker, Nathan Wiebe, Andrew C. Doherty, and Matthias Troyer, “The trotter step size required for accurate quantum simulation of quantum chemistry,” *Quantum Info. Comput.* **15**, 361–384 (2015).
- [57] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma, “Simulating hamiltonian dynamics with a truncated taylor series,” *Phys. Rev. Lett.* **114**, 090502 (2015).
- [58] “<https://www.allaboutcircuits.com/textbook/digital/chpt-3/logic-signal-voltage-levels/>,” .
- [59] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” *Nature* **299**, 802–803 (1982).
- [60] Peter W. Shor, “Scheme for reducing decoherence in quantum computer memory,” *Phys. Rev. A* **52**, 2493–2496 (1995).
- [61] D. Aharonov and M. Ben-Or, “Fault-tolerant quantum computation with constant error,” in *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*, STOC '97 (ACM, New York, NY, USA, 1997) pp. 176–188.
- [62] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek, “Resilient quantum computation: error models and thresholds,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **454**, 365–384 (1998).
- [63] Bryan Eastin and Emanuel Knill, “Restrictions on transversal encoded quantum gate sets,” *Phys. Rev. Lett.* **102**, 110502 (2009).
- [64] Sergey Bravyi and Alexei Kitaev, “Universal quantum computation with ideal clifford gates and noisy ancillas,” *Phys. Rev. A* **71**, 022316 (2005).
- [65] Panos Aliferis and John Preskill, “Fibonacci scheme for fault-tolerant quantum computation,” *Phys. Rev. A* **79**, 012332 (2009).
- [66] R Raussendorf, J Harrington, and K Goyal, “Topological fault-tolerance in cluster state quantum computation,” *New Journal of Physics* **9**, 199 (2007).
- [67] Clare Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter, “Surface code quantum computing by lattice surgery,” *New Journal of Physics* **14**, 123011

- (2012).
- [68] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, “Surface codes: Towards practical large-scale quantum computation,” *Phys. Rev. A* **86**, 032324 (2012).
 - [69] Daniel Gottesman, “The Heisenberg representation of quantum computers,” in *Group theoretical methods in physics. Proceedings, 22nd International Colloquium, Group22, ICGTMP’98, Hobart, Australia, July 13-17, 1998* (1998) pp. 32–43, arXiv:quant-ph/9807006 .
 - [70] Robert Raussendorf and Hans J. Briegel, “A one-way quantum computer,” *Phys. Rev. Lett.* **86**, 5188–5191 (2001).
 - [71] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, “Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels,” *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
 - [72] Xiao-Song Ma, Thomas Herbst, Thomas Scheidl, Daqing Wang, Sebastian Kropatschek, William Naylor, Bernhard Wittmann, Alexandra Mech, Johannes Kofler, Elena Anisimova, Vadim Makarov, Thomas Jennewein, Rupert Ursin, and Anton Zeilinger, “Quantum teleportation over 143 kilometres using active feed-forward,” *Nature* **489**, 269–273 (2012).
 - [73] Ji-Gang Ren, Ping Xu, Hai-Lin Yong, Liang Zhang, Sheng-Kai Liao, Juan Yin, Wei-Yue Liu, Wen-Qi Cai, Meng Yang, Li Li, Kui-Xing Yang, Xuan Han, Yong-Qiang Yao, Ji Li, Hai-Yan Wu, Song Wan, Lei Liu, Ding-Quan Liu, Yao-Wu Kuang, Zhi-Ping He, Peng Shang, Cheng Guo, Ru-Hua Zheng, Kai Tian, Zhen-Cai Zhu, Nai-Le Liu, Chao-Yang Lu, Rong Shu, Yu-Ao Chen, Cheng-Zhi Peng, Jian-Yu Wang, and Jian-Wei Pan, “Ground-to-satellite quantum teleportation,” (2017), arXiv:1707.00934 .
 - [74] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel, “Measurement-based quantum computation on cluster states,” *Phys. Rev. A* **68**, 022312 (2003).
 - [75] D. Gross and J. Eisert, “Novel schemes for measurement-based quantum computation,” *Phys. Rev. Lett.* **98**, 220503 (2007).
 - [76] D. Gross, S. T. Flammia, and J. Eisert, “Most quantum states are too entangled to be useful as computational resources,” *Phys. Rev. Lett.* **102**, 190501 (2009).
 - [77] Michael J. Bremner, Caterina Mora, and Andreas Winter, “Are random pure states useful for quantum computation?” *Phys. Rev. Lett.* **102**, 190502 (2009).
 - [78] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser, “Quantum computation by adiabatic evolution,” (2000), arXiv:quant-ph/0001106 .
 - [79] Dorit Aharonov, Wim van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and

- Oded Regev, “Adiabatic quantum computation is equivalent to standard quantum computation,” *SIAM Journal on Computing* **37**, 166–194 (2007).
- [80] E. Knill and R. Laflamme, “Power of one bit of quantum information,” *Phys. Rev. Lett.* **81**, 5672–5675 (1998).
- [81] Animesh Datta, Anil Shaji, and Carlton M. Caves, “Quantum discord and the power of one qubit,” *Phys. Rev. Lett.* **100**, 050502 (2008).
- [82] Aram W. Harrow and Ashley Montanaro, “Quantum computational supremacy,” *Nature* **549**, 203 EP – (2017).
- [83] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd, “Achieving quantum supremacy with sparse and noisy commuting quantum computations,” *Quantum* **1**, 8 (2017).
- [84] 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,” (2016), arXiv:1608.00263 .
- [85] David P. DiVincenzo, “The physical implementation of quantum computation,” *Fortschritte der Physik* **48**, 771–783 (2000).
- [86] Carlos A. Pérez-Delgado and Pieter Kok, “Quantum computers: Definition and implementations,” *Phys. Rev. A* **83**, 012303 (2011).
- [87] Jonathan A. Jones, “Quantum computing with nmr,” *Progress in Nuclear Magnetic Resonance Spectroscopy* **59**, 91 – 120 (2011).
- [88] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, and D. M. Lucas, “High-fidelity preparation, gates, memory, and readout of a trapped-ion quantum bit,” *Phys. Rev. Lett.* **113**, 220501 (2014).
- [89] Thomas Monz, Philipp Schindler, Julio T. Barreiro, Michael Chwalla, Daniel Nigg, William A. Coish, Maximilian Harlander, Wolfgang Hänsel, Markus Hennrich, and Rainer Blatt, “14-qubit entanglement: Creation and coherence,” *Phys. Rev. Lett.* **106**, 130506 (2011).
- [90] Naomi H. Nickerson, Ying Li, and Simon C. Benjamin, “Topological quantum computing with a very noisy network and local error rates approaching one percent,” *Nat. Comms.* **4**, 1756 (2013).
- [91] Bjoern Lekitsch, Sebastian Weidt, Austin G. Fowler, Klaus Mølmer, Simon J. Devitt, Christof Wunderlich, and Winfried K. Hensinger, “Blueprint for a microwave trapped ion quantum computer,” *Science Advances* **3** (2017).
- [92] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth,

- C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis, "Superconducting quantum circuits at the surface code threshold for fault tolerance," *Nature* **508**, 500–503 (2014).
- [93] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi, "A cmos silicon spin qubit," *Nat. Comms.* **7**, 13575 (2016).
- [94] M. Veldhorst, C. H. Yang, J. C. C. Hwang, W. Huang, J. P. Dehollain, J. T. Muhonen, S. Simmons, A. Laucht, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, "A two-qubit logic gate in silicon," *Nature* **526**, 410–414 (2015).
- [95] Leonard Mandel and Emil Wolf, *Optical coherence and quantum optics* (Cambridge university press, 1995).
- [96] Pieter Kok and Brendon W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge University Press, 2010).
- [97] C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2004) p. 332.
- [98] C. K. Hong, Z. Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
- [99] Sandu Popescu, "Knill-laflamme-milburn linear optics quantum computation as a measurement-based computation," *Phys. Rev. Lett.* **99**, 250501 (2007).
- [100] Stefan Scheel, "Permanents in linear optical networks," (2004), arXiv:quant-ph/0406127 .
- [101] Howard M Wiseman, "How many principles does it take to change a light bulb into a laser?" *Physica Scripta* **91**, 033001 (2016).
- [102] J G Rarity, P R Tapster, and R Loudon, "Non-classical interference between independent sources," *Journal of Optics B: Quantum and Semiclassical Optics* **7**, S171 (2005).
- [103] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd, "Gaussian quantum information," *Rev. Mod. Phys.* **84**, 621–669 (2012).
- [104] N. J. Cerf, C. Adami, and P. G. Kwiat, "Optical simulation of quantum logic," *Phys. Rev. A* **57**, R1477–R1480 (1998).
- [105] Hai-Woong Lee and Jaewan Kim, "Quantum teleportation and bell's inequality using single-particle entanglement," *Phys. Rev. A* **63**, 012305 (2000).
- [106] Barry C. Sanders, Stephen D. Bartlett, Terry Rudolph, and Peter L. Knight, "Photon-number superselection and the entangled coherent-state representation,"

- Phys. Rev. A **68**, 042329 (2003).
- [107] Isaac L. Chuang and Yoshihisa Yamamoto, “Simple quantum computer,” Phys. Rev. A **52**, 3489–3496 (1995).
- [108] G. J. Milburn, “Quantum optical fredkin gate,” Phys. Rev. Lett. **62**, 2124–2127 (1989).
- [109] Nobuyuki Matsuda, Ryosuke Shimizu, Yasuyoshi Mitsumori, Hideo Kosaka, and Keiichi Edamatsu, “Observation of optical-fibre kerr nonlinearity at the single-photon level,” Nat Photon **3**, 95–98 (2009).
- [110] Jeffrey H. Shapiro, “Single-photon kerr nonlinearities do not help quantum computation,” Phys. Rev. A **73**, 062305 (2006).
- [111] A. Javadi, I. Söllner, M. Arcari, S. Lindskov Hansen, L. Midolo, S. Mahmoodian, G. Kiršanskė, T. Pregnolato, E. H. Lee, J. D. Song, S. Stobbe, and P. Lodahl, “Single-photon non-linear optics with a quantum dot in a waveguide,” Nat. Comms. **6**, 8655 (2015).
- [112] Daniel J. Brod and Joshua Combes, “Passive cphase gate via cross-kerr nonlinearities,” Phys. Rev. Lett. **117**, 080502 (2016).
- [113] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn, “Linear optical quantum computing with photonic qubits,” Rev. Mod. Phys. **79**, 135–174 (2007).
- [114] E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics,” Nature **409**, 46–52 (2001).
- [115] J. Eisert, “Optimizing linear optics quantum gates,” Phys. Rev. Lett. **95**, 040502 (2005).
- [116] Michael A. Nielsen, “Optical quantum computation using cluster states,” Phys. Rev. Lett. **93**, 040503 (2004).
- [117] Daniel E. Browne and Terry Rudolph, “Resource-efficient linear optical quantum computation,” Phys. Rev. Lett. **95**, 010501 (2005).
- [118] Markus Michler, Klaus Mattle, Harald Weinfurter, and Anton Zeilinger, “Interferometric bell-state analysis,” Phys. Rev. A **53**, R1209–R1212 (1996).
- [119] K. Kieling, T. Rudolph, and J. Eisert, “Percolation, renormalization, and quantum computing with nondeterministic gates,” Phys. Rev. Lett. **99**, 130501 (2007).
- [120] Mercedes Gimeno-Segovia, Pete Shadbolt, Dan E. Browne, and Terry Rudolph, “From three-photon greenberger-horne-zeilinger states to ballistic universal quantum computation,” Phys. Rev. Lett. **115**, 020502 (2015).

- [121] Hussain A. Zaidi, Chris Dawson, Peter van Loock, and Terry Rudolph, “Near-deterministic creation of universal cluster states with probabilistic bell measurements and three-qubit resource states,” *Phys. Rev. A* **91**, 042301 (2015).
- [122] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, “Architectures for quantum simulation showing quantum supremacy,” (2017), arXiv:1703.00466 .
- [123] Seinosuke Toda, “Pp is as hard as the polynomial-time hierarchy,” *SIAM J. Comput.* **20**, 865–877 (1991).
- [124] L.G. Valiant, “The complexity of computing the permanent,” *Theoretical Computer Science* **8**, 189 – 201 (1979).
- [125] Leonid Gurvits, “On the complexity of mixed discriminants and related problems,” in *Mathematical Foundations of Computer Science 2005: 30th International Symposium, MFCS 2005, Gdansk, Poland, August 29–September 2, 2005. Proceedings* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005) pp. 447–458.
- [126] Larry Stockmeyer, “The complexity of approximate counting,” in *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC ’83 (1983) pp. 118–126.
- [127] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O’Brien, and T. C. Ralph, “Boson sampling from a gaussian state,” *Phys. Rev. Lett.* **113**, 100502 (2014).
- [128] Craig S. Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex, “Gaussian boson sampling,” (2016), arXiv:1612.01199 .
- [129] Levon Chakhmakhchyan and Nicolas J. Cerf, “Boson sampling with gaussian measurements,” (2017), arXiv:1705.05299 .
- [130] Scott Aaronson and Daniel J. Brod, “Bosonsampling with lost photons,” *Phys. Rev. A* **93**, 012335 (2016).
- [131] J B Spring, B J Metcalf, P C Humphreys, W S Kolthammer, X-M Jin, M Barbieri, A Datta, N Thomas-Peter, N K Langford, D Kundys, J C Gates, B J Smith, P G R Smith, and I A Walmsley, “Boson Sampling on a Photonic Chip,” *Science* **339**, 798–801 (2013).
- [132] Matthew A. Broome, Alessandro Fedrizzi, Saleh Rahimi-Keshari, Justin Dove, Scott Aaronson, Timothy C. Ralph, and Andrew G. White, “Photonic boson sampling in a tunable circuit,” *Science* **339**, 794–798 (2013).
- [133] Max Tillmann, Borivoje Dakic, Rene Heilmann, Stefan Nolte, Alexander Szameit, and Philip Walther, “Experimental boson sampling,” *Nat Photon* **7**, 540–544 (2013).

- [134] Andrea Crespi, Roberto Osellame, Roberta Ramponi, Daniel J. Brod, Ernesto F. Galvao, Nicolo Spagnolo, Chiara Vitelli, Enrico Maiorino, Paolo Mataloni, and Fabio Sciarrino, “Integrated multimode interferometers with arbitrary designs for photonic boson sampling,” *Nat Photon* **7**, 545–549 (2013).
- [135] Jacques Carolan, Jasmin D A Meinecke, Peter J Shadbolt, Nicholas J Russell, Nur Ismail, Kerstin Wörhoff, Terry Rudolph, Mark G Thompson, Jeremy L O’Brien, Jonathan C F Matthews, and Anthony Laing, “On the experimental verification of quantum complexity in linear optics,” *Nat. Photon.* **8**, 621 –626 (2014).
- [136] Hui Wang, Yu He, Yu-Huai Li, Zu-En Su, Bo Li, He-Liang Huang, Xing Ding, Ming-Cheng Chen, Chang Liu, Jian Qin, Jin-Peng Li, Yu-Ming He, Christian Schneider, Martin Kamp, Cheng-Zhi Peng, Sven Höfling, Chao-Yang Lu, and Jian-Wei Pan, “High-efficiency multiphoton boson sampling,” *Nat Photon* **11**, 361–365 (2017).
- [137] Ludovico Latmiral, Nicolò Spagnolo, and Fabio Sciarrino, “Towards quantum supremacy with lossy scattershot boson sampling,” *New Journal of Physics* **18**, 113008 (2016).
- [138] Alex Neville, Chris Sparrow, Raphal Clifford, Eric Johnston, Patrick M. Birchall, Ashley Montanaro, and Anthony Laing, “No imminent quantum supremacy by boson sampling,” (2017), [arXiv:1705.00686](https://arxiv.org/abs/1705.00686) .
- [139] Peter Clifford and Raphal Clifford, “The classical complexity of boson sampling,” (2017), [arXiv:1706.01260](https://arxiv.org/abs/1706.01260) .
- [140] Fabio Dell’Anno, Silvio De Siena, and Fabrizio Illuminati, “Multiphoton quantum optics and quantum state engineering,” *Physics Reports* **428**, 53 – 168 (2006).
- [141] W. P. Grice and I. A. Walmsley, “Spectral information and distinguishability in type-II down-conversion with a broadband pump,” *Phys. Rev. A* **56**, 1627–1634 (1997).
- [142] A. L. Migdall, D. Branning, and S. Castelletto, “Tailoring single-photon and multiphoton probabilities of a single-photon on-demand source,” *Phys. Rev. A* **66**, 053805 (2002).
- [143] Fumihiro Kaneda, Bradley G. Christensen, Jia Jun Wong, Hee Su Park, Kevin T. McCusker, and Paul G. Kwiat, “Time-multiplexed heralded single-photon source,” *Optica* **2**, 1010–1013 (2015).
- [144] Gabriel J. Mendoza, Raffaele Santagati, Jack Munns, Elizabeth Hemsley, Mateusz Piekarek, Enrique Martín-López, Graham D. Marshall, Damien Bonneau, Mark G. Thompson, and Jeremy L. O’Brien, “Active temporal and spatial multiplexing of photons,” *Optica* **3**, 127–132 (2016).
- [145] Wolfgang Mauerner, Malte Avenhaus, Wolfram Helwig, and Christine Silberhorn, “How colors influence numbers: Photon statistics of parametric down-conversion,”

- Phys. Rev. A **80**, 053815 (2009).
- [146] Silverstone J. W., Bonneau D., OhiraK., SuzukiN., YoshidaH., IizukaN., EzakiM., NatarajanC. M., TannerM. G., HadfieldR. H., ZwillerV., MarshallG. D., RarityJ. G., O’BrienJ. L., and ThompsonM. G., “On-chip quantum interference between silicon photon-pair sources,” *Nat Photon* **8**, 104–108 (2014).
 - [147] B. Darquié, M. P. A. Jones, J. Dingjan, J. Beugnon, S. Bergamini, Y. Sortais, G. Messin, A. Browaeys, and P. Grangier, “Controlled single-photon emission from a single trapped two-level atom,” *Science* **309**, 454–456 (2005).
 - [148] P. Maunz, D. L. Moehring, S. Olmschenk, K. C. Younge, D. N. Matsukevich, and C. Monroe, “Quantum interference of photon pairs from two remote trapped atomic ions,” *Nat Phys* **3**, 538–541 (2007).
 - [149] B. Lounis and W. E. Moerner, “Single photons on demand from a single molecule at room temperature,” *Nature* **407**, 491–493 (2000).
 - [150] Hannes Bernien, Lilian Childress, Lucio Robledo, Matthew Markham, Daniel Twitchen, and Ronald Hanson, “Two-photon quantum interference from separate nitrogen vacancy centers in diamond,” *Phys. Rev. Lett.* **108**, 043604 (2012).
 - [151] Sonia Buckley, Kelley Rivoire, and Jelena Vukovi, “Engineered quantum dot single-photon sources,” *Reports on Progress in Physics* **75**, 126503 (2012).
 - [152] Jake Iles-Smith, Dara P. S. McCutcheon, Jesper Mørk, and Ahsan Nazir, “Limits to coherent scattering and photon coalescence from solid-state quantum emitters,” *Phys. Rev. B* **95**, 201305 (2017).
 - [153] Hui Wang, Z.-C. Duan, Y.-H. Li, Si Chen, J.-P. Li, Y.-M. He, M.-C. Chen, Yu He, X. Ding, Cheng-Zhi Peng, Christian Schneider, Martin Kamp, Sven Höfling, Chao-Yang Lu, and Jian-Wei Pan, “Near-transform-limited single photons from an efficient solid-state quantum emitter,” *Phys. Rev. Lett.* **116**, 213601 (2016).
 - [154] Xing Ding, Yu He, Z.-C. Duan, Niels Gregersen, M.-C. Chen, S. Unsleber, S. Maier, Christian Schneider, Martin Kamp, Sven Höfling, Chao-Yang Lu, and Jian-Wei Pan, “On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar,” *Phys. Rev. Lett.* **116**, 020401 (2016).
 - [155] G. Reithmaier, M. Kaniber, F. Flassig, S. Lichtmannecker, K. Mller, A. Andrejew, J. Vukovi, R. Gross, and J. J. Finley, “On-chip generation, routing, and detection of resonance fluorescence,” *Nano Letters* **15**, 5208–5213 (2015).
 - [156] Anthony Laing, Alberto Peruzzo, Alberto Politi, Maria Rodas Verde, Matthaeus Halder, Timothy C. Ralph, Mark G. Thompson, and Jeremy L. O’Brien, “High-fidelity operation of quantum photonic circuits,” *Applied Physics Letters* **97**, 211109 (2010).

- [157] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. F. Matthews, M. G. Thompson, and J. L. O’Brien, “Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit,” *Nat Photon* **6**, 45–49 (2012).
- [158] Andrea Crespi, Roberta Ramponi, Roberto Osellame, Linda Sansoni, Irene Bongioanni, Fabio Sciarrino, Giuseppe Vallone, and Paolo Mataloni, “Integrated photonic quantum gates for polarization qubits,” *Nat. Comms.* **2**, 566 (2011).
- [159] Benjamin J. Metcalf, Justin B. Spring, Peter C. Humphreys, Nicholas Thomas-Peter, Marco Barbieri, W. Steven Kolthammer, Xian-Min Jin, Nathan K. Langford, Dmytro Kundys, James C. Gates, Brian J. Smith, Peter G. R. Smith, and Ian A. Walmsley, “Quantum teleportation on a photonic chip,” *Nat Photon* **8**, 770–774 (2014).
- [160] Justin B. Spring, Paolo L. Mennea, Benjamin J. Metcalf, Peter C. Humphreys, James C. Gates, Helen L. Rogers, Christoph Söller, Brian J. Smith, W. Steven Kolthammer, Peter G. R. Smith, and Ian A. Walmsley, “Chip-based array of near-identical, pure, heralded single-photon sources,” *Optica* **4**, 90–96 (2017).
- [161] Julien Claudon, Joel Bleuse, Nitin Singh Malik, Maela Bazin, Perine Jaffrennou, Niels Gregersen, Christophe Sauvan, Philippe Lalanne, and Jean-Michel Gerard, “A highly efficient single-photon source based on a quantum dot in a photonic nanowire,” *Nat Photon* **4**, 174–177 (2010).
- [162] Robert H. Hadfield, “Single-photon detectors for optical quantum information applications,” *Nat Photon* **3**, 696–705 (2009).
- [163] Adriana E. Lita, Aaron J. Miller, and Sae Woo Nam, “Counting near-infrared single-photons with 95% efficiency,” *Opt. Express* **16**, 3032–3040 (2008).
- [164] Alex D. Semenov, Gregory N. Goltsman, and Alexander A. Korneev, “Quantum detection by current carrying superconducting film,” *Physica C: Superconductivity* **351**, 349 – 356 (2001).
- [165] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J. Russell, Joshua W. Silverstone, Peter J. Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, Graham D. Marshall, Mark G. Thompson, Jonathan C. F. Matthews, Toshikazu Hashimoto, Jeremy L. O’Brien, and Anthony Laing, “Universal linear optics,” *Science* **349**, 711–716 (2015).
- [166] Masato Tanida, Ryo Okamoto, and Shigeki Takeuchi, “Highly indistinguishable heralded single-photon sources using parametric down conversion,” *Opt. Express* **20**, 15275–15285 (2012).
- [167] Peter Shadbolt, *Complexity and Control in Quantum Photonics* (Springer International Publishing, 2016).

- [168] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani, “Experimental realization of any discrete unitary operator,” *Phys. Rev. Lett.* **73**, 58–61 (1994).
- [169] A Hurwitz, “Über die Ereugung der Invarianten durh Integration,” *Nachr. Akad. Wiss. Gött. I. Math. Phys. Kl.* , 71 (1897).
- [170] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer, and Ian A. Walmsley, “Optimal design for universal multiport interferometers,” *Optica* **3**, 1460–1465 (2016).
- [171] Fulvio Flamini, Nicol Spagnolo, Niko Viggianiello, Andrea Crespi, Roberto Osellame, and Fabio Sciarrino, “Benchmarking integrated photonic architectures,” (2017), arXiv:1705.09211 .
- [172] Brian Borchers, “CSDP, A C library for semidefinite programming,” *Optim. Method. Softw.* **11**, 613–623 (1999).
- [173] Robin Blume-Kohout, “Optimal, reliable estimation of quantum states,” *New Journal of Physics* **12**, 043034 (2010).
- [174] Robin Blume-Kohout, “Hedged maximum likelihood quantum state estimation,” *Phys. Rev. Lett.* **105**, 200504 (2010).
- [175] M. Ringbauer, C. J. Wood, K. Modi, A. Gilchrist, A. G. White, and A. Fedrizzi, “Characterizing quantum dynamics with initial system-environment correlations,” *Phys. Rev. Lett.* **114**, 090402 (2015).
- [176] T. C. Ralph, N. K. Langford, T. B. Bell, and A. G. White, “Linear optical controlled-not gate in the coincidence basis,” *Phys. Rev. A* **65**, 062324 (2002).
- [177] J. L. O’Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, “Demonstration of an all-optical quantum controlled-not gate,” *Nature* **426**, 264–267 (2003).
- [178] T. C. Ralph, A. G. White, W. J. Munro, and G. J. Milburn, “Simple scheme for efficient linear optics quantum gates,” *Phys. Rev. A* **65**, 012314 (2001).
- [179] 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,” *Phys. Rev. A* **77**, 062316 (2008).
- [180] Jaewoo Joo, Peter L. Knight, Jeremy L. O’Brien, and Terry Rudolph, “One-way quantum computation with four-dimensional photonic qudits,” *Phys. Rev. A* **76**, 052326 (2007).
- [181] Barbara M. Terhal and Paweł Horodecki, “Schmidt number for density matrices,” *Phys. Rev. A* **61**, 040301 (2000).

- [182] Mirko Lobino, Dmitry Korystov, Connor Kupchak, Eden Figueroa, Barry C. Sanders, and A. I. Lvovsky, “Complete characterization of quantum-optical processes,” *Science* **322**, 563–566 (2008).
- [183] Saleh Rahimi-Keshari, Artur Scherer, Ady Mann, A T Rezakhani, A I Lvovsky, and Barry C Sanders, “Quantum process tomography with coherent states,” *New Journal of Physics* **13**, 013006 (2011).
- [184] Saleh Rahimi-Keshari, Matthew A. Broome, Robert Fickler, Alessandro Fedrizzi, Timothy C. Ralph, and Andrew G. White, “Direct characterization of linear-optical networks,” *Opt. Express* **21**, 13450–13458 (2013).
- [185] Anthony Laing and Jeremy L. O’Brien, “Super-stable tomography of any linear optical device,” (2012), arXiv:1208.2868 .
- [186] Richard Sinkhorn, “A relationship between arbitrary positive matrices and doubly stochastic matrices,” *Ann. Math. Statist.* **35**, 876–879 (1964).
- [187] Richard Sinkhorn and Paul Knopp, “Concerning nonnegative matrices and doubly stochastic matrices.” *Pacific J. Math.* **21**, 343–348 (1967).
- [188] Jacob Mower, Nicholas C. Harris, Gregory R. Steinbrecher, Yoav Lahini, and Dirk Englund, “High-fidelity quantum state evolution in imperfect photonic integrated circuits,” *Phys. Rev. A* **92**, 032322 (2015).
- [189] B. Ulmann, *Analog Computing* (Oldenbourg Wissenschaftsverlag, 2013).
- [190] Constantin Brif, Raj Chakrabarti, and Herschel Rabitz, “Control of quantum phenomena: past, present and future,” *New Journal of Physics* **12**, 075008 (2010).
- [191] Michael L. Hause, Y. Heidi Yoon, and F. Fleming Crim, “Vibrationally mediated photodissociation of ammonia: The influence of nh stretching vibrations on passage through conical intersections,” *The Journal of Chemical Physics* **125**, 174309 (2006).
- [192] Daan Brinks, Fernando D. Stefani, Florian Kulzer, Richard Hildner, Tim H. Taminiau, Yuri Avlasevich, Klaus Mullen, and Niek F. van Hulst, “Visualizing and controlling vibrational wave packets of single molecules,” *Nature* **465**, 905–908 (2010).
- [193] A.S. Alnaser, M. Kbel, R. Siemering, B. Bergues, Nora G Kling, K.J. Betsch, Y. Deng, J. Schmidt, Z.A. Alahmed, A.M. Azzeer, J. Ullrich, I. Ben-Itzhak, R. Moshhammer, U. Kleineberg, F. Krausz, R. de Vivie-Riedle, and M.F. Kling, “Subfemtosecond steering of hydrocarbon deprotonation through superposition of vibrational modes,” *Nat. Comms.* **5** (2014), 10.1038/ncomms4800.
- [194] Xin Tong, Alexander H. Winney, and Stefan Willitsch, “Sympathetic cooling of molecular ions in selected rotational and vibrational states produced by threshold photoionization,” *Phys. Rev. Lett.* **105**, 143001 (2010).

- [195] M. McDonald, B. H. McGuyer, F. Apfelbeck, C. H. Lee, I. Majewska, R. Moszynski, and T. Zelevinsky, “Photodissociation of ultracold diatomic strontium molecules with quantum state control,” *Nature* **534**, 122–126 (2016).
- [196] Tyler L Cocker, Dominik Peller, Ping Yu, Jascha Repp, and Rupert Huber, “Tracking the ultrafast motion of a single molecule by femtosecond orbital imaging,” *Nature* **539**, 263–267 (2016).
- [197] B. Wolter, M. G. Pullen, A.-T. Le, M. Baudisch, K. Doblhoff-Dier, A. Senftleben, M. Hemmer, C. D. Schröter, J. Ullrich, T. Pfeifer, R. Moshhammer, S. Gräfe, O. Vendrell, C. D. Lin, and J. Biegert, “Ultrafast electron diffraction imaging of bond breaking in di-ionized acetylene,” *Science* **354**, 308–312 (2016).
- [198] M. Born and R. Oppenheimer, “Zur quantentheorie der molekeln,” *Annalen der Physik* **389**, 457–484 (1927).
- [199] Daniel H. Friese, Christof Httig, and Jrg Komann, “Analytic molecular hessian calculations for cc2 and mp2 combined with the resolution of identity approximation,” *Journal of Chemical Theory and Computation* **9**, 1469–1480 (2013).
- [200] Christoph R. Jacob and Markus Reiher, “Localizing normal modes in large molecules,” *The Journal of Chemical Physics* **130**, 084106 (2009).
- [201] D J Clouthier and D A Ramsay, “The spectroscopy of formaldehyde and thioformaldehyde,” *Annual Review of Physical Chemistry* **34**, 31–58 (1983).
- [202] Ove Christiansen, Henrik Koch, and Poul Jargensen, “The second-order approximate coupled cluster singles and doubles model cc2,” *Chemical Physics Letters* **243**, 409 – 418 (1995).
- [203] Florian Weigend and Reinhart Ahlrichs, “Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy,” *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
- [204] M. Gruebele and P. G. Wolynes, “Vibrational energy flow and chemical reactions,” *Accounts of Chemical Research* **37**, 261–267 (2004).
- [205] David M. Leitner, “Energy flow in proteins,” *Annual Review of Physical Chemistry* **59**, 233–259 (2008).
- [206] Dvira Segal and Bijay Kumar Agarwalla, “Vibrational heat transport in molecular junctions,” *Annual Review of Physical Chemistry* **67**, 185–209 (2016).
- [207] Maja Kobus, Phuong H. Nguyen, and Gerhard Stock, “Coherent vibrational energy transfer along a peptide helix,” *The Journal of Chemical Physics* **134**, 124518 (2011).
- [208] A. Arkhipov and G. Kuperberg, “The bosonic birthday paradox,” *Geometry & Topology Monographs* **18**, 1–7 (2012).

- [209] Sander Woutersen and Huib J. Bakker, “Resonant intermolecular transfer of vibrational energy in liquid water,” *Nature* **402**, 507–509 (1999).
- [210] Jrg Lindner, Peter Vhringer, Maxim S. Pshenichnikov, Dan Cringus, Douwe A. Wiersma, and Maxim Mostovoy, “Vibrational relaxation of pure liquid water,” *Chemical Physics Letters* **421**, 329 – 333 (2006).
- [211] Krupa Ramasesha, Luigi De Marco, Aritra Mandal, and Andrei Tokmakoff, “Water vibrations have strongly mixed intra- and intermolecular character,” *Nat Chem* **5**, 935–940 (2013).
- [212] Roger A Horn and Charles R Johnson, *Topics in Matrix Analysis* (Cambridge University Press, 1991).
- [213] R. D. Amos, N. C. Handy, W. H. Green, D. Jayatilaka, A. Willetts, and P. Palmieri, “Anharmonic vibrational properties of CH_2F_2 : A comparison of theory and experiment,” *The Journal of Chemical Physics* **95**, 8323–8336 (1991).
- [214] Wataru Mizukami and David P. Tew, “A second-order multi-reference perturbation method for molecular vibrations,” *The Journal of Chemical Physics* **139**, 194108 (2013).
- [215] D. Kraft, *A Software Package for Sequential Quadratic Programming*, Tech. Rep. (DFVLR-FB 88-28, 1988).
- [216] Eric Jones, Travis Oliphant, Pearu Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” (2001–).
- [217] Joonsuk Huh, Gian Giacomo Guerreschi, Borja Peropadre, Jarrod R. McClean, and Alán Aspuru-Guzik, “Boson sampling for molecular vibronic spectra,” *Nat Photon* **9**, 615–620 (2015).
- [218] Andreas Kohn and Christof Httig, “Analytic gradients for excited states in the coupled-cluster model cc_2 employing the resolution-of-the-identity approximation,” *The Journal of Chemical Physics* **119**, 5021–5036 (2003).
- [219] John A. Nelder and Roger Mead, “A simplex method for function minimization,” *Computer Journal* **7**, 308–313 (1965).
- [220] Alex Arkhipov, “Bosonsampling is robust against small errors in the network matrix,” *Phys. Rev. A* **92**, 062326 (2015).
- [221] Nicholas J Russell, Levon Chakhmakhchyan, Jeremy L O’Brien, and Anthony Laing, “Direct dialling of haar random unitary matrices,” *New Journal of Physics* **19**, 033007 (2017).
- [222] C. M. Wilkes, X. Qiang, J. Wang, R. Santagati, S. Paesani, X. Zhou, D. A. B. Miller, G. D. Marshall, M. G. Thompson, and J. L. O’Brien, “60db high-extinction

- auto-configured mach–zehnder interferometer,” *Opt. Lett.* **41**, 5318–5321 (2016).
- [223] Nicholas C. Harris, Gregory R. Steinbrecher, Mihika Prabhu, Yoav Lahini, Jacob Mower, Darius Bunandar, Changchen Chen, Franco N. C. Wong, Tom Baehr-Jones, Michael Hochberg, Seth Lloyd, and Dirk Englund, “Quantum transport simulations in a programmable nanophotonic processor,” *Nat Photon* **11**, 447–452 (2017).
- [224] Michael Varnava, Daniel E. Browne, and Terry Rudolph, “Loss tolerance in one-way quantum computation via counterfactual error correction,” *Phys. Rev. Lett.* **97**, 120501 (2006).
- [225] Sean D. Barrett and Thomas M. Stace, “Fault tolerant quantum computation with very high threshold for loss errors,” *Phys. Rev. Lett.* **105**, 200502 (2010).
- [226] Terry Rudolph, “Why i am optimistic about the silicon-photonics route to quantum computing,” *APL Photonics* **2**, 030901 (2017).
- [227] Jonathan Crickmore, Jonathan Frazer, Scott Shaw, and Pieter Kok, “Effect of component variations on the gate fidelity in linear optical networks,” *Phys. Rev. A* **94**, 022326 (2016).
- [228] Adrian J. Menssen, Alex E. Jones, Benjamin J. Metcalf, Malte C. Tichy, Stefanie Barz, W. Steven Kolthammer, and Ian A. Walmsley, “Distinguishability and many-particle interference,” *Phys. Rev. Lett.* **118**, 153603 (2017).
- [229] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters, “Mixed-state entanglement and quantum error correction,” *Phys. Rev. A* **54**, 3824–3851 (1996).
- [230] Peter P. Rohde and Timothy C. Ralph, “Error models for mode mismatch in linear optics quantum computing,” *Phys. Rev. A* **73**, 062312 (2006).
- [231] Fabian Ewert and Peter van Loock, “3/4-efficient bell measurement with passive linear optics and unentangled ancillae,” *Phys. Rev. Lett.* **113**, 140403 (2014).
- [232] W. P. Grice, “Arbitrarily complete bell-state measurement using only linear optical elements,” *Phys. Rev. A* **84**, 042331 (2011).
- [233] Michael Varnava, Daniel E. Browne, and Terry Rudolph, “How good must single photon sources and detectors be for efficient linear optical quantum computation?” *Phys. Rev. Lett.* **100**, 060502 (2008).
- [234] M. Hein, J. Eisert, and H. J. Briegel, “Multiparty entanglement in graph states,” *Phys. Rev. A* **69**, 062311 (2004).
- [235] Koji Azuma, Kiyoshi Tamaki, and Hoi-Kwong Lo, “All-photonics quantum repeaters,” *Nature Communications* **6**, 6787 (2015).
- [236] Mihir Pant, Hari Krovi, Dirk Englund, and Saikat Guha, “Rate-distance tradeoff

- and resource costs for all-optical quantum repeaters,” *Phys. Rev. A* **95**, 012304 (2017).
- [237] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani, “Device-independent security of quantum cryptography against collective attacks,” *Phys. Rev. Lett.* **98**, 230501 (2007).
- [238] Jaromír Fiurášek, Raúl García-Patrón, and Nicolas J. Cerf, “Conditional generation of arbitrary single-mode quantum states of light by repeated photon subtractions,” *Phys. Rev. A* **72**, 033822 (2005).
- [239] Mikkel Heuck, Mihir Pant, and Dirk Englund, “Temporally and frequency multiplexed single photon source using quantum feedback control for scalable photonic quantum technologies,” (2017), arXiv:1708.08875 .
- [240] W. H. P. Pernice, C. Schuck, O. Minaeva, M. Li, G. N. Goltsman, A. V. Sergienko, and H. X. Tang, “High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits,” *Nat. Comms.* **3**, 1325 (2012).
- [241] J. W. Silverstone, R. Santagati, D. Bonneau, M. J. Strain, M. Sorel, J. L. O’Brien, and M. G. Thompson, “Qubit entanglement between ring-resonator photon-pair sources on a silicon chip,” *Nat. Comms.* **6**, 7948 (2015).
- [242] Simon Laibacher and Vincenzo Tamma, “From the physics to the computational complexity of multiboson correlation interference,” *Phys. Rev. Lett.* **115**, 243605 (2015).
- [243] 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,” *Nat. Comms.* **6**, 5873 (2015).
- [244] K. Tanizawa, K. Suzuki, S. Suda, H. Matsuura, K. Ikeda, S. Namiki, and H. Kawashima, “Silicon photonic 32 x 32 strictly-non-blocking blade switch and its full path characterization,” in *2016 21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)* (2016) pp. 1–3.
- [245] Tae Joon Seok, Niels Quack, Sangyoon Han, Richard S. Muller, and Ming C. Wu, “Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers,” *Optica* **3**, 64–70 (2016).
- [246] Martin Schwarz and Maarten Van den Nest, “Simulating quantum circuits with sparse output distributions,” (2013), arXiv:1310.6749 .
- [247] Takahiro Inagaki, Yoshitaka Haribara, Koji Igarashi, Tomohiro Sonobe, Shuhei Tamate, Toshimori Honjo, Alireza Marandi, Peter L. McMahon, Takeshi Umeki, Koji Enbutsu, Osamu Tadanaga, Hirokazu Takenouchi, Kazuyuki Aihara, Ken-ichi

- Kawarabayashi, Kyo Inoue, Shoko Utsunomiya, and Hiroki Takesue, “A coherent ising machine for 2000-node optimization problems,” *Science* (2016), 10.1126/science.aah4243.
- [248] Peter L. McMahon, Alireza Marandi, Yoshitaka Haribara, Ryan Hamerly, Carsten Langrock, Shuhei Tamate, Takahiro Inagaki, Hiroki Takesue, Shoko Utsunomiya, Kazuyuki Aihara, Robert L. Byer, M. M. Fejer, Hideo Mabuchi, and Yoshihisa Yamamoto, “A fully-programmable 100-spin coherent ising machine with all-to-all connections,” *Science* (2016), 10.1126/science.aah5178.
- [249] Mercedes Gimeno-Segovia, Hugo Cable, Gabriel J Mendoza, Pete Shadbolt, Joshua W Silverstone, Jacques Carolan, Mark G Thompson, Jeremy L O’Brien, and Terry Rudolph, “Relative multiplexing for minimising switching in linear-optical quantum computing,” *New Journal of Physics* **19**, 063013 (2017).
- [250] Sam Morley-Short, Sara Bartolucci, Mercedes Gimeno-Segovia, Pete Shadbolt, Hugo Cable, and Terry Rudolph, “Physical-depth architectural requirements for generating universal photonic cluster states,” (2017), arXiv:1706.07325 .
- [251] Ying Li, Peter C. Humphreys, Gabriel J. Mendoza, and Simon C. Benjamin, “Resource costs for fault-tolerant linear optical quantum computing,” *Phys. Rev. X* **5**, 041007 (2015).

Appendix A

Universal Linear Optics

A.1 Transfer Matrix Reconstructions

Here we provide the transfer matrix reconstructions referred to in Chapter 3.

The circuit for the Bell state generator:

$$M_B = \begin{pmatrix} 0.680 & 0.662 & 0.026 & 0.006 & 0.006 & 0.006 \\ 0.384 & -0.400 + 0.063i & 0.570 & 0.548 & 0.013 & 0.013 \\ 0.384 & -0.385 + 0.047i & -0.254 - 0.281i & -0.275 + 0.306i & 0.416 & 0.388 \\ 0.393 & -0.390 + 0.036i & -0.309 + 0.293i & -0.263 - 0.285i & -0.413 & -0.370 + 0.037i \\ 0.012 & 0.0181 & 0.539 & -0.549 & -0.075 - 0.428i & -0.029 - 0.405i \\ 0.009 & 0.0125 & 0.017 & 0.012 & 0.684 & -0.734 + 0.018i \end{pmatrix} \quad (\text{A.1})$$

The circuit for the KLM CNOT gate:

$$M_{\text{KLM}} = \begin{pmatrix} 0.468 & 0.612 & 0.430 & 0.402 & 0.009 & 0.009 \\ 0.606 & -0.478 & 0.015 & 0.035 & 0.624 & 0.022 \\ 0.365 & 0.061 & 0.292 & -0.672 & -0.394 & 0.329 \\ 0.367 & 0.020 & -0.686 + 0.150i & 0.266 & -0.352 + 0.075i & 0.315 + 0.217i \\ 0.010 & 0.605 & -0.4182 + 0.042i & -0.466 + 0.068i & 0.461 & 0.048 \\ 0.286 & 0.029 & -0.174 + 0.114i & -0.121 + 0.103i & -0.323 + 0.040i & -0.794 + 0.288i \end{pmatrix} \quad (\text{A.2})$$

The circuit for postselected CNOT gate:

$$M_{\text{post}} = \begin{pmatrix} 0.559 & 0.773 & 0.029 & 0.000 & 0.006 & 0.006 \\ 0.783 & -0.571 + 0.055i & 0.022 & 0.000 & 0.008 & 0.011 \\ 0.009 & 0.021 & 0.556 & 0.558 & 0.579 & 0.006 \\ 0.010 & 0.025 & 0.577 & 0.078 & -0.566 + 0.139i & 0.531 \\ 0.010 & 0.019 & 0.568 & -0.556 + 0.085i & 0.066 & -0.533 + 0.105i \\ 0.009 & 0.011 & 0.016 & 0.551 & -0.566395 & -0.587 - 0.111i \end{pmatrix} \quad (\text{A.3})$$

Appendix B

Simulating the Vibrational Quantum Dynamics of Molecules with Integrated Photonics

B.1 Molecule details

Here we include all additional computed vibrational frequencies and local basis transformations.

B.1.1 HFHF

The frequencies of the normal modes are calculated to be

$$\{4068.29, 3981.51, 578.25, 464.46, 207.95, 148.93\}\text{cm}^{-1} \quad (\text{B.1})$$

and the transformation between the normal modes and localised modes is given by

$$U_L = \begin{pmatrix} 0.768 & 0.147 & 0.281 & 0. & -0.410 & -0.377 \\ -0.178 & 0.920 & 0.265 & 0. & 0.220 & -0.046 \\ 0.064 & -0.303 & 0.844 & 0. & 0.357 & 0.254 \\ 0. & 0. & 0. & 1. & 0. & 0. \\ 0.607 & 0.135 & -0.351 & 0. & 0.462 & 0.525 \\ -0.076 & 0.145 & 0.123 & 0. & -0.665 & 0.718 \end{pmatrix} \quad (\text{B.2})$$

B.1.2 HNCO

The frequencies of the normal modes are calculated to be

$$\{3694.87, 2260.22, 1241.85, 805.38, 605.69, 552.52\}\text{cm}^{-1} \quad (\text{B.3})$$

and the transformation between the normal modes and localised modes is given by

$$U_L = \begin{pmatrix} 0.826 & 0.116 & -0.198 & -0.438 & 0. & 0.272 \\ -0.066 & 0.982 & 0.120 & 0.001 & 0. & -0.129 \\ 0.157 & -0.090 & 0.971 & -0.085 & 0. & 0.133 \\ 0.537 & -0.037 & 0.039 & 0.667 & 0. & -0.514 \\ 0. & 0. & 0. & 0. & 1. & 0. \\ 0.028 & 0.111 & -0.050 & 0.597 & 0. & 0.792 \end{pmatrix} \quad (\text{B.4})$$

B.1.3 N₄

The frequencies of the normal modes are calculated to be

$$\{1366.42, 1015.85, 947.85, 916.73, 412.35, 366.86\}\text{cm}^{-1} \quad (\text{B.5})$$

and the transformation between the normal modes and localised modes is given by

$$U_L = \begin{pmatrix} 0.191 & -0.356 & -0.611 & 0.494 & 0. & 0.469 \\ 0.658 & 0.611 & -0.356 & -0.011 & 0. & -0.258 \\ 0.658 & -0.611 & 0.356 & -0.011 & 0. & -0.258 \\ 0.191 & 0.356 & 0.611 & 0.494 & 0. & 0.469 \\ 0. & 0. & 0. & 0. & 1. & 0. \\ 0.244 & 0. & 0. & -0.716 & 0. & 0.654 \end{pmatrix} \quad (\text{B.6})$$

B.1.4 P_4

The frequencies of the normal modes are calculated to be

$$\{608.25, 460.20, 460.20, 460.20, 366.27, 366.25\} \text{cm}^{-1} \quad (\text{B.7})$$

and the transformation between the normal modes and localised modes is given by

$$U_L = \begin{pmatrix} 0.5 & -0.676 & -0.096 & -0.533 & 0. & 0. \\ 0.5 & 0.727 & 0.074 & -0.465 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0.853 & -0.522 \\ 0.5 & 0.059 & -0.691 & 0.519 & 0. & 0. \\ -0.5 & 0.110 & -0.713 & -0.480 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0.522 & 0.853 \end{pmatrix} \quad (\text{B.8})$$

B.1.5 SO_3

The frequencies of the normal modes are calculated to be

$$\{1334.97, 1334.55, 966.96, 497.57, 497.52, 469.50\} \text{cm}^{-1} \quad (\text{B.9})$$

and the transformation between the normal modes and localised modes is given by

$$U_L = \begin{pmatrix} 0.608 & -0.346 & -0.577 & -0.218 & 0.362 & 0. \\ 0. & 0.701 & -0.579 & 0.417 & 0. & 0. \\ 0.608 & 0.346 & 0.577 & 0.218 & 0.362 & 0. \\ 0. & -0.519 & -0.012 & 0.855 & 0. & 0. \\ -0.511 & 0. & 0. & 0. & 0.859 & 0. \\ 0. & 0. & 0. & 0. & 0. & 1. \end{pmatrix} \quad (\text{B.10})$$

B.1.6 NMA

The vibrational frequencies of NMA are calculated to be

$$\{3589.68, 3143.77, 3113.38, 3111.68, 3060.23, 3024.31, 3000.1, 1808.96, 1572.7, 1484.02, 1481.06, 1469.21, 1452.61, 1427.76, 1391.94, 1279.31, 1170.4, 1150.05, 1120.42, 1053.5, 995.522, 876.075, 637.311, 628.305, 475.592, 436.109, 291.874, 176.452, 96.9364, 39.7879\} \text{cm}^{-1} \quad (\text{B.11})$$

and the transformation used is $U_L \oplus \mathbb{1}_{24}$ where we select the subset

$$\{1279.31, 1170.4, 1120.42, 995.522, 876.075, 628.305\} \quad (\text{B.12})$$

to be the first six modes.

$$U_L = \begin{pmatrix} 0.796 & -0.007 & -0.401 & -0.345 & 0.050 & -0.290 \\ -0.115 & 0.794 & 0.141 & -0.178 & 0.500 & -0.233 \\ 0.518 & -0.085 & 0.807 & 0.139 & 0.156 & 0.172 \\ 0.211 & 0.147 & -0.312 & 0.889 & 0.214 & -0.012 \\ -0.190 & -0.583 & 0.006 & -0.031 & 0.704 & -0.357 \\ 0.059 & -0.011 & -0.267 & -0.197 & 0.427 & 0.839 \end{pmatrix} \quad (\text{B.13})$$

B.1.7 H₂O

The vibrational frequencies of the normal modes in the harmonic model are

$$\{3914.92, 3787.59, 1627.01\}\text{cm}^{-1}$$

and the transformation between normal and local modes is

$$U_L = \begin{pmatrix} 0.70711 & -0.70709 & -0.00502 \\ 0.70711 & 0.70709 & 0.00502 \\ 0. & -0.00710 & 0.99997 \end{pmatrix} \quad (\text{B.14})$$

Now, restricting to the subspace of the asymmetric and symmetric stretch modes and adjusting the energy levels for anharmonic potentials via perturbation theory results in single excitation energy levels of

$$\{3740.05, 3619.68\}\text{cm}^{-1}.$$

With two excitations the $\{20, 02, 11\}$ energy levels become

$$\{7391.43, 7154.35, 7206.46\}\text{cm}^{-1}$$

and the normal to local transformation for this subspace is given by

$$U_L = \begin{pmatrix} 0.70711 & -0.70711 & 0. \\ 0.70711 & 0.70711 & 0. \\ 0. & 0. & 1. \end{pmatrix} \quad (\text{B.15})$$

B.1.8 NH₃

The vibrational frequencies of the ground electronic state are given by

$$\{3646.13, 3646.09, 3503.94, 1678.57, 1678.4, 1065.59\}\text{cm}^{-1} \quad (\text{B.16})$$

and the frequencies of the excited state are

$$\{2590.27, 2452.95, 2451.97, 1052.62, 1050.93, 750.124\}\text{cm}^{-1}. \quad (\text{B.17})$$

The transformation between the two sets of modes is given by

$$U_{\text{GE}} = \begin{pmatrix} -0.085 & -0.289 & -0.216 & 0.090 & 0.913 & -0.147 \\ -0.058 & -0.011 & 0.096 & 0.023 & 0.169 & 0.979 \\ 0.935 & -0.023 & -0.341 & -0.031 & 0.017 & 0.087 \\ 0.131 & 0.859 & 0.313 & -0.118 & 0.358 & -0.072 \\ 0.308 & -0.336 & 0.819 & 0.331 & 0.0694 & -0.0856 \\ -0.052 & 0.256 & -0.244 & 0.931 & -0.071 & 0.014 \end{pmatrix} \quad (\text{B.18})$$

and the transformation between normal and local coordinates in the excited state is given by

$$U_L = \begin{pmatrix} 0.575 & -0.542 & -0.576 & -0.058 & 0.200 & 0. \\ 0.579 & 0.767 & -0.185 & 0.195 & -0.062 & 0. \\ 0.578 & -0.223 & 0.757 & -0.163 & -0.131 & 0. \\ 0.012 & -0.139 & 0.212 & 0.857 & 0.448 & 0. \\ -0.010 & 0.221 & 0.126 & -0.444 & 0.859 & 0. \\ 0. & 0. & 0. & 0. & 0. & 1. \end{pmatrix} \quad (\text{B.19})$$

B.2 Additional data

Here we include plots displaying the full data from our experiments. All plots represent the experimentally estimated probabilities for measuring patterns ordered top to bottom as $\{|100000\rangle, |010000\rangle, \dots\}$ for one photon data, $\{|110000\rangle, |101000\rangle, \dots\}$ for two photon data without number resolving, $\{|20000\rangle, |020000\rangle, \dots, |110000\rangle, \dots\}$ for two photon data with number resolving, $\{|111000\rangle, |110100\rangle, \dots\}$ for three photon data and $\{|22000\rangle, |202000\rangle, \dots, |211000\rangle, \dots, |111100\rangle\}$ for four photon data. The mode ordering is defined by the vibrational frequencies defined in the above Molecule details section.

B. Simulating the Vibrational Quantum Dynamics of Molecules with Integrated Photonics

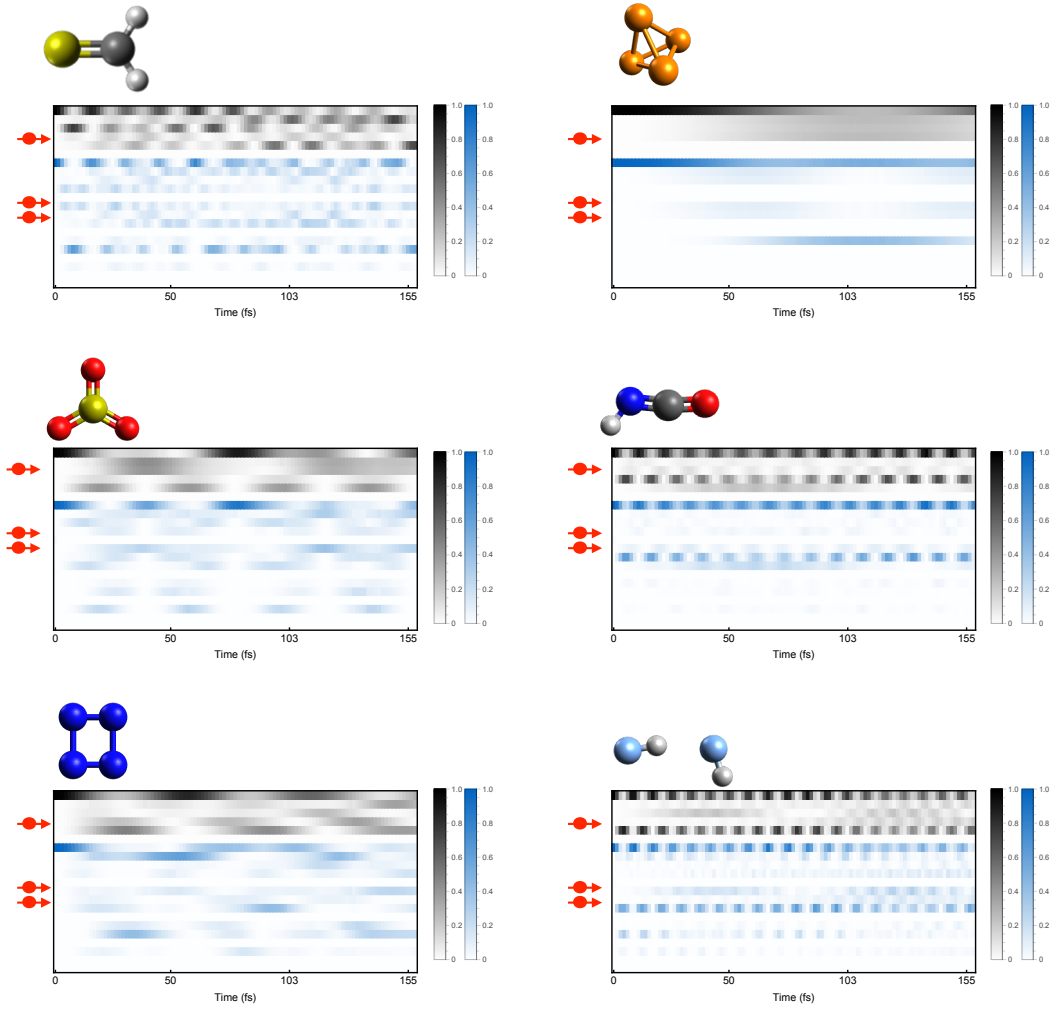


Figure B.1: Four atom molecule data. Full data sets for $|1\rangle$ (black) and $|1\rangle|1\rangle$ (blue) input experiments for H_2CS , P_4 , SO_3 , HNCO , N_4 and HFHF .

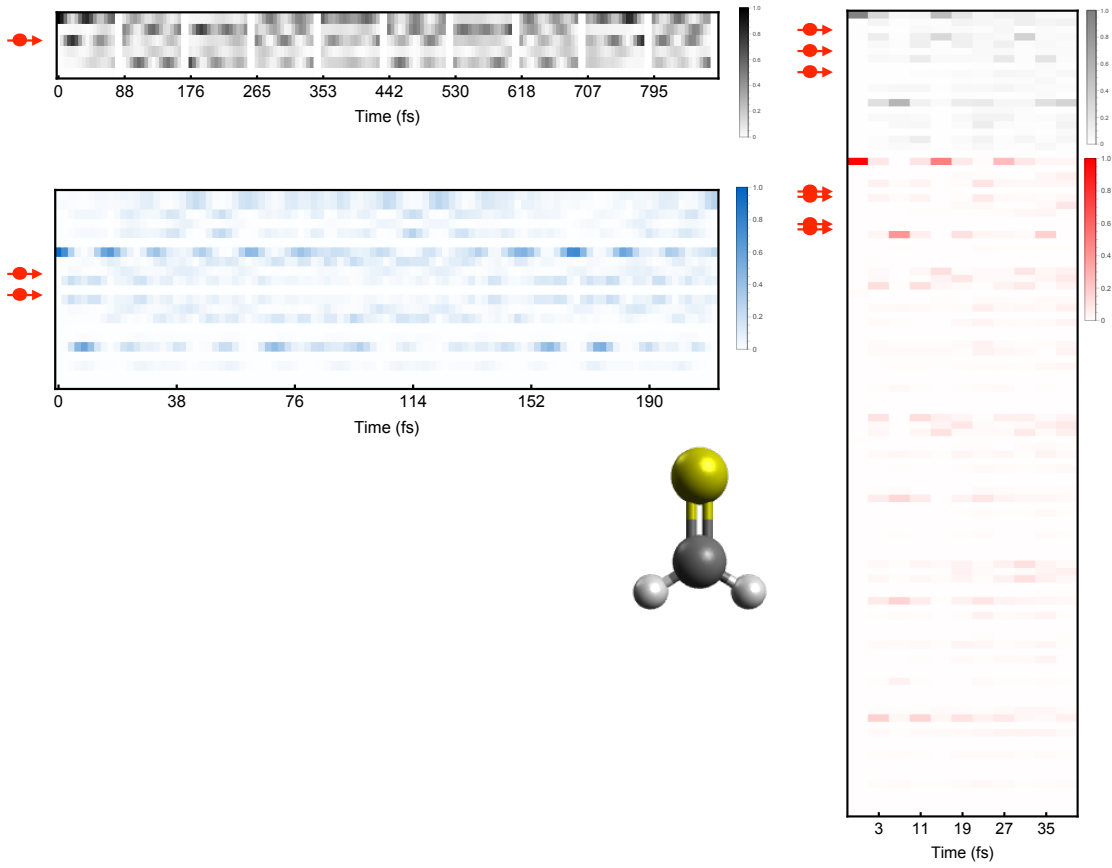


Figure B.2: H₂CS data. Full data sets for $|1\rangle$ (black), $|1\rangle|1\rangle$ (blue), $|1\rangle|1\rangle|1\rangle$ (grey) and $|2\rangle|2\rangle$ (red) input experiments.

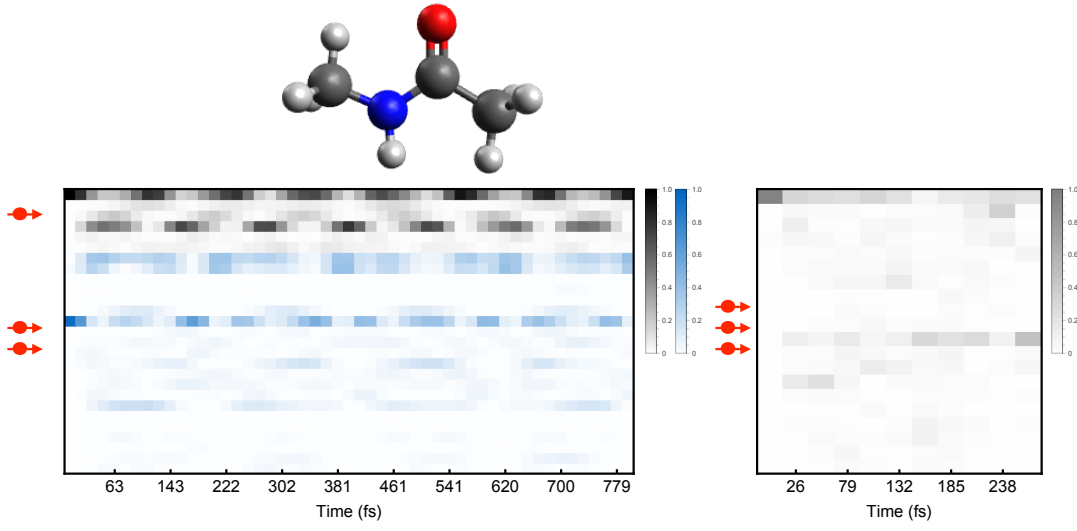


Figure B.3: NMA data. Full data sets for $|1\rangle$ (black), $|1\rangle|1\rangle$ (blue) and $|1\rangle|1\rangle|1\rangle$ (grey) input experiments.

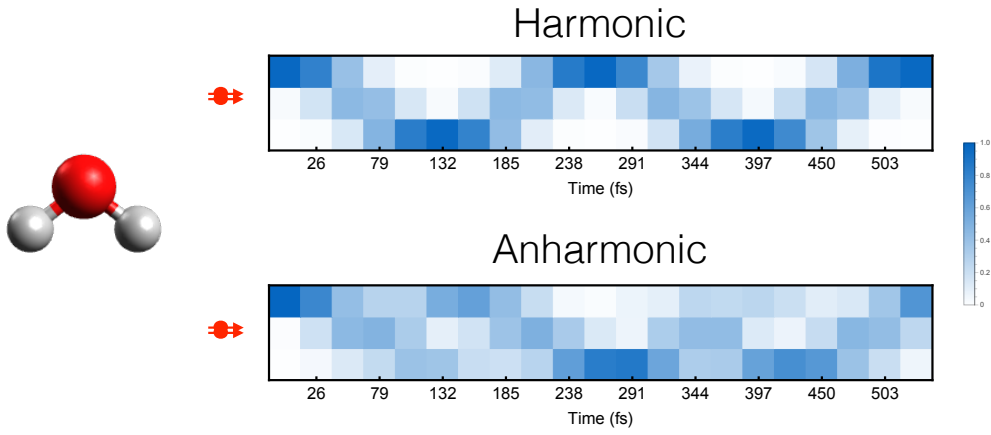


Figure B.4: H_2O data. Full data sets for $|2\rangle$ and $|2\rangle|1\rangle$ input experiments for harmonic and anharmonic. States are ordered top to bottom as $\{|20\rangle, |11\rangle, |02\rangle\}$

B. Simulating the Vibrational Quantum Dynamics of Molecules with Integrated Photonics

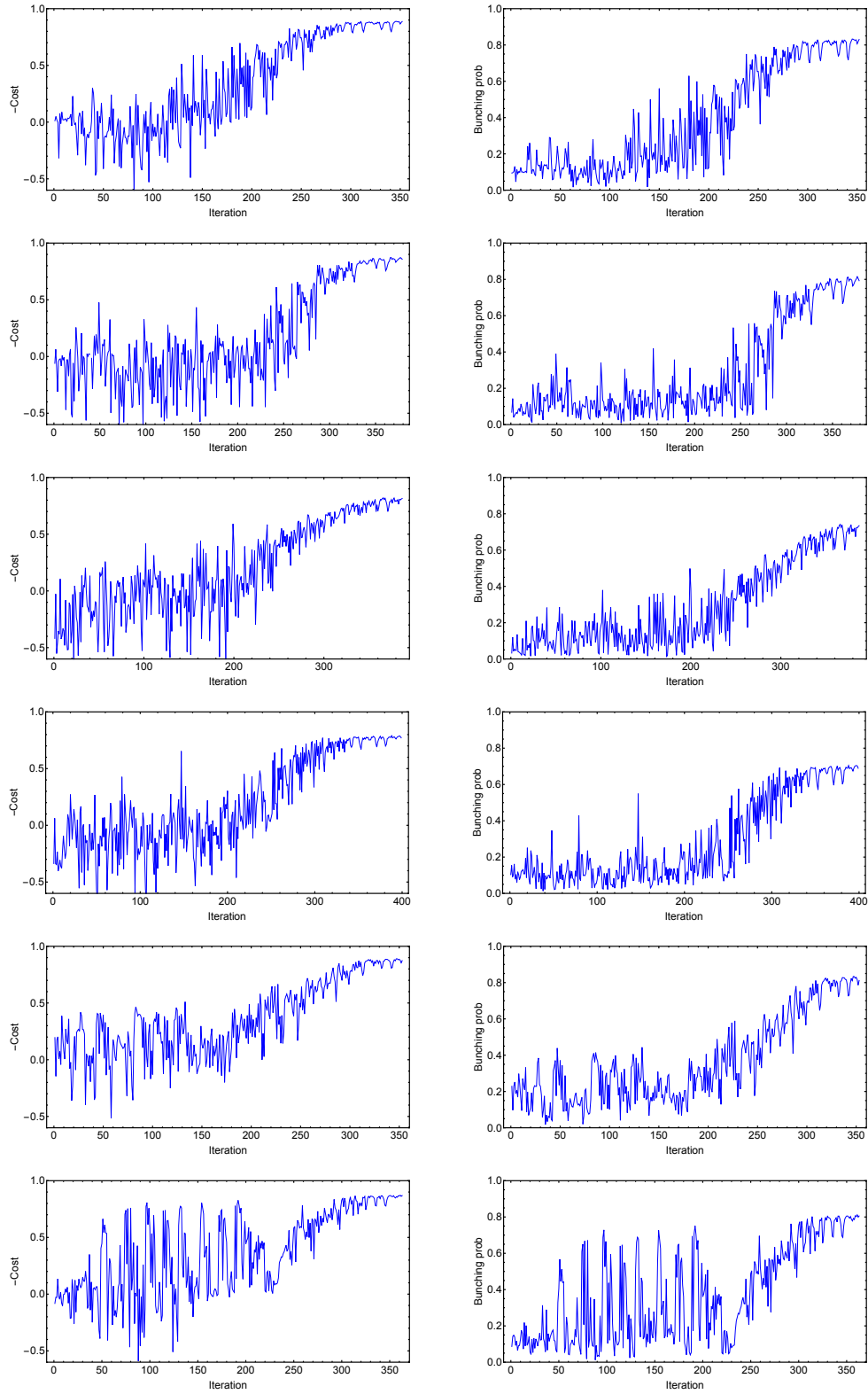


Figure B.5: Cost function (left) and probability of bunching in an NH stretch mode (right) for a series of random initial states in the NH_3 dissociation experiment.

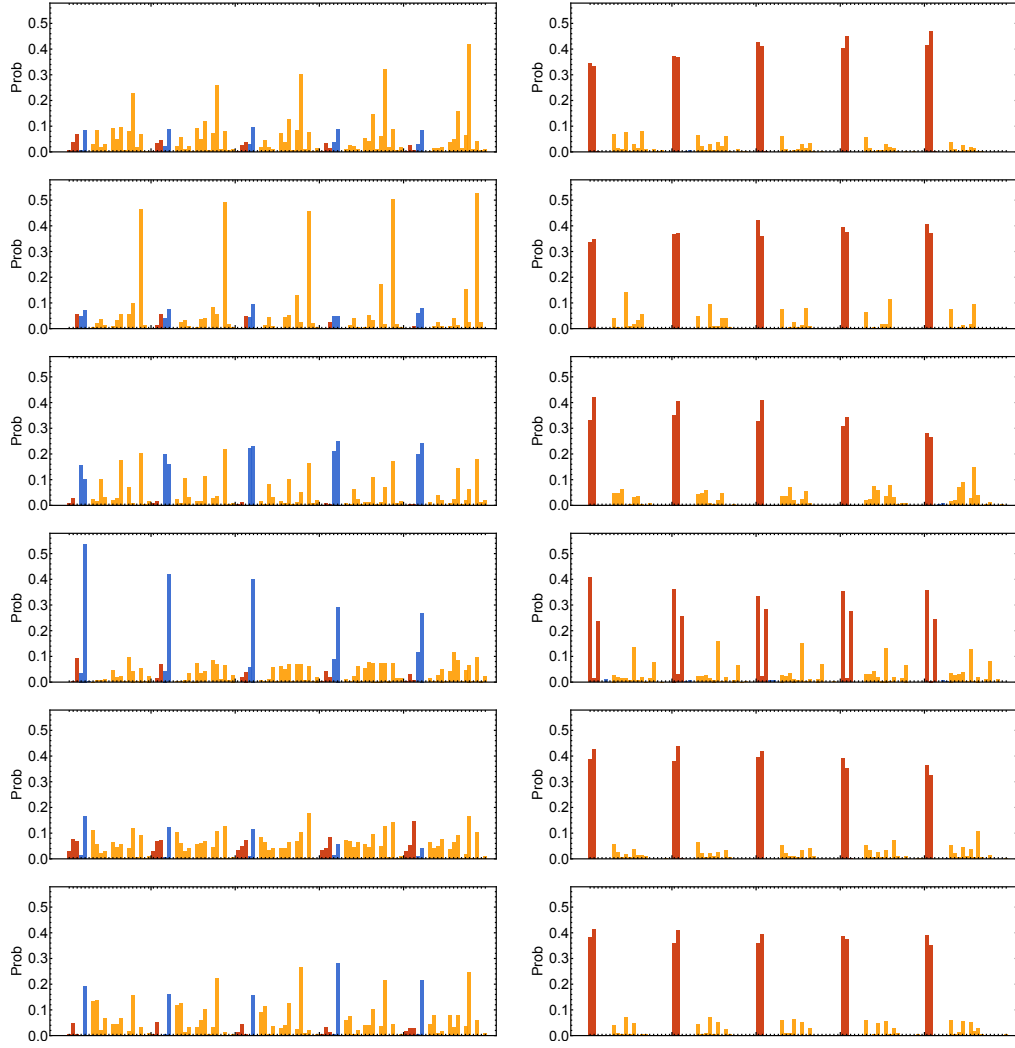


Figure B.6: Full distributions for initial (left column) and final (right column) states over the five time steps for the dissociation experiments summarised in Fig. B.5. Red bars are probability of bunching in an NH stretch mode, blue bars are remaining bunched states and yellow are anti-bunched states.

Appendix C

Linear Optical Quantum Computing with Partially-Distinguishable Photons

C.1 Trace inequality proof

We note that for two positive matrices ϱ_i, ϱ_j :

$$\mathrm{Tr}(\varrho_i \varrho_j) = \sum_k [s_k(\sqrt{\varrho_i} \sqrt{\varrho_j})]^2 \quad (\text{C.1})$$

where $s_k(A)$ is the k^{th} largest singular value of the matrix A . Proof: Due to the cyclic property of the trace $\mathrm{Tr}(\varrho_i \varrho_j) = \mathrm{Tr}(\sqrt{\varrho_i} \sqrt{\varrho_j} \sqrt{\varrho_j} \sqrt{\varrho_i})$ and then using the singular value decomposition of $\sqrt{\varrho_i} \sqrt{\varrho_j} = U D V$ and that $(\sqrt{\varrho_i} \sqrt{\varrho_j})^\dagger = \sqrt{\varrho_j} \sqrt{\varrho_i}$ (relies on positivity of ϱ 's), $\mathrm{Tr}(\varrho_i \varrho_j) = \mathrm{Tr}(U D V V^\dagger D U^\dagger) = \mathrm{Tr}(D^2)$.

Using von-neumans trace inequality:

$$\begin{aligned} \left| \mathrm{Tr} \left(\prod_{i=1}^n \varrho_i \right) \right| &= \left| \mathrm{Tr} (\sqrt{\varrho_1} \sqrt{\varrho_2} \sqrt{\varrho_2} \sqrt{\varrho_3} \sqrt{\varrho_3} \dots \sqrt{\varrho_n} \sqrt{\varrho_1}) \right| \\ &\leq \sum_k s_k(\sqrt{\varrho_1} \sqrt{\varrho_2}) s_k(\sqrt{\varrho_2} \sqrt{\varrho_3}) \dots s_k(\sqrt{\varrho_n} \sqrt{\varrho_1}) \end{aligned} \quad (\text{C.2})$$

Now we consider vectors $|v(i, j)\rangle$ with components given by the singular values $[v(i, j)]_k = s_k(\sqrt{\varrho_i}\sqrt{\varrho_j})$ and a vector $|\mathcal{V}\rangle$ with components $[\mathcal{V}]_k = s_k(\sqrt{\varrho_1}\sqrt{\varrho_2})s_k(\sqrt{\varrho_2}\sqrt{\varrho_3})\dots s_k(\sqrt{\varrho_{n-1}}\sqrt{\varrho_n})$ such that:

$$\begin{aligned} \sum_k s_k(\sqrt{\varrho_1}\sqrt{\varrho_2})s_k(\sqrt{\varrho_2}\sqrt{\varrho_3})\dots s_k(\sqrt{\varrho_n}\sqrt{\varrho_1}) \\ = \langle \mathcal{V} | v(n, 1) \rangle \\ \leq \sqrt{\langle \mathcal{V} | \mathcal{V} \rangle \langle v(n, 1) | v(n, 1) \rangle} \\ \leq \sqrt{\langle v(1, 2) | v(1, 2) \rangle \langle v(2, 3) | v(2, 3) \rangle \dots \langle v(n, 1) | v(n, 1) \rangle} \\ = \sqrt{\text{Tr}(\varrho_1 \varrho_2) \text{Tr}(\varrho_2 \varrho_3) \dots \text{Tr}(\varrho_n \varrho_1)}. \end{aligned} \tag{C.3}$$

Hence:

$$\left| \text{Tr} \left(\prod_{i=1}^n \varrho_i \right) \right| \leq \sqrt{\text{Tr}(\varrho_1 \varrho_2) \text{Tr}(\varrho_2 \varrho_3) \dots \text{Tr}(\varrho_n \varrho_1)} \tag{C.4}$$

C.2 Bound on real term

Aim is to upper and lower bound the expression:

$$\Re [\text{Tr}(\varrho_1 \varrho_2 \varrho_4 \varrho_3) - \text{Tr}(\varrho_1 \varrho_2 \varrho_3 \varrho_4)] = \Re [\text{Tr}(\varrho_1 \varrho_2 [\varrho_4, \varrho_3])]. \tag{C.5}$$

First we note that $[\varrho_4, \varrho_3]$ is anti-hermitian. Next we will make use of the fact that: the trace of the product of an anti-hermitian operator A and hermitian operator H , $\text{Tr}(AH)$, is imaginary. Proof:

$$\begin{aligned} \overline{\text{Tr}(AH)} &= \text{Tr}[(AH)^\dagger] \\ &= \text{Tr}[H^\dagger A^\dagger] \\ &= -\text{Tr}[HA] \\ &= -\text{Tr}[AH] \end{aligned} \tag{C.6}$$

where we have used the cyclic property and linearity of the trace. Similarly it can be found that $\text{Tr}(H_1 H_2)$ is real for two hermitian matrices H_1 and H_2 . Therefore we can

split $\varrho_1\varrho_2$ into a hermitian part: $\frac{1}{2}\{\varrho_1, \varrho_2\}$, and an anti-hermitian part: $\frac{1}{2}[\varrho_1, \varrho_2]$ together with the linearity of the trace this implies:

$$\Re \{ \text{Tr} (\varrho_1\varrho_2[\varrho_4, \varrho_3]) \} = \frac{1}{2} \text{Tr}([\varrho_1, \varrho_2][\varrho_4, \varrho_3]) . \quad (\text{C.7})$$

We can now bound the magnitude of this expression using the Hilbert-Schmidt inner product ($\langle A|B \rangle = \text{Tr}[A^\dagger B]$)

$$\begin{aligned} |\text{Tr} ([\varrho_1, \varrho_2][\varrho_4, \varrho_3])| &= |\langle [\varrho_2, \varrho_1][\varrho_4, \varrho_3] \rangle| \\ &\leq \|[\varrho_2, \varrho_1]\| \|[\varrho_4, \varrho_3]\|. \end{aligned} \quad (\text{C.8})$$

For pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$ by decomposing $|\phi\rangle = \alpha|\psi\rangle + \beta|\psi_\perp\rangle$ using gram schmidt we can compute $[|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|] = \alpha\bar{\beta}|\psi\rangle\langle\psi_\perp| - \bar{\alpha}\beta|\psi_\perp\rangle\langle\psi|$ and therefore:

$$\begin{aligned} \| [|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|] \| &= 2|\alpha\beta| \\ &= 2\sqrt{V(1-V)} \end{aligned} \quad (\text{C.9})$$

where $V = \text{Tr}(|\psi\rangle\langle\psi||\phi\rangle\langle\phi|) = |\alpha|^2$. Therefore, for two states $\varrho_\psi = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ and $\varrho_\phi = \sum_i q_i |\phi_i\rangle\langle\phi_i|$, using the convexity of the norm:

$$\begin{aligned} \|[\varrho_\psi, \varrho_\phi]\| &\leq \sum_{i,j} p_i q_j \| [|\psi_i\rangle\langle\psi_i|, |\phi_j\rangle\langle\phi_j|] \| \\ &= 2 \sum_{i,j} p_i q_j \sqrt{V_{ij}(1-V_{ij})} \\ &\leq 2\sqrt{\mathcal{V}(1-\mathcal{V})} \end{aligned} \quad (\text{C.10})$$

where $\mathcal{V} \equiv \sum_{i,j} p_i q_j V_{ij} = \text{Tr}(\varrho_\psi \varrho_\phi)$. The bound on the last line is due to Jensen's inequality since the function $f(x) = \sqrt{x(1-x)}$ is concave for $0 \leq x \leq 1$. Using equations

(C.7), (C.8) and (C.10) we observe:

$$\begin{aligned}
 & -2\sqrt{\mathcal{V}_{12}(1-\mathcal{V}_{12})\mathcal{V}_{34}(1-\mathcal{V}_{34})} \\
 & \leq \Re[\text{Tr}(\varrho_1\varrho_2[\varrho_4, \varrho_3])] \\
 & \leq 2\sqrt{\mathcal{V}_{12}(1-\mathcal{V}_{12})\mathcal{V}_{34}(1-\mathcal{V}_{34})}
 \end{aligned} \tag{C.11}$$

where $\mathcal{V}_{ij} = \text{Tr}(\varrho_i\varrho_j)$.

C.3 Additional transfer matrices

The transfer matrix for the GHZ generator circuit in Fig. 6.3c is given by

$$U_{GHZ} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & \sqrt{2} & -\sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \end{pmatrix}$$

The transfer matrix for the boosted fusion circuit in Fig. 6.3d is given by

$$U_F = \frac{1}{2} \begin{pmatrix} 1 & i & i & 0 & -1 & 0 & 0 & 0 \\ i & 1 & 0 & i & 0 & -1 & 0 & 0 \\ i & -1 & 1 & 0 & i & 0 & 0 & 0 \\ -1 & i & 0 & 1 & 0 & i & 0 & 0 \\ 0 & 0 & i & 0 & 1 & 0 & i & -1 \\ 0 & 0 & 0 & i & 0 & 1 & -1 & i \\ 0 & 0 & -1 & 0 & i & 0 & 1 & i \\ 0 & 0 & 0 & -1 & 0 & i & i & 1 \end{pmatrix}$$

C.4 Equivalence of pre- and post-fusion corrections

Here we briefly explain why local post-fusion correcting Pauli unitaries can be replaced with pre-fusion unitaries in the case of a graph-basis diagonal error model, leading to the simplified averaged measurement operators provided at the end of Sec. 6.6.3.

The average successful post-fusion state for a logical measurement with operators labelled by i acting upon the first two qubits, and associated Pauli correction operator K_i acting on the remaining qubits is

$$\rho_{\text{out}} = \frac{1}{\mathcal{N}} \sum_i K_i \text{Tr}_{1,2}[E^i \rho_{\text{in}}] K_i \quad (\text{C.12})$$

where $\mathcal{N}$ is a normalisation constant. Ideally, all of the measurement operators can be decomposed as $E^i = \tilde{K}_i E \tilde{K}_i$ where $\tilde{K}_i$ are local Paulis acting on the measured qubits. The corrections K_i are then chosen such that their combined operation with $\tilde{K}_i$ stabilises the ideal pre-fusion graph state $(\tilde{K}_i \otimes K_i)|G\rangle = |G\rangle \forall i$, ensuring that any successful measurement outcome leads to the same post-fusion state, $\rho_{\text{out}} = \frac{1}{\mathcal{N}} \text{Tr}_{1,2}[E|G\rangle\langle G|]$.

If the initial logical state is noisy but graph-basis diagonal, $\rho_{\text{in}} = \sum_j p_j |G^j\rangle\langle G^j|$, then

$$(\tilde{K}_i \otimes K_i)|G^j\rangle = \pm |G^j\rangle \forall i, j. \quad (\text{C.13})$$

Therefore for any measurement operators $\{E^i\}$,

$$\rho_{\text{out}} = \frac{1}{\mathcal{N}} \sum_i K_i \text{Tr}_{1,2}[E^i \rho_{\text{in}}] K_i \quad (\text{C.14})$$

$$= \frac{1}{\mathcal{N}} \sum_i \text{Tr}_{1,2}[\tilde{K}_i E^i \tilde{K}_i \rho_{\text{in}}] \quad (\text{C.15})$$

$$= \frac{1}{\mathcal{N}} \text{Tr}_{1,2} \left[\sum_i \tilde{K}_i E^i \tilde{K}_i \rho_{\text{in}} \right] \quad (\text{C.16})$$

$$= \frac{1}{\mathcal{N}} \text{Tr}_{1,2}[\bar{E} \rho_{\text{in}}] \quad (\text{C.17})$$

C.5 Fusion Operations

A feature of the scheme introduced in [120] is that the connectivity of the graph state is optimised when the Bell measurements fail. This requires performing a set of single qubit rotations before and after each Bell measurement. The procedure is depicted in Fig. C.1 for both the fusion operations which form five-qubit microclusters and those which then fuse these clusters together.

For the first case, shown in Fig. C.1a), we can understand its operation via tracking the stabilizers of the state. The stabilizer generators of the initial state are given by

$$\langle X_1 Z_2 A, X_2 Z_1, X_3 Z_4 Z_5, X_4 Z_3 B, X_5 Z_3 C \rangle \quad (\text{C.18})$$

where A, B and C represent Z on all other qubits connected to qubits 1, 4 and 5 respectively. Performing the Clifford rotations HRH to qubit 2 and HR to qubit 3, where H is the Hadamard and $R = (X - Y)/\sqrt{2}$, results in the updated generators

$$\langle X_1 Y_2 A, -X_2 Z_1, Y_3 Z_4 Z_5, -X_4 X_3 B, -X_5 X_3 C \rangle. \quad (\text{C.19})$$

If the measurement is successful, then qubits 2 and 3 are projected onto a Bell state e.g. $|\phi^+\rangle$ (if a different Bell outcome is received then additional correcting Paulis are required) which is stabilized by $\langle X_2 X_3, Z_2 Z_3 \rangle$. The updated generators (omitting qubits 2 and 3)

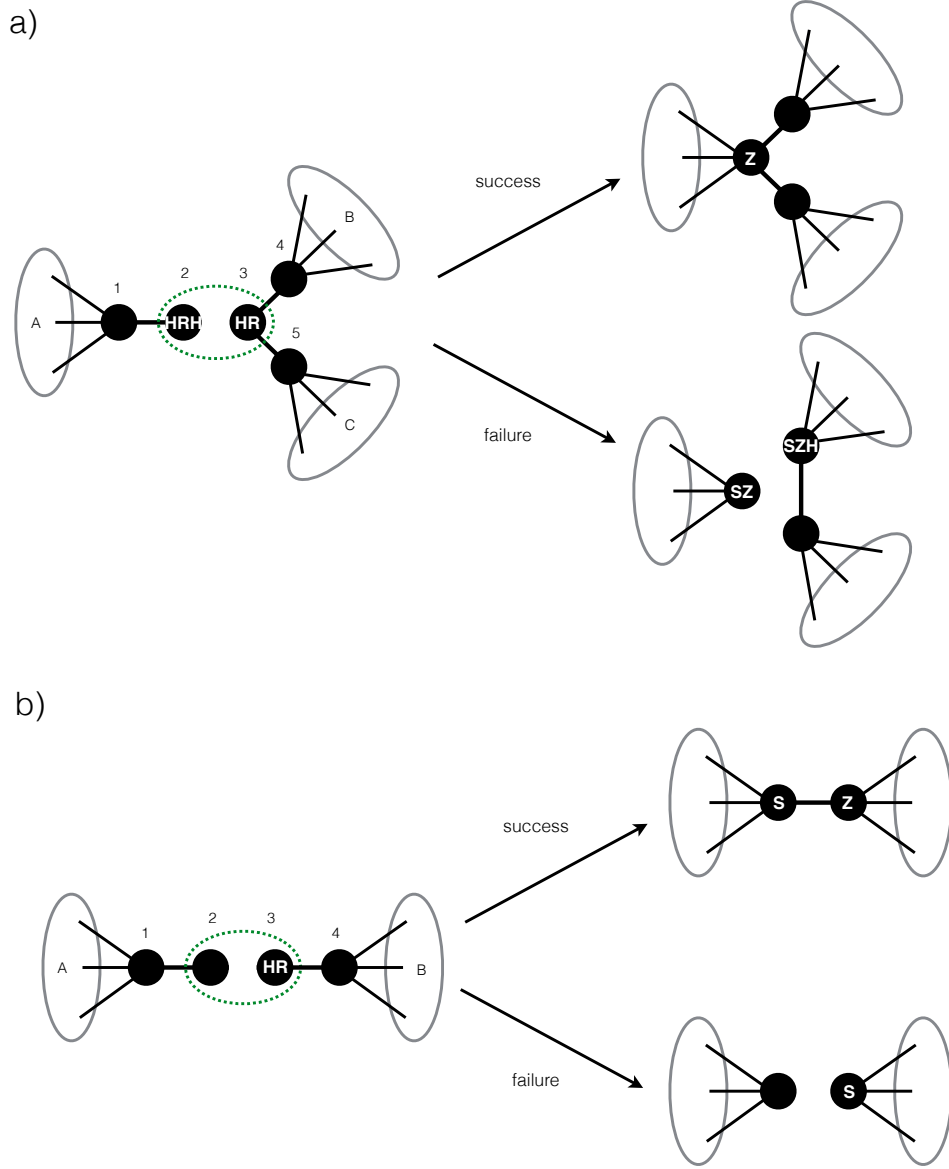


Figure C.1: Fusion operations including local unitaries. a) Fusions within microclusters. Shown are the local rotations required, before the Bell measurement circuit is used. If fusion is successful, the two parts of the graph state are joined, if it is not successful then the B and C sections are instead connected. Shown also are the local correcting unitaries required to recover a graph state. b) The same shown for the fusions between microclusters.)

then become

$$\langle -X_1 Z_4 Z_5 A, Z_1 X_4 B, Z_1 X_5 C \rangle \quad (\text{C.20})$$

a Z_1 rotation then results in the connected graph state desired.

If the measurement is unsuccessful it projects qubits 2 and 3 onto the state $|00\rangle$ (or $|11\rangle$) which is stabilized by $\langle Z_2, Z_3 \rangle$. The updated stabilizer generators in this case are

$$\langle Y_1 A, -Y_4 Z_5 B, -Z_4 Y_5 C \rangle \quad (\text{C.21})$$

performing the local rotations SZ on qubit 1 and SZH on qubit 4, where $S = \sqrt{Z}$, results in

$$\langle X_1 A, X_4 Z_5 B, X_5 Z_4 C \rangle \quad (\text{C.22})$$

So we recover a graph state where a bond has been added between qubits 4 and 5 and qubit 1 remains connected to its neighbours.

For the secondary fusion operations which connect microclusters, we can consider the simple case shown in Fig. C.1b). The initial state can be represented as

$$\langle X_1 Z_2 A, X_2 Z_1, X_3 Z_4, X_4 Z_3 B \rangle \quad (\text{C.23})$$

where A and B are Z 's on all other qubits connected to qubits 1 and 4 respectively. Performing the rotation HR on qubit 3 results in the pre-measurement state

$$\langle X_1 Z_2 A, X_2 Z_1, Y_3 Z_4, -X_4 X_3 B \rangle. \quad (\text{C.24})$$

Successful measurement results in the state given by

$$\langle -Y_1 Z_4 A, -X_4 Z_1 B \rangle \quad (\text{C.25})$$

where $S_1 Z_4$ can be used to produce a graph state with a bond between qubits 1 and 4.

Unsuccessful Bell measurement results in the state

$$\langle X_1 A, -Y_4 B \rangle \quad (\text{C.26})$$

where S_4 recovers a graph state with no bond between qubits 1 and 4.

C.6 Additional measurement operators

First, to evaluate the resultant state when we perform fusion between a qubit from two microclusters, where only one may contain a distinguishable photon, we sum the two cases where one of the ancilla photons on one side of the bell measurement circuit is distinguishable. This measurement is given by the POVM with elements $\{\sum_{j=1}^2 \langle \Psi_j | \mathbf{E}_i | \Psi_j \rangle\}$, $\{\sum_{j=5}^6 \langle \Psi_j | \mathbf{E}_i | \Psi_j \rangle\}$ resulting in the measurement $\{E_{A'}\}$

$$\begin{aligned} E_{A'}^1(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|) + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, & E_{A'}^2(\phi^+) &= \frac{1}{8}|\phi^+\rangle\langle\phi^+| \\ E_{A'}^3(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), & E_{A'}^4(\phi^-) &= \frac{1}{8}|\phi^-\rangle\langle\phi^-| \\ E_{A'}^5(\phi^-) &= \frac{1}{16}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|) + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, & E_{A'}^6(\psi^+) &= \frac{1}{4}|\psi^+\rangle\langle\psi^+| \\ E_{A'}^7(\psi^+) &= \frac{1}{8}|\psi^+\rangle\langle\psi^+| + \frac{1}{4}|00\rangle\langle 00|, & E_{A'}^8(\psi^+) &= \frac{1}{8}|\psi^+\rangle\langle\psi^+| + \frac{1}{4}|11\rangle\langle 11|, \\ E_{A'}^9(\psi^-) &= \frac{1}{2}|\psi^-\rangle\langle\psi^-|, & E_{A'}^{10}(00) &= \frac{3}{16}|00\rangle\langle 00|, & E_{A'}^{11}(11) &= \frac{3}{16}|11\rangle\langle 11| \end{aligned}$$

and $\bar{E}_{A'} = \bar{E}_A = (11|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + 3|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|)/8$. So the average post-fusion state is the same as when we sum over all the ancilla photons.

We now consider the cases where a bad photon from each of two microclusters meet at a fusion measurement. There are three ways in which this can happen; the first case is where two ancilla photons are bad (one on each side). This can be described by the operators $\{\sum_{i=1}^2 \sum_{j=5}^6 \langle \Psi_{i,j} | \mathbf{E}_i | \Psi_{i,j} \rangle\}$ and results in the measurement $\{E_{AA}\}$

$$\begin{aligned} E_{AA}^1(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|) + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, & E_{AA}^2(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|) \\ E_{AA}^3(\phi^-) &= \frac{1}{16}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|) + \frac{1}{8}|\psi^+\rangle\langle\psi^+|, & E_{AA}^4(\psi^+) &= \frac{1}{8}|\psi^+\rangle\langle\psi^+| + \frac{1}{4}|00\rangle\langle 00| \\ E_{AA}^5(\phi^+) &= \frac{1}{2}|\psi^-\rangle\langle\psi^-| \end{aligned}$$

the measurement is successful with a probability of 7/16. $\bar{E}^{AA} = (7|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| +$

$$3|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|)/8.$$

Next is the case where both logical photons are bad. This case involves interference between the bad photons and therefore we must further specify the structure of the input states ϱ . We will consider the two extremal cases where the internal state of the i th input photon ϱ_i is two-dimensional, $\varrho_i^2 = (1-\epsilon)|\xi_0\rangle\langle\xi_0| + \epsilon|\xi_1\rangle\langle\xi_1|$, hence the bad photons interfere perfectly with each other and the high-dimensional case, equivalent to the OBBM, $\varrho_i^\infty = (1-\epsilon)|\xi_0\rangle\langle\xi_0| + \epsilon|\xi_i\rangle\langle\xi_i|$, where the bad photons are orthogonal. These two cases can then be evaluated via $\{\text{Tr}_{\text{int}}[\mathbf{E}_i(\mathbb{1} \otimes |\xi_0\rangle\langle\xi_0|^{\otimes 2} \otimes |\xi_1\rangle\langle\xi_1|^{\otimes 2} \otimes |\xi_0\rangle\langle\xi_0|^{\otimes 2})]\}$, $\{\text{Tr}_{\text{int}}[\mathbf{E}_i(\mathbb{1} \otimes |\xi_0\rangle\langle\xi_0|^{\otimes 2} \otimes |\xi_3\rangle\langle\xi_3| \otimes |\xi_4\rangle\langle\xi_4| \otimes |\xi_0\rangle\langle\xi_0|^{\otimes 2})]\}$, resulting in the measurements $\{E_{LL,\varrho^2}\}, \{E_{LL,\varrho^\infty}\}$

$$\begin{aligned} E_{LL,\varrho^2}^1(\phi^+) &= \frac{1}{4}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), & E_{LL,\varrho^2}^2(\phi^-) &= \frac{1}{4}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|) \\ E_{LL,\varrho^2}^3(\psi^+) &= |\psi^+\rangle\langle\psi^+|, & E_{LL,\varrho^2}^4(\psi^-) &= |\psi^-\rangle\langle\psi^-| \end{aligned}$$

the measurement is successful with a probability of $3/4$. $\bar{E}_{LL,\varrho^2} = (5|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|)/2$.

$$\begin{aligned} E_{LL,\varrho^\infty}^1(\phi^+) &= \frac{1}{8}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), & E_{LL,\varrho^\infty}^2(\phi^-) &= \frac{1}{8}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|) \\ E_{LL,\varrho^\infty}^3(\psi^+) &= \frac{1}{2}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|), & E_{LL,\varrho^\infty}^4(\psi^-) &= \frac{1}{2}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) \end{aligned}$$

the measurement is successful with a probability of $5/8$. $\bar{E}_{LL,\varrho^\infty} = \bar{E}^L = 5/4(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|)$.

Finally we have the case where one ancilla and one logical photon (from opposite sides) are bad. Again, we can calculate the POVMs corresponding to the low and high dimensional extremes $\{E_{AL,\varrho^2}\}, \{E_{AL,\varrho^\infty}\}$

$$\begin{aligned}
 E_{AL,\varrho^2}^1(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), \\
 E_{AL,\varrho^2}^2(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^2}^3(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + 2|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^2}^4(\phi^-) &= \frac{1}{32}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|), \\
 E_{AL,\varrho^2}^5(\phi^-) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + 2|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^2}^6(\phi^-) &= \frac{1}{32}(|\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+|), \\
 E_{AL,\varrho^2}^7(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{16}|01\rangle\langle 01|, \\
 E_{AL,\varrho^2}^8(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{16}|10\rangle\langle 10|, \\
 E_{AL,\varrho^2}^9(\psi^+) &= \frac{1}{16}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{8}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{10}(\psi^+) &= \frac{1}{16}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{8}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{11}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|01\rangle\langle 01|) + \frac{3}{64}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{12}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|10\rangle\langle 10|) + \frac{3}{64}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{13}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|01\rangle\langle 01|) + \frac{3}{64}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{14}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|10\rangle\langle 10|) + \frac{3}{64}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{15}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|01\rangle\langle 01|) + \frac{1}{64}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{16}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|10\rangle\langle 10|) + \frac{1}{64}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{17}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|01\rangle\langle 01|) + \frac{1}{64}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{18}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 3|10\rangle\langle 10|) + \frac{1}{64}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{19}(\psi^-) &= \frac{1}{16}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|), \\
 E_{AL,\varrho^2}^{20}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{32}|01\rangle\langle 01|, \\
 E_{AL,\varrho^2}^{21}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{32}|10\rangle\langle 10|, \\
 E_{AL,\varrho^2}^{22}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{32}|00\rangle\langle 00|,
 \end{aligned}$$

$$\begin{aligned}
 E_{AL,\varrho^2}^{23}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{1}{32}|11\rangle\langle 11|, \\
 E_{AL,\varrho^2}^{24}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{3}{32}|00\rangle\langle 00|, \\
 E_{AL,\varrho^2}^{25}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+|) + \frac{3}{32}|11\rangle\langle 11|
 \end{aligned}$$

the measurement is successful with a probability of $9/16$. $\bar{E}_{AL,\varrho^2} = (9|\phi^+\rangle\langle\phi^+| + 11|\phi^-\rangle\langle\phi^-| + 9|\psi^+\rangle\langle\psi^+| + 7|\psi^-\rangle\langle\psi^-|)/16$.

$$\begin{aligned}
 E_{AL,\varrho^\infty}^1(\phi^+) &= \frac{1}{32}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|), \\
 E_{AL,\varrho^\infty}^2(\phi^+) &= \frac{1}{64}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^\infty}^3(\phi^+) &= \frac{1}{64}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + 3|\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^\infty}^4(\phi^-) &= \frac{1}{32}(|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|), \\
 E_{AL,\varrho^\infty}^5(\phi^-) &= \frac{1}{64}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + 3|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^\infty}^6(\phi^-) &= \frac{1}{64}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| + 3|\psi^-\rangle\langle\psi^-|), \\
 E_{AL,\varrho^\infty}^7(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{16}|01\rangle\langle 01|, \\
 E_{AL,\varrho^\infty}^8(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|) + \frac{1}{16}|10\rangle\langle 10|, \\
 E_{AL,\varrho^\infty}^9(\psi^+) &= \frac{1}{64}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + |10\rangle\langle 10| + 3|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{10}(\psi^+) &= \frac{1}{64}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + |01\rangle\langle 01| + 3|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{11}(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + \frac{1}{64}(|10\rangle\langle 10| + 5|11\rangle\langle 11|)), \\
 E_{AL,\varrho^\infty}^{12}(\psi^+) &= \frac{1}{32}(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + \frac{1}{64}(|01\rangle\langle 01| + 5|00\rangle\langle 00|)), \\
 E_{AL,\varrho^\infty}^{13}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|10\rangle\langle 10| + 9|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{14}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|01\rangle\langle 01| + 9|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{15}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|10\rangle\langle 10| + 7|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{16}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|01\rangle\langle 01| + 7|11\rangle\langle 11|),
 \end{aligned}$$

$$\begin{aligned}
 E_{AL,\varrho^\infty}^{17}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|10\rangle\langle 10| + 5|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{18}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|01\rangle\langle 01| + 5|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{19}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|10\rangle\langle 10| + 11|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{20}(\psi^-) &= \frac{1}{256}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|01\rangle\langle 01| + 11|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{21}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 4|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{22}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 4|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{23}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|00\rangle\langle 00|), \\
 E_{AL,\varrho^\infty}^{24}(\psi^-) &= \frac{1}{64}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 2|11\rangle\langle 11|), \\
 E_{AL,\varrho^\infty}^{25}(\psi^-) &= \frac{15}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 5|01\rangle\langle 01|), \\
 E_{AL,\varrho^\infty}^{26}(\psi^-) &= \frac{15}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 5|10\rangle\langle 10|), \\
 E_{AL,\varrho^\infty}^{27}(\psi^-) &= \frac{2}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 7|01\rangle\langle 01|), \\
 E_{AL,\varrho^\infty}^{28}(\psi^-) &= \frac{2}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 7|10\rangle\langle 10|), \\
 E_{AL,\varrho^\infty}^{29}(\psi^-) &= \frac{7}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + |01\rangle\langle 01|), \\
 E_{AL,\varrho^\infty}^{30}(\psi^-) &= \frac{7}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + |10\rangle\langle 10|), \\
 E_{AL,\varrho^\infty}^{31}(\psi^-) &= \frac{1}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 5|01\rangle\langle 01|), \\
 E_{AL,\varrho^\infty}^{32}(\psi^-) &= \frac{1}{512}(|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| + 5|10\rangle\langle 10|).
 \end{aligned}$$

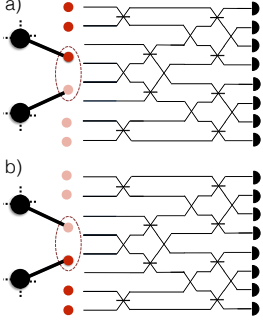
the measurement is successful with a probability of $9/16$. $\bar{E}_{AL,\varrho^\infty} = (5|\phi^+\rangle\langle\phi^+| + 5|\phi^-\rangle\langle\phi^-| + 4|\psi^+\rangle\langle\psi^+| + 4|\psi^-\rangle\langle\psi^-|)/8$.

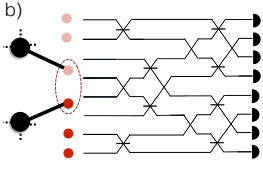
C.7 Second order fusion error maps

Here we investigate the difference between the state produced by putting a distinguishable photon in each input state which produces a final lattice qubit and the case where we simply compose the error map resulting from a single distinguishable photon at each lattice

site. The only place where these models differ when considering errors on neighbouring lattice sites is when the two distinguishable photons meet in a fusion measurement.

Fig. C.2 shows the error maps Λ_G^1, Λ_G^2 recovered from a model where the input state Eq. (6.22) is applied to the three photons on each side of the fusion measurement. In addition we then calculate the full error map in the case where we take second-order errors into account, i.e. we use Eq. (6.22) applied to all six photons. To analyse this case we require the extra measurement operators found in appendix C.6. We use the POVMs $\bar{E}_{AL,\rho^\infty}$ and $\bar{E}_{LL,\rho^\infty}$, corresponding to the OBB model as this will provide the largest discrepancy in the two output maps. This second-order map can be compared to the composition of the first-order model for each input qubit as shown in Fig. C.2. We find that the probability of second-order logical error is only slightly modified. Furthermore, as described in the main body, since both models produce the same error maps to $\mathcal{O}(\epsilon)$ and higher order terms do not result in higher weight errors, these differences may be reasonably neglected.

a)  $\Lambda_G^1 \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) = (1-\epsilon)^3 \begin{array}{c} \bullet \\ \bullet \end{array} + \epsilon(1-\epsilon)^2 \left\{ \frac{4}{3} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{1}{2} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{1}{4} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{1}{4} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$

b)  $\Lambda_G^2 \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) = (1-\epsilon)^3 \begin{array}{c} \bullet \\ \bullet \end{array} + \epsilon(1-\epsilon)^2 \left\{ \frac{4}{3} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{1}{2} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{1}{4} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{1}{4} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$

c) $\Lambda_G^2 \circ \Lambda_G^1 \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) = (1-\epsilon)^6 \begin{array}{c} \bullet \\ \bullet \end{array} + \epsilon(1-\epsilon)^5 \left\{ \frac{8}{3} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{3}{4} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{3}{4} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{1}{2} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$
 $+ \epsilon^2(1-\epsilon)^4 \left\{ \frac{301}{144} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{19}{16} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{19}{16} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{47}{48} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$

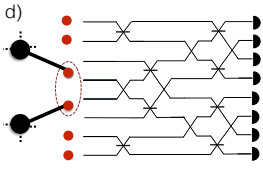
d)  $\Lambda_G^{12} \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) = (1-\epsilon)^6 \begin{array}{c} \bullet \\ \bullet \end{array} + \epsilon(1-\epsilon)^5 \left\{ \frac{8}{3} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{3}{4} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{3}{4} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{1}{2} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$
 $+ \epsilon^2(1-\epsilon)^4 \left\{ \frac{53}{24} \begin{array}{c} \bullet \\ \bullet \end{array} + \frac{31}{24} \begin{array}{c} \text{Z} \\ \bullet \end{array} + \frac{31}{24} \begin{array}{c} \bullet \\ \text{Z} \end{array} + \frac{11}{8} \begin{array}{c} \text{Z} \\ \text{Z} \end{array} \right\}$

Figure C.2: Error maps for second order errors in fusion measurements. a) Right: the effective error map on the remaining qubits when one of the upper three photons (dark red) shown on left may be distinguishable. b) Right: the effective error map when one of the lower three photons (dark red) shown on left may be distinguishable. c) The error map obtained by composing those from a) and b). d) Right: the effective error map when one of the upper three photons and one of the lower three photons shown on left may be distinguishable.

Appendix D

Mitigating Distinguishability Errors

The logical state produced by the B_6 Bell state generator with an RSM input, $\varrho^{\otimes 4}$, is:

$$\rho_L^{\varrho} = \frac{1}{216} \times \begin{pmatrix} (5 + 8\text{Tr}[\varrho^3] - 5\mathcal{V}) & -\sqrt{3}(1 + 2\text{Tr}[\varrho^3] - 3\mathcal{V}) & \frac{1}{\sqrt{3}}(5 + 18\text{Tr}[\varrho^3] - 14\text{Tr}[\varrho^4] - 11\mathcal{V} + 2\mathcal{V}^2) & (-1 - 4\text{Tr}[\varrho^3] + 8\text{Tr}[\varrho^4] + 5\mathcal{V}) \\ -\sqrt{3}(1 + 2\text{Tr}[\varrho^3] - 3\mathcal{V}) & 3(1 - \mathcal{V}) & (-1 - 4\text{Tr}[\varrho^3] + 2\text{Tr}[\varrho^4] + 5\mathcal{V} - 2\mathcal{V}^2) & \sqrt{3}(1 + 2\text{Tr}[\varrho^3] - 3\mathcal{V}) \\ \frac{1}{\sqrt{3}}(5 + 18\text{Tr}[\varrho^3] - 14\text{Tr}[\varrho^4] - 11\mathcal{V} + 2\mathcal{V}^2) & (-1 - 4\text{Tr}[\varrho^3] + 2\text{Tr}[\varrho^4] + 5\mathcal{V} - 2\mathcal{V}^2) & \frac{1}{3}(17 + 48\text{Tr}[\varrho^3] - 36\text{Tr}[\varrho^4] - 41\mathcal{V} + 12\mathcal{V}^2) & \frac{-1}{\sqrt{3}}(5 + 18\text{Tr}[\varrho^3] - 14\text{Tr}[\varrho^4] - 11\mathcal{V} + 2\mathcal{V}^2) \\ (-1 - 4\text{Tr}[\varrho^3] + 8\text{Tr}[\varrho^4] + 5\mathcal{V}) & \sqrt{3}(1 + 2\text{Tr}[\varrho^3] - 3\mathcal{V}) & \frac{-1}{\sqrt{3}}(5 + 18\text{Tr}[\varrho^3] - 14\text{Tr}[\varrho^4] - 11\mathcal{V} + 2\mathcal{V}^2) & (5 + 8\text{Tr}[\varrho^3] - 5\mathcal{V}) \end{pmatrix} \quad (\text{D.1})$$